



NTNU – Trondheim
Norwegian University of
Science and Technology

A new algorithm for remote sensing mesopause temperatures using the hydroxyl airglow

Endre Asheim Hennem

Master of Science in Physics and Mathematics

Submission date: June 2013

Supervisor: Patrick Joseph Espy, IFY

Norwegian University of Science and Technology
Department of Physics

Acknowledgements

This report is the product of TFY4900, Master Thesis in Physics at the Norwegian University of Science and Technology. I would like to thank Prof. Patrick J. Espy for sharing his knowledge on signal processing, instrumentation and atmospheric physics. I would also like to thank Daniel Skåre, the staff engineer at section for applied physics, for technical help with the instrument.

Abstract

Ground based measurements of the hydroxyl airglow started in Trondheim (63.4° N, 10.4° E) 20. November 2012 using a linear InGaAs PDA detector covering the \sim 1450-1650 nm wavelength range. The instrument have successfully recorded data with a time resolution of 1 minute through the whole winter until 27th April 2013. A complete analysis scheme has been developed to derive rotational temperatures from the (3-1) and (4-2) P-branch transitions. The method is based on least squares fitting a full synthetic model of the OH bands to the recorded spectrums.

The analysis scheme has been used to analyse data for the whole measurement period during winter 2012/2013. Average temperatures for the period November-March are found to be 209.1 K and 207.6 K for the (3-1) and (4-2) band respectively. The nightly averaged temperatures generally follow the expected climatological trend of warm winter temperatures and cooling towards summer. In January, the temperatures rise to \sim 230 K, which is a large deviation from predictions calculated with the MSISE model, but same pattern is also observed over Andøya (69.3 °N). The difference of 1.5 K between the (3-1) and (4-2) temperature is consistent with model predictions of the intensity distribution of the $v = 4$ vibrational state being higher in altitude and therefore colder.

Results from three nights have been analysed in more detail, revealing large variations in the OH airglow layer. The observed oscillations have periods ranging from \sim 30 minutes to several hours, and variations in rotational temperature up 47 K. Relative changes in the intensity is found to be typically 4-5 times larger than in the temperature. Both the temperature and the intensity show a high level of covariation between the (3-1) and (4-2) bands, but the results are generally more accurate for the (3-1) transition, because the (4-2) P-branch is affected by a lower instrument sensitivity. Relative errors in the (3-1) temperature is typically between 1-2%.

A detailed analysis of the analysis scheme is given, and some unresolved issues and possible improvements are suggested. The instrument and it's operation is also described.

Sammendrag

Bakkebaserte nattlige målinger av hydroxyl airglow startet i Trondheim 20. November 2012. Instrumentet som brukes er en lineær InGaAs fotodiode-detektor som er konfigurert til å dekke bølgelengdeområdet 1450-1650 nm. Målinger gjennom vinteren inntil 27. april er gjennomført med en tidsoppløsning på 1 min. En komplett analysemetode for å beregne rotasjonstemperaturer fra datene er utviklet. Metoden er basert på å tilpasse en full syntetisk modell, basert på minste kvadraters prinsipp, til de målte spektrumene.

Analysemetoden har blitt brukt til å analysere data for hele målingsperioden vinteren 2012/2013. Gjennomsnittelige verider for perioden november-mars er 209.1 K og 207.6 K for henholdsvis (3-1) og (4-2) overgangene. Nattlige temperaturgjennomsnitt følger generelt den forventede trenden med varme vintertemperaturer og kaldere temperaturer mot sommeren. I midten av januar stiger temperaturen til ~ 230 K, som er en stor forskjell fra prediksjonene beregnet med MSISE modellen, men det samme mønsteret er observert over Andøya (69.3 °N). Forskjellen på 1.5 K mellom (3-1) og (4-2) temperaturene er konsistent med at modellberegninger av intensitetsfordelingen av $v = 4$ vibrasjonsnivået ligger høyere i atmosfæren og dermed er kaldere.

Mer detaljerte analyser av tre netter i målingsperioden, avdekker store variasjoner i OH airglow laget. De observerte variasjonene har perioder fra 30 minutter og opptil flere timer, og variasjoner i rotasjonstemperaturen på opp til 47 K. Relative endringer i intensiteten er funnet til å være 4-5 ganger høyere enn i temperaturen. Både temperaturen og intensiteten viser en høy grad av samvariasjon for de to båndene, men resultatene er generelt dårligere for (4-2) båndene, fordi P-greinen er påvirket av en dårligere instrumentsensitivitet i bølgelengdeområdet. Relative feilestimer i (3-1) temperaturen ligger typisk mellom 1-2%.

Analyseskjemaet er beskrevet i detalj, og noen uløste problemer og mulige forbedringer er foreslått. Konfigurasjonen av instrumentet er også beskrevet.

Contents

1	Introduction	5
2	Theory	7
2.1	Atmospheric structure	7
2.2	The hydroxyl layer	9
2.3	Atmospheric gravity waves	10
2.4	Short introduction to spectroscopy of diatomic molecules. . .	11
2.4.1	Rotational spectrum	11
2.4.2	Vibrational spectrum	13
2.4.3	Rotational vibrational spectrum	14
2.4.4	Angular momentum of the electrons	15
2.5	The Meinel lines	15
2.6	Thermal distribution of quantum states	18
2.7	Intensities of rotational lines	20
2.8	OH airglow response to gravity waves	23
3	Instrument	24
3.1	Spectrometer	24
3.2	Calibration and noise	25
3.3	Operation	27
3.4	Instrument slit function	28
4	Data analysis	31
4.1	Dark signal	32
4.2	Baseline removal	33
4.3	Gauss-Newton algorithm	35
4.4	Cloud level indicator	37
4.5	Data products	40
4.6	Computer algorithm	41
4.6.1	The main routine	42
4.6.2	Result of the Gauss-Newton iterations	46
4.6.3	Calculating nocturnal statistics	50

5	Results	52
5.1	Nightly averaged temperatures and intensities	52
5.2	Errors	55
5.3	Gravity waves	57
5.3.1	29-30th November 2012	57
5.3.2	7-8th February 2013	60
5.3.3	19-20th February	63
5.3.4	Variance of wave activity	66
6	Discussion	67
6.1	The results	67
6.2	The analysis scheme	69
7	Conclusion	73
A	Tables	78
B	Plots	79
C	Code	85
C.1	Main routine: calc_ohdata	85
C.2	get_temp_subbg	89

Chapter 1

Introduction

Airglow is a photochemical luminescence originating from the upper atmosphere confined to a layer between 80-120 km. The hydroxyl airglow emissions was identified by A.B Meinel in 1950 [11]. Since then, this phenomena has been used to investigate many aspects of the mesopause [18]. The mesopause is ranging from 85-95 km above ground (midlatitude). This region of the atmosphere is too high to be reached by balloons and too low to be effectively sensed by satellites. Ground based remote sensing is therefore the optimal method to study this part of the atmosphere.

Gravity waves generated in the troposphere propagate all the way up to the upper and middle atmosphere, where they eventually deposit their energy. The interaction between the waves and the atmosphere modulate the airglow emissions, and therefore, gravity waves can be detected by continuously monitoring the airglow. It is now understood that most of the energy which is driving the general circulation in the atmosphere, is originating in the troposphere, and most of the energy is transported by gravity waves [12]. The short scale of important parts of the wave spectrum makes them difficult to incorporate in climate models. To be able to include the important effect of gravity waves in future models, more knowledge about the links between the energy flux by gravity waves and the dynamics of the middle and upper atmosphere is needed.

There is growing evidence for a cooling of the mesosphere, caused by a warmer climate on the earth. However, the mesosphere is a region of the atmosphere governed by large naturally variations, which can hide long time scale variations such as climate change. Longer term measurements of the mesospheric temperature is needed to reveal any additional climate effects.

This project is a step towards getting more knowledge about the mesosphere by collecting data which can be used for both short and long term analysis. There is now a network of ground-based airglow observations associated with the Network for the Detection of Mesospheric change. Comparison of observations from Trondheim with observations from all over the

world, may lead to a better global picture of the impact of gravity waves and the change in mesospheric temperature.

Chapter 2

Theory

2.1 Atmospheric structure

The atmosphere is divided into different layers with respect to their temperature gradient (figure 2.1). From the earth's surface, the first 13-15 km of the atmosphere is the troposphere. This is where most weather phenomena occur, as precipitation, thunder storms and clouds. The temperature decreases with height, because the distance to the heat source, the earth, increases. In the region between 15-20 km, which is called the tropopause, the decline in temperature stops at a local minimum of about $-60\text{ }^{\circ}\text{C}$. From 20 km and up to 45 km is the stratosphere. In this region, the vast majority of the earth's ozone is contained. Molecular oxygen and ozone are dissociated by ultraviolet radiation from the sun. The excess energy in these photochemical reactions results in heating, and the temperature continues to rise up to the stratopause, which is located between $\sim 45\text{-}50\text{ km}$. The next region is the mesosphere. Here, the temperature gradient is negative because carbon dioxide, and other infrared active constituents, radiates away the heat. Temperatures in the mesosphere can reach as low as $-130\text{ }^{\circ}\text{C}$ and is the naturally coldest place in the atmosphere. The mesopause starts at $\sim 85\text{ km}$ and above is the thermosphere. In this region, the air is heated by the sun directly because the density is so low that collision times for the dissociated molecules are very long. Temperatures can reach 1000 K , but the heat content is very low because of the low density.

Variation with latitude and season in atmospheric temperature profiles arises because of the varying incoming solar radiation. In the summer, the troposphere and stratosphere are logically warmer than in winter, but this is not true for the mesosphere. The temperature in the middle atmosphere is determined by a large scale circulation from the summer pole to the winter pole, which is wave driven [1][5]. Waves generated at low latitudes propagate upward in the atmosphere with growing amplitudes as they encounter lower

and lower density. Eventually, they break and the deposit momentum. The friction force which arises, is anti-parallel to the zonal winds, and they are therefore slowed by the breaking waves. The result is that the middle atmosphere is driven out of radiative equilibrium. Winds are kicked polewards in the winter, and away from the pole in summer, and creates a flow from the summer pole which expands and cools, and to the winter pole, which compresses and heats.

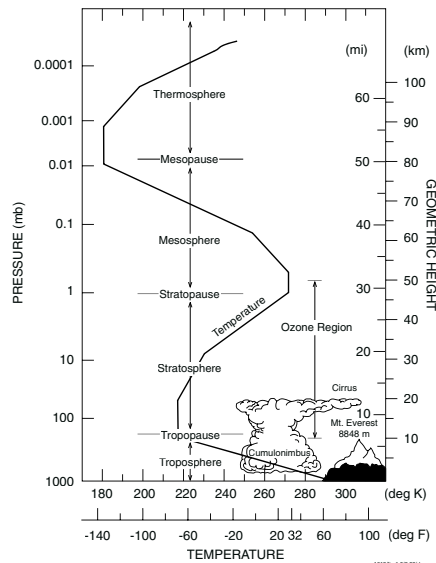


Figure 2.1: Atmospheric temperature profile. Source: [5].

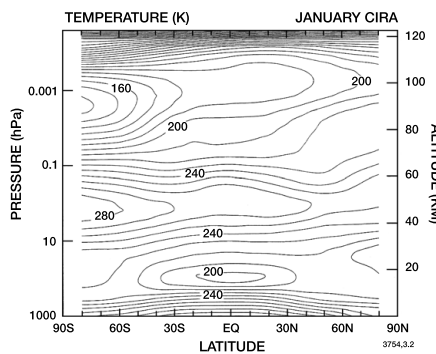


Figure 2.2: Zonally averaged temperatures for the Earth in January up to 120 km. The summer pole mesosphere (south) is the coldest place in the atmosphere with temperatures reaching 120-140 K. The winter pole (north) is much warmer than in the summer, reaching temperatures up to 220-230 K (latitude dependent). Source: [5].

2.2 The hydroxyl layer

The OH layer is a thin layer of hydroxyl radicals located in the mesopause in the upper atmosphere. During nighttime, the concentration peaks at 87 ± 3 km and has a half width of 7 ± 3 km [2]. The main source of excited hydroxyl molecule is the reaction [18]



At night, ozone is produced by recombination with atomic oxygen and losses are through reactions with atomic hydrogen [15]. Balance between loss and production gives that the ozone concentration is given by

$$[\text{O}_3] = \frac{k_{\text{O}+\text{O}_2+\text{M}}[\text{O}][\text{O}_3][\text{M}]}{k_{\text{H}+\text{O}_3}[\text{H}]}. \quad (2.2)$$

From (2.1), the production rate of OH^* is

$$P = k_{\text{H}+\text{O}_3}[\text{H}][\text{O}_3], \quad (2.3)$$

and upon substitution of $[\text{O}_3]$ into this equation, one sees that the production of OH^* is directly proportional to the atomic oxygen concentration.

At daytime, ozone is destroyed by sunlight and disassociates according to the reaction



O_2 and O quickly recombine to form O_3 after the sun has set leading to a very high ozone concentration. The production of OH^* (2.1) and the intensity of the OH^* emissions therefore often peaks in the beginning of the night if the layer is not perturbed. A typical OH nightly variation is shown in figure 2.3

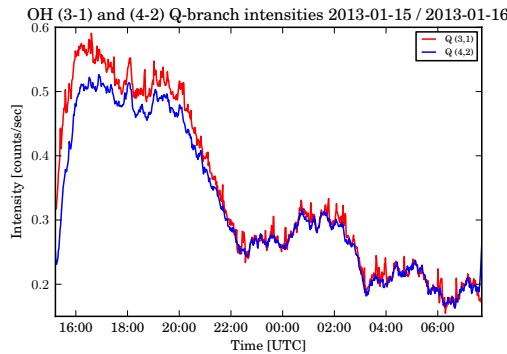


Figure 2.3: Typical OH intensity variation during a night. The intensity is often maximum in the beginning of the night because the ozone production rate is high right after sunset.

2.3 Atmospheric gravity waves

Gravity waves in the atmosphere are analogous to waves on the water surface. Water is denser than air, so the air in a wave trough is lighter than the surrounding water and will experience a buoyant force. Water in a wave crest is heavier than the surrounding air and will tend to fall. Waves in the atmosphere arises in a similar manner. The density of air is dependent on pressure and temperature. Temporal and spatial variations in these parameters will create density differences which the force of gravity will try to smooth out. While waves on the water surface can be characterized by a single horizontal wavelength, atmospheric gravity waves will have a horizontal wavelength and a vertical wavelength.

There are many sources of gravity waves in the atmosphere. Waves generated by air flow over mountains, cities and valleys are referred to as *mountain waves* or *lee waves*. Gravity waves can also be generated by convective activity in the atmosphere and are often observed over thunderstorms [12]. The

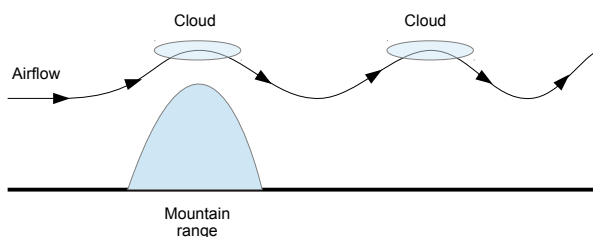


Figure 2.4: Simple schematic figure of a mountain wave and lee wave. Clouds may form if the water vapor condenses in the rising air. The water evaporates as the air descends and the clouds dissolve [1].

simplest possible way to illustrate gravity waves is by considering uniform flow over a surface corrugation. On the windward side of the corrugation, the air is pushed upwards. The opposite will happen on the lee side, where the air is accelerated downwards. Because the air is slowed down on the windward side, the pressure will increase. On the lee side, the pressure decreases. The result is vertical columns of air oscillating up and down. In the column of upward moving air, the air will cool down because of adiabatic expansion. The opposite is true for the downwards moving column. The air is compressed and heated as it is displaced downwards.

The wave pattern is stationary with respect to the ground, but if you were sitting in a warm air balloon and blowing with the wind speed over a corrugated land surface, you would experience waves just as you would, sitting in a boat in the ocean and feeling the waves pass by.

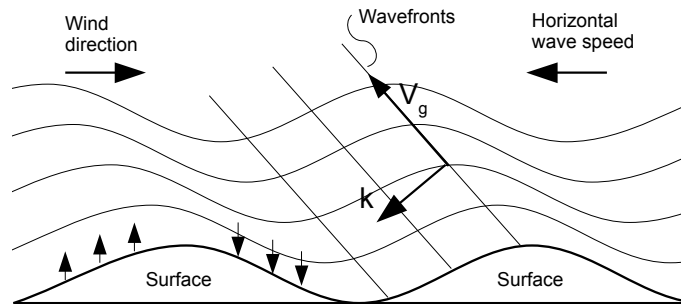


Figure 2.5: Schematic illustration of how gravity waves are generated by a surface corrugation. On the windward side, the air is slowed down and pushed upwards. On the lee side, the air is accelerated downwards. The horizontal wave speed will be equal to wind speed, but in opposite direction.

2.4 Short introduction to spectroscopy of diatomic molecules.

This section will give the reader a basic understanding of the molecular spectroscopy of diatomic spectroscopy in order to understand the principles behind the measurements in this project. It is heavily based on [10] but the most of the results presented here can be found in any standard textbook on quantum mechanics such as [4].

The energy of a molecule can roughly be divided into three parts

$$E = E_e + E_v + E_r, \quad (2.5)$$

where E_e is the total binding energy of the electrons to the nuclei, E_v is the energy associated with the vibrational modes of the molecule, and E_r is the rotational energy. For ordinary macroscopical objects, the principles of classical mechanics describes the system with indefinite precision. But for a small object, such as a molecule, one must treat the system quantum mechanically. This means that the energy levels of the system must be found by solving the Schrödinger equation. In quantum mechanics, the energy levels of a bound system cannot take any values but are discrete.

2.4.1 Rotational spectrum

Classically, one can picture the rotation of a diatomic molecule as two masses, m_1 and m_2 , linked together at a distance r . This system is called the rigid rotator. The rotational energy for this system is

$$E_r = \frac{1}{2}I\omega^2,$$

or expressed in terms of angular momentum $P = I\omega$, where I is the moment of inertia and ω is the angular frequency

$$E_r = \frac{P^2}{2I}. \quad (2.6)$$

This system can also be formulated in terms of a one mass system by introducing the *reduced mass*:

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \quad (2.7)$$

The energy levels given by quantum mechanics for this system is

$$E_r = \frac{h^2}{8\pi I} J(J+1). \quad (2.8)$$

J is called the rotational quantum number. These energy levels are *degenerate*, there are $2J+1$ states with the energy E_r . If one divide equation 2.8 with hc , it has the units of cm^{-1} . These are called the term values:

$$F(J) = \frac{E_r}{hc} J(J+1) \quad (2.9)$$

The prefactor is the rotational constant, defined as

$$B = \frac{h}{8\pi c I}. \quad (2.10)$$

It is already mentioned that a molecule also can vibrate, and therefore a diatomic molecule is not strictly speaking a rigid rotator. A correction term must be added because the centrifugal force is stretching the molecule, and the term values becomes

$$F(J) = BJ(J+1) + DJ^2(J+1)^2 + \dots \quad (2.11)$$

where D is

$$D = \frac{4B^3}{\omega^2}. \quad (2.12)$$

From classical electrodynamics, radiation of light is associated with a changing electric dipole moment. For two unlike atoms, there will always be a permanent dipole moment with direction perpendicular to the axis of rotation. In a molecule consisting of two like atoms, the center of positive and negative charges will coincide, and therefore there will be no permanent dipole moment, and therefore also no rotational spectrum. According to quantum theory, absorption or emission of a photon can happen only through transition between the rotational energy levels. The wavenumber of the absorbed or emitted photon will be the the difference between the energy levels in a transition between an upper state ($'$) and a lower state ($''$)

$$\nu = \frac{E'}{hc} - \frac{E''}{hc} = F(J') - F(J''). \quad (2.13)$$

The change in rotational quantum number in the transition, ΔJ , cannot take any value but are restricted to $\Delta J = \pm 1$.

2.4.2 Vibrational spectrum

The simplest model of the vibration of a diatomic molecule is the harmonic oscillator. As for the rigid rotator, it can be reduced to a one mass system with mass μ given by equation (2.7). Classically, the energy of this system is the sum of the kinetic energy and the potential energy

$$E_v = \frac{1}{2}\mu\dot{x}^2 + \frac{1}{2}k(r - r_e)^2 \quad (2.14)$$

Where k is the force constant and $(r - r_e)$ is the displacement from the equilibrium position. Inserting the potential into the Schrödinger equation and solving for the energy levels, the result is

$$E_v = h\nu(v + \frac{1}{2}). \quad (2.15)$$

Here ν is the frequency of the oscillation and v is the vibrational quantum number. These energy levels are non-degenerate. There is only one state with energy $E_v(v)$. Again, after division with hc , (2.15) can be written as term values:

$$G(v) = \omega(v + \frac{1}{2}). \quad (2.16)$$

The factor ω is the vibrational frequency ν/c with the units cm^{-1} .

A diatomic molecule is not exactly a harmonic oscillator. The "spring force" which arises because of the attraction and repulsion around the equilibrium spacing of the atoms, does not increase indefinitely with increasing displacement from equilibrium, but reaches a constant value. This is called an *anharmonic oscillator*. The potential can be approximated by adding higher order terms to the quadratic classical potential

$$U = f(r - r_e)^3 - g(r - r_e)^3 + \dots \quad (2.17)$$

When $g \ll f$ (small anharmonicity) the term values is approximately

$$G_0(v) = \omega_0 v - \omega_0 x_0 v^2 + \omega_0 y_0 v^3 - \dots, \quad (2.18)$$

with the point of zero energy equal to $v=0$. x_0 and y_0 is generally much smaller than one. In contrast to the quadratic potential, the spacing of the energy levels will decrease with increasing v .

If the diatomic molecule has a dipole moment, it will change if the internuclear distance changes. If the dipole moment changes linearly with the internuclear distance, classical electrodynamics gives that a photon with frequency equal to the mechanical vibration frequency will be emitted or absorbed by the molecule. The wavenumber of the photon will be

$$\nu = \frac{E(v')}{hc} - \frac{E(v'')}{hc} = G_0(v') - G_0(v''). \quad (2.19)$$

For the harmonic oscillator, the allowed energy transitions are only those with $\Delta v = \pm 1$. For the anharmonic oscillator transitions with $\Delta v = \pm 2, \pm 3 \dots$ also appear, but the intensity will rapidly decrease with increasing Δv .

