Reconstruction of Solar Spectral Irradiance back to the Maunder Minimum

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

MICHA SCHÖLL

MSc., St Andrews

born 4 April 1983

citizen of Germany

accepted on the recommendation of

Prof. Dr. M. Carollo, examiner
Prof. Dr. W. Schmutz, co-examiner
Prof. Dr. S. Solanki, co-examiner

2011
Contents

1 Introduction 5

I Radiative Transfer 7

2 Statistical Equilibrium 9
  2.1 Thermodynamic Equilibrium 9
  2.2 Local Thermodynamic Equilibrium 10
  2.3 Non Local Thermodynamic Equilibrium 10

3 Radiative Transfer and Spectral Synthesis 11
  3.1 Model Atmospheres 11
  3.2 Radiative Transfer 11
    3.2.1 Numerical Convergence 12
    3.2.2 Accelerated Lambda Iterator 12
    3.2.3 Ambipolar Diffusion 14
    3.2.4 Numerical Stability 15
  3.3 Spectral Synthesis Code 23
    3.3.1 Line Formation 23
    3.3.2 Quiet Sun Spectrum 25

4 NLTE Solar Irradiance Modelling with COSI 31
  4.1 Introduction 32
  4.2 Measured solar spectral irradiance 32
  4.3 Molecules in the COSI code 33
    4.3.1 Chemical equilibrium 33
    4.3.2 Main molecular bands in the solar spectrum 34
  4.4 NLTE calculations 36
    4.4.1 Calculating of the UV radiation 37
    4.4.2 Calculation of the visible and IR radiation 39
    4.4.3 Comparison with measurements and discussion 42
  4.5 Irradiance variations 43
  4.6 Conclusions 45
# Solar Irradiance Reconstruction

## Introduction

## Data

- **6.1 Solar Modulation Potential**
  - **6.1.1 Neutron Monitor**
- **6.2 Sunspots**
- **6.3 Filling Factors**
- **6.4 Total Solar Irradiance**
- **6.5 Solar Spectrum**
- **6.6 Lyman-alpha**

## Total Solar Irradiance

## Model Based Reconstruction

- **8.1 Introduction**
- **8.2 Filling Factors**
  - **8.2.1 The Network**
  - **8.2.2 Reconstructing Filling Factors**
  - **8.2.3 Two Component Quiet Sun**
- **8.3 A new approach to long-term reconstruction of TSI**
  - **8.3.1 Abstract**
  - **8.3.2 Introduction**
  - **8.3.3 Effects of solar radiative forcing on climate**
  - **8.3.4 Methods**
  - **8.3.5 Results and Discussion**
  - **8.3.6 Conclusions**
  - **8.3.7 Acknowledgements**
  - **8.3.8 Online Material**
  - **8.3.9 Quiet Sun**
  - **8.3.10 Active Sun**
- **8.4 Discussion**
  - **8.4.1 Uncertainties**
  - **8.4.2 Extrapolation – Interpolation**

## Spectral Reconstruction

- **9.1 Active Area Expansion**
- **9.2 Lyman-α**
  - **9.2.1 Results**
- **9.3 Total Solar Irradiance**
- **9.4 Spectral Variability**
- **9.5 Conclusion**
III Conclusion & Outlook

10 Discussion of Thesis Results

10.1 COSI Improvements

10.2 UV Opacities

10.3 1-D Modelling

10.4 Variability over Time

10.5 TSI Reconstruction

10.6 Spectral Reconstruction

10.6.1 Active Area Enhancement

10.6.2 Lyman-α

11 Outlook

11.1 Three Component Quiet Sun – Stretched Distribution Model

IV Appendix

A H-Minus Manual

A.1 H-Minus

A.1.1 Input

A.1.2 Output

A.2 FIOSS

A.2.1 Input

A.2.2 Control file

A.2.3 Output

A.2.4 Helper programs

A.2.5 Opacity

A.3 Module Dependencies

B Acronyms

Bibliography

Acknowledgements

Curriculum Vitae
Abstract

The aim of this thesis is to reconstruct the ultraviolet solar radiation back to the Maunder minimum, especially the Lyman-\(\alpha\) due to its relevance in the field of climatology.

The solar spectrum is calculated with COSI, the COde for Solar Irradiance, a state of the art numerical radiative transfer code. It solves both the radiative transfer equation and the statistical equilibrium and implements opacity distribution functions under NLTE assumptions. In this thesis COSI is extended to account for physical effects in the transition region and validated against observations. The extensions include ambipolar diffusion and the accelerated lambda iterator. The ambipolar diffusion mainly improves on numerical stability while the accelerated lambda iterator decreases the necessary number of iteration-steps. Validation of the calculated spectrum is done against both ATLAS 3 and SORCE spectra, with special emphasis given to the Lyman-\(\alpha\) spectrum, which is compared to full-disk SUMER images.

Solar total and spectral irradiance is reconstructed back in time by use of a modified five component SATIRE model. A five component model, incorporating a quiet sun, quiet and active network, plages regions and sunspot models is used to describe both long-term and short-term trends of solar activity by use of different solar proxies. The modification of the SATIRE model consists of an “active area enhancement”, which is implemented due to magnetic field expansion from the photosphere towards the transition region. The implementation is based on observations of active regions in the photosphere, the transition region and coronal active regions. It results in an increased UV variability. It also causes the filling factors to become dependent on both time and wavelength.

The short-term variability is assumed to follow the change of active network, sunspot and plage area. All of those are coupled to the sunspot number.

The long-term trend of solar irradiance back to the Maunder minimum is estimated using cosmogenic radionuclides for the quiet-sun variability, splitting the observed quiet sun into two components, one termed “very quiet” sun and the other being the quiet network, with the cosmogenic radionuclides determining the ratio of those two models.

In the outlook another approach to reconstruct the long-term trend is presented based on observations of the distribution of the magnetic field and the assumption that only the parameters for the distribution change over time, but not the form of it. The distribution used is a Voigt profile and is based on Stenflo and Holzreuter (2003).

Computed Lyman-\(\alpha\) irradiance is compared to SORCE-SOLSTICE and SUSIM measurements and the reconstruction by Woods et al. (2000) for the current-time. The long-term variability depends on the method used and is found to vary from 1 mW/m\(^2\) to 3 mW/m\(^2\), which translates to a 33% to 100% variability with respect to the solar cycle variability of 3 mW/m\(^2\).

The total solar irradiance variability is found to be between 1 W/m\(^2\) and 6 W/m\(^2\), again depending on the model used, with the distribution model resulting in the lower variability. The spectral solar irradiance is presented with a 1 nm resolution, ranging from 90 nm to 2 \(\mu\)m, and in time from 1610 CE to 2010 CE.
Zusammenfassung

Das Ziel dieser Dissertation ist die Rekonstruktion der Ultravioletten Solaren Stahlung zurück zum Maunoder Minimum mit dem Schwerpunkt der Lyman-α Linie.


COSI wird erweitert um Physikalische Auswirkungen der Übergangsregion zu berücksichtigen. Desweiteren wird das Programm mithilfe von spektralen Beobachtungen validiert, gegen ATLAS 3 und SORCE Spektren und hochauflösissten Lyman-α Spektren von SUMER Aufnahmen. Desweiteren wird die Lyman-α Linie mit der Rekonstruktion von Woods et al. (2000) verglichen. Der Langzeitrend der Lyman-α Strahlung hängt von der benutzten Methode ab und variiert zwischen 1 mW/m² und 3 mW/m², vergleichbar mit der Solar-Zyklus Variabilität von 3 mW/m².

Die TSI Variabilität liegt zwischen 1 W/m² und 6 W/m², wiederum abhängig vom benutzten Modell, wobei das Distributionsmodell zu der kleineren Variabilität führt.


Zusätzliche Erweiterungen beinhalten die ambipolare Diffusion und die Implementation des beschleunigten Lambda Iterators, wobei die ambipolar Diffusion hauptsächlich der numerischen Stabilität dient, während der Lambda Iterator die Anzahl an benötigten Iterationsschritte verringert.


Die Abwandlung des SATIRE Modells besteht aus der “Aktiven Flächen Erweiterung” welche die magnetische Feldexpansion von der Photosphere in die Übergangsregion beschreibt und basiert auf Beobachtungen aktiver Regionen in der Photosphäre, der Übergangsregion und der Corona. Das Ergebnis ist eine erhöhte Variabilität im UV.

Der langfristige Trend zurück in das Maunoder Minimum wird mithilfe kosmogener Radionuklide berechnet, welche das Verhalten der ruhigen Sonne beschreiben. Dafür wird die ruhige Sonne in zwei Bereiche aufgeteilt. Ein sehr ruhiger Anteil und ein aktiver Teil, wobei die radionuklide das Verhältniss dieser beiden Teile bestimmen.

Chapter 1

Introduction

“On a former occasion I have shewn, that we have great reason to look upon the sun as a most magnificent habitable globe; and, from the observations which will be related in this Paper, it will now be seen, that all the arguments we have used before are not only confirmed, but that we are encouraged to go a considerable step farther, in the investigation of the physical and planetary construction of the sun. The influence of this eminent body, on the globe we inhabit, is so great, and so widely diffused, that it becomes almost a duty for us to study the operations which are carried on upon the solar surface. Since light and heat are so essential to our well-being, it must certainly be right for us to look into the source from whence they are derived, in order to see whether some material advantage may not be drawn from a thorough acquaintance with the causes from which they originate.” (Herschel, 1801)

It is claimed that Herschel founded the research into solar physics with the above-quoted paper and its sun-earth connection due to a claim in this paper that the wheat price is influenced by solar activity and due to his importance in his time. While his findings of the influence of solar activity on the wheat price is doubtful, and most of his arguments are rejected nowadays, among others that sunspots allow us to see through the “glowing clouds” to “its own solid body\(^1\)”, his original assertion that light and heat are so essential and that we should look into the source from whence they are derived is most certainly true.

This thesis is part of the ETH poly-project “Variability of the sun and global climate II”. It is a multi-disciplinary project, encompassing climatology, geology, solar physics and astronomy. The project aims to understand the connection between solar activity and the climate, the so called sun-earth connection. To achieve this, the impact of solar radiation on the climate has to be understood. For the climate, SOCOL (Egorova et al., 2005; Schraner et al., 2008), a chemical climate model has been developed and validated as part of this project, while the solar variability has been modelled using COSI (Haberreiter et al., 2008), a state of the art solar radiation code. As part of the first poly project, Rozanov et al. (2002) found that a crucial input for the climate model is the spectral variability, especially the Lyman-\(\alpha\) line, the hydrogen emission line at 121.5 nm.

Thus the goal of this thesis is to model the ultra-violet solar irradiance and its long-term behaviour back to the Maunder minimum with a special emphasis on the Lyman-\(\alpha\) line. Solar visible irradiance and near-UV has been investigated for the space-time (1978–present) by my predecessor, Margit Haberreiter in her thesis, Haberreiter (2006). She reconstructed the Lyman-\(\alpha\) variability within a factor of 2 while additionally obtaining the measured Lyman-\(\alpha\) minimum irradiance.

The Lyman lines are especially difficult to model due to their formation in the transition region, the part of the solar atmosphere above the chromosphere where the temperature raises from ten thousand Kelvin to a million Kelvin within a few hundred kilometres. Hence it is important to calculate the formation height, departure ratio, absorption and emission coefficients of the lines in this region with

\(^1\)Both in Herschel (1801, p. 267)
high precision, incorporating both physical and numerical difficulties, some of them are described in Part I.

The second goal is to go back in time to the Maunder minimum. This implies that the time-resolution and accuracy will decrease compared to e.g. Wenzler (2005), who successfully reconstructed the total solar irradiance for the space-time. In this work, the main-focus is on the long-term evolution of solar activity by establishing a link between proxy data and solar irradiance.

The thesis consists of two main parts, an outlook and an appendix. Part I, Radiative Transfer, describes the numerical calculation of solar irradiance. Here a short overview of theory of radiative transfer is given, followed by an introduction of the code used and the improvements and fixes done to enable the correct calculation of the transition region and hence the Lyman lines. Those fixes are validated against observations. Part II, Solar Irradiance Reconstruction, is mainly concerned with temporal variation of solar irradiance, with a special emphasis on the Lyman-\(\alpha\) line, where the concepts of the least-active sun and active area enhancement are introduced.
Part I

Radiative Transfer
Chapter 2

Statistical Equilibrium

First the equations of statistical equilibrium are explained as given in the standard literature. The most simple case is the case of Thermodynamic Equilibrium (TE), followed by Local Thermodynamic Equilibrium (LTE) which can be applied when the assumptions of TE are invalid globally, but they hold locally. Finally Non-Local Thermodynamic Equilibrium (NLTE) can be used to describe systems that are neither TE nor LTE e.g. where every particle may directly be influenced from any other particle in the system, as is the case in a optically thin atmosphere.

2.1 Thermodynamic Equilibrium

TE assumes an invariant system both in space and time with a constant temperature $T$ and pressure $p$. The particle velocity follows a Maxwell distribution. The population number depend only on temperature and the radiation field. For a given temperature $T$, the source function, defined as the ratio of emission to absorption,

$$S_{\nu} := \frac{\eta_{\nu}}{\kappa_{\nu}}$$

equals the Planck distribution,

$$B_{\nu}(T) := \frac{2\hbar\nu^3}{c^2} \frac{1}{e^{\frac{\hbar\nu}{kT}} - 1} \quad [W/m^2/sr/Hz]$$

where $\nu$ is the frequency, $h$ Planck’s constant, $c$ the speed of light and $k$ the Boltzmann constant. The ratio of the population of level $i$ to level $j$ of a population $n_{il}$ in ionization state $l$ follows Boltzmann’s law,

$$\frac{n_{ik}}{n_{jk}} = \frac{g_i}{g_j} e^{-\frac{\chi_{k} - \chi_{l}}{kT}} = \frac{g_i}{g_j} e^{-\frac{h\nu}{kT}}$$

with $g$ the statistical weight and $\chi$ the excitation energy. An extension of the Boltzmann equation is the Saha equation describing the ionization ratio of ion $k$ to $k + 1$,

$$\frac{\sum_i n_{il+1}n_{e}}{\sum_i n_{il}} = \frac{g_{i+1}}{g_i} \left( \frac{2\pi m_e k}{\hbar^2} \right)^{\frac{1}{2}} T^{\frac{1}{2}} e^{-\frac{\chi_{ion}}{kT}}.$$ 

Also the principle of detailed equilibrium is valid, i.e. the number of direct processes equals the number of inverse processes. Hence in case of TE the global temperature and the atomic data is sufficient to calculate the population numbers of the system. However, the condition of TE are only satisfied for a closed and stable system.
2.2 Local Thermodynamic Equilibrium

LTE assumes a local thermodynamic equilibrium. However, it does allow for a global change, i.e. a temperature and density fluctuation. Its implicit assumption is that every point is only influenced by its neighbours, as is the case when the collisions dominate over the radiative processes.

In this case the population numbers and radiation fields must be calculated for each point in space separately.

2.3 Non Local Thermodynamic Equilibrium

If the radiative rates dominate over the collisional rates, the assumption of LTE does not apply any more. In this case a particle can be influenced by a photon coming from anywhere in the atmosphere. Then the statistical equilibrium must be solved simultaneously with the radiative transfer equation for all points in space, hence for a steady-state, static atmosphere we must solve

\[ \forall i, k : n_{ik} \sum_{j \neq i} P_{ijk} = \sum_{j \neq i} n_j P_{ijk} + f \]

(2.5)

where \( n_{ik} \) is the population of element \( k \) at level \( i \) and \( P_{ijk} \) is the transition probability of \( n_{ik} \) to \( n_{jk} \) and \( f \) are additional processes, e.g. ambipolar diffusion (section 3.2.3). This equation couples all states with each other and hence it is numerically expensive to solve. The ratio of population numbers calculated in LTE to NLTE is described by the departure coefficient for each population \( n_{ik} \),

\[ d_{ik} = \frac{n_{ik}^{\text{LTE}}}{n_{ik}^{\text{NLTE}}} \]

(2.6)

For a detailed explanation see Mihalas (1978), chapter 5-4, “The Non-LTE Rate Equations”. For the remainder of the thesis the element \( k \) is often left out if the element in question can be derived from the text, e.g. hydrogen.
Chapter 3

Radiative Transfer and Spectral Synthesis

COde for Solar Irradiance (COSI) (Haberreiter, 2006; Haberreiter et al., 2008; Shapiro et al., 2010) is a program to solve the statistical equilibrium in NLTE under spherical symmetry. It consists of two parts, the radiative transfer code H-Minus developed by Hamann and Schmutz (1987) and Schmutz et al. (1989) and the synthetic spectrum program SYNSPEC, originally developed by Hubeny (1981) and extended by Hubeny and Lanz (1992). Shapiro et al. (2010) implemented molecular lines. Chapter 4 describes this addition, together with a comparison to observations.

3.1 Model Atmospheres

The temperature, electron and proton density and the turbulent velocities are input for COSI. Based on this and atomic data the height dependent population density is calculated for different parts of the sun. For this work Fontenlas model atmospheres from (Fontenla et al., 1999) are used (hereafter FAL-99). They are seven semi-empirical models describing a set of steady-state, one-dimensional solar atmospheres for different parts of the sun. The radial height of the models range from 100 km below photosphere into the transition region.

The models are derived by an analysis of observed intensities of different features together with the SRPM radiative transfer code (Fontenla et al., 2009a). Hence the atmospheric model depends on both observations and the radiative transfer code.

3.2 Radiative Transfer

The statistical equilibrium under NLTE assumption is solved by H-Minus, extended by Haberreiter et al. (2008) who introduced a new treatment of the so called line blanketing. Shapiro et al. (2010) also implemented molecules.

The first part solves the radiative transfer equation

$$\mu \frac{dI_{\mu \nu}}{d \tau_\nu} = I_{\mu \nu} - S_\nu$$  \hspace{1cm} (3.1)

where $\mu$ is the cosine of the angle between the ray and the solar surface, $\nu$ the frequency, $I_{\mu \nu}$ the intensity of the specific angle and frequency, $S_\nu$ the source function and the optical depth $\tau_\nu$. The formal solution can be written in operator form as

$$I_{\mu \nu} = \Lambda_{\mu \nu}[S_\nu]$$  \hspace{1cm} (3.2)
H-Minus calculates the population numbers and radiative field for a given set of elements based on a model atmosphere, containing the temperature, pressure and velocity profile together with stellar abundances. For a detailed explanation of the necessary input, see the manual at appendix A.

### 3.2.1 Numerical Convergence

“…the archetype scattering problem of the transfer equation, the iterative process then stabilizes to a spurious value without converging, and successive iterations differ but slightly, even though the current estimate is far from the true solution.” (Mihalas, 1978, § 7-5 – Solution by Iteration)

The above-quoted problem was encountered when examining the converged quiet sun hydrogen population with different Λ’s for the lines as described in the previous section. The different populations of the same model are shown in Figure 3.5. The transition region is numerically and physical unstable, resulting in two different solutions. One possibility to solve the internal discrepancy is to introduce a physical effect that acts as a forcing, e.g. ambipolar diffusion. The above mentioned slight difference is in some cases cyclic. For example, the Lyman-α profile as computed by Haberreiter (2006) converged to a correct quiet sun integrated Lyman-α, however this is an unstable value, that turned out to yield the correct result. Further iteration diverged from this result.

### 3.2.2 Accelerated Lambda Iterator

The lambda iteration are used to solve the radiative transfer problem. The radiative field is described by the monochromatic directional Λ operator and the source function $S_{\mu\nu}$ by substituting 3.2 into the state equation.

$$S = (1 - \epsilon)\Lambda[S] + \epsilon B \quad (3.3)$$

The lambda iteration consists of solving this equation by iteratively calculating the source function,

$$S^{(n+1)} = (1 - \epsilon)\Lambda[S^{(n)}] + \epsilon B \quad (3.4)$$

Equation 3.4 is known to both converge very slowly and may also result in spurious results, that is it may stabilize before the result is reached or it may also converge to a wrong result, that is a local minimum.

An improvement can be achieved by operator splitting as originally conceived by Jacobi, that is rewriting Λ as $\Lambda^* + (\Lambda - \Lambda^*)$. Hence equation 3.4 becomes

$$S^{(n+1)} = (1 - \epsilon)\left(\Lambda^*[S^{(n+1)}] + (\Lambda - \Lambda^*)[S^{(n)}]\right) + \epsilon B \quad (3.5)$$

$$\approx (1 - \epsilon)\left(\Lambda^*[S^{(n+1)}] + (\Lambda - \Lambda^*)[S^{(n)}]\right) + \epsilon B \quad (3.6)$$

It becomes more instructive when rewriting Equation 3.6 as

$$\delta S^{(n)} = S^{(n+1)} - S^{(n)} = (1 - \epsilon)\left(\Lambda^*[S^{(n+1)}] + (\Lambda - \Lambda^*)[S^{(n)}]\right) + \epsilon B - S^{(n)} \quad (3.7)$$

$$= (1 - \epsilon)\Lambda^*\left[S^{(n+1)} - S^{(n)}\right] + (1 - \epsilon)\Lambda[S^{(n)}] + \epsilon B - S^{(n)} \quad (3.8)$$

$$= (1 - \epsilon)\Lambda^*\delta S^{(n)} + S_{FS}^{(n+1)} - S^{(n)} \quad (3.9)$$

$$= \left(1 - (1 - \epsilon)\Lambda^*\right)^{-1}\left[S_{FS}^{(n+1)} - S^{(n)}\right] \quad (3.10)$$

with $S_{FS}^{(n+1)} = (1 - \epsilon)\Lambda[S^{(n)}] + \epsilon B$, the formal solution of the $n^{th}$ iteration.
3.2. Radiative Transfer

Hence each step of the Accelerated Lambda Iterator (ALI) is amplified by the factor $a > 1$, resulting in an improved lambda iteration. $\Lambda^*$ must be chosen such that $\Lambda^*[S_{n+1}]$ can be easily calculated. Several choices exist for $\Lambda^*$, Hubeny (2003) lists several ones, with COSI setting a constant $\Lambda^*$ to $1 - \epsilon$ or 0, depending on the optical depth and a free parameter $\Gamma$. This $\Gamma$ is constant over all depth points for the lines and iterations $n$ as an input option for continuum, resonance and other lines separately. This method works well for the photosphere and chromosphere, however, it does not guarantee a solution in the transition region where it tends to “get stuck” at different solutions, depending on $\Gamma$. Figure 3.1 is an example of the resulting Lyman-$\alpha$ and H-$\alpha$ spectrum by use of two different $\Gamma$’s ($\frac{1}{100}$ and $\frac{1}{1000}$), together with their formation height.

Figure 3.1: Top panels: The profiles of Lyman-$\alpha$ and H-$\alpha$ for two different values of $\Gamma$, $\frac{1}{10}$ (solid) and $\frac{1}{100}$ (dashed) with their corresponding formation height (dash-solid, right scale). The difference is significant in the Lyman-$\alpha$ case, with $\Gamma = \frac{1}{10}$ 3.3 times higher irradiance than $\Gamma = \frac{1}{100}$. For H-$\alpha$ the difference is at most 5% in the line centre due to the lower formation height which does not reach the numerical unstable region at 2700 km where the transition region begins and where the population changes by a factor of up to 4. Lower Panel: Population ratio for H I-2 (solid) and H I-3 (dashed), showing the regions that are sensitive to a change in $\Gamma$. 

(a) Lyman-$\alpha$ profile from two different converged models. Solid and dashed show the profile, dash-dotted the Lyman-$\alpha$ profile. Solid and dashed show the profile, dash-dotted the formation height (axis on the right)

(b) H-$\alpha$ profile from the same two model runs as the Lyman-$\alpha$ profile. Solid and dashed show the profile, dash-dotted the formation height (axis on the right)

(c) Population Number Ratio
Rybicki and Hummer (1991) present a method to calculate $\Lambda^\ast$ dynamical and independent for each depth point, based on the change of population w.r.t. the previous iteration. This does not change the result, for a converged model $S_{n+1} = S_n$ holds, however it does improve the rate of convergence significantly. An approximate lambda operator which in this case is the tri-diagonal of the true lambda operator is used. This lambda operator is calculated by a $O(n)$ tri-diagonal inversion routine. This has been implemented, resulting in an improvement of the convergence.

To gain the full expected improvement, the lines should also be considered using a modified version of the now implemented ALI for the continuum.

### 3.2.3 Ambipolar Diffusion

Fontenla et al. (1990) describe ambipolar diffusion, an exchange of protons with neutral hydrogen in a partially ionized gas due to different velocities. This is most important in the transition region with its steep temperature gradient. Because of problems with numerical stability the implementation differs slightly from the proposed one. COSI calculates the ambipolar influence only directly for neutral hydrogen level 2, HI-2 based on Equations Fontenla 4.4–4.10. The other hydrogen levels are not explicitly calculated because of numerical stability but follow the second level due to radiative forcing by level grouping (c.f. Hubeny and Lanz, 2003, p. 57).

\[
\frac{d}{dz} \left[ n_1 \left( \frac{n_1}{n_p} V_A + U \right) \right] + r n_1 = s \quad \text{(Fontenla 4.4)}
\]

with

\[
r = P_{1\kappa} + \sum_{l \neq 1} n_l P_{1l} \quad \text{(Fontenla 4.5)}
\]

\[
s = n_p P_{\kappa 1} + \sum_{l \neq 1} n_l P_{1l} \quad \text{(Fontenla 4.6)}
\]

\[
V_A = D_x \frac{d}{dz} \ln \left( \frac{n_p}{n_1} \right) + D_T \frac{d \ln T}{dz} \quad \text{(Fontenla 4.7)}
\]

(3.11)

where $n_l$ is the hydrogen population density for level $l$, with $\kappa$ referring to the continuum, $P_{ij}$ the transition probability per unit time and the numerical approximation of the ambipolar diffusion velocity $V_A$ is the same as given in Fontenla et al. (1990),

\[
D_x = 90.7 \frac{T^{1.76}}{p_x} \quad \text{(Fontenla 3.1)}
\]

\[
D_T = 64.1 \frac{T^{1.76}}{p_x} \left[ \frac{x + 2.57 - 4000/(T \sqrt{x})}{x + 0.02/x} \right] \quad \text{(Fontenla 3.2)}
\]

\[
V_A = D_x \frac{d \ln x}{dz} + D_T \frac{d \ln T}{dz} \quad \text{(Fontenla 3.3)}
\]

\[
\frac{d}{dz} (n_m V_m) = P_{km} n_p - P_{mn} n_m + \sum_l (P_{lm} n_l - P_{ml} n_m) \quad \text{(Fontenla 4.1)}
\]

\[
V_m = \frac{n_p}{n_H} V_A \quad \text{(Fontenla 4.2)}
\]

where $x = n_p/n_1$ is the ionization ratio and $D_x = D_x(T, p)$ and $D_T = D_T(T, p, x)$ are fitted fractional polynomial functions.