2.4.3 Rotational vibrational spectrum

Both rotation and vibration of a diatomic molecule will take place at the same time. This means that a transition between energy states of the molecule is associated with a Δv and a ΔJ . The fact that the rotational energy levels are much smaller than the vibrational energy levels leads to what is called "fine structure splitting" of the vibrational levels and therefore the *rotational vibrational spectrum* arises in the infrared.

When treating the vibration and the rotation of a molecule simultaneously, one must take into account the interaction between the vibration and the rotational motion. Since the internuclear distance is changing when the molecule is vibrating, the moment of inertia will change. Hence, the vibrational constant B will be different for each vibrational level v . It can be shown that the rotational constant for a vibrational state B_v can be expressed

$$B_v = B_e - \alpha_e \left(v + \frac{1}{2} \right) + \dots \quad (2.20)$$

where B_e is the rotational constant at equilibrium separation and α_e is a constant much smaller than B_e . Also, the constant representing the correction from the centrifugal force will be different at each vibrational level. Similarly, it becomes

$$D_v = D_e + \beta_e \left(v + \frac{1}{2} \right) + \dots \quad (2.21)$$

where again, β_e is much smaller than D_e .

The interaction between rotation and vibration does not influence the selection rules. As before, ΔJ can only be unity while Δv can take any integer value. If $\Delta v = 0$, the spectrum is referred to as the *pure rotational* spectrum. Adding the vibrational and the rotational part using (2.9), the wavenumber of a transition from an upper state J' to a lower state J'' is

$$\nu = \nu_0 + B'_v J'(J' + 1) - B''_v J''(J'' + 1). \quad (2.22)$$

Where ν_0 is the wavenumber of the *pure vibrational* ($\Delta J = 0$) transition given by (2.19).

The fact that ΔJ can be 0 or ± 1 leads to *branching* of the transitions. $\Delta J = 0$ is the Q-branch. By setting $J' = J'' = J$ and rearranging the terms, the wavenumber becomes

$$\nu_Q = \nu_0 + (B'_v - B''_v)J + (B'_v - B''_v)J^2. \quad (2.23)$$

If $\Delta J = 1$ the transitions are called the R-branch. Setting $J'' = J$ and $J' = J + 1$, the wave number is

$$\nu_R = \nu_0 + 2B'_v + (3B'_v - B''_v)J + (B'_v - B''_v)J^2. \quad (2.24)$$

Finally $\Delta J = -1$ gives the P-branch. Now, $J'' = J$ and $J' = J - 1$, the wave number becomes

$$\nu_P = \nu_0 - (B'_v + B''_v)J + (B'_v - B''_v)J^2. \quad (2.25)$$

Since $B''_v > B'_v$, ν_Q will always decrease with increasing J , but the $B'_v - B''_v$ is a small factor and the resulting lines in this branch will be close. The Q-branch is more complex, where the wave number increases for small J but decreases for large J . ν_P will always decrease for increasing J and the spacing between the lines is larger than for the Q-branch because of the $(B'_v + B''_v)$ -term.

2.4.4 Angular momentum of the electrons

Taking the angular momentum of the electrons into account, an additional splitting of the lines occurs because of the total electron *spin* vector. The spin of the individual electrons can take the values $\pm\frac{1}{2}$ and therefore the total spin vector can add up to be either integral or half integral. The *orbital* angular momentum of the electrons is usually denoted L and the corresponding component along the internuclear axis is Λ and can take the values

$$\Lambda = 0, 1, 2, 3, \dots, L. \quad (2.26)$$

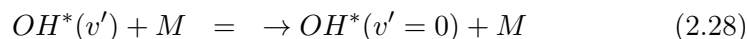
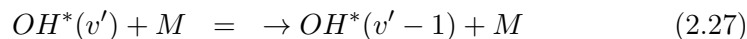
This gives rise to the molecular states $\Sigma, \Pi, \Delta, \Phi, \dots$ respectively. The component of S along the internuclear axis can take the values $\Sigma = S, S - 1, S - 2, \dots, -S$. Total angular momentum of the electrons is given by $L + S$ and the quantum number of the total angular momentum along the internuclear axis is $\Omega = |\Lambda + \Sigma|$. It is not defined for $\Lambda = 0$.

Thus, for a molecule with Λ different from zero, the state splits into a *multiplet* of $2S + 1$ components. If the number of electrons is odd, giving $S = \pm\frac{1}{2}$ and the molecule is in the Π state, this state is said to be *doublet*. So, in a rotational-vibrational transition, a molecule in the Π state can be in the states $\Pi_{\frac{3}{2}}$ and $\Pi_{\frac{1}{2}}$.

2.5 The Meinel lines

The Meinel lines were discovered by A. B. Meinel in 1950 [11]. The observed spectrum was first believed to be an electronic band structure, but was later determined to be the rotational-vibrational spectrum of the hydroxyl molecule. The ozone reaction (2.1) produces OH^* in vibrational states $v = 6$ up to $v = 9$ [18]. The molecule then radiates as it cascades down to lower vibrational levels. The lower states can also be reached by collisional quenching by other air molecules. This is believed to happen either a stepwise or

by 'sudden death' reaction [18]:



The Meinel bands (3-1) and (4-2) means a transition between the vibrational level $v = 3$ to $v = 1$ and $v = 4$ to $v = 2$. The lines in the each band are denoted $(branch)_i(J'')$, $i = 1, 2$, where (branch) is Q,P or R referring to the different branches as described in section 2.4.3. The index i represents the two possible values of the angular momentum in the ground state of the OH molecule which is $^2\Pi$. $i = 1$ refers to the $^2\Pi_{\frac{3}{2}}$ while $i = 2$ refers to the $^2\Pi_{\frac{1}{2}}$. The spectrum of the OH (3-1) and the (4-2) band is shown in figure 2.6. The original lines are only 0.01 nm wide but are broadened by the instrument. As mentioned in section 2.4.3, the Q-branch wave numbers is parabolic in J causing the lines to overlap. This makes the branch very narrow and the individual lines are not possible to distinguish with this resolution. The lines in the R-branch are also very closely spaced such that only their peaks can be separated from each other. The first line in the P-branch is the $P_2(2)$, (since $J = 0$ does not exist in the $^2\Pi$ state). Then follows $P_1(2), P_2(3)$ and so on. The P_1 lines are the most intense, since the $^2\Pi_{\frac{3}{2}}$ energy levels are lower and therefore more molecules will populate these states.

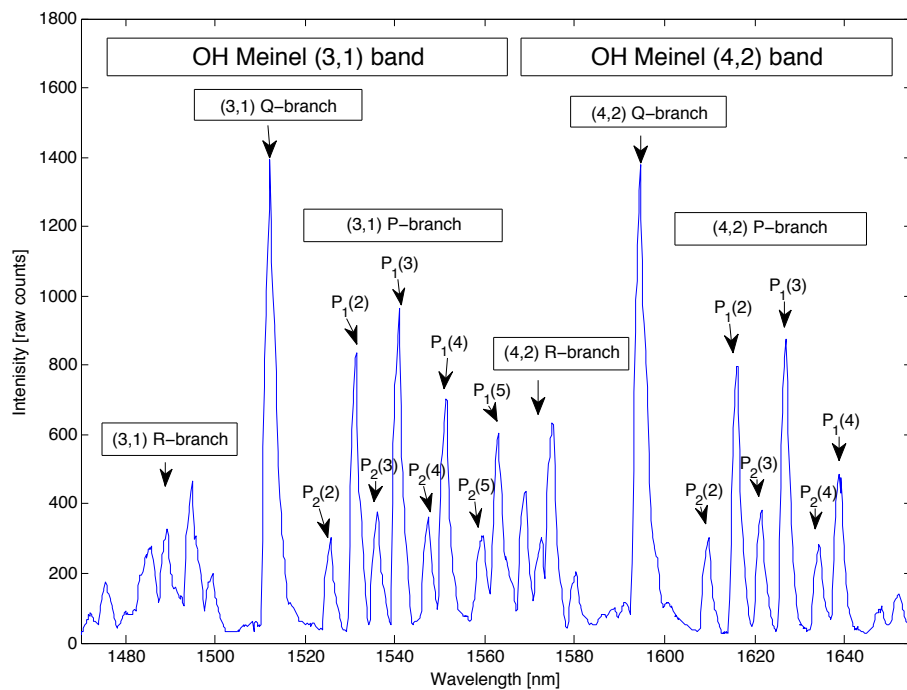


Figure 2.6: The plot shows the Meinel (3-1) and (4-2) bands. The spectrum is a night average of the night between 29-30 November 2012. This is a 'raw' spectrum, so the intensity of the (4,2) $P_2(4)$ and $P_1(4)$ are really almost twice as great.

2.6 Thermal distribution of quantum states

For a gas in thermal equilibrium with its surroundings, a well known result from statistical physics is that the relative number of molecules in two states with different energy is given by the Boltzmann factor

$$\frac{N_1}{N_2} = e^{-(E_1 - E_2)/kT}. \quad (2.29)$$

From this result follows that the number of molecules in vibrational state v is $\propto e^{-G_0(v)hc/kT}$. This is the number of molecules in state v relative to the ground state. To find the actual fraction of the total population in state v , the Boltzmann factor must be divided with the sum over all possible states,

$$N_v \propto \frac{e^{-G_0(v)hc/kT}}{Q_v}. \quad (2.30)$$

Q_v is called the state sum or the *partition function*, which in this case is

$$Q_v = 1 + e^{-G_0(1)hc/kT} + e^{-G_0(2)hc/kT} + \dots \quad (2.31)$$

The actual number of molecules in state v is the total population N multiplied with the fraction of molecules in state v

$$N_v = \frac{N}{Q_v} e^{-G_0(v)hc/kT}. \quad (2.32)$$

Since there are $2J + 1$ states with the same energy for each value of the rotational quantum number J , the distribution of rotational lines is not given only by the Boltzmann factor $e^{-E_J/k_B T}$, but are weighted by the degeneracy factor $2J + 1$. For the lowest vibrational state, the number of molecules in rotational state J will follow

$$N_J \propto (2J + 1)e^{-F(J)hc/kT} \quad (2.33)$$

N_J will therefore first go through a maximum before it decreases. The distribution for the OH molecule with $B = 18.867 \text{ cm}^{-1}$ is shown in figure 2.7 for a temperature 200 K. As before, the actual number of molecules in state J is the total population multiplied with the Boltzmann factor divided with the partition function, which for rotational states is

$$Q_r = 1 + 3e^{-2Bhc/kT} + 5e^{-6Bhc/kT} + \dots \quad (2.34)$$

The number of molecules in state J is then

$$N_J = \frac{N}{Q_r} (2J + 1)e^{-BJ(J+1)hc/kT}. \quad (2.35)$$

If the probabilities were the same for all transitions and keeping B constant, the intensity distribution would totally be determined by (2.35). This

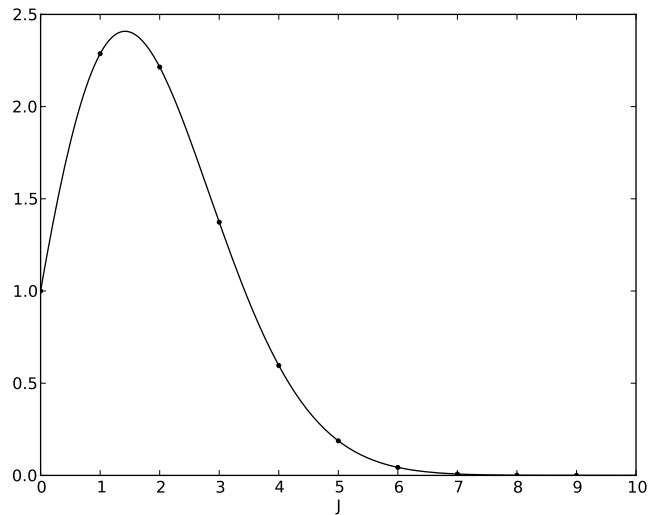


Figure 2.7: The distribution of rotational levels computed from (2.33) OH with $B = 18.867$ and $T = 200$ K

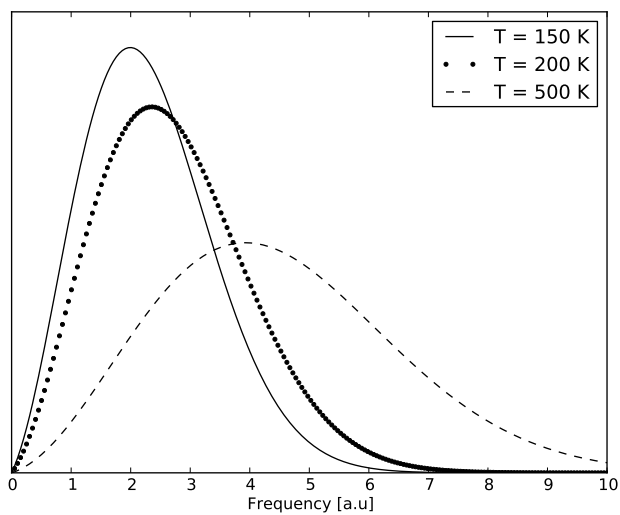


Figure 2.8: The intensity distribution (normalised) for a rotational-vibrational transition with $\Delta J = -1$ calculated with $B = 18.871$ cm^{-1} . The frequency units on the abscissa are arbitrary, but the numbers represent the line spacing.

theoretical intensity distribution is represented for different temperatures in figure 2.8 for a transition with $\Delta J = -1$. The effect of increasing temperature is that the intensity distribution stretches towards higher frequencies.

For higher vibrational levels, one must include both the rotational and the vibrational term values in the Boltzmann factor such that

$$N_J \propto (2J + 1)e^{-(G+F)hc/kT}. \quad (2.36)$$

2.7 Intensities of rotational lines

A spectral line is the signature of a certain transition between energy levels in a quantum system. The intensity of a spectral line, in terms of photons emitted per second is

$$I = N_n A_{nm} \quad (2.37)$$

where N_n is the number of molecules in the initial state. A_{nm} is the probability per unit time for the transition to spontaneously occur. In the dipole approximation of radiation [10], this transmission coefficient is

$$A_{nm} = \frac{64\pi^4 \nu_{nm}^3}{3h} |\mathbf{R}^{nm}|^2 \quad (2.38)$$

which is also called the *Einstein coefficient*. For transitions between rotational energy levels, the degeneracy factor $2J+1$ must be taken into account. The transmission coefficient becomes

$$A_{nm} = \frac{64\pi^4 \nu_{nm}^3}{3h} \frac{\sum |\mathbf{R}^{n_i m_k}|^2}{2J + 1} \quad (2.39)$$

where the indices i and k represents the degenerate sublevels of the upper and lower states and the sum is over all possible combinations of the sublevels of the initial and final state.

If a line within a given transition branch is denoted J and the initial energies of these lines ϵ_J , we see from (2.35) that the intensity of the line becomes

$$I_J = \frac{N}{Q_r} A_{nm} (2J + 1) e^{-\epsilon_J/k_B T} \quad (2.40)$$

$$= \frac{64\pi^4 \nu_{nm}^4 c N}{3Q_r} \left(\sum |\mathbf{R}^{n_i m_k}|^2 \right) e^{-\epsilon_J/k_B T} \quad (2.41)$$

This can again be written, in terms of photons per second by dividing with $h\nu$,

$$I_J = \frac{\rho C \nu^3}{Q_r} S_J e^{-\epsilon_J/k_B T}. \quad (2.42)$$

S_J is the *line strength* of the transition, C a constant and ρ is representing the initial concentration. The partition function can then be written as

$$Q_r = \sum_J S_J e^{-\epsilon_J/k_B T} \quad (2.43)$$

It is assumed that the OH molecules which are in the lower vibrational states are thermalized in the mesopause. This is because the relatively long lived OH* state ($^2\Pi, v' > 2$) undergoes at least 10 collisions before they emit a photon [18]. The rotational energy levels should therefore follow a Boltzmann distribution, and the intensity distribution of the transitions should be well represented with (2.42). The linestrengths, initial energies, and theoretical wavelengths for the OH (3-1) and (4-2) transitions have been calculated numerically [6] and are listed in table 2.1

Table 2.1: Table of the initial energies, line strengths and the theoretical wavelengths of the P-lines in the OH (3,1) and (4,2) band.

Line	e_J [cm ⁻¹]	S_J	Wavelength [nm]
(3,1)			
$P_2(2)$	10300.41	3.9839e+11	1518.70
$P_1(2)$	10172.30	4.9798e+11	1524.06
$P_2(3)$	10354.21	7.3932e+11	1528.76
$P_1(3)$	10247.07	9.0706e+11	1533.19
$P_2(4)$	10443.29	1.0852e+12	1539.51
$P_1(4)$	10352.45	1.2943e+12	1543.16
$P_2(5)$	10567.02	1.4404e+12	1550.94
$P_1(5)$	10488.78	1.6789e+12	1553.96
(4,2)			
$P_2(2)$	13377.61	3.4350e+11	1597.25
$P_1(2)$	13248.92	4.2855e+11	1603.05
$P_2(3)$	13429.00	6.3772e+11	1607.96
$P_1(3)$	13320.76	7.8114e+11	1612.81
$P_2(4)$	13514.12	9.3671e+11	1619.43
$P_1(4)$	13421.92	1.1156e+12	1623.47
$P_2(5)$	13632.41	1.2443e+12	1631.68
$P_1(5)$	13552.73	1.4469e+12	1635.05

2.8 OH airglow response to gravity waves

There are many reports on gravity waves in the mesosphere detected by OH airglow variations. Taylor et al. (1991)[20] detected a short period gravity wave by simultaneously measuring changes in intensity and the OH (3,1) rotational temperature. The temperature change was ≤ 10 K from a mean of 163 K and the relative intensity change was 8 times greater than for the temperature. The intensity and the temperature variation were found to be well correlated in space and time with the intensity variation slightly leading the temperature change. An observation of a large scale oscillation in airglow brightness and OH rotational temperature is described by Oznovich et al. (1995)[13]. The oscillation is believed to be caused by a gravity wave with period of about 8 hours. Peak to peak change in the derived temperature was 19 K, giving a relative change in the intensity of ~ 2.8 larger than the relative change in temperature. They also found that the temperature variation was ahead the variation in intensity. A large wave event in the OH airglow layer was observed by Taylor et al. (1995). The measured change in OH intensity was $> 50\%$ and the rotational temperature amplitude was ~ 20 K. The jump in temperature is reported to lead the intensity by almost 15 min.

Gravity wave induced changes in the airglow layer temperature is caused by adiabatic changes because of the compression and expansion of the air as the layer is corrugated by the gravity wave [20] (and references therein). The temperature variation will therefore be dependent on the background atmospheric temperature profile. Since the temperature is brightness weighted, the largest relative temperature changes occurs near the centroid height of the layer.

The induced changes in the intensity is a more complex process. Because the OH molecule is chemically active, the altitude profile will depend on the distribution of the chemical species which it depends on and also the temperature which will affect the reaction rates. A detailed analytical model of the OH (8-3) emissions was developed by Swenson and Gardener (1998) [19]. In this model, the OH volume emission rate peaks ~ 3.75 km below the centroid height of the OH layer because of the strong dependence of the volume emission rate on the atomic oxygen concentration. They also show that this layer difference will give a phase shift between the intensity and temperature variations, which is dependent on the vertical wavelength of the perturbing wave as well as the layer separation.

Chapter 3

Instrument

3.1 Spectrometer

The spectrometer is mainly a spectrograph in combination with a linear infrared InGaAs PDA detector. The spectrograph is equipped with an adjustable entrance slit and an external shutter. A liquid cooling system is installed to ensure instrument stability. The parts comprising the instrument system are listed in the table below.

Type	Description
Spectrograph	Andor Shamrock SR-163
Camera	Andor iDus InGaAs μ 1.7 DU491A
Shutter	Melles Griot 04 UTS 201
Shutter control	Melles Griot 04 ISC 850
Cooling	Koolance Exos-2
Grating	Andor SR1-GRT-0600
Slit	Andor SR1-ASM-0020

The Shamrock SR-163 is a benchtop spectrograph based on the Czerny-Turner layout. Its focal length is 163 mm and it has a numerical aperture of f/3.6. The gratings are interchangeable, and the angle of the grating, and hence the wavelength range, is changed with a micrometer drive. The grating in use has 600 lines/mm and a nominal dispersion of 7.11 nm/mm.

The Andor iDus InGaAs camera is a linear PDA detector with 1024 pixels. The wavelength coverage is 0.6-1.7 μ m and the minimum exposure time is 1.4 μ s. Because of the low the bandgap of the InGaAs semiconductor, it is subject to a very high level of dark current. The detector must therefore be cooled to keep the thermal noise at an acceptable level, and prevent the pixels from saturate. The pixels are vacuum insulated, and the camera has a built in thermo electric element which can cool down the chip to -90 °C, depending on the temperature of the surroundings. The Koolance Exos-2 liquid cooling system assures that heat is transported away effectively.

The shutter protects the detector from sunlight during daytime, and is used to update the dark signal level during measurements. The technical details of the camera, grating and spectrometer are listed in table 3.1.

Table 3.1: Instrument specification details. For the iDus camera, the specifications are listed for the high sensitivity mode. Source: [web, c].

Andor Shamrock SR-163	
Aperture	F/3.6
Focal length	163
Grating	Single, interchangeable
Slit height	3 or 6 mm
Size (H×B×L)	198 × 216 × 96 mm
Weight	3.5 kg
Andor SR1-GRT-0600-1900 grating	
Lines/mm	600
Blaze	1600 nm
Nominal dispersion	7.11 nm/mm
Max recommended wavelength	1715 nm
Andor iDus InGaAs μ1.7 DU491A	
Pixels	1 × 1024
Pixel size (W×H)	25 × 500 μ m
Max spectra pr. sec	97
Minimum exposure time	1.4 μ s
Wavelength range	600 nm - 1.7 μ m
Digitization	16 bit
Minimum temperature	-90 °C
Dark current (max cooling)	10.1 ke ⁻ /pixel/sec
Read noise	580 e ⁻
Pixel well depth	5 e ⁻
Sensitivity	90 e ⁻ /count
Linearity	>99%

3.2 Calibration and noise

The instrument's response and channel to wavelength relation were calibrated in the spring of 2011 and are described in detail in [3]. Before deployment in 2012, the systems response was re-calibrated following the same procedure and using the same equipment as earlier. A transfer source was also scanned, so that the system's response can be checked at a later time by comparing the response to this source.