The solution of equation Fontenla 4.4 depends on $r$ and $s$, the left and right hand side of the stationary statistical equilibrium equation $rn = s$. $r$ comes from the rate equation while $s$ is derived from the current population $n$ and $r$. This approach guarantees a self-consistent set of equations.
3.2. Radiative Transfer

From this, a new set of population \( n^{VA} \) is calculated with the difference \( \Delta^{VA}n_1 = n^{VA}_1 - n_1 \) used as a forcing for the radiative equilibrium equation, \( Rn = V_1 + \Delta^{VA}n_1 \). \( V_1 \) contains the mass and charge conversation. To induce numerical stability, \( \Delta^{VA}n_1 \) is smoothed by an three-element wide triangular filter, while the new populations \( n^{VA} \) are only calculated every fifth run. The process is numerically unstable and susceptible to a positive feedback loop. Because of this numerical problems, the ALI for lines must be lowered by a factor of 10. For a more detailed explanation of ALI see Section 3.2.2.

![Figure 3.2: Hydrogen ionization ratio for model A. The spike at \( \approx 9000 \text{ K} \) results from the temperature increase inside the photosphere](image)

The computed hydrogen population with and without ambipolar diffusion are compared in Figure 3.3, showing the influence of ambipolar diffusion on the hydrogen population number and the Lyman-\( \alpha \) profile. Figure 3.4 shows the ratio of the UV spectrum. While Lyman-\( \alpha \) profile does not show a significant difference, the UV continuum has a 4% increase from C I-2 (142.07 nm) ionization edge up to the C I-1 ionization edge (110.18 nm) and 10% from there to the hydrogen ionization edge O I (91.17 nm), at which point it returns to near unity (6% divergence). The ambipolar diffusion processes force the hydrogen population of the chromosphere and lower transition region into a new, more stable state. The upper transition region remains numerical unstable. See Subsection 3.2.2 for a method to stabilize the numerics and also Subsection 3.2.4 for a discussion of the numerical stability. The greater influence on the UV-continuum, compared to the Lyman-\( \alpha \) profile does not agree with Fontenla et al. (1990), but can be explained by the new, more stable chromospheric H I-1–10 populations.

3.2.4 Numerical Stability

Implementing both ambipolar diffusion and line-blanketing together results in numerical instability at the lower transition region due to different forcing of the line-blanketing increasing opacity in the chromosphere and ambipolar diffusion ionizing the lower transition region. This is shown in Figure 3.6 to 3.9. It is most profound in the ambipolar rate and the H-1 to H-11 rates. Numerical problems remain and have to be taken care of by carefully selecting the \( \Gamma \)'s. It also helped to only gradually introduce the ODF’s from Haberreiter (2006) by multiplying them with a factor that gradually increases to 1 over the iterations. This is implemented as a threshold. If the solution becomes numerical stable, add 0.1 to the factor until it reaches unity.

However, ambipolar diffusion does help in the sense that it invalidates one of the two possible states and forces the solution into one state. Figure 3.5 displays the effect of choosing different values for \( \Gamma \) on the hydrogen population without considering ambipolar diffusion.
(a) Both top panels display the relative neutral hydrogen population calculated for model A with (solid) and without (dashed) ambipolar diffusion. The right-hand panel displays the direct ambipolar diffusion effect from ≈2280.7 km to 2286 km as described by Fontenla while the left-hand panel shows the forcing of the solution into another state, suggesting that the numerical solution is unstable in this region.

(b) **Left Panel:** Neutral level one (solid), neutral level two (dashed) and ionized (dash-dot) hydrogen ratio of ambipolar diffusion. **Right Panel:** Effect of ambipolar diffusion on Lyman-α. The Lyman-α profile with (solid) and without (dashed) ambipolar diffusion show no significant difference. Also shown is the observed Lyman-α as described in Section 6.6. This profile also incorporates the artificial turbulent velocity as described in Section 3.3.2.

**Figure 3.3:** *Effect of ambipolar diffusion on hydrogen.* Top panel displays neutral hydrogen population with and without the effect of ambipolar diffusion. Lower panel displays the ratios for neutral hydrogen level one and two, and ionized hydrogen together with Lyman-α profile of the quiet sun model atmosphere FAL-C-99. Main result is the limited effect of ambipolar diffusion on Lyman-α. However, it does affect the continuum as shown in the following figure, Figure 3.4.

Lambda iteration is still slow and prone to divergence and numerical instability in the transition region, or more precisely where the hydrogen level changes state from mostly neutral state to dominantly ionized (Figures 3.6–3.9). This problem can be diminished by changes in $\Gamma$. Outside this region, both below and above this region, the convergence is stable. I expect that stability improve when ALI is also implemented for lines. One source of numerical instability is that the lines sometimes “lag behind” the continuum radiation field, causing an overcorrection the next step. On the other hand, this method rejects one of the two states obtained.
Figure 3.4: Ratio of spectrum with and without ambipolar diffusion. The spikes at 92 nm, 97 nm, 121 nm are due to flattening of Lyman-profiles due to ambipolar diffusion.

Figure 3.5: Illustration denotes difficulty of convergence. It displays the hydrogen populations ($\text{H I-1}$ (dashed) and ionized hydrogen (solid)) of the quiet sun model FAL-99-C, using different $\Gamma$’s (thin and thick). The difference in population is most pronounced in the chromosphere and lower transition region. The upper transition region is hot enough to fully ionize hydrogen and hence both departure coefficient approach unity.

Hubeny and Lanz (2003, p. 55–66) presents an extensive list of possible improvements to the numerical side of the problem, some of which are already implemtended, including Broyden, Ng, opacity distribution functions and (partly) ALI.
Figure 3.6: Net radiative brackets for depth points 15, 17, ..., 45 for hydrogen levels 1-2 (blue), 1-3 (light-blue) to 1-4 (yellow), 1-11 (red) and the ambipolar rate (black, dashed) for iteration steps 1 to 900 of the first run of model A, without line blanketing. Y-Axis describe net radiative rates, X-Axis is the number of iteration. Negative rates are dotted. The rates change over iterations, showing that they are not yet converged. The noise that is most visible at depth points 23–33 is due to numerics.
3.2. Radiative Transfer

Figure 3.7: The same figure as Figure 3.6, but for the second iteration, which is the first with line-blanketing enabled. The peaks show a change of the net radiative brackets brackets to absolute rates while the jump at iteration 400 is due to changing the gammas of the accelerated lambda iterator, accelerating the convergence at cost of stability.
Figure 3.8: The same figure as Figure 3.6, but for the second iteration with line-blanketing. The jump at iteration 400 is again due to changing the gammas of the ALI, causing a further change in the rates.
Figure 3.9: Fourth iteration with ambipolar diffusion enabled. While changing gammas do influence the rates (depth points 17–25 at iteration step 400, 600 and 800) it causes numerical instability (depth points 31–41)
Figure 3.10: Fifth iteration, this iteration is calculated without ambipolar diffusion, but based on the model described in Figure 3.9.
3.3 Spectral Synthesis Code

FIOSS is a spherical spectrum synthesizer, formerly called SYNSPEC, originally developed by Hubeny (1988); Hubeny et al. (1994). Based on the output from H-Minus and atomic data, FIOSS calculates a synthetic spectrum under LTE assumption but with NLTE populations from HMINUS. Most lines (3.6 millions) are calculated under LTE assumptions, while the most important ones, several hundred from thirty different elements, are calculated in NLTE. See Table 3.1 for a list of those elements.

Spherical geometry yields a more correct intensity at the limb compared to a plane-parallel geometry (Haberreiter et al., 2008; Haberreiter, 2006; Mihalas, 1978, p. 251 ff). Molecular lines where introduced by Shapiro et al. (2010).

3.3.1 Line Formation

The atomic line data are taken from Kurucz line lists. Almost all line profiles are calculated using the general Voigt profile while the most important ones like the Balmer, Lyman and Helium lines are calculated explicitly under NLTE assumptions.

Stark Broadening

The previous code uses Stark broadening tables by Vidal et al. (1971) for the Balmer lines up to \( n = 10 \) with the other lines using an approximate stark broadening. The Stark broadening for the first four hydrogen series Lyman, Balmer, Paschen and Brackett up to \( n = 22 \) are now supplied by Lemke (1997). They adapted Vidal et al. (1971) by changing the array dimensions to allow for the higher quantum numbers. The mathematical and physical description of the model has not been changed. An comparison of Lemke to the approximate stark broadening is given in Figure 3.11. The difference of the methods are generally very small, with the most pronounced effect in the wing near line centre with the Lemke profile causing a flattening of the line (i.e. lower radiance for emission lines and higher for absorption lines) and an increase in the far-wing, additionally flattening the line profile.

Figure 3.11: Comparison of Lyman \( \alpha \) (left) and Balmer \( \alpha \) (right) lines using Lemke (solid) and the approximate stark broadening (dashed). The difference is most pronounced in the wings near the line centre, but generally small, showing that the approximate stark broadening is sufficient.
Table 3.1: Table of all Elements that are calculated under full NLTE. All elements are implemented from neutral to the single-ionized state, with the exception of hydrogen, which also considers the negative charged state. Most elements only consider the ionization levels without the lines.

When the temperature profile of the atmosphere has a local minimum, as is the case in the sun, where the temperature increases above the photosphere, the spectrum synthesizer uses the minimum temperature for all points above the minimum for LTE lines. Otherwise all lines become emission lines. Hence important emission lines must be calculated explicitly.

**Formation Height**

The formation height of a line at frequency $\lambda$ is defined as the point where the optical depth $\tau(\lambda) = 1$ while the Rosseland optical depth is defined as $\tau(\lambda) = 2/3$, accounting for the spherical influence, assuming a grey atmosphere. See Figure 3.12 for the formation height of the spectrum of model A. Since the formation height depends on the frequency, one can obtain height dependent information from observations of different frequencies, as done in section 9.1.

**Additional Background Opacity**

The LTE calculations overestimate the UV flux up to 300 nm. This effect increases with NLTE calculations due to an increase in ionization of iron and other metals from the hotter part of the atmosphere to the photosphere, causing a decrease in neutral elements relative to the LTE case.

This can be alleviated by utilizing opacity distribution functions which includes the background line opacities into the NLTE calculations (Haberreiter, 2006; Haberreiter et al., 2008).

There still remains an increased UV flux up to 320 nm due to possible missing lines in the Kurucz line list and missing background opacities. Recently the method to account for the missing opacities in the UV was changed. Previously, Haberreiter (2006) broadened the wing of each with a wavelength dependent increase of the turbulent velocity. Shapiro et al. (2010) instead adds opacities to the continuum and hence is visible in the continuum calculations. The opacities are added by an iterative procedure. In both cases an increase of opacities is needed to fit the spectrum, indicating missing lines in the ultraviolet spectrum.

Both methods adjust most of the flux in the UV. The new method allows for a faithful reproduction of the line profiles present in the line-list. While this approach can rectify the flux of weak lines which are formed in the photosphere and the transition region, it does not influence the Lyman-$\alpha$ flux due to its formation in the transition region.
3.3. Spectral Synthesis Code

3.3.2 Quiet Sun Spectrum

Figure 3.13 shows the flux and formation height of Fontenla model C. The infra-red follows closely the Planck curve at about 5000 K while the $T_{\text{eff}}$ of the quiet sun is 5777 K. The minimum formation height of 20 km below the photosphere at 1629 nm (see Figure 3.12) is the minimum opacity due to the sum of the $H^{-}$ bound-free and $H^{-}$ free-free opacities.

Figure 3.13 displays the flux and formation height with and without correcting the missing fluxes as described in Chapter 4 for the continuum and the spectrum including lines. The flux is compared to the Spectral Solar Irradiance (SSI) composite from LASP Interactive Solar Irradiance Datacenter (LISIRD).

Figure 3.14 is the height dependent the departure ratio of hydrogen levels. The NLTE effect becomes significant at approximately the formation height of the visible part of the spectrum. The NLTE effect for H I-1 is up to $10^{10}$ in the upper transition region. Figure 3.15 compares the quiet sun spectrum to a black body spectrum with temperatures at 5000 K, 5777 K and the temperature at the formation height of the specific frequency.

To compensate for the missing corona in the FAL models, the turbulent velocities from the FAL models are increased by a factor of four in the upper transition region up to the formation height of Ly-$\alpha$. Below the formation height of Ly-$\alpha$ this factor decrease to one over one kilometer of height of the atmosphere, resulting in a better agreement with observation. Figure 3.16 shows the new and old turbulent velocities, together with the ratio. Figure 3.17 displays the effect on population and the Lyman-$\alpha$ profile. It displays a difference in population of up to 10% for the ionized hydrogen, up to 5% for neutral hydrogen at level one, and 25% difference for neutral level two hydrogen at 2100 km above photosphere, the formation height of the Lyman-$\alpha$ wing.

Comparison of COSI to ATLAS 3

Figure 3.18 compares the synthetic COSI spectrum to the ATLAS 3 spectrum. For this comparison the Wenzler (2005) approach is used with the three component model consisting of the quiet sun, plage and sunspots with contributions of 3.4% and 0.8 ppm for plage and sunspots, respectively with the remainder set to the quiet sun. While the near-infra-red difference is below 10%, the visible difference goes up to 20% due to missing lines and the opacity correction as explained in Section 4.4.1. The
Figure 3.13: Flux (top panel) and formation height (bottom panel) of the current day quiet sun model. Shown are the flux and formation height of the continuum with (solid) and without (dash-dotted) correction for missing opacities, the flux and formation height of the model, including all lines (dashed) and irradiance measurement from SOLSTICE (Spectral Composite, grey). The formation height is given as 100 km above $\tau_{500} = 1$. 
3.3. Spectral Synthesis Code

Figure 3.14: Departure coefficients of hydrogen for model C. The NLTE effect becomes significant from a depth of approximately 40 km above $\tau_{500\,\text{nm}} = 1$ outwards.

Figure 3.15: Irradiance (solid) and the Planck spectrum for temperatures of the infra-red (5000 K, crosses), the effective temperature of the quiet sun 5777 K (dots) and the Planck spectrum derived from the temperature at the formation height of the specific wavelength (circle). Bottom panel compares the flux to the temperature at formation height.
average difference for the spectrum from 200 nm into the infra-red is below 5%, mainly due to the opacity correction. The Asplund abundances (Scott et al., 2006, and references therein) are used for this work. They are calculated solar abundances resulting from a 3D NLTE code. The congruence of COSI with ATLAS 3 observation can be improved by using a mixture of the Asplund abundances with Greevese ones, as done in section 8.3, Shapiro et al. (2010), resulting in an 98% agreement with observation.

The missing NLTE lines become apparent in the UV, while the treatment of LTE lines as described in section 3.3.1 induce absorptions instead of emissions, causing the divergence to observations in this region. Also the difference at Ly-α is partly due different folding used and possibly partly due an uncertainty of the solar surface features.
3.3. Spectral Synthesis Code

(a) Relative change (old turbulent velocities divided by new ones) of hydrogen population in the chromosphere for neutral hydrogen level one (solid), level two (dashed) and ionized hydrogen (dash-dot).

(b) Effect on Lyman-α profile. The new (solid) profile is compared to both the old (dashed) and observation (dots). For a more detailed description and discussion of the observations, see Section 6.6. The position of the peaks and the gradient of the wings are parts that should match.

Figure 3.17: Effect of new turbulent velocity. Shown is the quiet sun model FAL-C-99 change in hydrogen population and how it affects the Lyman-α profile.
Figure 3.18: Comparison of the Atlas-3 reference spectrum to COSI. The contribution factors follow Wenzler’s method, with a sunspot contribution of 0.8 ppm and a plage part of 3.4%. For the remainder the quiet sun model FAL-C-99 is used. This contributions are explained in more detail in Section 8.
Chapter 4

NLTE Solar Irradiance Modelling with COSI

Published in A&A, Volume 517, July-2010

A. I. Shapiro, W. Schmutz, M. Schoell, M. Haberreiter, E. Rozanov

1 Physikalisch-Meteorologishes Observatorium Davos, World Radiation Center, 7260 Davos Dorf, Switzerland
2 Institute for Astronomy ETH, Zurich, Switzerland
3 Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80303, USA
4 Institute for Atmospheric and Climate science ETH, Zurich, Switzerland

Received 30 December 2009; accepted 29 March 2010

Abstract

The solar irradiance is known to change on time scales of minutes to decades, and it is suspected that its substantial fluctuations are partially responsible for climate variations. We are developing a solar atmosphere code that allows the physical modeling of the entire solar spectrum composed of quiet Sun and active regions. This code is a tool for modeling the variability of the solar irradiance and understanding its influence on Earth. We exploit further development of the radiative transfer code COSI that now incorporates the calculation of molecular lines. We validated COSI under the conditions of local thermodynamic equilibrium (LTE) against the synthetic spectra calculated with the ATLAS code. The synthetic solar spectra were also calculated in non-local thermodynamic equilibrium (NLTE) and compared to the available measured spectra. In doing so we have defined the main problems of the modeling, e.g., the lack of opacity in the UV part of the spectrum and the inconsistency in the calculations of the visible continuum level, and we describe a solution to these problems. The improved version of COSI allows us to reach good agreement between the calculated and observed solar spectra as measured by SOLSTICE and SIM onboard the SORCE satellite and ATLAS 3 mission operated from the Space Shuttle. We find that NLTE effects are very important for the modeling of the solar spectrum even in the visual part of the spectrum and for its variability over the entire solar spectrum. In addition to the strong effect on the UV part of the spectrum, NLTE effects influence the concentration of the negative ion of hydrogen, which results in a significant change of the visible continuum level and the irradiance variability.

Key Words: Line: formation – atomic data – Molecular processes – Sun: atmosphere – Sun: UV radiation – Radiative transfer
Chapter 4. NLTE Solar Irradiance Modelling with COSI

4.1 Introduction

The solar radiation is the main source of the input of energy to the terrestrial atmosphere, so that it determines Earth’s thermal balance and climate. Although it has been known since 1978 that the solar irradiance is not constant but instead varies on scales from several minutes to decades (cf. Fröhlich, 2005; Krivova and Solanki, 2008), the influence of this variability on the climate is not yet fully understood. Nowadays several datasets for the past spectral solar irradiance (SSI) based on the different reconstruction approaches (e.g. Lean et al., 2005; Krivova et al., 2009a) and satellite measurements are available. However, the remaining disagreements between these data lead to different atmospheric responses when they are used in the climate models (Shapiro et al., 2011b). The task of constructing a self-consistent physical model in order to reconstruct the past solar spectral irradiance (SSI) remains of high importance.

Modern reconstructions of the SSI are based on the assumption that the irradiance changes are determined by the evolution of the solar surface magnetic field (Foukal and Lean, 1988; Krivova et al., 2003; Domingo et al., 2009). Areas of the solar disk are associated to several components (e.g. quiet Sun, bright network, plage, and sunspot) according to the measured surface magnetic field and the contrast of the features. These components are represented by corresponding atmosphere structures (cf. Kurucz, 1991; Fontenla et al., 1999). The SSI is calculated by weighting the irradiance from each model with the corresponding filling factor.

The calculation of the emergent solar radiation, even from the atmosphere with known thermal structure, is a very sophisticated problem because consistent models have to account for the NLTE effects in the solar atmosphere. As the importance of these effects has become clear over the past several decades, several numerical codes have been developed. One of the first NLTE-codes, LINEAR, was published by Auer et al. (1972), who used a complete linearization method developed by Auer and Mihalas (1969, 1970). Later, the MULTI code was published by Carlsson (1986). This code is based on the linearization technique developed by Scharmer (1981) and Scharmer and Carlsson (1985a,b). More recently, the RH code, which is based on the MALI (Multi-level Approximate Lambda Iteration) formalism of Rybicki and Hummer (1991, 1992) has been developed by Uitenbroek (2001).

Combining the radiative transfer code by Hamann and Schmutz (1987) and Schmutz et al. (1989) and spectral synthesis code SYNSPEC by Hubeny (1981), Haberreiter et al. (2008) has developed the 1D spherical symmetrical COde for Solar Irradiance (COSI). The NLTE opacity distribution function (ODF) concept, which was implemented in this code, allows an indirect account for the NLTE effects in several million lines. This makes COSI especially suitable for calculating the overall energy distribution in the solar spectrum. In this paper we introduce a new version of COSI (version 2), describe the main modifications, and present first results.

In Sect. 4.2 we describe the datasets of the measured spectral irradiance, which were used for comparison with our calculated spectrum. In Sect. 4.3 we introduce the integration of molecular lines into COSI and show that it can solve the discrepancies to observations and to other codes. In Sect. 4.4 we present the resulting solar spectrum from NLTE calculations. Due to missing opacity in the UV, the flux appears to be significantly higher than measured (Busá et al., 2001; Short and Hauschildt, 2009). We solve this problem by introducing additional opacity to the ODF for selected spectral ranges (Sect. 4.4.1), while the problem of the NLTE visible continuum due to the deviations in the concentration of the hydrogen negative ion is addressed in Sect. 4.4.2. In Sect. 4.5 we present the synthetic spectra of the active regions and its implications for the solar variability study. Finally, we summarize the main results in Sect. 4.6.

4.2 Measured solar spectral irradiance

We compared our calculated solar spectrum with several available datasets. We used the observations taken with SOLSTICE (SOLar-STellar Irradiance Comparison Experiment, McClintock et al. (2005)) up to 320 nm and the Solar Irradiance Monitor (SIM) (Harder et al., 2005) from 320 nm onward instruments onboard the SORCE satellite (Rottman, 2005) obtained during the 23rd solar cycle minimum
(hereafter SORCE measurements). For the comparisons shown in this paper, we used the average of the observations from 21 April 2008 to 28 April 2008.

Our calculated spectrum was converted with a 1 nm boxcar profile for comparison with the SOLSTICE measurements and trapezoidal profile for comparison with the SIM measurements. The SIM resolution strongly depends on wavelength (FWHM is about 1.5 nm for the 310 nm and about 22 nm for the 800 nm) so the parameters of this trapezoidal profile are also wavelength dependent and were provided by Harder (2009).

We also used the SOLar SPECtral Irradiance Measurements (SOLSPEC) (Thuillier et al., 2004) during the ATLAS 3 mission in November 1994. For this comparison we convolved the calculated spectrum with a Gaussian with FWHM = 0.6 nm.

4.3 Molecules in the COSI code

COSI simultaneously solves the equations of statistical equilibrium and radiative transfer in 1D spherical geometry. All spectral lines in COSI can be divided into two groups. The first group comprises about one thousand lines that are the most prominent in the solar spectrum. These lines are explicitly treated in NLTE. The second group contains several million background lines provided by Kurucz (2006) and calculated under the assumption that their upper and lower levels are populated in LTE relative to the explicit NLTE levels. These background lines are taken into account in the spectral synthesis part of the code but also affect NLTE calculations via the ODF. The ODF is iteratively recalculated until it becomes self-consistent with populations of the NLTE levels (Haberreiter, 2006; Haberreiter et al., 2008). This allows us to indirectly account for the NLTE effects in the background lines.

Molecular lines were not included in the previous version of COSI (Haberreiter et al., 2008). However, molecular lines play an important role in the formation of the solar spectrum. In some spectral regions they are even the dominant opacity source. Therefore, a considerable amount of opacity was missing. This led to a discrepancy with observations. In this version of COSI molecular lines were taken into account. This requires computing molecular concentrations, preparing the molecular line lists, and incorporating molecular opacities and emissivities into the code.

4.3.1 Chemical equilibrium

Under the assumption of the instantaneous chemical equilibrium, the concentration of molecule AB is connected with the concentrations of atoms A and B by the Guldberg-Waage law (cf. Tatum, 1966):

\[
\frac{n_A}{n_B} = K_{AB}(T) = \left( \frac{2\pi mkT}{\hbar^2} \right)^{3/2} \frac{Q_A Q_B}{Q_{AB}} kT \exp\left(-\frac{D_0}{kT}\right),
\]

where \(K_{AB}\) is the temperature dependent equilibrium constant, \(Q\) internal partition function for the atoms A and B, \(m\) is the reduced molecular mass, and \(D_0\) the dissociation energy. The contribution to the opacity from the transitions between lower level \(i\) and upper level \(j\) of molecule AB is given by

\[
\kappa_{AB}^{ij}(\nu) = n_{AB} g_i \exp\left(-\frac{E_i}{kT}\right) \frac{\hbar \nu}{4\pi} (n_i B_{ij} - n_j B_{ji}) \phi(\nu - \nu_{ij}),
\]

where \(g_i\) is statistical weight of the lower level, \(B\)'s are the Einstein coefficients, and \(\phi(\nu)\) is the line profile normalized to one.

There are several available temperature polynomial approximations for the atomic and molecular partition functions, as well as for the equilibrium constants \(K_{AB}\) (cf. Tatum, 1966; Tsuji, 1973; Irwin, 1981; Sauval and Tatum, 1984; Rossi et al., 1985). Although molecular partition functions can contain significant errors due to the inaccurate values of the energy levels and even due to unknown electronic states (Irwin, 1981), this does not significantly affect the accuracy of molecular opacity calculations. This insensitivity exists because the overall molecular concentration is proportional to the partition
Figure 4.1: The ratio \( R \) of the number concentration of carbon and oxygen that are not attached to the molecules to their total element amount (upper panels) and the CN and CO concentrations as a function of height (lower panels) for three atmospheric models: FAL-A-99 (dashed curves), FAL-C-99 (continuous curves), and FAL-P-99 (dotted curves). The zero-height depth point in all models is defined as the radius at which the continuum optical depth at 5000 Å is equal to one.

function (see Eq. 4.1), while the opacity in a given line is inversely proportional to it (see Eq. 4.2). Therefore it is crucial to either calculate equilibrium constants directly from partition functions or use the same source for both approximations.

Tsuji (1973) shows that abundances of hydrogen, carbon, nitrogen, and oxygen can be determined quite accurately without taking into account their dependencies on other elements. Because we are only interested in molecular bands that can significantly contribute to the opacity over a broad spectral range, the system of equations like Eq. (4.1) were solved only for these four main elements and the diatomic molecules built from them.

Because a significant fraction of atoms can be associated with molecules the chemical equilibrium calculation also affects the atomic lines. In Fig. 4.1 we present the changes in the carbon and oxygen concentrations due to the association with molecules, together with CN and CO concentrations for three atmospheric models by Fontenla et al. (1999): the relatively cold supergranular cell center model FAL-A (hereafter FAL-A-99), the averaged quiet Sun model FAL-C (hereafter FAL-C-99), and the relatively warm bright network model FAL-P (hereafter FAL-P-99). Both molecular concentrations and deviations in atomic concentration show strong temperature sensitivities, and this is important for assessing of the solar spectral variability.

4.3.2 Main molecular bands in the solar spectrum

The line list for the spectrum synthesis was compiled using the molecular databases of Kurucz (1993) and Solar Radiation Physical Modeling (SRPM) database (e.g. Fontenla et al., 2009b), which is based on Gray and Corbally (1994) and HITRAN. Wavelengths and line strengths of the most significant lines of the CN violet system and CH G band were calculated based on the molecular constants by Krupp (1974), Knowles et al. (1988), and Wallace et al. (1999). The OH and CH continuous opacities were calculated according to Kurucz et al. (1987) under the LTE assumption.