The channel to wavelength relation for the instrument is

$$Wavelength[nm] = -0.86627 \cdot 10^6 \cdot SCN^2 + 0.2071 \cdot SCN + 1333.6, \quad (3.1)$$

where SCN is the super channel number, ranging from 1-2080. The super channel system consists of three overlapping regions where the middle one covers the OH (3-1) and (4-2) bands, and starts at SCN 625. This means that the wavelength range in this region is 1462.7-1672.8 nm. The micrometer setting for this region is 7.9. Because of the very small second order term in (3.1), the channel to wavelength range is very close to linear.

The system response to different detector temperatures is very non-linear. Figure 3.1 shows the thermal noise measured as the standard deviation of the residuals between subsequent background scans, taken with 1 second integration time. The plot shows that below -50 °C, the noise curve flattens out at the read-out level which amounts to ~ 15 counts. With an integration time of 60 s, the thermal noise was measured in a similar manner to be ~ 35 counts.

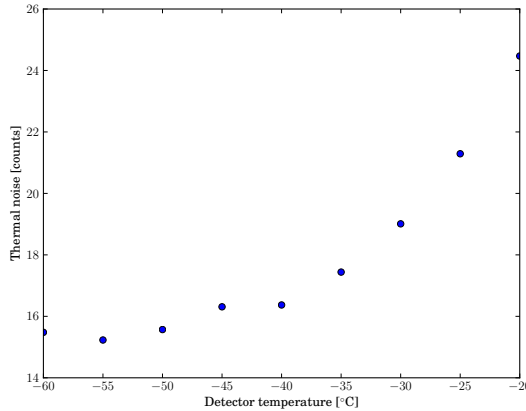


Figure 3.1: Thermal noise of the detector at different detector temperatures, measured with 1 second integration time.

As the chip temperature is reduced, more energy is needed to excite the electrons across the semiconductor bandgap. At -50 °C, the wavelength where the sensitivity is reduced to on half times the maximum, was found to be ~ 1643 nm. As a result, the OH (4-2) P-branch signal is affected by the sharp drop in sensitivity in this region. To keep sensitivity over the (4-2) P-branch, while keeping the noise relatively low, -50 °C was set as an operating temperature for the instrument.

The system response and slit settings were checked against the transfer source 12th January 2013, and the resulting response curve is plotted in figure 3.2. Originally, the response curve was converted to units of Rayleighs

$\left(\frac{\text{counts/sec}}{\text{kR/nm}}\right)$, but as only the relative intensities are of interest in the calculations, it can be normalised to maximum sensitivity, which is in the 1500 nm region. The sensitivity curve normalized to the value 3.7, is shown in figure 3.2. The theoretical wavelength of the (4-2) $P_1(4)$ line is 1623.47 nm, but is actually shifted to longer wavelengths on the instruments wavelength scale. Reducing the the detector temperature would have moved the (4-2) P-branch further into the region where sensitivity drops dramatically.

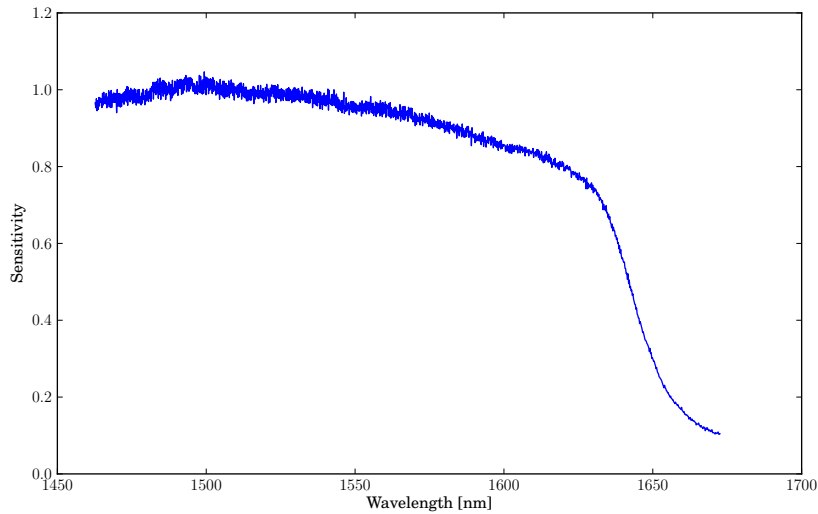


Figure 3.2: Instrument response normalised to 3.7 (counts/sec)/(kR/nm).

3.3 Operation

The instrument started operating on 20th November 2012 at Dragvoll NTNU where, the measurement site for the Atmospheric Physics Group at NTNU is located. It is mounted on a metal installation in a chimney, with it's optical axis straight up through a borosilicate window. All the walls of the chimney are covered with black paper to absorb incoming light. When in operation, the bottom of the chimney is sealed with a dark blanket. A fan blows air through a hose on the window from beneath, to prevent water from condensate/freeze on it. The metal ring which seals the borosilicate window on the top side, is designed in a way to allow water to escape from the window. Normally, ice will melt and the water wil pour off or evaporate. This solution seems to work well for most days, but with the changing weather in Trondheim, where freeze-thaw conditions occur often during winter, having the window covered with water, snow or ice is unavoidable on some nights. Water absorbs broadly in the (4-2) and (3-1) wavelength region, and only a

thin film of water is enough to filter out much of the OH signal. This can be detected as the shorter wavelengths are affected first. Whenever the (3-1) band is far below the (4-2), absorption by water is probably the reason. The installed instrument is shown in figure 3.3.

Automatic acquisition software is developed for this instrument and is also described in [3]. The instrument runs completely automatically, and no user interaction is necessary unless the settings are to be changed. When to take data is specified by setting a solar zenith angle interval. Standard integration time since start up have been 60 s except for the nights between 7-8. December 2012 where the integration time was set to 15 s for experimentation. The total integration can be read out in several accumulations, but the current setting is to read out in one accumulation. All the operation settings are summarized in table 3.2.

Table 3.2: Current standard operation settings for the Trondheim OH spectrometer

Detector temperature	-50 °C
Entrance slit	2 μm
Grating angle drive	7.9 μm
Background cycle	5
Integration time	60 s
Accumulations	1
SZA interval	above 95°



Figure 3.3: The spectrometer installation seen from above.

3.4 Instrument slit function

The instrument slit function is generally the bandpass function of the spectrometer. It is the convolution of the image of the entrance slit with the

image of the exit slit. In an ideal instrument, with no optical aberrations and negligible diffraction, rectangular slits results in the typical triangular or trapesiodal slit functions. Diffraction of light explains the rounding of the edges of real slit functions, while the often seen assymetry is caused by various optical phenomena, affecting the light on the way from the entrance slit to the exit slit. The instrument function is a result of a complicated process and not easy to approximate with at mathematical function. It was therefore decided to measure the spectrometer slit function by scanning a monochromatic source for a long enough time to yield a smooth function.

The Hg I ion emits a strong, single standing line at 1529.58 nm. This is well within the instruments wavelength range, and a source if this line is therefore suitable to measure the slit function. A lamp from Bentham Instruments was used for the measurements. The model name of the lamp is CAL_CL_HG, and the serial number 8728.

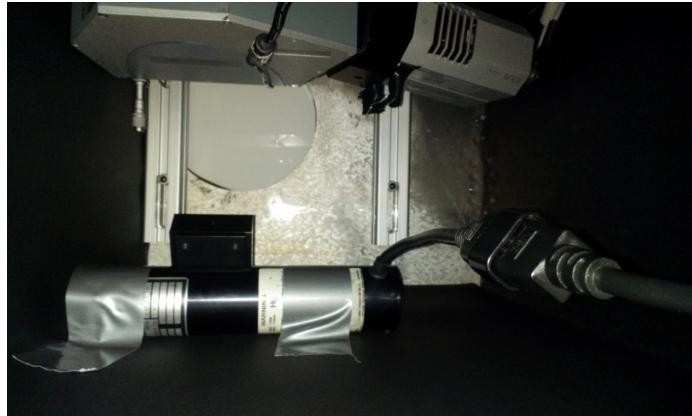


Figure 3.4: The setup used to measure the instrument slit function. The diffusive screen can be seen through the window and the mercury lamp is attached to the chimney wall.

A diffusive screen was placed on top of the window and covered with a dark blanket. The diffusive screen will ensure that the spectrometers entire field of view is filled. The Hg-lamp was then attached to the wall of the instrument chimney and directed towards the diffusive screen. The Andor Solis software was used to ensure optimal illumination and determine a suitable exposure time to avoid detector saturation. The chimney was then sealed to prevent daylight from entering. A picture of the setup is shown in figure 3.4.

The lamp was scanned 76 times with an integration time of 60 s and a background update every 5th scan. The spectra were then 3 point median filtered, divided with the response curve and averaged. A plot of the average spectrum is shown in figure 3.5.

The line offset was removed by fitting a straight line trough the points next to the line. The instrument function is taken to be the pixel interval

352-385. The normalised instrument function is shown in figure 3.6. When applying the instrument function, it is assumed that it is constant over the wavelength range, and that the wavelength scale is linear, which should be a good approximation according to the wavelength scale for the instrument (3.1).

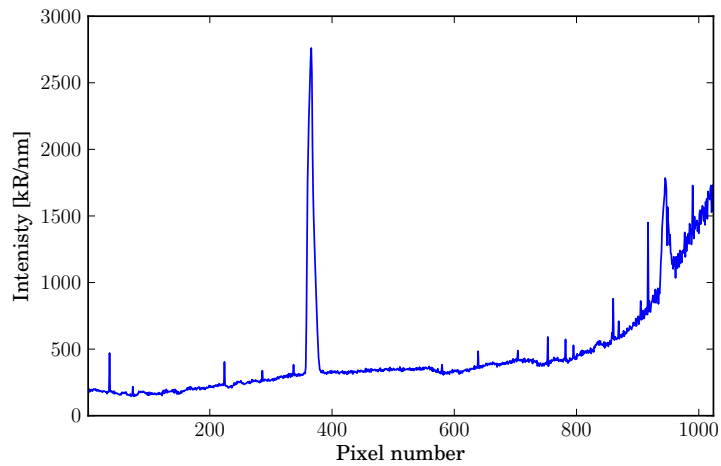


Figure 3.5: The measured Hg-lamp spectrum used to estimate the instrument function

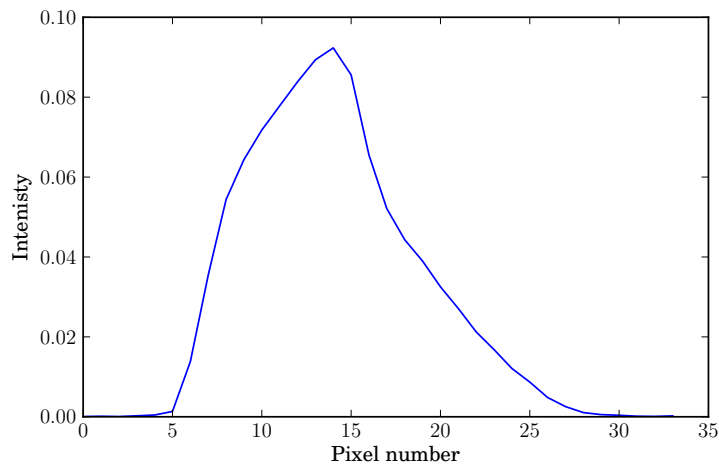


Figure 3.6: The normalised instrument function measured from with the Hg-lamp.

Chapter 4

Data analysis

The purpose of the measurements is to infer both the intensities and the rotational temperatures of the OH (3-1) and (4-2) P-branches. This is done by fitting a synthetic spectrum based on (2.42) to the observed data. The lines included in the model, are the six lines from P₂(2)-P₁(4) (see figure 2.6). Taking only the instrument lineshape into account, the OH signal at pixel i can be modeled as

$$M_i = \sum_J I_J w_{i,J}, \quad (4.1)$$

where I_J is equation (2.42) and $w_{i,J}$ is the instrument function for the J 'th line. In other words, the intensity at the i 'th pixel is the contribution from all the J lines, which is described by the instrument function. The instrument response can be taken into account in two ways. Either, the spectra can be divided with the response curve and converted to kR/nm, or the model can be multiplied with it to have the same relative intensities, measured in detector counts, as the observed data. The latter is preferred, since the response curve will blow up the noise in the long wavelength end of the spectrum. There is more in the measured signal than the hydroxyl emissions, and therefore a baseline B_i must be added to the model. Estimating this background will be detailed in section 4.2. Finally, the OH emissions are passing through nearly 90 km of atmosphere, and therefore some atmospheric absorption must be taken into account [8]. The transmission coefficients used are listed in appendix A.1 and A.2. The synthetic model of the observed data then becomes

$$M_i = R_i \sum_J \frac{I_J}{T_J} w_{i,J} + B_i, \quad (4.2)$$

with R_i being the response at pixel i and T_J the transmission coefficient for line J .

The model is fit by minimizing the sum of squared errors between the observed data, D_i , and the model (4.2). Since the model is non linear, this is done iteratively with the Gauss-Newton method.

4.1 Dark signal

As mentioned in chapter 3, the detector is subject to a high level of dark current because of the low semiconductor bandgap. For the detector temperature of $-50\text{ }^{\circ}\text{C}$, it is typically 80-140 counts/sec/pixel depending on the temperature instrument's surroundings. This is ~ 5 -10 times the strength of the hydroxyl emissions.

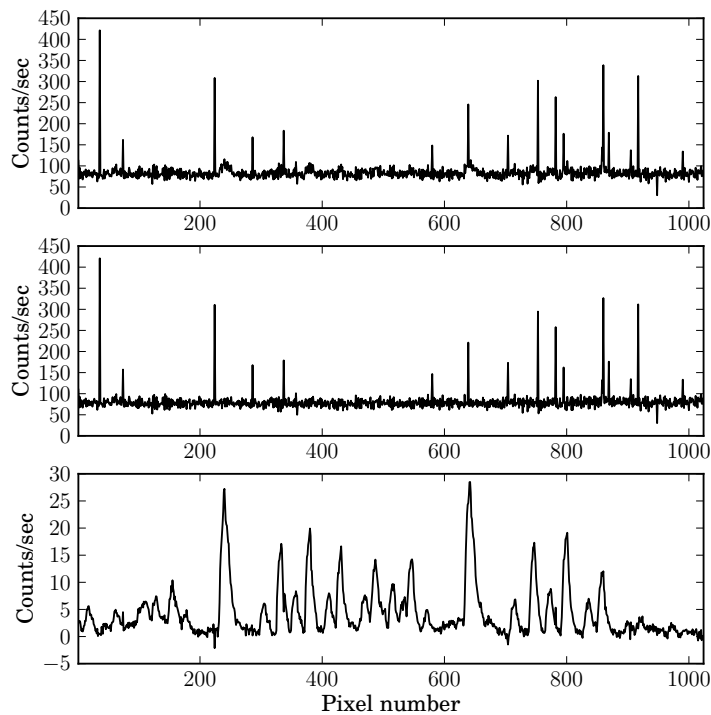


Figure 4.1: Upper: A raw scan. Middle: The nearest measured dark signal. Lower: Upper plot minus the dark signal.

Both the dark signal and the signal scans contains the same amount of thermal noise, so by subtracting the dark signal, one is left with double noise in the residual spectrum. Therefore, a method to estimate the dark signal in each pixel, described in [14], have been implemented get around this.

The method is based on the fact that the pixel to pixel variation in the dark signal is almost constant (fixed pattern noise). It is therefore a nearly linear relationship between the mean of the dark signal and the value at each pixel. An estimate of the actual dark signal in each scan can be made by averaging all the dark signals for each night, to reduce the influence of thermal noise. The average dark signal is then scaled with the mean value of the spectrum, and the dark signal in each scan can be found by multiplying

the mean with this spectrum. However this is not an accurate approximation for this instrument, as there is an underlying side to side structure which changes over the course of the night. The method of removing the dark signal is therefore as follows. For each night, all the dark signals are averaged. This spectrum is then highpass filtered, taking out all variations extending over more than 15 pixels. This is taken as the fixed dark signal pattern for the night. In each background scan, the underlying structure is found by lowpass filtering it, removing all variations shorter than 15 pixels. The method is complete by adding the fixed pattern to the underlying structure. Figure 4.2 shows how the dark signal for three different pixels vary through night, and the result of estimating the dark signal. This shows that the actual dark signal is smoother than recorded. By not taking into account the underlying variation and only scaling to match the average dark signal for the night, some pixels are over/under-estimated (pixel 18 and 512). The highpass-lowpass method estimates the dark signal better. It is not a perfect method, but the systematic errors have not been found to exceed 3-4 counts which is small compared to the thermal noise of ~ 35 counts.

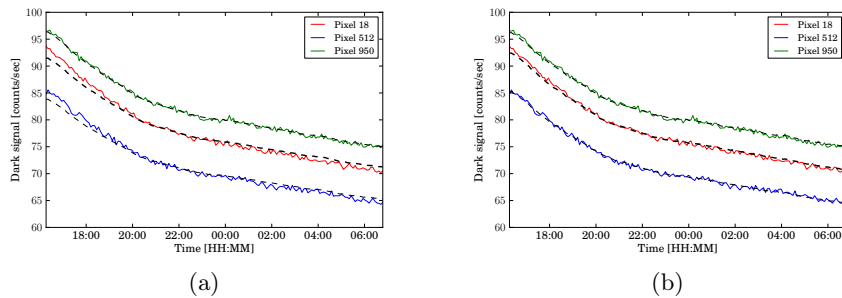


Figure 4.2: The dark signal for three different pixels through a night. The dashed black lines are the estimate of the actual dark signal. (a) is the mean estimate, while (b) is the result of the lowpass-highpass method.

4.2 Baseline removal

In addition to the OH emissions, there will always be a baseline of blackbody radiation, scattered light and a possible small offset because of the dark signal drift. Since the temperature is dependent on the relative intensities of the lines, incorrect offset or subtracting too much or too little of one of the lines will affect the temperature.

The instrument function reveals the pixels in the band where there are no contribution from the hydroxyl emissions. This is shown in figure 4.3. For the (3-1) band, this plot suggests that it is not sufficient to remove a constant offset for the band, but that a baseline profile must be subtracted

to bring all the no signal points to zero. The baseprofile in the (4-2) band is much straighter than the in the (3-1), but a baseprofile is also subtracted for this band.

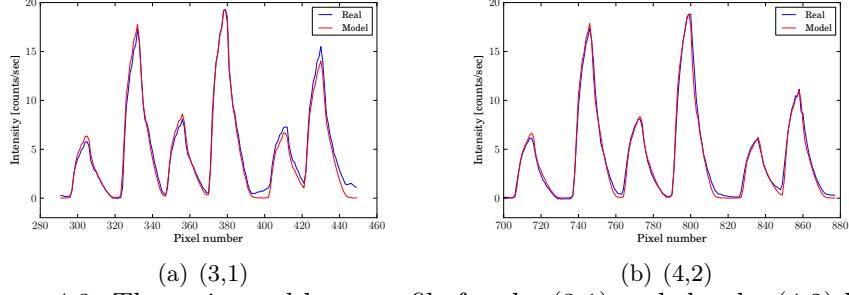


Figure 4.3: The estimated base profile for the (3,1) and the the (4,2) band.

The underlying baseprofiles are estimated by fitting a 3rd. degree polynomial through the no signal points. This is based on the assumption that the OH lines are sitting on a relatively smooth underlying atmospheric background profile. The no signal points are determined from the peaks, and are for both bands the points left to $P_2(2)$ -line, in between the $P_2(2)$ and $P_1(2)$ -line, between the $P_1(3)$ and the $P_2(4)$ -line and to the right of the $P_1(4)$ -line. The estimated profiles and the no signal points are plotted in figure 4.4. The result of subtracting these points for the same spectra are shown in figure 4.5.

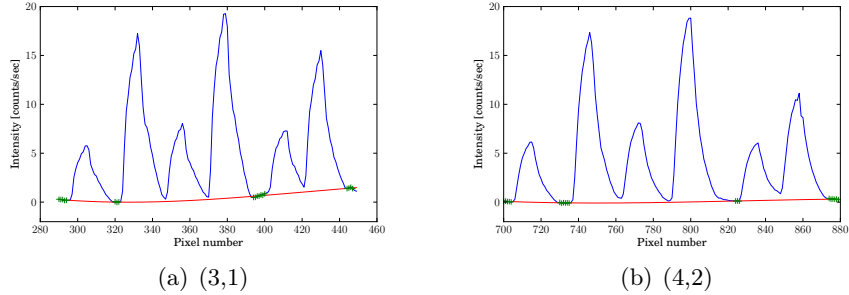


Figure 4.4: The estimated base profile (red line) and the no signal points (green points) for the (3-1) and the the (4-2) band.

As mentioned earlier, there can be a small offset because the mean value of the dark current drifts in time. Therefore, the minimum of the estimated baseprofiles is subtracted so that they have a zero point. A constant offset is instead included in the Gauss-Newton fit, such that the model in general becomes OH-model + baseprofile + offset. The baseprofile is treated in two different ways. For spectra near the sunrise, the shape of the baseprofile is included in the fitting routine. Assuming that it can be expressed as $\alpha \times b_i$, where α is a constant and b_i is the baseprofile at pixel i , the full model of

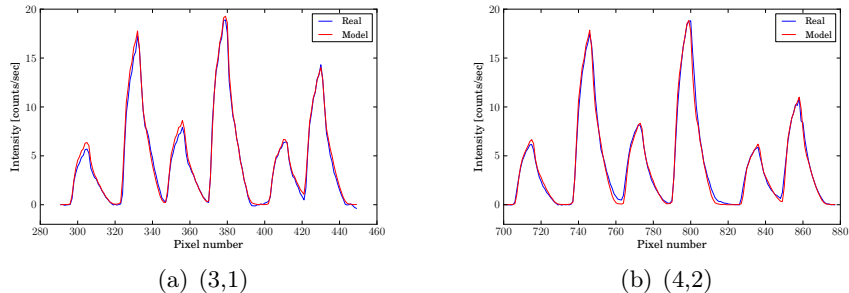


Figure 4.5: The estimated base profile (red line) and the no signal points (green points) for the (3-1) and the the (4-2) band.

the recorded signal can be expressed

$$M_i = R_i \sum_J \frac{I_J}{T_J} w_{i,J} + \alpha b_i + \beta, \quad (4.3)$$

where β is the constant offset. When the sun has set and the sky is free from residual sunlight, and under normal conditions, the baseprofile does not change. It can therefore be subtracted from the spectrum, and the model becomes

$$M_i = R_i \sum_J \frac{I_J}{T_J} w_{i,J} + \beta. \quad (4.4)$$

The experience is that this is a more stable way of handling the baseline, and is therefore preferred as the general method. Including the baseline shape in the fit, resulted in a higher level of random errors than (4.4). Fitting (4.3) gives reasonable values down to the 5-10 scans near the sunset/sunrise for many nights, but the fit is not always reliable. In analysis of the results, the 25 (30 minutes) first scans are disregarded.

It was first attempted to fit the baseprofile on a per scan basis, but as there is relatively few points to estimate the baseprofile with, it resulted in a too high level of random noise. Best results have been achieved by fitting the profile from the night averaged spectrum and assuming it is constant over the course of the night. Spectra where the cloud level is assumed to influence the signal are excluded when averaging the spectra, as this might influence the baseprofile. This will be explained in more detail in section 4.4

4.3 Gauss-Newton algorithm

The principle of the Gauss-Newton algorithm is to iteratively find the parameters which minimizes the sum of squared errors (*SSE*) between the data and the model. With i referring to the i 'th data point, *SSE* becomes

$$SSE = \sum_i (D_i - M_i)^2 = \sum_i r_i^2, \quad (4.5)$$

where r_i is the i 'th residual. To first order, a Taylor expansion of the residual function $r = D - M$ in the paramters β from a starting point β^k is

$$\mathbf{r}(\beta) \approx \mathbf{r}(\beta^k) + \mathbf{J}_r(\beta^k)\mathbf{\Delta} \quad (4.6)$$

where $\mathbf{\Delta} = \beta - \beta^k$ and \mathbf{J}_r is the Jacobi matrix with elements $J_{ij} = \frac{\partial r_i}{\partial \beta_j}$. Finding the $\mathbf{\Delta}$ which minimises the sum of square errors on the right hand side, is a linear least squares problem and results in the *normal equations* [16]

$$\mathbf{J}^T \mathbf{J} \mathbf{\Delta} = \mathbf{J}_r. \quad (4.7)$$

Starting from some initial paramaters and iteratively finding new parameters $\beta^{n+1} = \beta^n + \mathbf{\Delta}$ is the Gauss-Newton algorithm. Inserting the normal equations (4.7), the iterations can be written

$$\beta^{n+1} = \beta^n + (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{r}(\beta)^n \quad (4.8)$$

Switching to the parameter $\tau = \frac{1}{T}$ in (2.42) and merging the constant and ρ into 1 factor (ρ) representing the intensity, for easier calculations, the full model (4.3) becomes

$$M_i = R_i \frac{\rho}{Q_r} \sum_J \frac{S_J \nu_J^3}{T_J} w_{i,J} e^{-\tau \frac{\epsilon_J}{k_B}} + \alpha b_i + \beta. \quad (4.9)$$

Computing the partial derivatives of this model, the elements of the Jacobi matrix are

$$\frac{\partial M_i}{\partial \rho} = \frac{R_i}{Q} \sum_J \frac{S_J \nu_J^3}{T_J} w_{i,J} e^{-\tau \frac{\epsilon_J}{k_B}} \quad (4.10)$$

$$\frac{\partial M_i}{\partial \tau} = -R_i \frac{\rho}{Q_r} \sum_J \frac{S_J \nu_J^3}{T_J} w_{i,J} \left(\frac{\epsilon_J}{k_B} + \frac{1}{Q} \frac{\partial Q}{\partial \tau} \right) e^{-\tau \frac{\epsilon_J}{k_B}} \quad (4.11)$$

$$\frac{\partial M_i}{\partial \alpha} = b_i \quad (4.12)$$

$$\frac{\partial M_i}{\partial \beta} = 1$$

The jacobi matrix is a $(n \times 4)$ matrix where n is the number of pixels included in the model. For the model equation (4.4), the jacobi matrix is an $(n \times 3)$ matrix without the $\frac{\partial M_i}{\partial \alpha}$ -elements. With (2.43) the derivative of Q with respect to τ becomes

$$\frac{\partial Q_r}{\partial \tau} = - \sum_J S_J \frac{\epsilon_J}{k_B} e^{-\tau \frac{\epsilon_J}{k_B}}. \quad (4.13)$$

The matrix $(\mathbf{J}^T \mathbf{J})^{-1}$ is called the *covariance matrix* and the diagonal elements are the variances of the fitted parameters [16]. By combining the

variance of the residuals, σ_r^2 , and the covariance matrix in the last iteration, the uncertainties of the fitted parameters can be found by

$$\sigma_{\beta_j} = \sqrt{C_{jj}}\sigma_r. \quad (4.14)$$

With $T = \frac{1}{\tau}$, the error in the temperature becomes

$$\sigma_T = \sqrt{\left(\frac{\partial T}{\partial \tau}\sigma_\tau\right)^2} = \frac{1}{\tau^2}\sigma_\tau = T^2\sigma_\tau. \quad (4.15)$$

Since the model intensity is proportional to ρ , the error in the model can be expressed as

$$\sigma_I = \frac{\sigma_\rho}{\rho}I. \quad (4.16)$$

4.4 Cloud level indicator

Clouds will have a negative effect on the signal in two ways. Obviously, clouds will block the OH emissions which will result in poor signal to noise ratios. The other effect is that clouds will scatter light into the spectrometer. Among the light which is scattered into the spectrometer, is the light from street lights. These lamps are commonly sodium vapor lamps, which is a source of several closely spaced lines in the ~ 1637 - 1640 nm region [web, b]. These lines show up in the far wavelength end of the spectrum as one broad line on cloudy nights.

Night averaged spectrums from three nights with different cloud levels are shown in figure 4.6. The signal is plotted in raw counts per second, which means that the Na-signal in reality is many times stronger, as the instrument sensitivity in this region is very low. The middle plot shows a spectrum with a medium level of Na-intensity. This plot shows that the OH signal is unaffected by Na-intensity at this level by comparing it to the upper, which is Na-free. In the lower plot, the OH emissions are completely drowned by the high Na-intensity.

It is felt that above a certain cloud level, the spectrum is changed so much that the model fails to yield a good fit (see section 4.6.2). As a result, both the temperature and the error estimate are no longer reliable. It is therefore not possible to get rid of all the potential unreasonable fits with a limit on the error, and spectrums are therefore excluded with an additional limit on the Na-intensity. As figure 4.6 illustrates, although clouds will block OH emissions and reduce the signal, it is still possible to obtain useable temperatures from many of these. The goal is to keep the spectra where the model is a reasonable fit, and exclude most of the spectra where the model is not valid. As a parameter to decide this, the ratio of the integrated Na-intensity and integrated (4-2) Q-branch intensity is used. This will allow temperatures to be calculated from spectra with a high relatively Na-intensity but where the

OH intensity is high enough for the model to be reasonable. Also, as mentioned in section 4.2, the baseline will be estimated from the spectra with a Na-OH ratio beneath a certain level, to have a correct baseline as possible for the useable spectrums.

By looking on plots like figure 4.7, it is felt that the signal starts to be affected with an Na-OH ratio somewhere > 0.5 . Above a ratio of 2, one often see the intensity significantly dropping. Usable values of the Na-OH parameter therefore seems within these values.

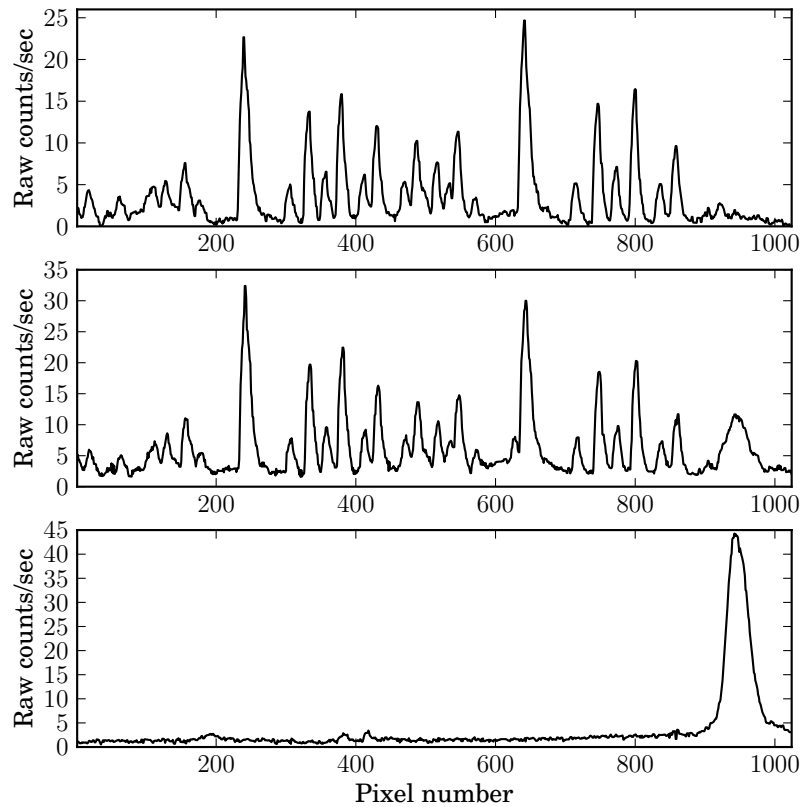
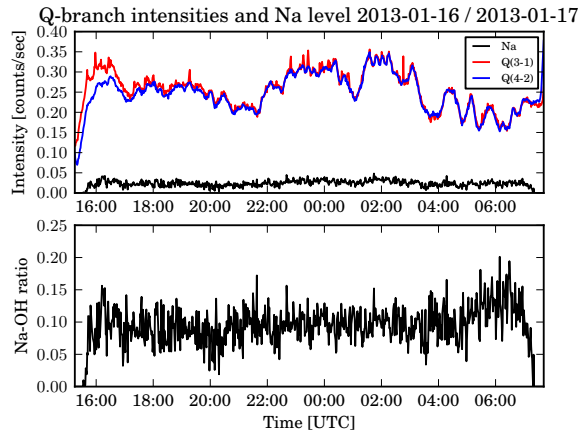
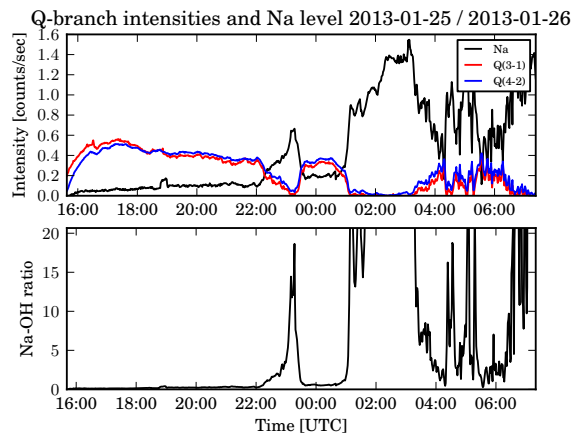


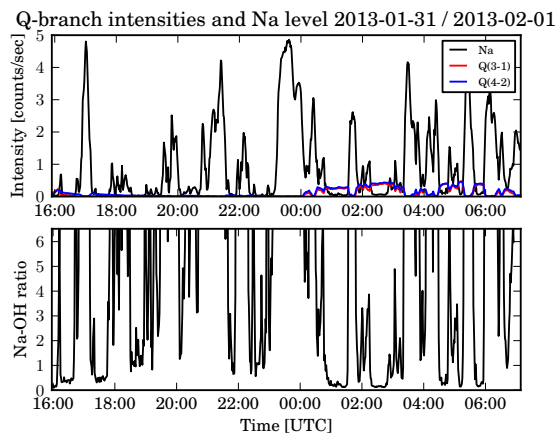
Figure 4.6: Spectra showing the effect of clouds scattering Na-light into the instrument. Upper panel: No clouds. Middle: Partly clouded. Lower: Completely clouded.



(a) No clouds.



(b) Partly clouded.



(c) Heavy clouds.

Figure 4.7: Plots showing the effect of the absolute Na-intensity and the Na-OH ratio the OH signal.

4.5 Data products

For every night, the weighted mean of the temperatures and intensities are calculated. The expression is

$$\langle T \rangle = \frac{\sum_i^n w_i T_i}{\sum_i^n w_i} \quad (4.17)$$

where the weighting w_i is the inverse square of the individual measurement errors σ_i ,

$$w_i = \frac{1}{\sigma_i^2}. \quad (4.18)$$

The standard deviation of the mean then becomes

$$\sigma_{mean} = \sqrt{\frac{1}{\sum_i^n w_i}}. \quad (4.19)$$

The same principle is used to smoothe the temperature series. Each temperature is then replaced with a weighted average of the nearby temperatures.

A rough estimate of the wavepower in the calculated time series of temperatures and intensities is obtained from the spectral power densities. The Lomb-Scargle method [16] is used to estimate the power spectrums because the time series consists of more or less unevenly spaced data.

The Lomb-Scargle periodogram is calculated for a minimum frequency of $\frac{1}{T}$, where T is the measurement duration. Maximum frequency calculated is 1.5 times the average Nyquist frequency

$$f_{max} = 1.5 \times \frac{N}{2T}. \quad (4.20)$$

where N is the total number of data points. The number of frequencies calculated is

$$N_p = \frac{4 \times 1.5}{2} N, \quad (4.21)$$

where 4 is the oversampling factor. These principles are described in detail in [16].

Total daily wavepower is found by integrating from the minimum gravity wave frequency, the Brunt-Väisälä frequency of 5 minutes [1], to 0.8 times the measurement duration. The reason for this is to exlude the changes with periods longer than the time series itself. The time series are windowed with a Hamming window. Noise is estimated as the mean power in the frequencies below 5 min and is subtracted from the wavepower integral.

Figure 5.17 shows a Lomb-Scargle periodogram where the red lines forming the large rectangle indicates the power that can be associated with gravity waves.

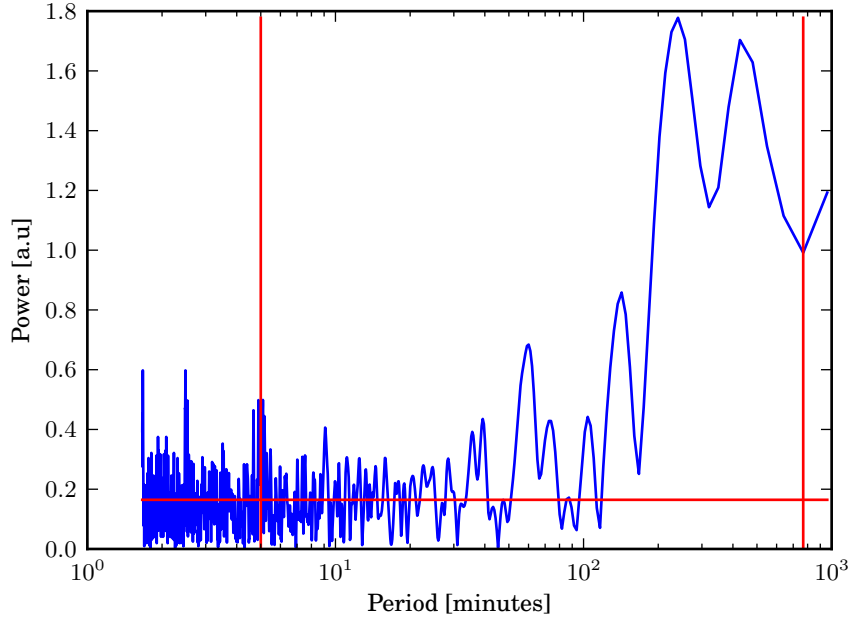


Figure 4.8: Lomb-Scargle periodogram of the (3-1) temperature series. The red lines mark period boundaries that are realistic for gravity waves.

4.6 Computer algorithm

The computer program which performs the above described calculations, and follows with this thesis, is written in Python 3.2. External libraries used NumPy [web, e] and SciPy [web, f] which have objects for fast matrix operations and basic routines such as polynomial fitting, filtering and fast fourier transforms. The library matplotlib [web, g] is used for plotting. The developed routine, is not a complete computer program. There is very little documentation written and error handling included. The focus in this section is on the essential sub algorithms, with the purpose of explaining how the results were achieved. The ideas and algorithms developed here also may serve as a good basis for a more complete routine.

The programs functions are divided into four files, "importtools", "spectools", "temptools" and "disptools", depending on their use in the main routine which is located in the "ohdata" file. The file importtool contains functions for extracting the signal from the raw datafile created by the acquisition program. Functions which perform operations on the spectra are located in the spectools file. Such functions are for instance filtering, averaging of spectra, peak locating and baseline estimation. Temperature calculations are performed with functions in the temptools. Finally, functions for statistics

and displaying are contained in the `disptools` file. The main routine `"oh-data"`, and the function which perform the model temperature calculation, `"get_temp_subbg"`, are shown in the appendix C.1 and C.2.

In the following, the main routine will be explained, and the essential sub-algorithms will be explained along with it.

4.6.1 The main routine

The main routine starts with extracting the signal from the data files produced by the acquisition program. Dark signal scans are picked with a flag in the data file. These are averaged, and the fixed pattern is found with a highpass filter as described in section 4.1, using the FFT (fast fourier transform) function in NumPy. The crucial part of the import signal routine is shown below. Two backgrounds are averaged for the signals in between.

```
#Averaging dark signal
avgds = numpy.zeros(1024)
for i in range(1024):
    avgds[i] = numpy.sum(darksig[:,i])/nds

fp = ohtools.highpass_vec(avgds) #The fixed pattern
...
elif(daymat[i,9] == 0 and i < numofscans-1):
    n = i+1
    while(daymat[n,9] == 1 and n != numofscans-1): #Finding next
        background. If it the last scan is not a background,
        take the last.
        n += 1
    if(n == numofscans-1 and daymat[n,9] == 1):
        n = i
        break
    ds = ohtools.lowpass_vec(0.5*(daymat[i,11:1035]+daymat[n
        ,11:1035]))+fp #Estimating dark signal. Taking average of
        two dark signals
```

After the signal is imported, all the individual spectra, excluding the 25 (~30 minutes) first and last scans, and all spectrums with Na-OH ratio above a certain limit, are averaged. This average spectrum is used to derive important information for the rest of the calculations.

Because the average peak positions were found to change from night to night, a peak location algorithm was implemented to avoid lineshape mismatching. The reason why the peak positions move, is most likely because the temperature of the instrument's surroundings change, as there is only a metal plate separating the instrument from the outdoors.

The basics of the peak location algorithm, is to move the instrument function from a predefined position to where the sum of the square residuals of the average spectrum and the instrument function is minimum. This

method was found to be more stable than just picking the position of the maximum within a given range. The essential code is as follows:

```
...
while True:

    k = start+1
    r = spectrum[k+x] - scaled_instr[x] #Residuals
    newsse = 0

    for i in r:
        newsse += i**2

    if newsse < sse: #If sse has decreased, proceed. Else stop.
        start += 1
        sse = newsse
    else:
        break

peaks[index] = start + 14 #Position of the peak
```

As the peak positions move, everything else on the spectrum will also move. Therefore, the peak positions are used to derive the no-signal points, used to estimate the baseline, and also the pixel interval to include in the fit (the band borders). The principle of finding these points, is to locate them from a fixed number of pixels relative to the peaks.

```
def band_borders(peaks):
    #BAND_BORDERS: Returns the borders of the (3,1) and (4,2) band in
    #the following order: left 31, right 31, left 42, right 42

    return [peaks[0]-14, peaks[5]+20, peaks[6]-15, peaks[11]+20]

def basepts(peaks):
    #BASEPTS: Returns the defined no signal points.

    #(3,1) Basepts:
    leftp22 = numpy.arange(peaks[0]-15, peaks[0]-10)
    rightp22 = numpy.arange(peaks[0]+15, peaks[0]+18)
    ...
    return basepts31, basepts42
```

The baseline is estimated by fitting a third degree polynomial through the derived no-signal points in the average spectrum. As already mentioned in section 4.2, the minimum of the fitted polynomial is subtracted since the offset is fit with the Gauss-Newton algorithm.

Next, all the individual spectra are filtered with a 3 point median filter. This will eliminate the values of any hot pixels, which often stick up in the spectrum, and also remove some of the random noise without changing the shape of the lines. The difference between an unfiltered and a filtered spectrum is shown in figure B.

The peaks and band borders are then used to construct instrument function matrices for the (3-1) and (4-2) band. It is the elements $w_{i,J}$, from (4.1) and can be thought of as a frame to construct the models from. The columns represent the distribution of each line intensity within the band, and is determined by positioning the instrument functions so that the peak positions match the ones found from the average spectrum. The algorithm for this is

```
w_31 = numpy.zeros((borders[1]-borders[0], 6))
w_42 = numpy.zeros((borders[3]-borders[2], 6))

for i in range(6):
    k = peaks[i]-borders[0]-14 #Start index for the line
    for j in range(lineshppts):
        w_31[k+j,i] = lineshape[j]
    ...

return w_31, w_42
```

All rows in w are next multiplied with the respective constant factors in the model equation. The mathematical expression for the matrix is

$$W_{ij} = R_i \frac{S_j \nu_j^3}{T_j} w_{ij}. \quad (4.22)$$

The wavelength dependency of the instrument response does not change, so the correct interval of the response curve must found. This is done by picking the interval according to the derived band borders. The code which calculates (4.22) is

```
w_31, w_42 = instrfunc(peaks, borders) #Model frames

for i in range(6):
    w_31[:,i] *= response_kr[borders[0]:borders[1]]*fr3s[i]/trc_winter[i]
    #Multiplies the frames with the constant factors

for i in range(6):
    w_42[:,i] *= response_kr[borders[2]:borders[3]]*fr3s[i+6]/trc_winter[i+6]

return w_31, w_42
```

The program now has all the information it needs to go through all the signal scans, and calculate the temperatures and intensities. Since the model equations are different for the two ways of handling the baseline, which was described in section 4.2, there are two different routines which perform the model fitting. The principles for these are the same, and the only difference is mainly that the routine to fit the baseline shape has an extra row of b_i elements in the jacobi matrix according to (4.13)

The fitting routine starts with constructing an initial model from the input temperature guess, the instrument function matrix described above

and the intensity of the input spectrum. The procedure is as follows. First a "spectrum shape" is calculated by multiplying each row in (4.22) with the respective $e^{-\tau \frac{\epsilon_J}{k_B}}$ values and then summing over each row. The calculation of the partition function is a stright forward routine. An intitial value for ρ is found by scaling the shape of the model to the match the maximum of the spectrum intensity minus the offset guess. The procedure is as follows

```
rho_initial = numpy.max(spectrum - betaguess)/numpy.max(specshape)

tauguess = 1/tguess
qr = partfunc(tauguess)
wmod = expwmat(w, tauguess, shift)/qr #Divide by partition function
specshape = ohmodel(1, wmod) #Summing over each row in wmod

ohband = rho_initial*specshape
model = ohband + betaguess
intens_initial = numpy.sum(ohband)
```

Here, "expwmat" is a function which just multiplies each column in w with the exponential factor. The summing over the rows is done with the seperate function "ohmodel".