In Fig. 4.2 we present the part of the solar spectrum where the LTE calculations with the previous version of COSI, which did not account for molecular lines, showed significant deviations from the calculations with the LTE radiative transfer code ATLAS12 carried out by Kurucz (2005). Both calculations used the same atmosphere structure and abundances by Kurucz (1991) (hereafter K91). One can see that introducing the molecular lines into COSI quite strongly affects the spectrum and significantly diminishes disagreements with the ATLAS 12 code. The most prominent features are the CH G band around 430 nm, CN violet G band around 380 nm, CN, NH, and OH bands between 300
4.3. Molecules in the COSI code

Figure 4.2: LTE calculation of molecular lines with COSI. Upper panel: Synthetic solar spectrum calculated with (solid line) and without (dashed line) molecular lines. Lower panel: Corresponding ratios to the irradiance calculated by the ATLAS12 code. All spectra are averaged with a 1-nm boxcar.

Figure 4.3: The irradiance calculated by COSI in LTE (thin solid line) and the ATLAS 12 (dashed line), using K91 atmosphere structure and abundances. The spectrum observed by SORCE is given by the thick solid line. All spectra are convolved with the instrument profile of SORCE.

and 350 nm. The remaining differences can possibly be attributed to the different photoionization cross sections implemented in COSI and ATLAS12.

In Fig. 4.3 we compare the spectra calculated by both codes with the SORCE measurements. One can see that both codes predict the correct value of the continuum in the red part of the visible spectrum (600 nm) but fail to reproduce the molecular bands and UV part of the spectrum. The LTE spectra calculated with the FAL-C-99 and two sets of abundances by Grevesse and Anders (1991) (hereafter G91) and by Asplund et al. (2005) (hereafter A05) are presented in Fig. 4.4. The calculations with the G91 abundances predict the correct value of the visible and IR continua. Furthermore, they give quite a good fit of molecular bands (see right panel of the picture where the calculation of CN violet system and CN G band are shown in more detail). However the calculated irradiance in the UV part is again much higher than observed for both abundance sets.

The metal abundances by A05 are significantly lower (up to about 2 times) than in G91. The lower metal abundance leads to a significantly lower electron density, consequently to a lower concentration of negative hydrogen ion. The latter is the main source of the continuum opacity in the solar atmosphere (cf. Mihalas, 1978, p. 102). Therefore the application of the A05 abundances results in a lower continuum opacity for the visible and IR wavelengths and as a consequence in a higher irradiance. This effect is especially prominent at longer wavelengths (from about 400 nm) where the number of strong lines is relatively low and the opacity of negative hydrogen dominates all other sources of opacity.
In turn, the lower metal abundance given by A05 leads to lower line opacities and photoionization opacities in the UV. As a result, the A05 abundances also lead to a higher irradiance in the UV than the G91 abundances. The FAL-C-99 model has so far mainly been used with G91 abundances.

### 4.4 NLTE calculations

Most of the solar radiation emerges from regions of the atmosphere with a steep temperature gradient. Therefore a significant part of the escaping photons is not in the thermodynamic equilibrium with the surrounding medium. This implies that a self-consistent radiative transfer model has to account for NLTE effects. It is well-known that these effects are extremely important for the strong lines and the UV radiation. In this section we present NLTE calculations of the solar spectrum with COSI and show that the NLTE effects are also important for the overall energy distribution in the solar spectrum, including the visible and infrared wavelength ranges. We want to emphasize here that the NLTE calculations presented in this paper were performed using temperature and pressure profiles of the solar atmosphere obtained by Fontenla et al. (1999). These profiles in turn depend on assumptions,
so all effects presented in this paper are relative to the assumed model of the solar atmosphere (see also the discussion at the end of Sect. 4.4.3).

In the left panel of Fig. 4.5, the NLTE calculations of the solar spectrum with the atmosphere structure FAL-C-99 and the abundance sets A05 and G91 are compared to the observed solar spectrum. The right panel of Fig. 4.5 shows the ratios of the NLTE to the LTE calculations. The most prominent overall NLTE effects are the strong increase in the irradiance in the UV and mild decrease of the irradiance in the red part of the spectrum. The first effect is discussed in detail in Sect. 4.4.1, and the second one in Sect. 4.4.2.

4.4.1 Calculating of the UV radiation

NLTE opacity effects

Independent of the solar atmosphere model and abundances, the LTE calculations overestimate the UV irradiance up to about 300 nm (see Figs. 4.3 and 4.4). This problem is even more severe in the NLTE calculations as they predict a UV irradiance that is higher than in the case of the LTE calculations (see left panel of Fig. 4.5).

The effect of the increase in the far UV flux in the NLTE calculations is actually well known (cf. Short and Hauschildt, 2009). The UV radiation incident from the higher and hotter parts of the solar atmosphere onto the photosphere causes a stronger ionization of iron and other metals relative to the LTE case. This decreases the populations of the neutral atoms and consequently weakens the strength of the corresponding line and photoionization cross section, which are the main sources of the opacity in the far-UV. Therefore it leads to a decrease in the opacity and consequently an increase in the flux in the wavelengths up to about 250 nm (red threshold of the photoionization from the Mg I 2 level). In the longer wavelengths the line and the photoionization opacities of the ionized metals become more important so the increase in their populations results in lower irradiance.

The effect of the increased UV flux in the NLTE calculations can be decreased significantly by the use of the opacity distribution function (see Fig. 4.6), which allows incorporation of the opacity from the background lines into the solution of statistical equilibrium equations (Haberreiter et al., 2008). This background opacity diminishes the amount of the penetrating UV radiation and decreases the degree of metal ionization. It leads to the relative decrease in the spectral irradiance up to 250 nm and an increase at higher wavelengths (in agreement with the above discussion). The complicated wavelength dependence of the spectrum alterations caused by the influence of the ODF can be explained by the

Figure 4.6: Left panel: NLTE calculations with and without iterated ODF, compared to the solar irradiance as observed by SORCE. Right panel: Ratio between the emergent flux from NLTE calculated with and without iterated ODF. All calculations are done with the FAL-C atmosphere model and G91 abundances.
large number of the strong metal lines and photoionization thresholds. Overall, introducing the ODF causes the redistribution of the solar irradiance over the entire spectrum as it decreases the far-UV part and increases the near-UV and visible.

Although the concept of the ODF allows the agreement between calculations and observations to be improved, the ODF based on the Kurucz linelist (Haberreiter et al., 2008) cannot solve the discrepancy with measurements. The most probable source of the problem is the inaccuracy of the atomic and molecular line list and possible missing continuum opacity. The production of a complete line list is an extremely laborious task. Considerable progress has been achieved during past decades thanks to the effort of the OPACITY project (Seaton, 2005) and, in particular, of Kurucz (1993) in calculating atomic data. Concerning the problem of the UV blanketing, it is important to keep in mind that about 99%(!) of the atomic and molecular lines are predicted theoretically and only 1% of all lines are observed in the laboratory (Kurucz, 2005). This leads to potential errors in the total opacity as lines could still be missing while wavelengths and oscillator strengths of the predicted lines could be inaccurate. An incompleteness of the atomic data is especially significant in the UV where the immense number of lines form the UV line haze.

The consistent account for the line blanketing effect is very important for the overall absolute flux distribution, especially for the UV (cf. Collet et al., 2005; Avrett and Loeser, 2008; Fontenla et al., 2009b). Busá et al. (2001) used version 2.2 of the MULTI code, which UV opacity were contributed only from several hundred calculated in the NLTE lines, bound-free and free-free processes. They showed that NLTE calculations without proper account for the line blanketing effect can lead to overestimating of the UV flux by six orders of magnitude in the case of cold metal-rich stars. To compensate for this, they multiplied the continuum opacity by a wavelength-dependent factor and gave an explicit expression for the factor parameterization. A similar approach is used by Bruls (1993) to calculate the NLTE populations of Ni I. Short and Hauschildt (2009) show that the use of different line lists for the line blanketing calculations can lead to significantly different spectral irradiance distributions. They also show that the problem of the excess of the UV flux in the 300-420 nm spectral region can be solved by increasing of the continuum absorption coefficient.

### Additional opacities in COSI

COSI takes the opacity from the several million lines into account (Haberreiter et al., 2008). However, the significant disagreement between the calculated and measured solar UV flux as discussed above indicates that the computations still miss an essential part of the opacity. To compensate for this missing opacity, Haberreiter et al. (2008) calculated the UV part of the solar spectrum with artificially increased Doppler broadening. This approach proves itself successful, but it cannot be used to fit high-resolution spectra. Therefore, here we have extended the approach by Busá et al. (2001) and Short and Hauschildt (2009) and included additional opacity in our NLTE-ODF scheme, as described below.

COSI overestimates the irradiance, i.e., shows an opacity deficiency in the spectral region between 160 nm and 320 nm. At shorter wavelengths, the opacity is dominated by continuum photoionization, while at longer wavelengths tabulated atomic and molecular lines are sufficient to reproduce the observations. In this wavelength region, we grouped the solar spectrum into 1 nm intervals (corresponding to the resolution of the SOLSTICE/SORCE spectrum). For each interval we empirically determined the coefficient \( f_c(\lambda) \) by which the continuum opacity has to be multiplied to reproduce the SOLSTICE measurements. The additional emission coefficient was calculated under the assumption that the missing opacity source has thermal structure, so its source function is equal to the Planck function. Because the correction of the ODF affects the statistical equilibrium of the populations, we iteratively solve the statistical equilibrium equations and the factors \( f_c(\lambda) \) using newly updated ODFs until we reached a self-consistent solution. Thus, the \( f_c(\lambda) \) factors were found iteratively. We stopped the iteration when the deviations to the observed spectrum was smaller than 10%. In the considered region there were very few wavelength points where the irradiance was underestimated. We did not apply any correction to these points. This underestimation can be explained by the inaccuracy of the line list because it can be caused by the small wavelength inaccuracy of a few strong lines.
4.4. NLTE calculations

The wavelength dependence of \( f_c(\lambda) \) is shown in Fig. 4.7 for the calculations with the FAL-C-99 atmosphere model and G91 and A05 sets of abundances. One can see that \( f_c(\lambda) \) is very close to 1 (which means little additional opacity) for wavelengths longer than 280 nm. There are also several prominent peaks that indicate a strong lack of opacity at specific wavelengths. Although different abundances lead to a change in the peak amplitudes, the overall profile of wavelength dependence stays the same for both sets of abundances, G91 and A05, respectively. The averaged line opacity is higher by several orders of magnitude than the continuum opacity, so the ratio of the additional opacity, which was needed to achieve agreement with the observed solar spectrum, to the total opacity that was already included into COSI is for most wavelength below 5 % (see lower panel of Fig. 4.7). Therefore, the correction of the overestimation of the emergent UV flux can be achieved by a moderate modification of the input line list.

Calculations with the G91 abundances lead to a lower continuum than calculations with A05 abundances. Therefore the flux overestimation is stronger in the case of calculations with A05 abundances. Consequently one has to use larger correction factors \( f_c(\lambda) \) for A05 abundances. The correction factor strongly depends on the chosen line list. We tested our calculations using the Vienna Atomic Line Database (VALD) (Kupka et al., 1999, 2000) and retrieving the lines with known radiative damping parameter. Because this database contains fewer lines than provided by Kurucz (2006), a larger factor \( f_c(\lambda) \) is needed to achieve agreement with the observations.

4.4.2 Calculation of the visible and IR radiation

NLTE effects decrease the visible and near IR irradiance by about 10 % (see Fig. 4.5), because they influence the concentration of electrons and the negative hydrogen ions that determines the continuum opacity, as well as the continuum source function.
**Figure 4.8:** NLTE effects in electron and H\(^-\) concentrations. Upper panel: depth dependence of the total electron concentration and the electron concentrations resulting from the ionization of elements as indicated. Second panel: ratios of the change of the electron concentrations due to the NLTE effects and total electron concentration calculated in LTE. Plotted are total changes with and without applying ODF and resulting from the ionization of particular elements change without ODF. Third panel: depth dependence of the ratio of negative hydrogen and electron concentration calculated in NLTE and LTE. Lower panel: Contribution functions to the intensity for the 5000 Å (dotted lines) continuum radiation and temperature on depth dependence (solid curve) for the FAL-C-99 atmosphere model. The contribution functions are plotted for eight values of the cosine of the angle between the propagation direction of radiation and the local solar radius \(\mu\). From right to left: \(\mu = 0.05\) (almost solar limb), 0.1, 0.2, 0.4, 0.6, 0.8, 0.9, 1.0 (solar disk center). All calculations are carried out for the FAL-C-99 atmosphere model and A05 abundances.

**NLTE effects in electron and negative hydrogen concentrations**

Departures from LTE in the electron and hydrogen negative ion concentrations are presented in Fig. 4.8 for the calculations with FAL-C-99 atmosphere structure and A05 abundances. The upper panel illustrates the main electron sources in the solar atmosphere. In the lower part of the photosphere and in the chromosphere, hydrogen is ionized and provides most of the electrons; however, in the photosphere, which is responsible for the formation of the visible and near-IR irradiance, the temperature is too low for ionizing hydrogen. Therefore, here the electrons are mainly provided by the metals with low ionization potential, such as iron, silicon, and magnesium.

As one can see from the second upper panel of Fig. 4.8, the ionization of silicon and iron is strongly affected by the NLTE effects. This is basically the same effect of NLTE “overionization” as already discussed in Sect. 4.4.1, where the main emphasis was on the deviations in the concentration of the neutral atoms. Applying the ODF significantly decreases the NLTE effects on the electron concentration as it strongly increases the UV opacity. In the last part of Fig. 4.8, the contribution functions for the continuum radiation are plotted for different \(\mu\)-positions. The contribution functions show where the emergent intensity originates (see, e.g. Gray, 1992, p. 151), so one can see that, although the continuum is mainly formed in the lower part of the photosphere where the NLTE effects in the
electron concentration are not so prominent, they still can be important, especially for the radiation emitted close to the solar limb.