With the initial model, an initial sum of squared errors is calculated, and the Gauss-Newton iterations are ready to start. The iterations proceeds until a desired accuracy limit is reached, which is measured as the change in SSE , until a maximum number of iterations, or the normal equations cannot be solved. The jacobi matrix from section 4.3 is constructed in the following way

```
jacobi = numpy.zeros((speclen,3)) #Initialising Jacobi matrix
jacobi[:,0] = specshape
jacobi[:,1] = params[0]*jacobitau(wmod, params[1], qr, shift)
jacobi[:,2] = 1
```

The $\frac{\partial M_i}{\partial \rho}$ elements are identical to the already calculated "specshape" vector. Furtherly, the function "jacobitau" calculates the $\frac{\partial M_i}{\partial \tau}$ -elements according to (4.12) by multiplying the "wmod" matrix with the factor

$$\left(\frac{\epsilon_J}{k_B} + \frac{1}{Q} \frac{\partial Q}{\partial \tau} \right), \quad (4.23)$$

and again summing over each row. By using already calculated factors, the CPU time of the total routine is greatly reduced. Especially saving many calls to the exponential function, makes the algorithm much more computationally efficient.

The normal equations are solved with the function "solve" from the linear algebra library in NumPy.

```
jtj = jacobi.T.dot(jacobi) #Calculating (J^T*J)
try:
```

```

#Calculating (J^T*J)^-1*J^Tr:
invjtjtr = numpy.linalg.solve(jtj,jacobi.T.dot(residuals))
except numpy.linalg.LinAlgError:
    err_message_gn = 'Normal eqn. failure'
    break

```

The changes are added to the parameters as described in section 4.3 and a new model is constructed with the new parameters in the same way as described above. If the Gauss-Newton iterations successfully converged, the errors are calculated by finding the inverse of the $\mathbf{J}^T\mathbf{J}$ with the "inv" function from NumPy, as shown in the following code outline.

```

try:
    covmat = numpy.linalg.inv(JTJ) #Computing covariance matrix
except numpy.linalg.LinAlgError:
    err_message = 'Covariance fail'
    return uncrt, err_message

sigma = numpy.std(residuals) #Std of residuals

temp = 1/params[1]
dtau = numpy.sqrt(covmat[1,1])*sigma #Error in tau
dt = temp**2*dtau
reldt = dt/temp*100
...

```

Errors and values are then checked to be within decided limits. If not, "NaN" values are returned. The Q-branch and Na-intensities are found by integrating under certain pixel intervals and subtracting a constant offset, which is found from the pixels next to the lines. The pixels used are listen in table 4.1.

Table 4.1: Pixels use to integrate Q-branch intensities and Na-intensity.

Branch	Intensity	Base
Q(3-1)	232-275	197-231
Q(4-2)	632-669	592-631
Na	911-1001	877-901

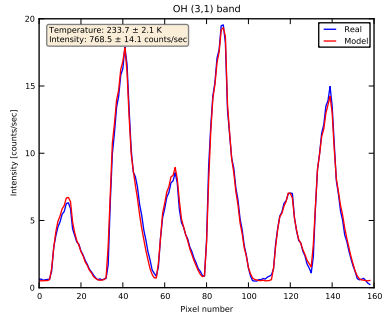
4.6.2 Result of the Gauss-Newton iterations

Figure 4.9 shows the the result of fitting the synthetic model (4.4) to the average spectrum from the night between 14-15th January 2013. As the real spectrum is an average of nearly 750 scans, it is practically a noise free and should therefore give a good representation of the best possible fit and the limits on the accuracy.

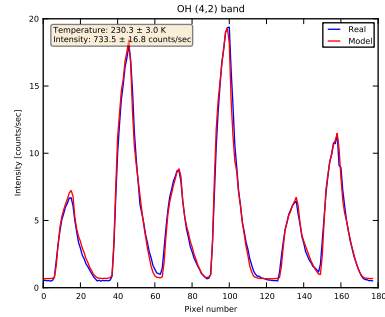
Figure 4.9 shows fitted models to the OH (3-1) and (4-2) spectra. The measured instrument lineshape seems to model lines well, and the response

curve reproduces the smaller details around the peaks. Some mismatch is also clearly seen, and the distribution of these are shown in the plot of the residuals. Especially, the model lie slightly above the measured $P_2(2)$ line for both bands. Between the lines, one sees that the baseline estimation is not 100% correct. The over all mismatch is likely to be caused by contributions from all the factors going into the model (4.1). The error estimate of 2.1 K in the (3-1) spectrum is a typical value of the error estimate for night averaged spectrums. The (4-2) error is typically 0-1 K higher. An error of ~ 2 K, or around $\sim 1\%$ is therefore the estimated best achievable accuracy in the temperature. The statistics of the iterations are shown in table 4.2 and 4.3. For the average spectra, the algorithm successfully converges in the temperature interval 110 to 1000 K. The standard initial value of 200 K should therefore be safe in all cases. Initial guess on β is set to 1.0. With these initial values and the temperature guess updated from the last calculations, the Gauss-Newton algorithm converges in the vast majority of spectrums with 2-3 iterations.

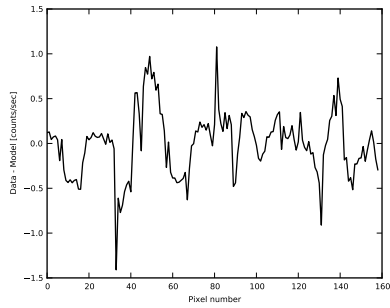
It generally seems as the temperature changes much faster than the error estimate. For example, calculating the temperature of the same (3-1) band without subtracting any baseline results (figure 4.10 in a temperature of 251.3 K and an error estimate 3.6, which is far from the true value. The error estimate is proportional to the standard deviation, which does not take into account that there is a linear trend over the band. It seems therefore only to be a reasonable estimate if the residuals are straight (as in figure 4.9) over the band. This analysis shows that one should look up for effects which can cause this, as incorrect baselines, and some of the lines being affected by H_2O absorption, which was mentioned in section 3.3.



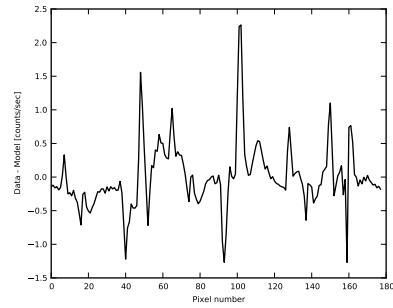
(a) (3-1) result



(b) (4-2) result



(c) (3-1) residuals



(d) (4-2) residuals

Figure 4.9: Plots showing the result of fitting the model and the residuals between the model and the data.

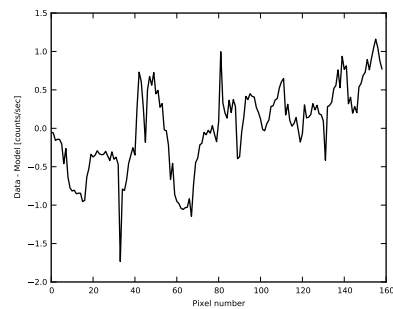
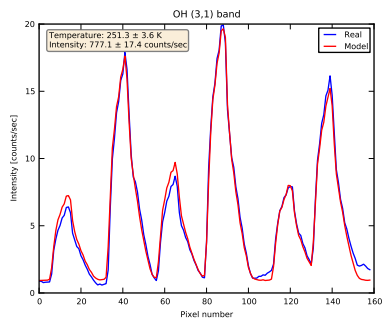


Figure 4.10: Result of calculating the temperature of the same (3-1) band in figure 4.9 without the baseline subtracted. The temperature changes much more than is indicated with the error estimate. The plot of the residuals shows that there is a linear trend over the band. This will not necessarily increase the standard deviation.

Table 4.2: Result of the Gauss-Newton fit on the nightly averaged (3-1) P-branch 14-15th January 2013.

Before iteration:	
Initial temperature:	200.00 K
Initial intensity:	889.40 counts/sec
Initial rho:	3.23e-41
Initial beta:	1.0
Initial SSE:	72.83
After iteration:	
New Temperature:	233.69 \pm 2.13, acc: 0.91% K
Real intensity:	768.47 \pm 14.07 counts/sec, acc: 1.83%
Model intensity:	768.47 \pm 6.85 counts/sec, acc: 0.89%
New rho:	3.635e-41 \pm 3.240e-43, acc: 0.89 %
New beta:	0.503 \pm 0.042, acc: 8.40 %
New SSE:	20.93
Iterations:	3
CPU time:	0.027046 s

Table 4.3: Result of the Gauss-Newton fit on the night averaged (4-2) P-branch 14-15th January 2013.

Before iteration:	
Initial temperature:	200.00 K
Initial intensity:	868.00 counts/sec
Initial rho:	4.70e-41
Initial beta:	1.0
Initial SSE:	70.99
After iteration:	
New Temperature:	230.29 \pm 3.04, acc: 1.32% K
Real intensity:	733.45 \pm 16.75 counts/sec, acc: 2.28%
Model intensity:	733.45 \pm 9.61 counts/sec, acc: 1.31%
New rho:	5.355e-41 \pm 7.018e-43, acc: 1.31 %
New beta:	0.666 \pm 0.050, acc: 7.57 %
New SSE:	20.93
CPU time:	0.030163 s

4.6.3 Calculating nocturnal statistics

The weighted means (4.17) and standard deviation of the means (4.19) are computed with the following code

```
def wgt_mean(values, errors):
    #WGT_MEAN: Calculates the weightheted

    nofvals = len(values)
    sumofw = 0
    wgmean = 0

    for i in range(nofvals):
        if math.isnan(values[i]): pass
        else:
            w = 1/errors[i]**2
            sumofw += w
            wgmean += w*values[i]

    if(sumofw != 0 and wgmean != 0):
        return wgmean/sumofw, numpy.sqrt(1/sumofw)
    else:
        return float('nan'), float('nan')
```

The wavepower calculation routine starts with counting the number of 'nan' values in the data set. If the number of 'nan' values is greater than 1/3 of the total number of points, the routine does not proceed. The next step is to assign a minute number to the each datapoint, with the first usable data point as the zeroth minute. Time jumps in the data set is chekced by calculating the time difference between each data point as the code below shows.

```
minutes = [] #Number of minutes since start
newdata = [] #New datavector
m = 0 #Minute count

for i, time in enumerate(timevec):

    if i < len(timevec)-1:
        td = timevec[i+1]-time #Time between data points

    if td.seconds >= 120: #Check for a background scan between
        m += 1

    if math.isnan(data[i]):
        pass
    else:
        minutes.append(m)
        newdata.append(data[i])

    m += 1
```

The mean is then subtracted from dataset, and is next windowed with a Hamming window to remove the "foots" at the beginnning and the end of

the dataset. This is done with the function "hamming" in NumPy. The frequencies to be calculated is determined by the equation in section 4.5. They are converted to angular frequencies and periods with the code below.

```
freq = numpy.linspace(minfreq, maxfreq, nout) #Linear spacing of "nout"  
      frequencies bewtween the maximum and minumum frequency  
w = 2*numpy.pi*freq  
periods = 1/freq[::-1] #Reversing frequency vector
```

The Lomb-Scargle periodogram is calculated with the "lombscargle" routine in the ScyPi signal module. This function returns the non normalised periodogram. To normalise it, each value must be transformed by $\sqrt{4 \times pdg_i / N}$, where pdg_i is the value of non normalised periodogram and N is the length of the dataset.

Chapter 5

Results

5.1 Nightly averaged temperatures and intensities

Figure 5.1 and 5.2 shows the calculated nightly average rotational temperatures and intensities for the OH (3-1) and the (4-2) transitions in the period 24th November 2012 - 27th April 2013. The results are obtained by excluding every calculated temperature with errors exceeding 5% or 10 K and 10% or 50 counts/sec on the intensities. Also, every spectrum where the Na-OH-ratio is greater than 1.5 is disregarded.

The measurement period starts with a relatively warm period with temperatures around ~ 220 K in the last part of November before it cools down to ~ 200 K for the most of December. During the days between 10-14th January 2013, the temperature rises 37 K from below 200. It reaches the winter record on 25. December with a (3-1) temperature of 237 K and 235 K (4-2) temperature. After this warm period, the temperature decreases steadily towards summer with the (3-1) temperature reaching 175 K on 27. April 2012. A straight line fit between the dates from 1. Jan - 27. April yields a daily decline of 0.38 K for the (3-1) temperature and 0.35 K for the (4-2). A plot of the line fits is shown in figure B.1. The average winter temperature (24. Nov - 31. March) is 209.1 K for the (3-1) and 207.6 K for the (4-2). Standard deviations in the same period are 11.7 K and 10.8 K respectively.

For the vast majority of the days, the (3-1) and the (4-2) temperature follow each other closely with the (3-1) typically 1-4 K higher. The systematic exception seems to be the last part of April where most of the days have a slightly higher (4-2) temperature.

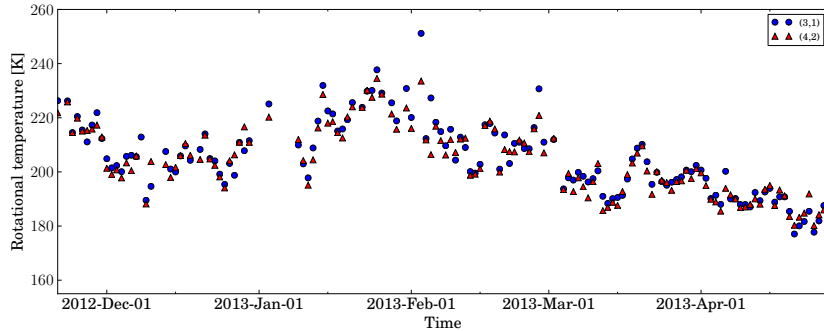


Figure 5.1: Nightly averages of the OH (3-1) and (4-2) P-branch rotational temperatures between 24th Nov 2012 and 27th April 2013.

The nightly averaged OH intensities show many of the same features as the temperature plot. It is relatively calm in December before the intensities rises again in January. In the period until March the intensity shows huge day to day variations. During March and April, the intensity varies more smoothly, and seem to follow more or less the same pattern as the temperature variations, but it does not show the same decline towards summer. The (4-2) intensity is steadily below the (3-1) intensity except for some nights.

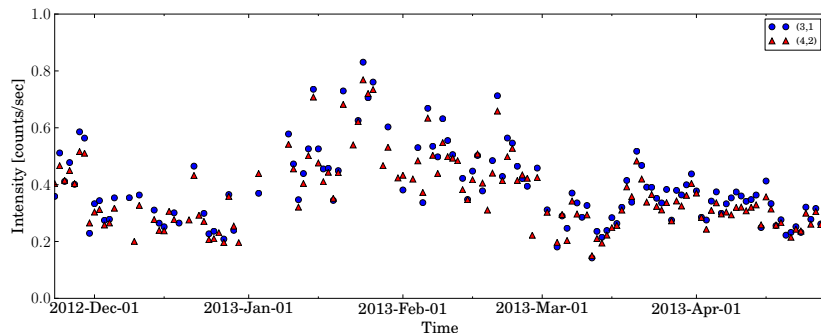


Figure 5.2: Nightly averages of the OH (3-1) and (4-2) P-branch intensities between 24th Nov 2012 and 27th April 2013. The intensity scale is 0-1000 counts/sec.

In figure 5.3 is the (3-1) temperatures plotted together with temperatures calculated with the MSISE 1990 model [9]. The model is empirical and describes the neutral temperature and densities in the Earth's atmosphere. Temperature is calculated with a height resolution of 1 km and a time resolution of 2 minutes. It is averaged in altitude by weighting with a Gaussian profile centered at 87 km with about 8 km FWHM to simulate the distribution of OH with altitude. The temperatures are calculated for the

year 2012 and are repeated into the beginning of 2013. From the plot, one sees that the derived OH (3-1) temperature follows the trend of decreasing temperatures in December, but the model shows no trend of the large increase in the January temperatures. After the warm period, the OH (3-1) declines in relatively good agreement with the MSISE model.

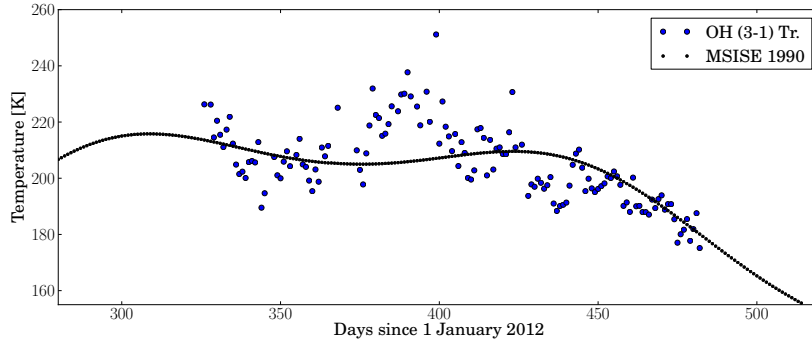


Figure 5.3: The derived OH (3-1) rotational temperatures plotted together with neutral atmospheric temperatures calculated with the MSISE 90 model for the year 2012.

A photochemical equilibrium model has been used to calculate altitude distributions for the volume emission rates for of different OH vibrational states. From this model, which is plotted in figure B.3, it is derived that the altitude difference of the $v = 4$ and $v = 3$ vibrational states emission rates is ~ 1 km. Figure 5.4 shows altitude profiles of the atmospheric neutral temperature calculated with the MSISE profile for 14th January and 25th April 2013 at 01:00 UTC. The temperature gradient in the 85-90 km region is approximately -2.1 K/km and -1.2 K respectively (linear fit between 86 and 89 km) and also thereby giving the estimates of the temperature difference between the measured (3-1) and (4-2) rotational temperatures.

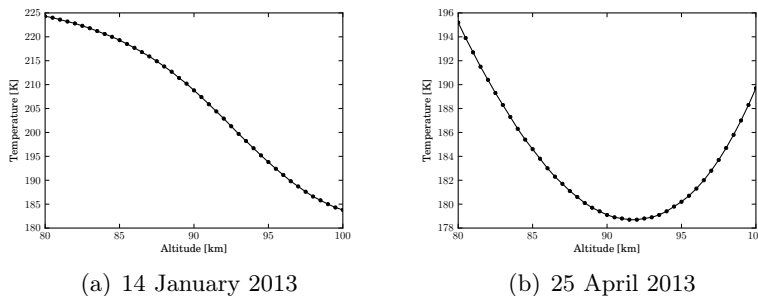


Figure 5.4: Temperature profiles for two different days calculated with the MSISE model for the altitude range 80-100 km.

5.2 Errors

Typical errors of the nightly averaged temperature is 0.12-0.20 K mid winter when most of the scans from each night are usable. At the end of April, these errors have increased to 0.20-0.40 K as the measurement duration decreases towards summer.

Relative error in the (3-2) temperature is plotted against intensity for all the calculated temperatures used to produce the nocturnal means. The plot shows that the intensities are lie in the range 200- 800 counts/sec with an errors in the range 1-2%. It also shows that the error never reaches zero but hits a floor just above 1%. This agrees well with the error calculated from night averaged spectrums in section 4.6.2, where the noise is negligible. The cluster at 500 counts/sec and 3% error is caused by the the two nights between 7-8. December 2012 where the instrument was run with 15 sec integration time which produced around 2500 spectra each night. The same plot with these days excluded is shown in figure B.2.

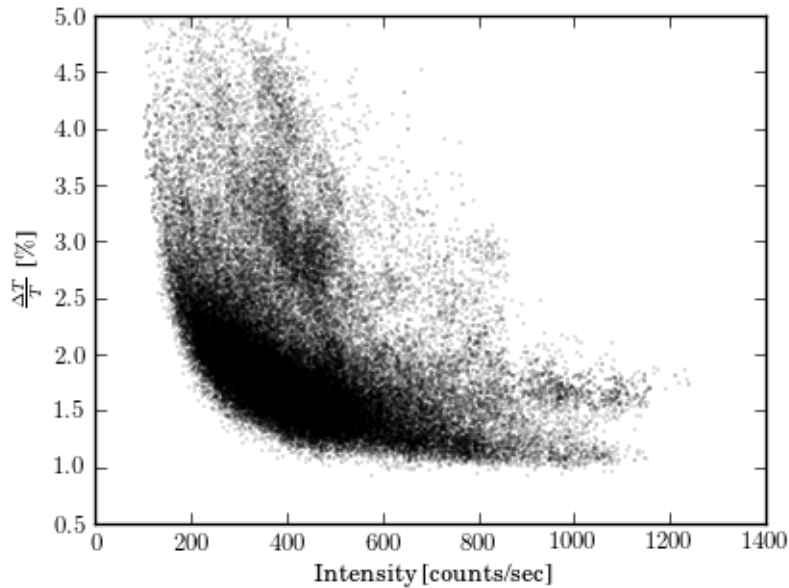


Figure 5.5: Relative error in the (3,1) P-branch rotational temperature as a function of intensity. All individual calculation from all days between 24th November 2012 - 27th April 2013 are included.

The same plot with logarithmic axes is shown in figure 5.6. This emphasises the effect of fitting a non-linear model with errors increasing exponentially. The cut off at 5% is clearly seen. Both figure 5.5 and 5.6 shows that there is not much data with errors higher than 5% and that these are spread on all intensities.

Figure 5.7 shows the same type of plot with data from similar measurements on Rothera, Antarctica. The integration time on this instrument was 15 minutes. It shows the same non-linear relation between the intensity and the temperature error.

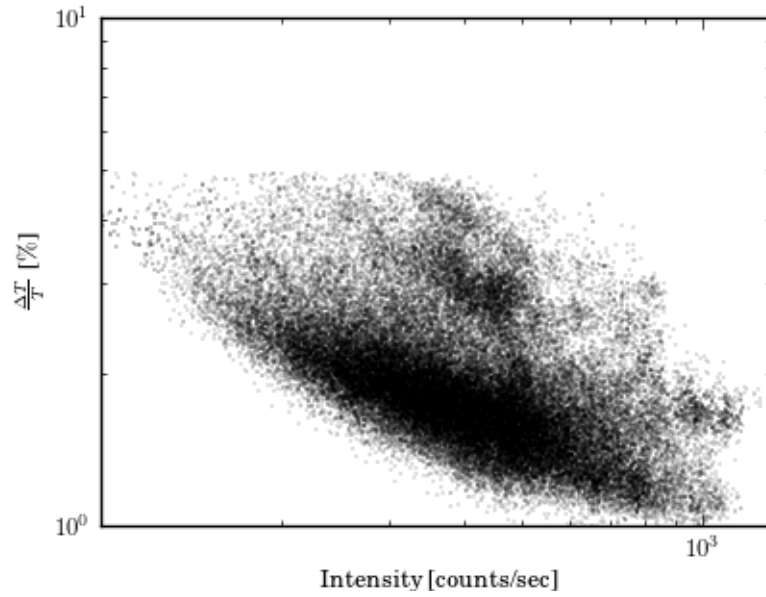


Figure 5.6: Log-log plot of the relative error in the (3,1) temperature as a function of intensity.

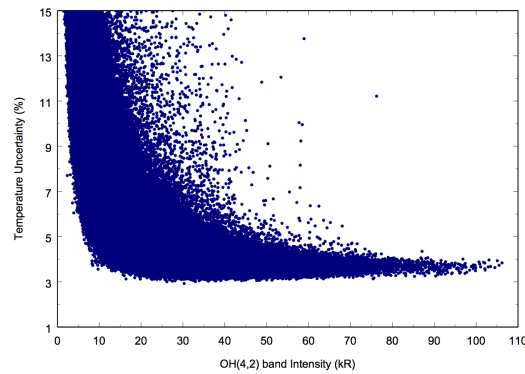


Figure 5.7: Uncertainty in temperature vs. Intensity for measurements of the OH (4,2) P-branch rotational temperature on Rothera, Antarctica. The integration time on the instrument was 15 minutes.

5.3 Gravity waves

In this section, three nights of OH emission measurements are described in more detail with a focus on gravity waves. In figure 5.8, 5.11 and 5.14 the OH (3-1) rotational temperature (blue) is shown together with the intensities of the OH (3-1) and (4-2) P- (blue, red) and Q-branches (green, purple) and the Na intensity (black). The black curve in (3-1) the temperature plot is 5 minute smoothed with weighted averages. The (3-1) and (4-2) temperature are plotted together in figure 5.9, 5.12 and 5.15 to show the covariation of the two temperatures and intensities. Relevant period intervals of the Lomb-Scargle periodogram for the nights are shown in figure 5.10, 5.13 and 5.16. Mentioned periods are marked with red lines. Also, the statistics for each night are listed in table 5.1, 5.2 and 5.3. Relative change in intensity over relative change in temperature is calculated as

$$\eta = \frac{(I_{max} - I_{min})/(I_{max} + I_{min})}{(T_{max} - T_{min})/(T_{max} + T_{min})}. \quad (5.1)$$

5.3.1 29-30th November 2012

This night is belongs to the relatively warm period at the end of November 2012. It shows clear periodic variations with long and short periods in both the intensity and the rotational temperature. The intensity variation is dominated by a large scale 2 hour period oscillation which the temperature in some sense seem to follow. Special for this night are the very distinct wave pattern in the rotational temperature seen in both bands between 19:00 and 21:00. The variations have amplitude of ~ 10 K and period of ~ 30 min. In the same time interval, a similar clear wave pattern cannot be seen in the intensity, but there is some structure and their peaks and periods seem to match the temperature variation. In the last large variation between 22:45 and 00:45 the relative intensity change is nearly 5 times greater in the intensity than in the temperature.

Band	(3,1)	(4,2)
$\langle T \rangle$	221.87 K	217.26 K
σ_T	0.13 K	0.17 K
$\langle \Delta T \rangle$	3.13	4.08
$\langle I \rangle$	563.73 c/s	510.52 c/s
Std. of T	6.67 K	6.42 K
Std. of I	76.27 c/s K	65.62 c/s

Table 5.1: Nocturnal statistics 29-30. February 2012

The Lomb-Scargle periodogram of the temperature series is plotted in the 1-300 minute period region. Two periods have been marked at 33 and 196

min. As expected, most of the power is comes from the ~ 3 hour oscillation, but the ~ 30 minute is also clearly seen.

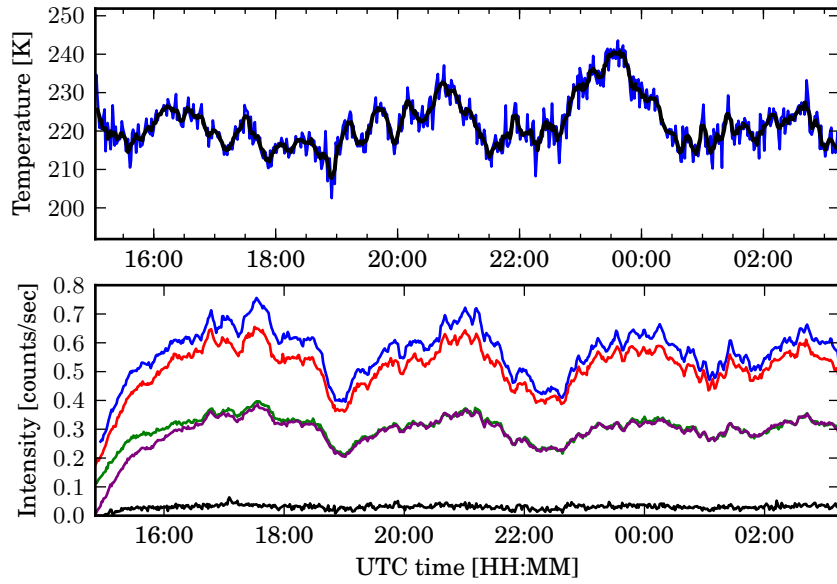


Figure 5.8: Plots of the OH (3-1) rotational temperature and the corresponding intensity variations during the night of 29-30. November 2012.

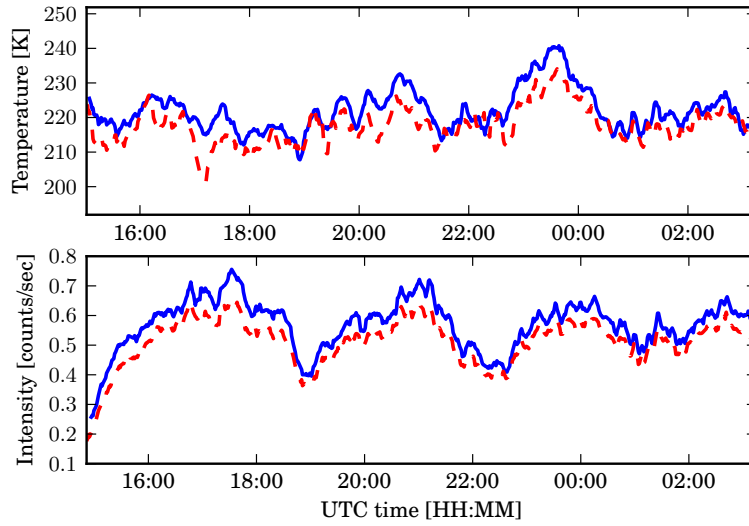


Figure 5.9: 5 minute OH smoothed (3-1) and (4-2) temperature series and their corresponding intensity variation during the night of 29-30th November 2012.

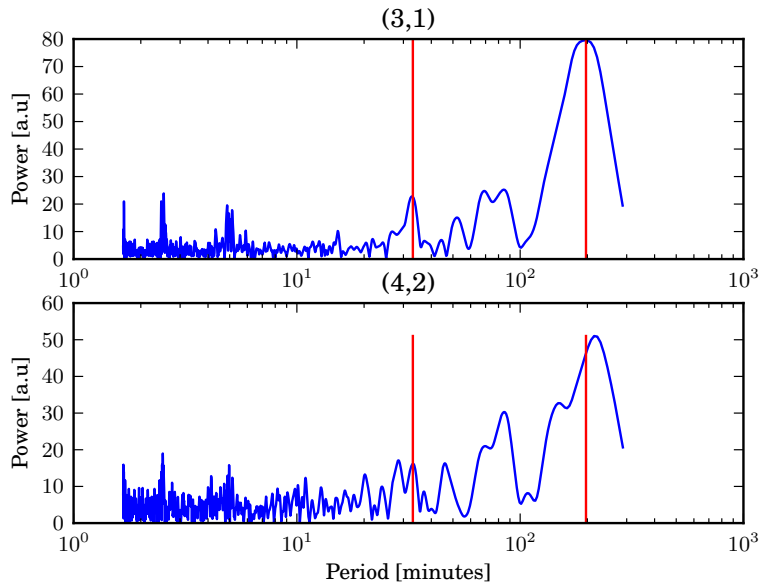


Figure 5.10: Lomb-Scargle periodogram of the OH (3-1) and (4-2) rotational temperature 29-30th November 2012 in the 1-300 minute region. The two marked periods are 33 and 196 minutes.

5.3.2 7-8th February 2013

A long period large amplitude variation in both rotational temperature and the intensity shows up in the OH emissions this night. For the largest of the variations between 21:00 and 03:00 the temperature rises ~ 47 K from 188 K while the intensity rises from 350 to 900 detector counts/sec. This gives a relative change in the intensity ~ 4 times greater than for the rotational temperature. The intensity and the temperature seem in general to be co-varying, but the temperature starts to rise almost 2 hours before the intensity in the largest of the variations.

Band	(3,1)	(4,2)
$\langle T \rangle$	214.88 K	211.53 K
σ_T	0.16 K	0.17 K
$\langle \Delta T \rangle$	4.42	5.04
$\langle I \rangle$	534.96 c/s	503.23 c/s
Std. of T	12.62 K	10.66 K
Std. of I	172.76 c/s	143.77 c/s

Table 5.2: Nocturnal statistics 7.-8. February

The periodogram is plotted for periods between 10 and 600 minutes. It is obviously dominated by the large scale variations mentioned above. The (3-1) power peaks at a period of 410 minutes. A peak of 138 minutes is also marked, corresponding to the smallest of oscillations seen in the rotational temperatures.

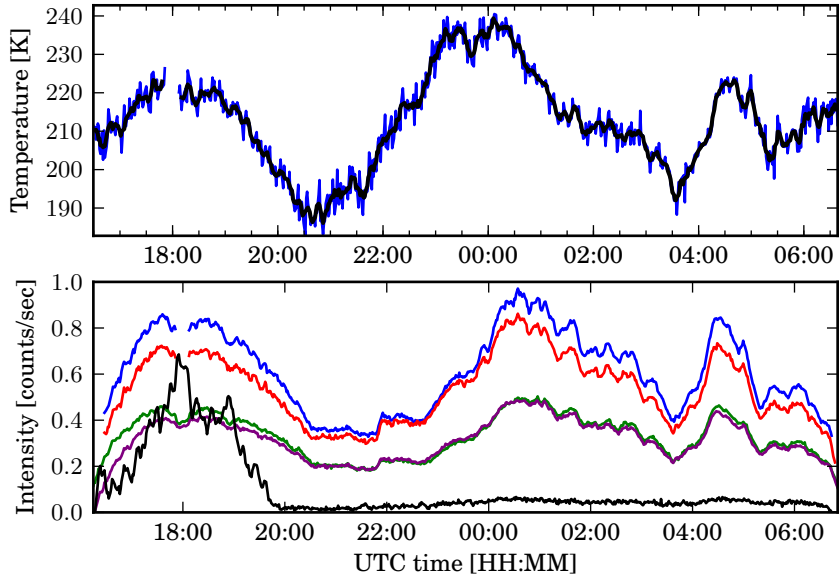


Figure 5.11: Plots of the OH (3,1) P-branch rotational temperature and the corresponding intensity variations during the night of 7-8th February 2013.

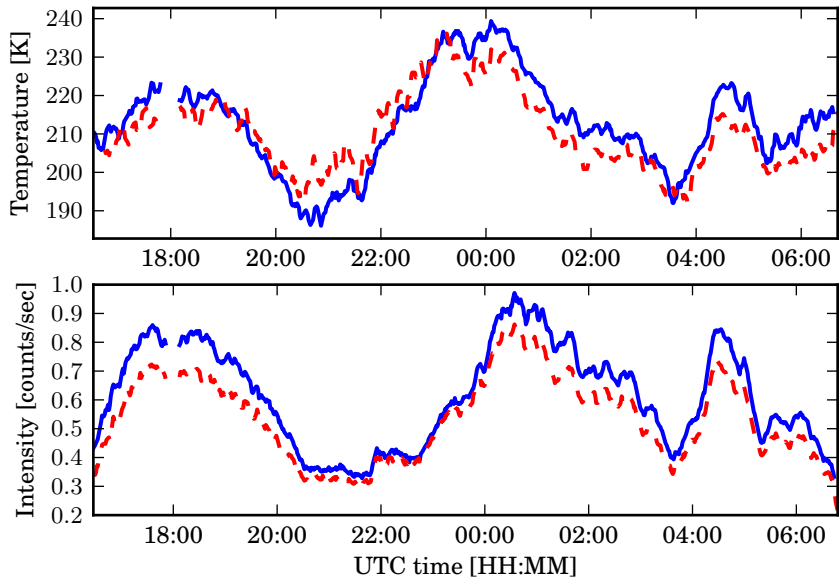


Figure 5.12: 5 minute smoothed OH (3-1) and (4-2) temperature series and their corresponding intensity variation during the night of 7-8th February 2012.

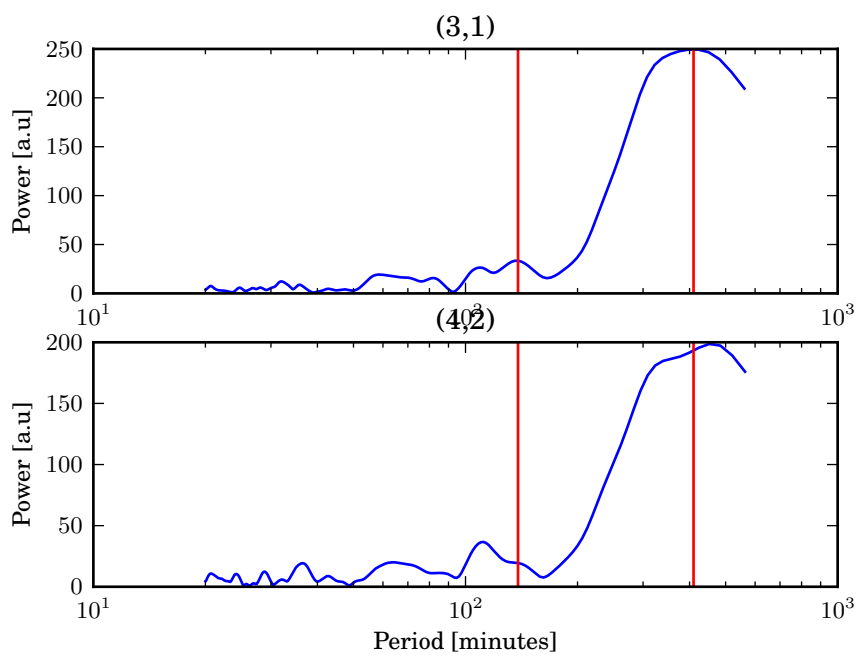


Figure 5.13: Periodogram of the OH (3-1) and (4-2) P-branch rotational temperature 7-8. February 2013. The red lines mark periods of 138 minutes and 410 minutes.

5.3.3 19-20th February

A clear, persistent wave pattern can be seen in the OH airglow emissions this night. The oscillations start around 22:00 and last until 03:00. The period of the variations are ~ 1 hour. They start with a relatively large amplitude of ~ 15 K in the temperature and ~ 250 counts/sec in the intensity and then gradually fade away. This gives a relative change in the intensity around 5.5 times greater than in the temperature. The timeseries plots show that the temperature variations are clearly in phase with the intensity. The co-plot of the (3,1) and (4,2) temperatures shows that the (4,2) temperature follows the same pattern, but it is not so clear for the last oscillations.

Band	(3,1)	(4,2)
$\langle T \rangle$	201.01 K	199.94 K
σ_T	0.12 K	0.13 K
$\langle \Delta T \rangle$	3.01	3.33
$\langle I \rangle$	484.53 c/s	440.41 c/s
Std. of T	7.63 K	5.82 K
Std. of I	94.52 c/s	83.24 c/s

Table 5.3: Nocturnal statistics 7.-8. February

The ~ 60 minute period sticks well up in the periodogram, but it also shows that there is almost as much power coming from a 30 minute oscillation. Most of the power comes from a ~ 85 minute variation.

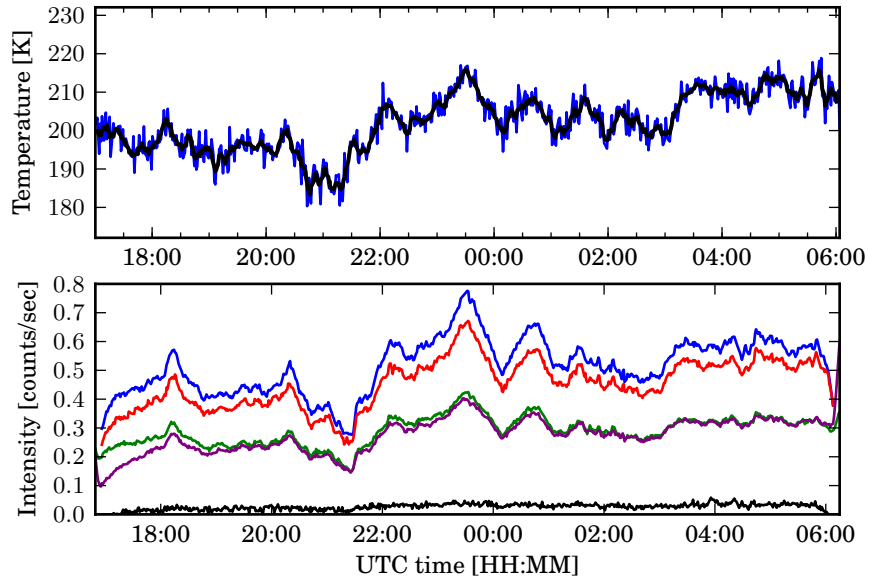


Figure 5.14: Plots of the OH (3-1) rotational temperature and the corresponding intensity variations during the night of 19-20th February 2013.

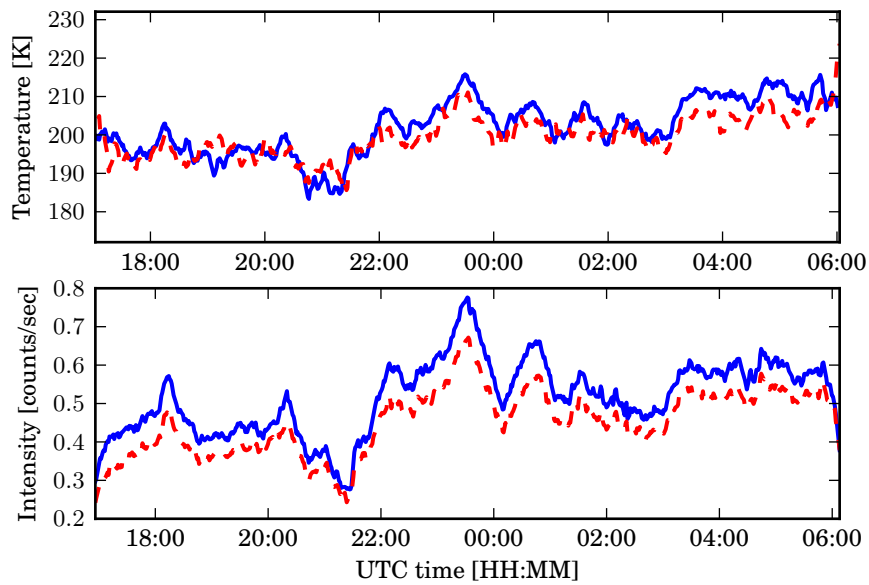


Figure 5.15: 5 minute OH (3-1) and (4-2) smoothed temperature series and their corresponding intensity variation during the night of 19-20th February 2012.

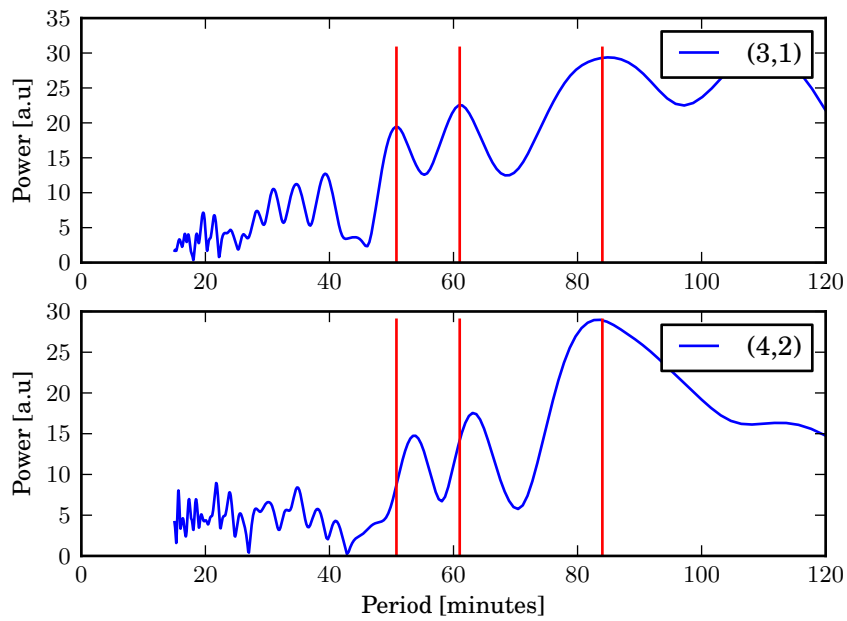


Figure 5.16: Periodogram of the OH (3-1) and (4-2) P-branch rotational temperatures for periods between 15 and 120 minutes 19-20th February 2013. The red lines mark periods of 32, 62 and 85 minutes.

5.3.4 Variance of wave activity

Figure 5.17 shows the development of the wavepower estimated with the method described in section 4.5. Both the wavepower in the temperature and the intensity shows very large day to day variations but also a clear trend with decreasing wave activity towards summer.