The photoionization cross section of negative hydrogen has its threshold at about 16 500 Å, and it reaches its maximum at about 8 500 Å (cf. Mihalas, 1978, p. 102). Although the photosphere is optically thick for the UV and radiation in the strong atomic or molecular lines, the visible and IR photons can easily escape, even from the lower part of the photosphere. (The optical depth of the zero point in the continuum radiation is about one, depending on the wavelengths.) This means that all over the photosphere there is a lack of photons that are able to photoionize the negative hydrogen. Thus, in the NLTE case, the radiative destruction rate of H\(^-\) is decreased, which in turn increases the concentration of the negative hydrogen ion compared to the LTE equilibrium. This is illustrated in the third panel of Fig. 4.8. The effect is especially strong in the layers where the continuum radiation is formed (see the contribution functions in the lower panel of Fig. 4.8). Moreover, in contrast to effects on the electron concentration, applying the ODF does not decrease the NLTE deviations of the negative hydrogen concentration since UV radiation does not make any significant contribution to its ionizations. The effect of negative hydrogen underionization was first pointed out by Vernazza et al. (1981), who found that photospheric departure coefficients of the negative hydrogen (ratios between NLTE and LTE concentrations) are greater than one.

Thus, compared to the LTE case, NLTE ionization of hydrogen and metals result in higher electron concentration. The NLTE ionization of negative hydrogen also leads to the increase in its concentrations. This enlarges the continuum opacity so that continuum radiation comes from higher and cooler layers of the photosphere. In addition, there is also a decrease in the NLTE continuum source function, which contributes to the decrease in the continuum emission.

The effects described above significantly decrease the level of the emergent flux in the visible and near IR. Although the NLTE treatment is physically more consistent than LTE, it leads to a severe deviation from the observed solar spectrum (see Fig. 4.5).

Collisions with negative hydrogen

One possible cause of the inability of the NLTE calculations to reproduce the measurements is the inadequate treatment of collisional rates for the negative hydrogen ion. The present version of COSI accounts for the electron detachment through collisions with electrons and neutral hydrogen, as well as charge neutralization with protons as given by Lambert and Pagel (1968). We do not treat collisions of H\(^-\) with heavy elements, which in principle could lead to a strong underestimation of the collisional rates. We found that increasing the collisional rates involving H\(^-\) by a factor of ten solves the disagreement between the measured and calculated continuum flux for the models using A05 abundances. However, when using increased collisional rates, the molecular bands (especially CH G band) appear to be weaker than the observed ones. To make the molecular lines consistent with observations we changed the A05 carbon and nitrogen abundances and used them as given by G91. This allows us to reach a very good agreement with measurements (see Sect. 4.4.3).

Let us emphasize that an increase in the collisional rates for H\(^-\) cannot help adjust the NLTE calculations performed with G91 abundances, because the considered increase in collisions can only enforce an LTE population ratio between negative hydrogen and H I 1 level. The latter is, however, in an NLTE regime mainly thanks to the NLTE “overionization”. As the calculations with G91 abundances give good agreement in a purely LTE regime (see Fig. 4.4), the considered increase in the collisional rates for H\(^-\) is not enough to enforce full LTE agreement. An NLTE treatment of the solar atmosphere yields population numbers of negative hydrogen that strongly deviate from LTE.

Change in abundances

The effect of using of different abundances on the emergent synthetic spectrum is illustrated in Fig. 4.9. One can see that the abundance change of elements like iron, magnesium, and silicium strongly affects the UV radiation (mainly due to the photoionization opacities), but also the visible and infrared continuum, because these elements are the main donors of the photospheric electrons. In contrast,
a change in nitrogen and oxygen abundances has no effect on the continuum and only affects the spectrum via the molecular bands. The carbon abundance slightly influences the continuum, but the main effect of carbon in the visible spectral range comes from molecular systems. The CN violet system at about 380 nm and CH G band at about 430 nm are the most prominent features. Thus, the carbon and nitrogen abundance changes from the A05 to G91 values almost do not affect the continuum level (see Fig. 4.9) and lead to good agreement with the measurements (see Fig. 4.10). Therefore, we can conclude that A05 abundances should be given preference for calculating the NLTE continuum, except for carbon and nitrogen, for which the G91 abundances lead to better reproduction of the molecular bands. The favored combination of abundances leads to good overall agreement with the observed solar spectrum.

4.4.3 Comparison with measurements and discussion

In Fig. 4.10 we present the solar spectrum calculated with the FAL-C-99 atmosphere model compared with the SORCE measurements. The agreement in the 160-320 nm spectral region is automatically good (see Sect. 4.4.1). At shorter wavelengths, the observed flux is determined by several strong emission lines, which are treated in LTE in the current version of COSI. This leads to a strong underestimation of the irradiance. In a future version of COSI, these lines will be treated in NLTE. The Ly α line has already been computed in NLTE, and we obtain an overestimation of its flux. In a forthcoming paper we will investigate it in more detail (Schöll et al., 2010).

The detailed comparison of our synthetic spectrum with SOLSPEC measurements is given in Figs. 4.13 and 4.14 in the Online Material. One can see that starting from 160 nm two spectra are in remarkable agreement with each other. Although several strong emission lines are currently not properly calculated in shorter wavelengths, the continuum level there is also consistent with SOLSPEC measurements. This can only be achieved with NLTE calculations.

Thus the assumption of enhanced negative hydrogen collisions and the use of the combined G91 and A05 abundance set provide us with the possibility of successfully modeling the entire solar spectrum.
4.5. Irradiance variations

Another possible cause of the problem that synthetic NLTE models have difficulty fitting the observed solar spectrum is that the FAL-C-99 atmosphere model was developed to fit the visible and IR continuum with the radiative transfer code PANDORA (see Fontenla et al., 1999), which uses a different approximate approach to treat the NLTE effects, different sets of the atomic and molecular data as well as different sets of abundances. The Fontenla et al. (1999) atmosphere structures have been used successfully in various applications (cf. Penza et al., 2004a,b; Vitas and Vince, 2005) in which, however, the continuum flux was calculated in LTE. The quest to construct a self-consistent atmosphere model that fully accounts for all NLTE effects will be addressed in a future investigation.

4.5 Irradiance variations

It is generally accepted that most of the solar variability is introduced by the competition between the solar flux decrease due to the dark sunspots and a flux increase due to the bright magnetic network and plage (e.g. Willson and Hudson, 1991; Krivova et al., 2003). It is therefore important to consistently calculate the solar irradiance from these active components of the solar atmosphere.

For our calculations we used the Fontenla et al. (1999) atmosphere model FAL-C-99 for the quiet Sun, model FAL-F-99 for the bright network, model FAL-P-99 for plage, and model FAL-S-99 for the sunspot. Both NLTE and LTE spectra of the quiet Sun, active network, plage, and sunspot are presented in Fig. 4.11. The bright network spectrum is hardly distinguishable from the quiet Sun spectra. The NLTE spectra produce a lower visible and UV flux and, accordingly, lower total solar irradiance (TSI). This effect for the quiet Sun irradiance was already discussed in Sect. 4.4.2.

Another interesting detail is the decrease in the contrast between the quiet Sun, plage, and bright network in the NLTE calculations. The main reason for this is that the NLTE effects that decrease the continuum source function are stronger in the hotter FAL-P model as it corresponds to the lower particle density (and consequently lower collisional rates).
Figure 4.11: NLTE (left panel) and LTE (right panel) calculations of the solar irradiance from the quiet Sun, active network, plage, and sunspot.

Figure 4.12: The flux differences between plage and quiet Sun (upper panel), and bright network and quiet Sun (lower panel) calculated in NLTE and LTE.
Table 4.1: *TSI differences with FAL-C model per 1% area (in W/m²).*

<table>
<thead>
<tr>
<th></th>
<th>FAL-F</th>
<th>FAL-P</th>
<th>FAL-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTE G91</td>
<td>0.28</td>
<td>1.48</td>
<td>-10.7</td>
</tr>
<tr>
<td>LTE A05</td>
<td>0.25</td>
<td>1.27</td>
<td>-10.82</td>
</tr>
<tr>
<td>NLTE G91</td>
<td>0.18</td>
<td>0.76</td>
<td>-11.02</td>
</tr>
<tr>
<td>NLTE comb.</td>
<td>0.15</td>
<td>0.99</td>
<td>-11.05</td>
</tr>
</tbody>
</table>

Figure 4.12 shows the flux differences between the active components of the solar atmosphere and the quiet Sun calculated in LTE with A05 abundances, and NLTE with enhanced collisions and combined A05 and G91 abundances (see Sect. 4.4). The most prominent peak at about 3890 Å corresponds to the CN violet system. Although the differences in the CH G band around 4300 Å are also clearly visible, they are less pronounced than in the CN violet system. This is caused mainly by the differences in the CH and CN dissociation energies (3.465 eV and 7.76 eV), so the CN concentration is more sensitive to temperature variations (see Sect. 4.3.1). Furthermore, the contrast between the active components and the quiet Sun is significantly decreased in the NLTE calculations. As a consequence, the NLTE calculations reduce the solar variability in the visible and IR, shifting it to the UV.

The differences between the spectral fluxes for active components and for the quiet Sun integrated from 900 Å to 40 000 Å are presented in Table 4.1 for the LTE (A05 and G91 abundances), for the NLTE with G91 abundances and NLTE with enhanced collisions and combined A05 and G91 abundances (NLTE comb.). We conclude that the NLTE calculations (both with and without ODF) significantly reduce the variations in the TSI.

4.6 Conclusions

We have presented a further development of the radiative transfer code COSI. The code accounts for the NLTE effects in several hundred lines, while the NLTE effects in the several million other lines are indirectly included via iterated opacity distribution function. The radiative transfer is solved in spherical symmetry. The main conclusions can be summarized as follows.

- The inclusion of the molecular lines in COSI allowed us to reach good agreement with the SORCE observations in the main molecular bands (especially in the CN violet system and CH G band). It has also solved the previous discrepancies between the LTE calculations with the COSI code and ATLAS 12 calculations. We showed that their strong temperature sensitivity allows molecular lines to significantly contribute to the solar irradiance variability.

- We introduced additional opacities into the opacity distribution function. It allowed us to solve the well-known problem of overestimating the UV flux in the synthetic spectrum. We should emphasize, however, that the magnitude and behavior of the additional opacity strongly depend on the applied model.

- We have shown that the concentration of negative hydrogen is strongly affected by NLTE effects as explained in Sect. 4.4.2. It decreases the level of the visible and infrared continuum and leads to a discrepancy with the measured level if the calculations are done with the current models of the quiet Sun atmosphere.

- We presented calculations of the total and spectral solar irradiance changes due to the presence of the active regions and showed that NLTE effects can strongly affect both of these values.

Acknowledgements

These investigations have benefited from stimulating discussions at meetings organized at the ISSI in Bern, Switzerland. The research leading to this paper received funding from the European Com-
munity’s Seventh Framework Program (FP7/2007-2013) under grant agreement N 218816 (SOTERIA project, www.soteria-space.eu). We thank J. Fontenla for supplying the SRPM database.
Figure 4.13: The spectra calculated with COSI (solid line) vs. SOLSPEC measurements (dotted line).
Figure 4.14: As Fig. 4.13. Continuation.
Part II

Solar Irradiance Reconstruction
Chapter 5

Introduction

The modelling of the variable solar irradiance can be mostly separated into two parts, modelling of the total solar irradiance and the spectral components of the irradiance. For the active area expansion described in section 9.1 both are used. For the current time, that is since the advent of space-based measurements, the two factors sunspot-darkening and faculae-brightening have the most significant role in the change of solar activity. As a sunspot traverses the solar surface, it causes a darkening of the sun. However, as the number of sunspots increase, so do the corresponding faculae. This overcompensates the sunspot darkening and causes an increase of about 0.1% in the solar irradiance during the maximum of the sunspot cycle.

Several models exist that use proxies for both the darkening and brightening of the sun, utilizing statistical analysis, e.g. Chapman et al. (1996) and Fröhlich and Lean (1998). Dudok de Wit et al. (2009) compares nine different proxies, from UV measurements through visible proxies like the International Sunspot Number (ISN) to radio-sources to select the best proxies to reconstruct the UV band. They show that for all wavelength but the FUV (130 nm–170 nm) band, no single proxy can be used for a waveband for all time scale with the MUV (220 nm–270 nm) not being fitted well by any combination of the nine proxies.

Wenzler (2005) employs magnetograms and white-light images from the Kitt Peak Vacuum Tower (KPVT) to extract magnetic active regions, namely the filling factors for sunspots and plages. Several reconstructions exist based on those filling factors, the most used is the SATIRE (Spectral and Total Irradiance REconstruction) model (Krivova et al., 2003) with one free parameter used to reconstruct both the Total Solar Irradiance (TSI) and Spectral Solar Irradiance (SSI). This model can be extended to go back in time by using proxies for the filling factors.

Examples of total solar irradiance reconstruction based on proxies are e.g. Schöll et al. (2007) utilizing sunspot count and neutron monitor data. Beer et al. (2004) uses Beryllium 10 and Steinhilber et al. (2008) reconstruction of the TSI is based on the Interplanetary Magnetic Field (IMF) reconstructed from Beryllium 10 and the geomagnetic field strength.

Solanki et al. (2000) introduces a magnetic decay lasting up to several years. Based on this work Krivova et al. (2010) develops a physical magnetic flux model using the group sunspot number (Hoyt and Schatten, 1993) and sunspot area (Balmaceda et al., 2009) to reconstruct the solar active regions. This model is again based on the SATIRE model.