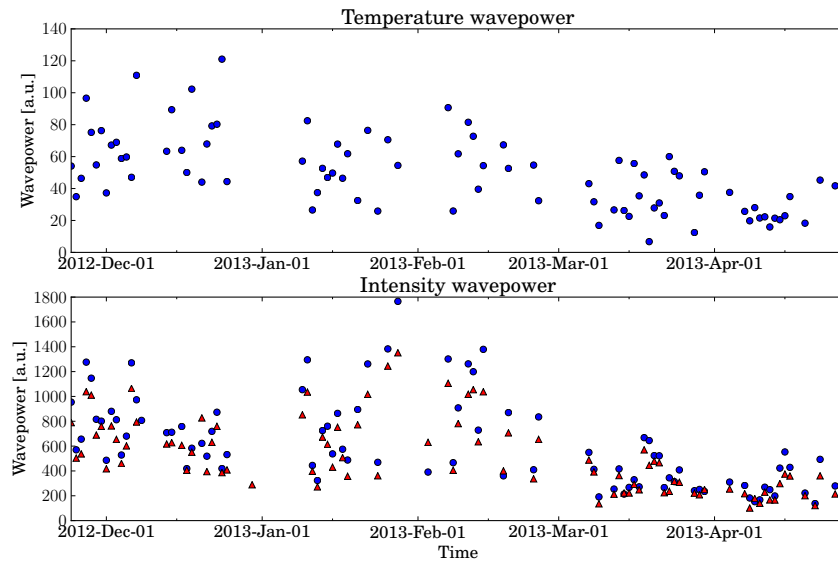


Figure 5.17: Estimated wavepower in the OH (3-1) rotational temperature and in the OH (3-1) and (4-2) P-branch intensities from 24th Nov 2012 - 27th April 2013.

Chapter 6

Discussion

6.1 The results

Although the derived OH temperatures is following the climatological trend with warm winter temperatures and cooling towards summer, in agreement with theory and the MSISE model, there are also large and important differences. The observed large day to day variations, and the high January temperatures are not predicted by the model. On the 6th of January occurred a large stratospheric warming, which are associated with a cooling of the mesosphere [21]. Because of bad sky conditions, there are no usable data in the period 1-9 January, and the possible effect is therefore not observed in the mesosphere over Trondheim. If this event could be linked up with high January temperatures, is an interesting study. High January temperatures (220-230 K) have been observed earlier [7][17], and are also seen in the results from Andøya (69.3°N), which is shown figure 6.1 (C. Schmidt, personal communication).

The (4-2) temperature being slightly lower than the (3-1) is consistent with the model predicted centroid height of the $v = 4$ emission distribution being higher in altitude and therefore colder, according to temperature profiles calculated with the MSISE model. The difference between the mean (3-1) and (4-2) temperatures of 1.5 K seems therefore not to be too far from reality. As the baseline is fitted individually for the (3-1) and (4-2) P-branches, close values of the two temperatures is a good sign that they are reasonable values, with respect to the discussion in section 4.6.2.

The nightly variations in the OH airglow layer presented in section 5.3, show a wide range of periods and amplitudes in both the intensity and the temperature. In general the, covariation of the intensity and the temperature are well correlated in time. The observed phase differences between the variations does not show any consistent pattern on the phase lag between the temperature and the intensity variations.

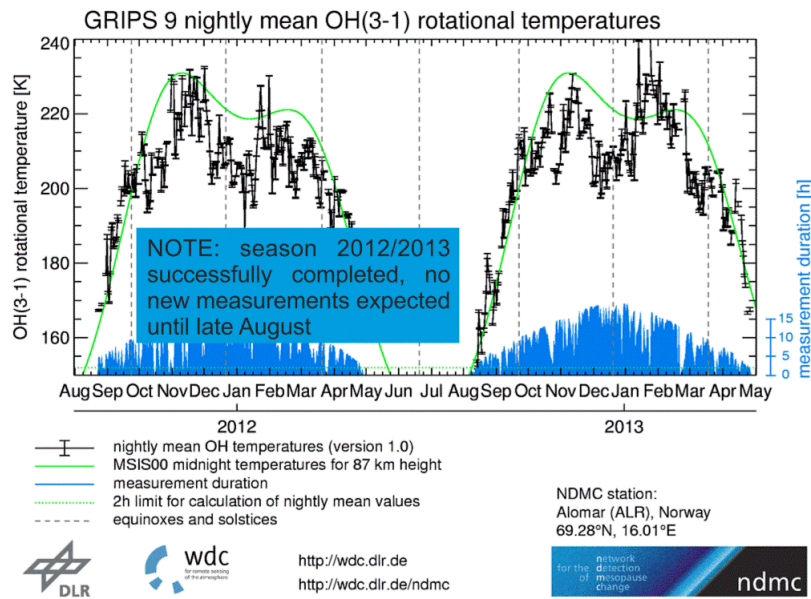


Figure 6.1: Results from measurements of the OH (3-1) rotational temperature over Andøya for the two seasons 2011/2012 and 2012/2013.

The amplitude of ~ 47 K peak to peak in the night between 7-8 February, is very large compared to what is observed on the other nights and compared to the observations mentioned in section 2.8, but the corresponding large intensity variation gives a value for the relative change between the intensity and temperature about the same as for the other nights, which is in the order ~ 4 -5.

The time series plots show that the (4-2) band experiences the same dynamics as the (3-1), in good agreement with the two layers being close in altitude. The intensity variation of the two bands are almost identical, although their relative difference varies during the night. The fitted integrated intensities of the P-branches also match the raw integrated Q-branch intensities. The temperature series show for all three nights good agreement between the (3-1) and (4-2) rotational temperature, but they are not so easy to compare on the smaller variations because the errors are higher for the (4-2). This is the effect of the (4-2) P-branch being affected by the lower instrument sensitivity. Therefore, the (3-1) temperature is a better source for future detailed analyses, but the (4-2) serves as a good backup for the observations.

The wavepower results is a very rough estimate, and is not discussed further here. It is stated here as a reference for future calculations and observations.

6.2 The analysis scheme

The Na-OH ratio used to filter out data that are probably affected by clouds is very debatable. The parameter value of 1.5 used to obtain the data in the results presented here, is decided from general experience and not from a systematic analysis. Some usable data is probably filtered out and some data affected by clouds is likely to have been included. The intensity of the (3-1) P-branch is plotted against Na-intensity in figure 6.2 for a data set without cloud filter. As can be seen, there is no clear relation between the Na-intensity and the effect on the (3-1) band except from the higher intensities clearly dropping off when the Na-intensity increases. The reason why the intensities are distributed along the very steep, distinct line is because the Na-lines sits in the middle of an OH R-branch. Therefore, counts in this region increases linearly with increasing OH intensity. For the dataset used in the results section, the (3-1) intensity is plotted vs. the Na-OH ratio in figure 6.3. Again, it is not obvious on which value of the Na-OH ratio to start excluding data. A plot of nocturnal temperatures obtained from the dataset with no cloud filter and a temperature error limit of 7.5% is shown in figure 6.4. The Na-intensity is not a perfect way of filtering out unreliable spectrums, but by comparing this plot with the one in the results section, one sees that most of the obvious outliers have effectively been filtered out while most of the other data is unchanged. Clouds are by far the most disturbing phenomena for the measurements, and the Na-filter therefore effectively excludes most of the bad spectrums. The remaining obvious outliers, which stand out with very high difference between the (3-1) and (4-2) temperature, are the dates 3, 5 and 26 February. Common for these nights is that the raw integrated (4-2) Q-branch intensities lie above the (3-2), indicating absorption of the (3-1) by water. Intensity plots of these nights are shown in figure B.4.

As both the temperature and the intensity shows large variations over the course of a night, it is obvious that when data is excluded, the calculated averages can change sharply from night to night and explain some of the large day to day variation.

The error limit on the temperature, which was set to 5%, was decided from a point of view that typical errors lie within the range 3-4 K and three times these values gives errors from 9-12 K. The additional limit of 10 K was set on the temperature because higher temperatures are often associated with higher intensities, leading to better signal to noise ratios. By setting this additional absolute limit, the limits on the relative error is gradually increased as the temperature goes beyond 200 K. As described in section 4.6, there are situations where Gauss-Newton fit is not reliable, and the error estimate from the covariance matrix and the residuals obviously do not give a sensible picture of the actual error. This is why the temperatures calculated with the algorithm 3 February, which reached ~ 260 K and with

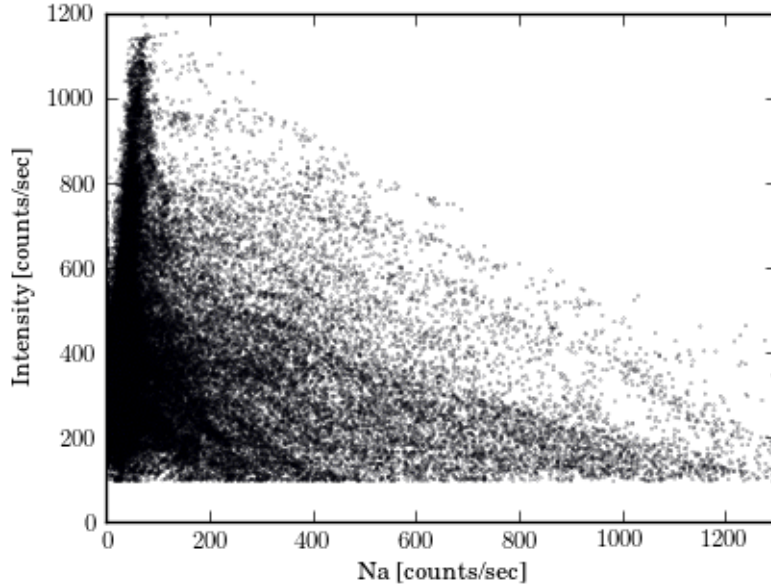


Figure 6.2: OH (3-1) P-branch intensity as a function of Na intensity (cloud indicator).

error estimates of around 6 K, were not covered by the analysis scheme. Figure 5.5 shows that above $\sim 4\%$, the errors are spread more evenly on the intensities and that there is not much data above 5%. The lower limit on valid intensities was set to 100 counts/sec. Figure 6.2 and 6.3 shows that intensities below this value is unlikely when spectrums with high Na-level are excluded.

An issue which was described in section 4.6, was that the peaks were found to change position from night to night on the detector, most likely because the temperature of the instrument surroundings changes. This means that when the peaks are found from the average spectrum for each night, they might be systematically wrong for periods of the night. Perhaps could this be solved by making the algorithm more dynamic by updating the peak positions during the night. Ideally, the surroundings temperature should have been constant. Whether or not it is an option or if it at all makes sense to for instance insulate the instrument chimney is a relevant discussion in the future of these measurements.

The analysis scheme is developed to deal with more or less constant baselines for reasons that were explained in section 4.2. If the baseline changes significantly during the night because of increase in scattered blackbody radiation, for instance by moonlight, the potential errors are large (section 4.6.2). The method of subtracting the offset from the spectrums described in section

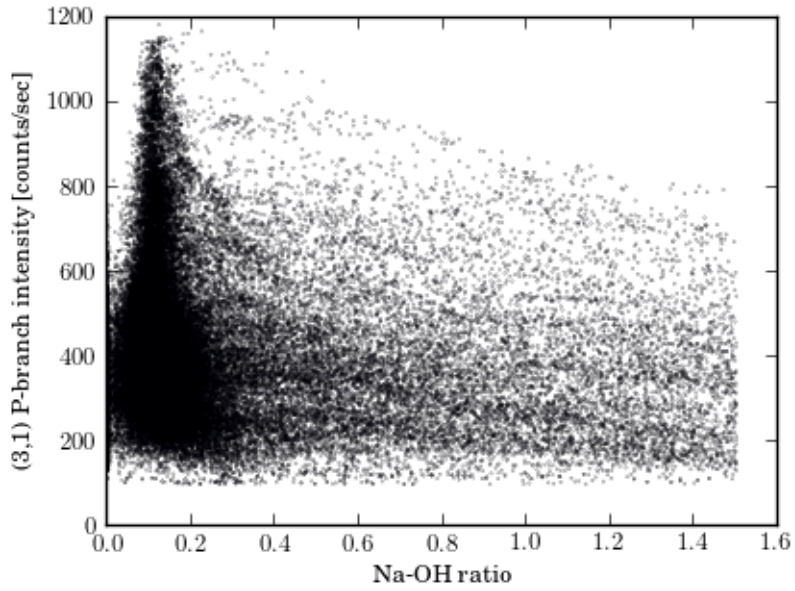


Figure 6.3: Intensity of the (3-1) P-branch vs. the Na-OH ratio.

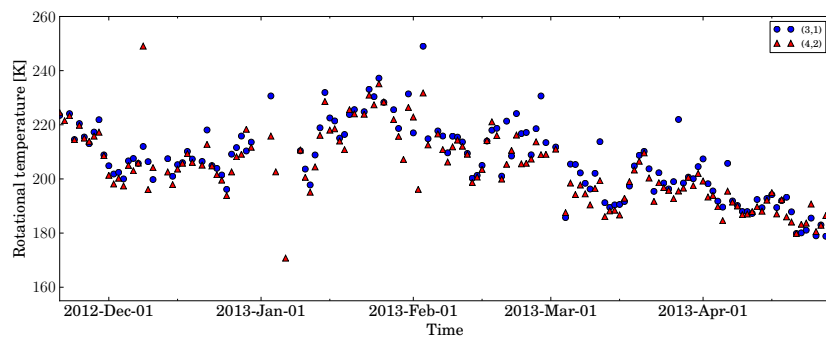


Figure 6.4: Nightly averaged temperatures obtained from a dataset with no cloud filter. Compared to figure 5.1, more outliers are produced.

4.2 is only based on the no-signal points from the instrument function and the assumption that the underlying atmospheric profile changes slowly over the band. A more general way of dealing with the offset, would be to obtain an estimate of the actual underlying background profile. A way to do this could be to scan the atmosphere with the spectrometer just after the sunset or before sunrise, when the spectrum is nearly OH free. Perhaps could this make the model more adaptable to changes in the baseline, which would make it possible to get usable fits closer to sunrise/sunset and to account for changes in the background of scattered blackbody radiation.

Chapter 7

Conclusion

Ground based measurements of the hydroxyl airglow over Trondheim (63.4° N, 10.4° E) in the ~ 1450 -1650 nm region started 20th November 2012. Measurements for the season 2012/2013 successfully completed on 27. April and a complete analysis scheme has been developed to derive the rotational temperatures of the (3-1) and (4-2) transitions from the recorded spectrums. The method is based on fitting a synthetic model based on a measurement of the instrument lineshape, using a non linear least squares fitting routine.

The analysis scheme has been used to analyse data for the whole measurement period. Comparisons to model and theory, show that the nightly averaged temperatures follow the expected climatological pattern of warm winter temperatures and cooling towards summer. The most profound exception from the model, are the high January temperatures of ~ 230 K, but the same high temperatures is also observed at Andøya (69°N). On a general basis, the (3-1) and (4-2) temperatures have approximately the same values. The average difference of 1.5 K colder (4-2) temperatures agrees well with model predictions of the $v = 4$ vibrational state intensity distribution being higher in altitude and therefore colder.

Temperature and intensity series from three nights have been analysed, revealing large fluctuations in the hydroxyl layer. A very large peak to peak temperature variation of ~ 47 K was observed on 7-8 February, but the magnitude of the corresponding intensity variation is consistent with the observations from the two other nights, with relative changes in the intensity ~ 4 -5 times larger than in the temperature. The high degree of covariation of the (3-1) and (4-2) variations is in good agreement with the two layers being close in altitude and should experience the same dynamics.

The relative error in the derived (3-1) temperature is for the vast majority of the individual measurements in the range 1-2%, this is a good result compared to the maximum precision found to be $\sim 1\%$. The results are not so good for the (4-2), because the P-branch is affected by a lower instrument sensitivity.

Although the results indicate that the algorithm produces realistic results and to a good precision, this work also reveals that there are some unresolved issues. These are related to the removal of the baseline, which the derived temperature seem to be very sensitive to. It is suggested that the baseline is removed with an estimate of the actual atmospheric background profile. This will probably make the model more adaptive to changes caused for instance by moonlight. It is also pointed out that the wavelength scale of the instrument is changing, most likely caused by the varying temperature in the instrument's surroundings. The effect on the derived temperatures, and if there are possible solutions, should be investigated in the future of these measurements. With these issues solved, the algorithm has the potential of being a very robust method to remote sense mesospheric temperatures.

Bibliography

- [1] Andrews, D.G., (2012) An Introduction to Atmospheric Physics. 2nd. edn., Cambridge University Press, UK.
- [2] Baker, D.J., Stair, A.T. Jr., (1988) Rocket Measurements of the Altitude Distribution of the Hydroxyl Airglow. *Physica Scripta*, 37, 611-622.
- [3] Berge, F.T., (2011) Development of a spectrometer system to remotely sense mesospheric temperature. MSc. Thesis, Norwegian University of Science and Technology.
- [4] Bransden B.H., Joachain C.J., (2000) Quantum Mechanics. 2nd. edn. Prentice Hall, UK.
- [5] Brasseur G. P., Solomon S., (2005) Aeronomy of the Middle Atmosphere. 3rd. rev, Atmospheric and Oceanographic Sciences Library, 32.
- [6] Espy, P. J., (1986) A Spectroscopic Investigation of the Infrared Molecular Band Systems of N_2 and N_2^+ Resulting From Low Energy Electron Impact Excitation. PhD thesis, Utah State University, USA.
- [7] Espy P.J., Stegman J., (2002) Trends and variability of mesospheric temperature at high latitudes. *Physics and Chemistry of the Earth*, 5, 543-553.
- [8] Espy P.J., Hammond M.R., (1995) Atmospheric transmission coefficients for hydroxyl rotational lines used in rotational temperature determinations. *Journal of Quantum Spectroscopy and Radiative Transfer*, 54, 879-889.
- [9] Hedin A.E. (1991), Extension of the MSIS Thermospheric Model into the Middle and Lower Atmosphere, *J. Geophys. Res.* 96, 1159.
- [10] Herzberg, G., (1950) Spectra of Diatomic Molecules. Van Nostrad Reinhold, New York.
- [11] Meinel, A.B. (1950) OH emissions bands in the spectrum of the night sky I, *Astrophys J.*, 111, 555.

- [12] Nappo, C.J., (2012) An Introduction to Atmospheric Gravity Waves. 2nd edn. Academic Press Inc. International Geophysics, 102.
- [13] Oznovich, I., McEwen, D.J., Sivjee, G.G., (1995) Temperature and airglow brightness oscillations in the polar mesosphere and lower thermosphere, *Planet. Space Sci.*, 43, 1121-1130.
- [14] Schmidt C., Höppner K., Bittner M., (2013), A ground-based spectrometer equipped with an InGaAs array for routine observations of OH (3-1) rotational temperatures in the mesopause region. *Journal of Atmospheric and Solar-Terrestrial Physics*, 102, 125-139.
- [15] Marsh, D.R., Smith, A.K., Mlynczak, M.G., Russell III, J.M. (2006) SABER observations of the OH Meinel airglow variability near the mesopause. *J. Geophys. Res.*, 111.
- [16] Press, W.H., Teukolsky, S.A., Vetterling W.T. (2007) *Numerical Recipes*. 3rd edn. Cambridge: Cambridge University Press.
- [17] Sigernes F., Shumilov N., Deehr C.S., Nielsen K.P., SvenÅrve T., Havnes O., (2003) Hydroxyl rotational temperature record from the auroral station in Adventdalen, Svalbard (78ÅrN, 15ÅrE), *Journal of Geophysical Research*, **108**, NO. A9, 1342
- [18] Sivjee, G.G., (1992) Airglow Hydroxyl Emissions. *Planet. Space Sci.*, 40, 235-242.
- [19] Swenson, G.R., Gardner C.S. (1998) Analytical models for the responses for the responses of the mesospheric OH* and Na layers to atmospheric gravity waves, *Journal of Geophysical Research*, 103, 6271-6294.
- [20] Taylor, M.J., Espy, P.J., Baker, D.J., Sica, R.J., Neal, P.C., Pendelton Jr., W.R. (1991), Simultaneous intensity, temperature and imaging measurements of short period wave structure in the OH night glow emission, *Planet. Space Sci.*, 39, 1171-1188.
- [21] Walterscheid, R.L., Sivjee, G.G., Roble R.G., (2000) Mesospheric and Lower Thermospheric Manifestations of a Stratospheric Warming Event Over Eureka, Canada (80°N), *Geophysical Research Letters*, 27, 2897-2900.
- [web, b] <http://www.nist.gov/pml/data/handbook/index.cfm>
- [web, c] Andor technology website, <http://www.andor.com/>
- [web, d] Python programming language official website, <http://www.python.org>
- [web, e] NumPy website, <http://www.numpy.org>

[web, f] SciPy website, <http://www.scipy.org>

[web, g] Matplotlib, <http://matplotlib.org>

Appendix A

Tables

Table A.1: Atmospheric transmission coefficients for the OH (3,1) P-branch lines

Line	High Latitude Winter	High Latitude Summer
$P_1(2)$	0.995	0.977
$P_1(3)$	0.992	0.976
$P_1(4)$	0.994	0.978
$P_1(5)$	0.996	0.980
$P_2(2)$	0.994	0.959
$P_2(3)$	0.970	0.819
$P_2(4)$	0.994	0.978
$P_2(5)$	0.995	0.978

Table A.2: Atmospheric transmission coefficients for the OH (4,2) P-branch lines

Line	High Latitude Winter	High Latitude Summer
$P_1(2)$	0.985	0.970
$P_1(3)$	0.992	0.974
$P_1(4)$	0.996	0.979
$P_1(5)$	0.993	0.972
$P_2(2)$	0.985	0.936
$P_2(3)$	0.990	0.976
$P_2(4)$	0.995	0.979
$P_2(5)$	0.996	0.979

Appendix B

Plots

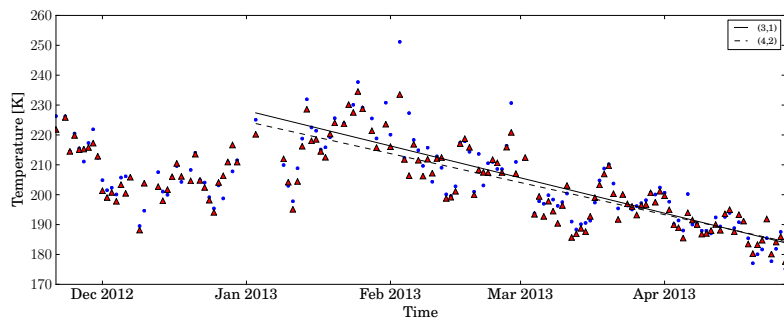


Figure B.1: Straight line fits of the OH (3,1) and (4,2) temperatures from 1. Jan - 27. April 2013. The fits yield a decline of 0.38 K/day for the (3,1) temperature and 0.35 for the (4,2) temperature.