In this work I assume that all irradiance variation is based on a change of surface features, following the SATIRE model (Solanki et al. (2005); Krivova et al. (2011)). Whether this is the case is still open for discussion, c.f. Steinhilber (2008, Chapter 7).
Chapter 6

Data

In this chapter I will give a short overview of the data used to reconstruct and verify the total and spectral solar irradiance. First the long-term proxies for solar modulation, Beryllium 10 is described, followed by sunspots, the filling factor by Wenzler, and the TSI and SSI data. Also, since I am especially interested in the Lyman-α, measurements by SUMER are described in chapter 6.6

6.1 Solar Modulation Potential

The solar modulation function $\Phi$ describes the modulation of cosmic ray particles while passing through the heliosphere where they produce cosmogenic radionuclides. The radionuclides may either be stored in the Arctic ice as is the case for e.g. Beryllium 10 or become part of the carbon cycle and stored in trees, as for $^{14}$C. They may also be directly measured by neutron monitors. The radionuclide signal is also influenced by factors such as the location of the solar system in the interstellar neighbourhood, the transport through the heliosphere, the geomagnetic field strength, the atmospheric transport and the transport in the ice archive for the case of Beryllium 10. For a description of the correction used see Steinhilber et al. (2008). These effects have to be taken into account when one wants to obtain the solar modulation. For the time of interest for this thesis the stellar neighbourhood is assumed to be constant, while all other effects do play a role and are accounted for in the data used. McCracken et al. (2004) provides two $\Phi$’s, based on two different Beryllium 10 archives, the Dye 3, Greenland and the South Pole archive. The data is provided with a 22-year running average. This is done to decrease the noise to 1.5%, caused by a low sampling rate from the South Pole data. The data shows a correlation of 0.87 or 0.91 when some outliers where excluded. The difference is claimed to be due to the “statistical, systematic, and meteorological variations in both the Greenland and South Pole data and also any timing errors between the two data sets”. $\Phi$ extracted from those archives are available from 1600 CE to 1957 CE for the Dye 3 archive and from 1500 CE to 1971 CE for the South Pole archive.

6.1.1 Neutron Monitor

The neutron monitor also measure cosmogenic radionuclides, from which the solar modulation can be inferred. The advantage of the neutron monitor is that they are not affected by the atmospheric transport and their high-frequency measurements. However, they are only available starting 1952. For this work this data is used to extend $\Phi$ based on Beryllium 10 into the present. The solar modulation potential derived from neutron monitor data is provided by Usoskin et al. (2005).
6.2 Sunspots

Sunspots have been recorded at least since 165 BCE in China (Strom, 2008). However, only with the invention of the telescope in 1609 by Galileo, sunspots were continuously observed, starting in 1610. Several sunspot records exist. For this work, two are used. The Group Sunspot Number (GSS) and the International Sunspot Number (ISN).

The GSS is a re-evaluation of the ISN with daily, monthly and yearly means available from 1610 until 1995. This record has included observations not present in the ISN. Hoyt and Schatten (1997a) define the GSS as a weighted average of sunspot observations, normalised to match the ISN from 1874 to 1976 (see Equation 6.1, where $k_i$ is the correction factor and $g_i$ the observed number of sunspot group by observer $i$. The 12.08 is the normalisation factor).

$$R_g = \frac{12.08}{N} \sum_{i=1}^{N} k_i g_i$$ (6.1)

This record contains data until 1995. To obtain a complete record over the whole time, the ISN is used to complete the sunspot count to the present. The ISN contains sunspot maxima and minima from 1610 to present, yearly averages from 1700 to present, monthly averages from 1749 to present and daily averages from 1818 to present. The difference between GSS and ISN is minimal for the period of 1818 to present. Hence to obtain an homogeneous record for the present time, the ISN is used for 1818 to present, instead 1995 to present, despite the fact that GSS is less noisy. No corrections are made to either record, for the GSS is already normalized to fit the ISN.

All sunspot data is obtained from NOAA national geophysical data centre, http://www.ngdc.noaa.gov/stp/solar/ssndata.html.

6.3 Filling Factors

The filling factors describe the fraction of the solar surface occupying a specific solar feature. E.g. we use a model with five filling factors, sunspots, quiet and active network, faculae and plage regions. They are described in detail in the following chapter. The filling factors used to calibrate the model are obtained from Wenzler (2005). He extracted them from SPM images for part of solar cycle 23, from February 1996 to December 2001. They are also compared to SOHO/MDI images. The averaged noise level of the MDI magnetograms is estimated to be $\sigma = 9G$ for five minute integrated magnetogram, varying with $1/\sqrt{T}$, $T$ the integration time in minutes.

6.4 Total Solar Irradiance

Several space based measurements of the Total Solar Irradiance (TSI) exist, starting with the Hickey-Frieden (HF) Hickey et al. (1980) radiometer aboard the NOAA/NASA satellite Nimbus 7. The next mission was the ACRIM-I on SMM, followed by ACRIM II and SOLSTICE on Upper Atmosphere Research Satellite (UARS), VIRGO (Variability of Solar Irradiance and Gravity Oscillations) on SoHO and the Total Irradiance Monitor at SORCE Kopp and Lawrence (2005).

To generate a TSI composite is not trivial for all instruments suffer degradation and none of the instruments data is available for the whole period. Hence several composites are available. From the ACRIM team (Willson and Mordvinov, 2003b), the Institut Royal Meteorologique Belge (IRMDB) (Dewitte et al., 2004) and also from the PMOD/WRC (Fröhlich, 2006).

While all composites use the same data, they treat the degradation differently. The most notable difference is how they treat the ACRIM gap, the time between the switch-off of ACRIM I and launch of ACRIM II. While other instrument data (HF, ERBE) exist, they are of lower accuracy. The HF instrument displays a 0.4 W/m² increase after being switched off and back on in September 1989. The
long term trend exist, depending on whether to take it into account (Fröhlich, 2006) or not (Willson and Mordvinov, 2003b).

The TSI used for this work is the PMOD/WRC composite unless otherwise stated. It is a composite of ACRIM-I, ACRIM-II, HF and VIRGO data, corrected for degradation of HF and ACRIM-I. These measurements are calibrated against data from ERBE and ACRIM-III as well as empirical models. A description of the composite is given in Fröhlich (2006, 2009a).

6.5 Solar Spectrum

For comparison both SUSIM, version 22 and SOLSTICE (Solar Stellar Irradiance Comparison Experiment), level 3, version 10 instruments aboard the UARS and SORCE (Solar Radiation and Climate Experiment) platform are used, available from LISIRD (LASP Interactive Solar Irradiance Datacenter). The spectral data differ significantly from each other. While the average spectra only differ within 10%, mainly in the FUV continuum, as shown in figure 6.1(a), the correlation matrix of those two data sets differ significantly. SOLSTICE displays a very strong correlation between any two lines in the FUV with a reversal in the MUV. SUSIM (Solar Ultraviolet Spectral Irradiance Monitor) displays a spectrum with a lower “cross dependency”. Hence when comparing a modelled spectrum with measurements, it may be interesting to also look at this cross dependency.

Also the correlation between the three instruments is sometimes as low as zero or negative for some wavelength using a weekly averaging. Figure 6.1(b) displays the correlation coefficient for all wavelength available from all three instruments using a time average of a week. The weekly averaging is needed to compensate for the time-shift of twelve hours between SUSIM and the SOLSTICE instruments. The red curve is the correlation of SUSIM flux to the SUSIM flux of the next day, showing that the weekly averaging is sufficient to guarantee an sufficient inter-instrument correlation of the flux.

The inter instrument correlation of the fluxes for any wavelength has an maximum of 0.9 for the fluxes measured by SUSIM to SORCE-SOLSTICE fluxes in the wavelength region at 150nm to 200nm. However those two instruments show opposite trends at the region around 300nm. The SOLSTICE instrument aboard UARS has an maximum correlation of 0.77 at the Lyman-α line, relative to the SOLSTICE instrument aboard SORCE. For all but a few lines the correlation is below 0.5, again displaying different instrumental behaviour.

6.6 Lyman-α

Lemaire et al. (1998) provide a high-resolution spectrum of the Ly-α line at disc-centre. Figure 6.3 compares this measurements to a mixture of 6.25% of quiet network, 2.5% plage and 6 ppm of sunspots, consistent with the measured plage and sunspot filling factors by Wenzler (2005). The calculated Ly-α is described in detail in Chapter 8.2 and Section 9.2. In summary, this is the profile of the reconstructed Ly-α at the date of measurement using the original filling factors and not applying Active Area Expansion (AAE) as described in Section 9.1.

The integrated flux from -1Å to 1Å around the line-centre is 79.6W/m²/sr for the measurement and 67.0W/m²/sr for the mixture. The calculated Lyman-α line shows a stronger wing with a lower maximum flux. This is possibly due to the approximate line-broadening method applied. However, the total energy output is conserved.

2http://lasp.colorado.edu/lisird/index.html (as of 11-2010)
(a) Average SOLSTICE (green) and SUSIM (blue) spectra. They agree well within the error of measurement.

(b) Intercorrelation of SUSIM and SOLSTICE fluxes using a weekly averaging. Shown is the correlation of the measured flux by SUSIM to the SOLSTICE measurement onboard UARS (green, dashed), the measured flux by SUSIM to the SOLSTICE measurement onboard SORCE (blue, solid) and the correlation of SUSIM for each wavelength to its previous day (red, dash-dotted, axis on the right).

Figure 6.1: Average SOLSTICE and SUSIM spectra and their inter-correlation coefficients.

Figure 6.2: SORCE-SOLSTICE and SUSIM intra-correlation matrix for April 2003 to July 2005. Each point \((x, y)\) shows the correlation at wavelength \(x\) to the wavelength \(y\). Negative values are given in grey shades, while the positive correlation is a colour shade from red to blue, with red translating to a correlation of one and blue/black describing a correlation of zero. The four-block structure of SOLSTICE, with the edge at 180nm is due to the two SOLSTICE instruments A and B.
Figure 6.3: High-resolution spectrum of the Lyman-α line from Lemaire et al. (1998) (dashed) and the modelled Lyman-α line of the atmosphere model C (solid). The integrated flux from -1Å to 1Å around the line-centre is 79.6 W/m²/sr and 67.0 W/m²/sr respectively.
Total Solar Irradiance

Total Solar Irradiance (TSI) is defined as integrated irradiance over wavelength,

$$\text{TSI}(t) = \int_0^\infty \text{SSI}(t, \lambda) d\lambda.$$ 

There are different methods to reconstruct TSI, one possibility is to look at solar proxies and extend them to the past. Several approaches exist, Dudok de Wit et al. (2008, 2009) analyses several proxies to use for reconstruction of the current TSI and SSI. The current solar activity up to 2009 does not show a significant long-term trend and hence methods based only on correlation have a high uncertainty in the possible existence of a long-term trend of the past. One can assume a correlation between long-term and short-term trends, however this assumption is rejected by e.g. Fröhlich (2011).

Hence physical-proxy based hybrid models may produce a more faithful reconstruction of both TSI and SSI. In (Schöll et al., 2007) we describe such an approach to reconstruct the long-term solar activity utilizing neutron monitor data. One implicit assumption in this work is, contrary to Fröhlich (2011), that the long-term trend of TSI is coupled with the Neutron Monitor (NM) part and that it is equivalent to the short-term trend. This assumption is replaced in Chapter 8 by assuming that the long-term trend is coupled to the long-term evolution of $^{10}\text{Be}$ isotopes and that the minimum sun corresponds to the minimum observed quiet sun.
Chapter 8

Model Based Reconstruction

8.1 Introduction

Krivova et al. (2003) model for both TSI and SSI is SATIRE, utilizing reconstructed filling factors. It assumes that all solar variation can be explained by surface changes using filling factors as described in chapter 8.2, and by its corresponding TSI or SSI irradiance model. The four component SATIRE model used by Krivova et al. (2003), utilizing sunspot umbra, penumbra, plage and quiet sun, results in a $r_c = 0.96$ for TSI when compared to VIRGO. The SATIRE model cited does not take into account the possibility of a long-term change, for the current quiet sun, model C, defines the lower limit. However, $^{10}$Be measurements (see Figure 9.9(d)) and their reconstructed solar modulation factor (Steinhilber, 2008) do show a long-term, with the solar modulation factor of 1690 at 30% of today’s value.

The assumption in this work is that today’s measured quiet sun as modelled by FAL-99-C (table 8.1, model C) is composed of an extreme quiet sun, that is a region that emits less radiance than the average quiet sun and a more active region, the quiet network. The extreme quiet sun is described by model FAL-99-A, which is assumed to correspond to a non-magnetic sun, while model FAL-99-E describes the quiet network. Also a linear relationship of the magnetic activity and the covering fraction of model FAL-99-E is assumed.

8.2 Filling Factors

The solar surface can be divided into regions of different level of activity. Wenzler (2005) extracted three different regions on the sun, sunspots, faculae and the quiet sun for the time of available SPM data and compared it to SOHO/MDI data. They are extracted based on intensity on continuum images and magnetograms. A dark region on a continuum image is classified as a sunspot, while faculae are defined by area where magnetic activity is above three sigma of average activity. Fontenla et al. (2009b) divides the solar surface in seven different regions according to their magnetic field strength. A description of the seven models is in table 8.1, with the contrast functions shown in Figures 8.1 and 8.2.

8.2.1 The Network

The filling factors extracted by (Wenzler, 2005, p. 96ff.) do not display a significant long-term trend because of the two active factors, plage and sunspots decreasing to zero during the current-time solar minima. Hence using a SATIRE based, time-independent, direct reconstruction of the TSI or SSI based only on those filling factors do not show a long-term trend. Another surface phenomena on the sun is the network which has two components, classified as “quiet” and “active” network (Woods et al., 2000; Fontenla et al., 2009b).
Figure 8.1: Contrast of the quiet sun model A, plage and sunspots to the quiet sun model C.