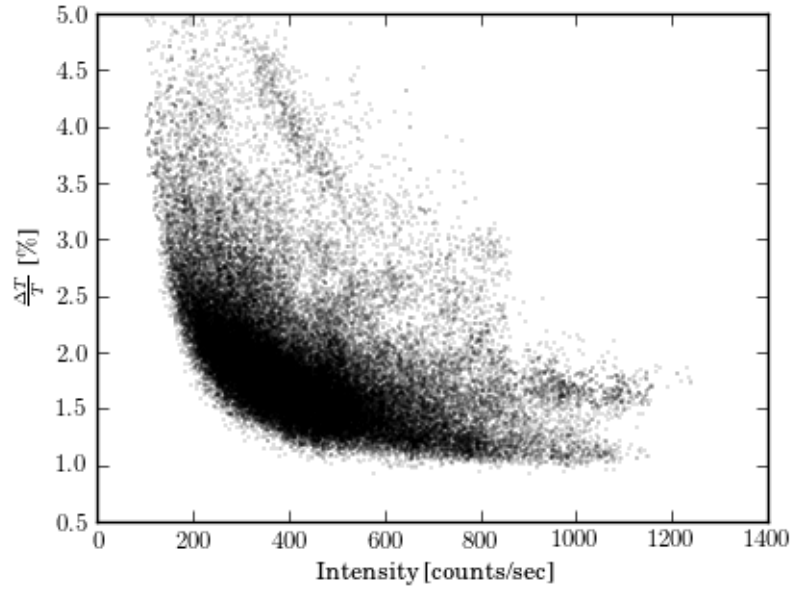
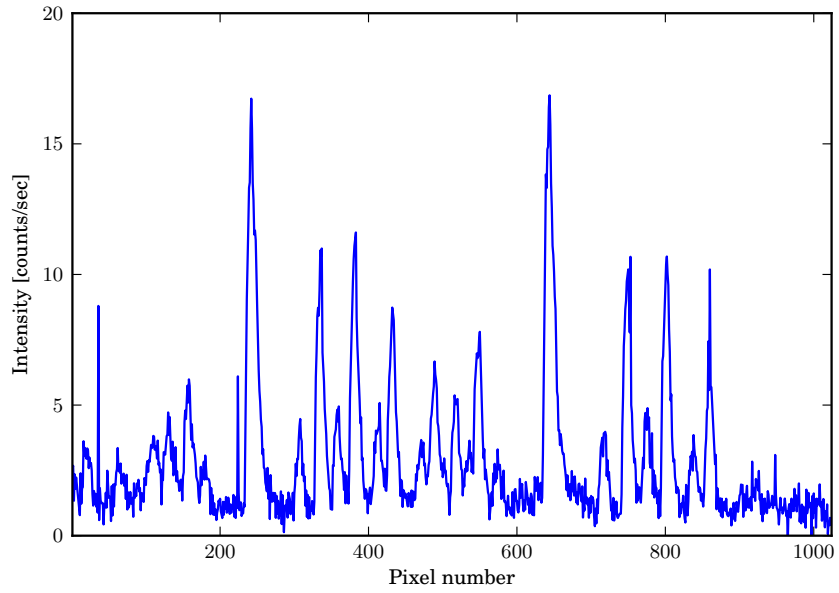
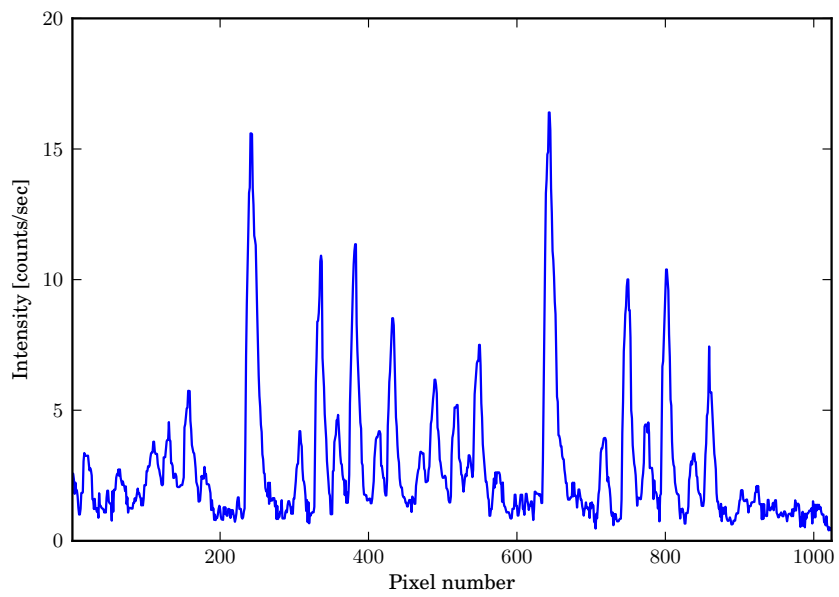


Figure B.2: Relative error in the (3,1) Temperature as a function of intensity. Same plot as figure 5.5 but with data from nights between 6-7. and 7-8. December 2012 excluded where the instrument was run with 15 seconds integration time. The cluster which was commented in section 5.2 is gone.



(a) Unfiltered spectrum with hot pixels.



(b) 3 point median filtered spectrum. Hot pixels removed

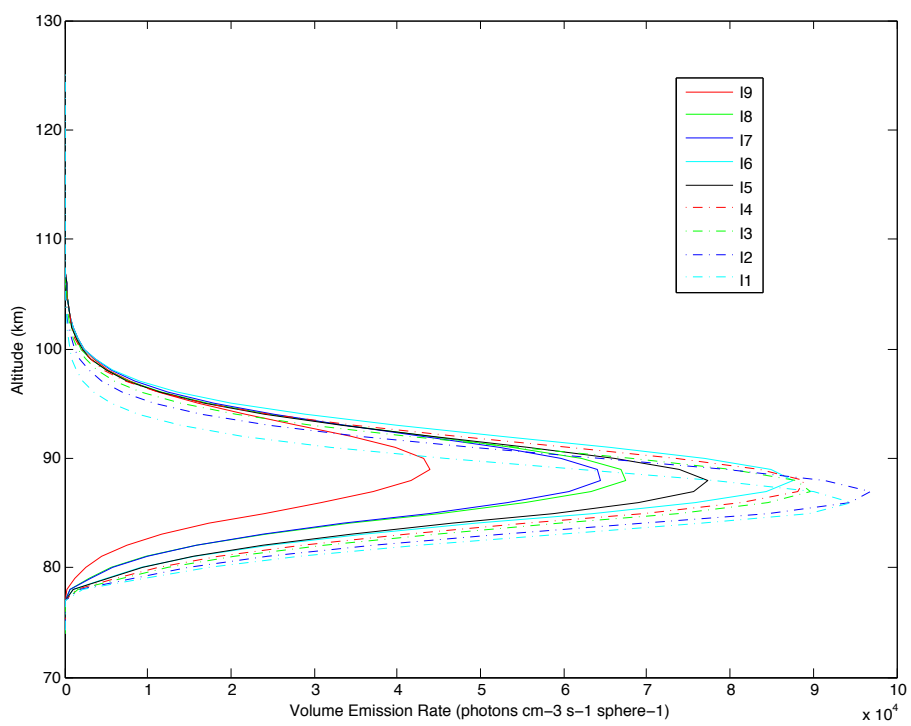
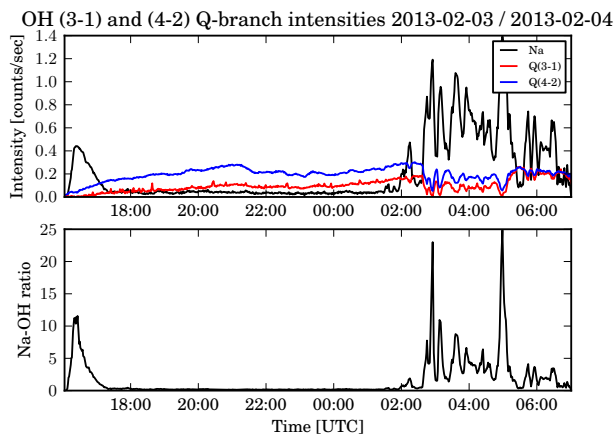
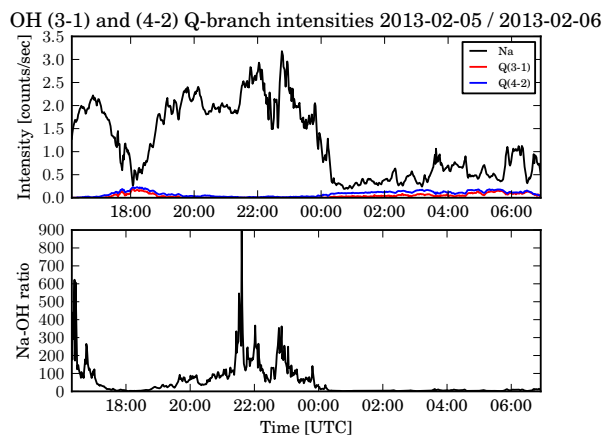


Figure B.3: The altitude distribution of the different vibrational states of the OH* molecule calculated with a photochemical equilibrium model.



(a)



(b) a

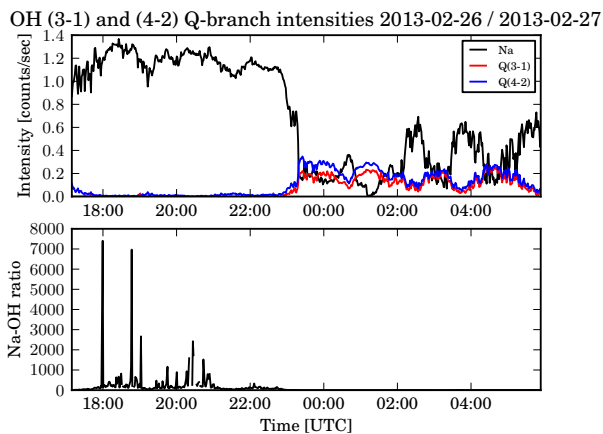


Figure B.4: Intensity plots of the nights causing the outliers in the nightly averaged temperature series. Common for these nights are that the (4-2) Q-branch intensity is higher than than for the (3-1), indicating water absorption of the OH signal. Probably because instrument window was covered with water, snow or ice.

Appendix C

Code

C.1 Main routine: calc_ohdata

```
import ohtools
from datetime import datetime
import numpy
import math
import warnings

warnings.filterwarnings('ignore')

def calc_ohdata(year, day):

    #importing signal:
    signal_raw = ohtools.import_day(year, day)

    try:
        if signal_raw == 0:
            return 0
    except ValueError:
        pass

    NA_LIMIT = 1.5

    filename_31 = ohtools.mkfilename(signal_raw)
    resp_kr = numpy.loadtxt('ohtools/response_kr.dat')
    numofsig = signal_raw.shape[0]
    avgspec = ohtools.avgspec_nafilt(signal_raw, 25, numofsig-25, NA_LIMIT
    )
    avgspec_kr = avgspec/resp_kr

    peaks = ohtools.locate_peaks(avgspec_kr)

    borders = ohtools.band_borders(peaks)
    oh31pts = numpy.arange(borders[0], borders[1])
    oh42pts = numpy.arange(borders[2], borders[3])
    w_31, w_42 = ohtools.instrfunc_kr(peaks, borders)
```

```

#Indices on which to start and stop
START = 0
STOP = numofsig
TG = 200

avg31band = avgspec[oh31pts]
avg42band = avgspec[oh42pts]

#Calculating baseshapes from average spectrums
basepts31, basepts42 = ohtools.basepts(peaks)
baseshape_31 = ohtools.basecurve(avgspec, basepts31, borders[0],
    borders[1], 0)
baseshape_42 = ohtools.basecurve(avgspec, basepts42, borders[2],
    borders[3], 0)
baseshape_31 -= numpy.min(baseshape_31)
baseshape_42 -= numpy.min(baseshape_42)
avg31band -= baseshape_31
avg42band -= baseshape_42

#Computing temperatures of average spectrum
mt31, mmi31, mie31, mit31, ms31, err31 = ohtools.get_temp_subbg(
    avg31band, TG, w_31, 0, 0)
mt42, mmi42, mie42, mit42, ms42, err31 = ohtools.get_temp_subbg(
    avg42band, TG, w_42, 1, 0)

sig_indices = numpy.arange(START,STOP)
numofres = len(sig_indices)

print('Filtering signals...')
signal_raw = ohtools.filtsig(signal_raw)
print('Done')

print('Calculating temperatures:')

filepath = '/Users/endre/Dropbox/Dokumenter/NTNU/10.semester/oh data/
    oh data/' + filename_31
f = open(filepath, 'w')

temps_31 = numpy.zeros(numofres)

k = 0
for i in sig_indices:

    if(k > 4 and k%5 == 0):
        TG_new = ohtools.mean(temps_31[k-5:k])
        if(math.isnan(TG_new) != True):
            TG = TG_new

    current_sig = signal_raw[i,11:1035]
    q31, q42 = ohtools.qbranch(current_sig)
    na = ohtools.sodiumlev(current_sig)

    try:
        naohratio = na/q42

```

```

except ZeroDivisionError:
    naohratio = float('nan')

solspecrat = ohtools.solspecratio(current_sig)

current_band_31 = current_sig[oh31pts]-baseshape_31
current_band_42 = current_sig[oh42pts]-baseshape_42

time = ohtools.get_time(signal_raw[i,:])

if naohratio < NA_LIMIT:

    if i < 25 or i > numofsig-25:

        t31, et31, i31, ei31, it31, err31 = ohtools.get_temp_fitbg
            (current_band_31, TG, w_31, 0, 0, baseshape_31)
        t42, et42, i42, ei42, it42, err42 = ohtools.get_temp_fitbg
            (current_band_42, TG, w_42, 1, 0, baseshape_42)

    else:

        t31, et31, i31, ei31, it31, err31 = ohtools.get_temp_subbg
            (current_band_31, TG, w_31, 0, 0)
        t42, et42, i42, ei42, it42, err42 = ohtools.get_temp_subbg
            (current_band_42, TG, w_42, 1, 0)

    else:

        t31 = float('nan')
        et31 = float('nan')
        i31 = float('nan')
        ei31 = float('nan')
        it31 = 0
        err31 = 'High Na-OH ratio'

        t42 = float('nan')
        et42 = float('nan')
        i42 = float('nan')
        ei42 = float('nan')
        it42 = 0
        err42 = ''

temps_31[k] = t31

ohtools.write_scan_result(t31, et31, i31, ei31, t42, et42, i42,
    ei42, q31, q42, na, time, f)

k += 1

f.close()

print('Iterations complete.')
a = ohtools.get_ohstats(filepath, 1)

```

```
print('Filename: {}'.format(filename_31))
```

C.2 get_temp_subbg

```
def get_temp_subbg(spectrum, tguess, w, bandflag, infoflag):

    err_message_gn = '' #Error message from Gauss-Newton iteration
    err_message = ''

    speclen = spectrum.size
    betaguess = 1.0
    uncrt = numpy.zeros((4,2))

    #Deciding model function from the bandflag
    if(bandflag == 0):
        shift = 0
        plot_title = 'OH (3,1) band'
        partfunc = partfunc_31
        delpartfunc = delpartfunc_31
    elif(bandflag == 1):
        shift = 6
        plot_title = 'OH (4,2) band'
        partfunc = partfunc_42
        delpartfunc = delpartfunc_42

    start = time.clock()
    tauguess = 1/tguess
    qr = partfunc(tauguess)
    wmod = expwmat(w, tauguess, shift)/qr #Divide by partition function
    specshape = ohmodel(1, wmod)

    #Scaling the band to match the measured spectrum:
    rho_initial = numpy.max(spectrum - betaguess)/numpy.max(specshape)

    ohband = rho_initial*specshape
    model = ohband + betaguess
    intens_initial = numpy.sum(model)

    residuals = numpy.zeros(speclen)
    for i in range(speclen):
        residuals[i] = spectrum[i] - model[i]

    sse_initial = 0 #Initial sum of square erros
    for i in residuals:
        sse_initial += i**2

    ds = 1e6
    itcount = 0

    sse = sse_initial
    params = numpy.array([rho_initial, tauguess, betaguess])

    #Gauss-Newton algorithm
    while(ds > ACCLIM and itcount < ITMAX):
```

```

jacobi = numpy.zeros((speclen,3)) #Jacobi matrix
jacobi[:,0] = specshape
jacobi[:,1] = params[0]*jacobitau(wmod, params[1], qr, shift)
jacobi[:,2] = 1

#jacobi[:,1] *= params[0] #Multiplying with rho

jtj = jacobi.T.dot(jacobi) #Calculating (J^T*J)

try:
    #Calculating (J^T*J)^-1*J^Tr:
    invjtjtr = numpy.linalg.solve(jtj,jacobi.T.dot(residuals))
except numpy.linalg.LinAlgError:
    err_message_gn = 'Normal eqn. failure'
    break

params[0] += invjtjtr[0] #Adding changes
params[1] += invjtjtr[1]
params[2] += invjtjtr[2]

qr = partfunc(params[1])
wmod = expwmat(w, params[1], shift)/qr #New model with the
    adjusted parameters
specshape = ohmodel(1, wmod)
ohband = params[0]*specshape
model = ohband + params[2]

oldsse = sse
residuals = numpy.zeros(speclen)
for i in range(speclen):
    residuals[i] = spectrum[i] - model[i]

sse = 0
for i in residuals:
    sse += i**2

ds = abs(oldsse - sse) #Change in sum of square errors
itcount += 1
elapsed = time.clock()-start
if(len(err_message_gn) == 0):
    #If err_message_gn != 0 Gauss-Newton algortim succeeded. Check values
    and calculate
    #uncertainties

    newt = 1/params[1] #New temperature
    intens_mod = numpy.sum(ohband)
    intens_real = numpy.sum(spectrum-params[2])

    #Calculating errors:
    uncrt, errmess_cov = uncertainties_subbg(jtj, residuals, params,
        intens_real, speclen)

    intens_mod_err = intens_mod*uncrt[1,1]/100

```



```

#Checking result:
if(newt < templim_low or newt > templim_upper): #Temperature
    limits
    newt = float('nan')
    uncrt[0,1] = float('nan')
    uncrt[0,0] = float('nan')
    intens_mod_err = float('nan')
    err_message += 'Not valid temp, '

if(intens_mod < intenslim_low or intens_mod > intenslim_upper): #
    Intensity limits
    intens_real = float('nan')
    intens_mod = float('nan')
    intens_mod_err = float('nan')
    uncrt[2,:] = float('nan')
    uncrt[1,:] = float('nan')
    err_message += 'Not valid intensity, '

#Checking errors:
if(uncrt[0,0] > temperr_lim or uncrt[0,1] > rel_temperr_lim): #
    Temperature error (set to 25 K)
    uncrt[0,:] = float('nan')
    newt = float('nan')
    err_message += 'High T error, '

if(intens_mod_err > intenserr_lim or uncrt[1,1] >
    rel_intenserr_lim): #Intensity error
    intens_real = float('nan')
    intens_mod = float('nan')
    intens_mod_err = float('nan')
    err_message += 'High I error'

else:
    newt = float('nan')
    intens_real = float('nan')
    intens_mod = float('nan')
    uncrt[:,:] = float('nan')
    intens_mod_err = float('nan')
    err_message = err_message_gn #Setting error message from Gauss-
        Newton calculation

if(infoflag == 1):
    print('-----Before iteration-----')
    print('Initial temperature: {:.4.2f}'.format(tguess))
    print('Initial intensity: {:.6.2f}'.format(intens_initial))
    print('Initial rho: {:.6e}'.format(rho_initial))
    print('Initial beta: {:.3f}'.format(betaguess))
    print('Initial SSE: {:.6.2f}'.format(sse_initial))
    print('-----After iteration-----')
    print('New temperature: {:.4.2f} Ås {:.3.2f}, acc: {:.4.2f}%'.format(
        newt, \
            uncrt[0,0], uncrt[0,1]))
    print('Real intensity: {:.5.2f} Ås {:.3.2f}, acc: {:.2.2f}%'.format(
        intens_real, \

```

```

        uncrt[2,0], uncrt[2,1]))
print('Model intensity: {:.52f} Ås {:.2f}, acc: {:.22f}%'.format(
    intens_mod, \
    intens_mod_err, uncrt[1,1]))
print('New rho:         {:.3e} Ås {:.3e}, acc: {:.22f} %'.format(
    params[0], uncrt[1,0], uncrt[1,1]))
print('New beta: {:.93f} Ås {:.63f} acc: {:.22f} %'.format(params
    [2], uncrt[3,0], uncrt[3,1]))
print('SSE: {:.62f}'.format(sse))
print('Number of iterations: {:.2d}'.format(itcount))
print('CPU time: {:.26f}'.format(elapsed))
print('Error: {}'.format(err_message))

offset = numpy.zeros(speclen)
offset[:] = params[2]

fig1 = pyplot.figure(figsize=(4.5,3.5))
ax = fig1.add_subplot(111)
ax.plot(spectrum,'b', label = 'Real')
ax.plot(model,'r', label = 'Model')
pyplot.ylabel('Intensity [counts/sec]')
pyplot.xlabel('Pixel number')
pyplot.title(plot_title)
pyplot.legend()
boxtext = 'Temperature: {:.31f} Ås {:.31f} K \n'.format(newt,
    uncrt[0,0])
boxtext += 'Intensity: {:.41f} Ås {:.41f} counts/sec'.format(
    intens_real, uncrt[2,0])
props = dict(boxstyle='round', facecolor='wheat', alpha=0.5)
ax.text(0.03, 0.97, boxtext, transform=ax.transAxes,
    verticalalignment='top', bbox=props)

fig2 = pyplot.figure(figsize=(4.5,3.5))
pyplot.plot(spectrum-model, 'k')
pyplot.xlabel('Pixel number')
pyplot.ylabel('Data - Model [counts/sec]')
pyplot.show()

return newt, uncrt[0,0], intens_mod, intens_mod_err, itcount,
    err_message

```