Figure 8.2 displays the contrast of the quiet and active network with respect to the quiet sun model C.

Figure 8.2: Contrast of the quiet and active network to the quiet sun model C. The dips in the UV are at e.g. 97.2 nm, 102.6 nm, 121.6 nm and correspond to the core of the Lyman series.

Fontenla et al. (1999) divides the solar surface into seven different regions (Table 8.1). For the reconstruction of the solar spectrum one of two models are used, further described in this Section and Section 11. The first uses the atmospheres A, E, F, P and S. Model A describes the very quiet solar surface, while model E describes the quiet network. The average quiet sun model C is replaced by a mixture of the models A and E. Atmospheric models P and S are used the same way as done by Wenzler (2005). The second method additionally uses model C and utilizes a change in the distribution, as described in Section 11.1.
8.2. Filling Factors

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Description</th>
<th>( \int \text{Irradiance} ) (W/m(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Faint super granule cell interior</td>
<td>1357.23</td>
</tr>
<tr>
<td>C</td>
<td>Average super granule cell interior</td>
<td>1365.5</td>
</tr>
<tr>
<td>E</td>
<td>Average network</td>
<td>1371.13</td>
</tr>
<tr>
<td>F</td>
<td>Bright network/faint plage</td>
<td>1379.91</td>
</tr>
<tr>
<td>H</td>
<td>Average plage</td>
<td>\text{-N.A.-}</td>
</tr>
<tr>
<td>P</td>
<td>Bright plage</td>
<td>1464.07</td>
</tr>
<tr>
<td>S</td>
<td>Sunspot umbra</td>
<td>260.40</td>
</tr>
</tbody>
</table>

Table 8.1: The seven solar atmosphere structures as defined by Fontenla et al. (1999) and their integrated irradiance, as calculated by COde for Solar Irradiance (COSI) for the wavelength interval from 90 nm to 2 \( \mu \text{m} \), normalized to the PMOD composite. There are no COSI model calculations of model H.

8.2.2 Reconstructing Filling Factors

The filling factors for sunspot area, plage, quiet network and active network is based on the international monthly sunspot number for 1749–2010, the group sunspot data for 1610–1748 and Beryllium 10 data. The plage-area is calculated by linear regression w.r.t. sunspots, e.g. \( \alpha_p(t) = c_p SS(t) : \min_{c_p} ||c_p SS(t) - \alpha_p \text{Wenzler}(t)||_2 \) with \( c_p \) the linear regression constant. The same method is used for the active network. The quiet network describes the long-term variation and is coupled to Beryllium 10. Equation 8.1–8.3 describes the filling factors of each region.

\[
[\alpha_S, \alpha_P, \alpha_F] = [c_S, c_P, c_F] SS \quad (8.1)
\]

\[
\alpha_E = c_E \text{Beryllium} \quad (8.2)
\]

\[
\alpha_A = 1 - (\alpha_E + \alpha_F + \alpha_P + \alpha_S) \quad (8.3)
\]

with \( c_X \) constant and SS the sunspot composite as described in Section 6.2. Equation 8.1 describes the short-term variability, while Equation 8.2 models the long-term. Equation 8.3 is the remainder of the solar disk, the quiet sun.

In Schöll et al. (2007) the total solar irradiance is reconstructed using the sunspots for the short-term variation and neutron monitor data for the long-term, i.e. for the eleven year variability. The neutron monitor data, available back to 1930 was used because of lack of an available IMF time series which would allow us to extend the record back in time. For this work the \( \Phi \) provided by McCracken et al. (2004) is used.

The network is calculated as described in Section 8.2.1. This results in the time dependent filling factor \( \alpha'(t)_{P,S,F} \) shown in Fig. 8.3.

Sunspots and faculae are derived from the composite sunspot number.

8.2.3 Two Component Quiet Sun

Splitting the quiet sun into two parts, atmospheres A and E, introduces filling factors \( \alpha_A \) and \( \alpha_E \), s.t.

\[
\alpha_{\text{quiet}} = c_m \alpha_E + (1 - c_m) \alpha_A \quad (8.4)
\]

with \( c_m = \frac{I_{\text{Observed}} - I_A}{I_E - I_A} \quad (8.5) \)

The constant \( c_m \) results from the assumption that the current quiet sun is a composite of the two models A and E. Using the values given in table 8.1 and under the assumption that model C describes the quiet sun for the current time (1996), the mixing constant \( c_m \) has a value of \( c_m = 0.56 \). For a comparison of the Model C spectrum to observation see Shapiro et al. (2010). However, it would have
been also possible to use observation directly. Also for that the TSI values $I_X$ are normalised w.r.t. the PMOD composite, the $c_m$ derived above is based on measurement. The difference becomes important when comparing the spectra, as done below.

The main assumption of the long-term reconstruction is that the time dependency of the active network is coupled to $\Phi$ based on Beryllium 10 measurements with a $\Phi$ of zero corresponding to a non-magnetic sun, and the current-time averaged $\Phi$ over 22 years corresponding to the current-time quiet sun magnetic activity. Furthermore, it is assumed that the model $A$ describes the hypothetical minimal sun, corresponding to a $\Phi$ of zero.

Hence

$$\alpha_E = \frac{\langle \Phi >_{22yr} c_m}{\langle \Phi_{t_0} >_{22yr}} \alpha_{quiet}$$

(8.6)

$$\alpha_A = \left(1 - \frac{\langle \Phi >_{22yr} c_m}{\langle \Phi_{t_0} >_{22yr}}\right) \alpha_{quiet}$$

(8.7)

where the solar minimum of 1996 is used as $t_0$. Since the solar minima of the last three cycles where about the same, this is a good representation of the current day quiet sun.

The mixing ratio for the current-day quit sun $c$ is derived from the integrated values $I_{A,C,E}$. This is also a good approximation for the spectral component from 300 nm upwards, as shown in Figure 8.4. The differences in the UV are possibly due to missing lines in the UV. It should also be noted that the $L_Y - \alpha$ is in good agreement with the stated assumptions. The upper Lyman lines $\beta, \gamma, \ldots$ diverge from C by up to 20%. That is still below the UV divergence which diverges from model C up to 100%. This approach of reconstructing the two different kind of networks is the one used in Shapiro et al. (2010), reprinted in full in Section 8.3.

Applying the techniques described in Subsection 8.2.3 lead to a reconstruction of both TSI and SSI, described in the following paper, Shapiro et al. (2011a).
Figure 8.4: The ratio of model A, E and the best fit mixture of E and A with respect to model C. The differences to Figure 8.9 is partly due to numerics and partly due to the use of AAE, described in Section 9.1.

8.3 A new approach to long-term reconstruction of the solar irradiance leads to large historical solar forcing

Published in A&A, Volume 529, May-2011

A. I. Shapiro, W. Schmutz, E. Rozanov, M. Schoell, M. Haberreiter, A. V. Shapiro, S. Nyeki

8.3.1 Abstract

The variable Sun is the most likely candidate for natural forcing of past climate change on time scales of 50 to 1000 years. Evidence for this understanding is that the terrestrial climate correlates positively with solar activity. During the past 10'000 years, the Sun has experienced substantial variations in activity and there have been numerous attempts to reconstruct solar irradiance. While there is general agreement on how solar forcing varied during the last several hundred years — all reconstructions are proportional to the solar activity — there is scientific controversy on the magnitude of solar forcing. We present a reconstruction of the Total and Spectral Solar Irradiance covering 130 nm–10 µm from 1610 to the present with annual resolution and for the Holocene with 22-year resolution. We assume that the minimum state of the quiet Sun in time corresponds to the observed quietest area on the present Sun. Then we use available long-term proxies of the solar activity, which are $^{10}$Be isotope concentrations in ice cores and 22-year smoothed neutron monitor data, to interpolate between the present quiet Sun and the minimum state of the quiet Sun. This determines the long-term trend in the solar variability which is then superposed with the 11-year activity cycle calculated from the sunspot number. The time-dependent solar spectral irradiance from about 7000 BC to the present is then derived using a state-of-the-art radiation code. We derive a total and spectral solar irradiance that was substantially lower during the Maunder minimum than observed today. The difference is remarkably larger than other estimations published in the recent literature. The magnitude of the
8.3.2 Introduction

The Sun is a variable star whose activity varies over time-scales ranging from minutes to millennia. Over the last thirty years the solar irradiance was measured by numerous space missions. Measurements of the Total Solar Irradiance (TSI) became available with the launch of the NIMBUS 7 mission in 1978 (Hoyt et al., 1992). Since then TSI was measured by several consecutive instruments. Each of them suffered from degradation and individual systematic effects, so the direct comparison of the measurements is impossible. Three TSI composite based on the available data were constructed by three groups: PMOD (Fröhlich, 2006), ACRIM (Willson and Mordvinov, 2003a), and IRMB (Dewitte et al., 2004). These composites give quite different values of TSI, especially before 1980 and during the so-called ACRIM gap between June 1989 (end of ACRIM I observations) and October 1991 (beginning of ACRIM II observations). The most striking detail is the increase of TSI between the minima 1986 and 1996 in the ACRIM composite and the absence of such increase in the PMOD and IRMB composites. Although significant progress was made during the last few years and the increase of the TSI in the ACRIM composite was strongly criticized (Fröhlich, 2009b; Krivova et al., 2009a), the question of construction of the unique self-consistent TSI composite for the satellite epoch is still open.

Measurements of the Spectral Solar Irradiance (SSI) are even more difficult and instrumental problems prevent the construction of essential composites (Krivova et al., 2011; Domingo et al., 2009). Recently (Krivova et al., 2009b) the theoretical SATIRE (Spectral And Total Irradiance REconstruction) (Krivova et al., 2003; Krivova and Solanki, 2008) model and SUSIM (Solar Ultraviolet Spectral Irradiance Monitor) measurements were used to reconstruct the solar UV irradiance back to 1974.

Taking into account the problems of the reconstruction of TSI and SSI during the recent period of satellite observations and combining it with the fact that the solar dynamo, which is believed to drive all activity manifestations, is still not fully understood (Charbonneau, 2010), one recognizes the difficulty of reconstructing TSI and SSI to the past, when no direct measurements were available.

Long-term changes in solar irradiance were suspected as early as the mid-nineteenth century (C. P. Smyth, 1855). One of the first quantitative estimates of its magnitude as well as past solar irradiance reconstructions was obtained by using the observations of solar-like stars (Lean et al., 1995). It was concluded that TSI during the Maunder minimum was about 3–4 W/m$^2$ less than at present which translates into a solar radiative forcing\footnote{Solar radiative forcing is a direct energy source to the Earth and is related to the change in TSI by $\Delta F = \Delta \text{TSI} \cdot (1 - A)/4$, where $A$ is the Earth’s albedo.} $\Delta F_{P-M} \sim 0.5$–0.7 W/m$^2$. However, these results were not confirmed by large surveys of solar-like stars and are no longer considered to be correct Hall and Lockwood (2004). Recent reconstructions based on the magnetic field surface distribution (Wang et al., 2005; Krivova et al., 2007) and on extrapolation of the assumed correlation between TSI and the open magnetic flux (Lockwood et al., 1999) during the last three minima (Steinhilber et al., 2009; Fröhlich, 2009b) resulted in a low solar forcing value within the range $\Delta F_{P-M} \approx 0.1$–0.2 W/m$^2$.

8.3.3 Effects of solar radiative forcing on climate

Variations on time-scales up to the 27-day rotational period have an important influence on space weather but not on terrestrial climate. The effects of the 11-year solar cycle are clearly detected in the atmosphere and are widely discussed in the literature (e.g., Egorova et al., 2004; Haigh, 2007), while the imprints from variations on longer time-scales are more subtle (Gray et al., 2010). Analysis of historical data suggests a strong correlation between solar activity and natural climate variations on centennial time-scales, such as the colder climate during the Maunder (about 1650–1700 AD) and Dalton (about 1800–1820 AD) minima as well as climate warming during the steady increase in solar activity in the first half of the twentieth-century (Siscoe, 1978; Hoyt and Schatten, 1997b; Solomon et al., 2007; Gray et al., 2010). Numerous attempts to confirm these correlations based on different
climate models have shown that it is only possible if either the applied perturbations of direct solar radiative forcing are large (consistent with a direct solar radiative forcing from the present to Maunder minimum $\Delta F_{P-M} \sim 0.6-0.8 \text{ W/m}^2$) or the amplification of a weak direct solar forcing is substantial. Because the majority of recent $\Delta F_{P-M}$ estimates (see Sect. 8.3.2) are only in the range $0.1-0.2 \text{ W/m}^2$, and amplification processes have not been identified, the role of solar forcing in natural climate change remains highly uncertain (Hoyt and Schatten, 1997b). In this paper we show that the solar forcing may be significantly larger than reported in the recent publications.

8.3.4 Methods

In this paper we present a new alternative technique, avoiding calibration of our model with presently observed TSI variations and extrapolation to the past. We assert that the amount of magnetic energy that remains present (de Wijn et al., 2009) at the surface of a spotless (i.e. quiet) Sun is the main driver of solar irradiance variability on centennial time scales. The main concept of our technique is to determine the level of the magnetically enhanced contribution to the irradiance of the present quiet Sun. Then, for the reconstruction to the past, this magnetically enhanced component has to be scaled with the proxies for the quiet Sun activity. However a proxy for the long-term activity of the quiet Sun does not yet exist. A good candidate for such a proxy is the small-scale turbulent magnetic fields accessed with the Hanle effect (Stenflo, 1982). However the consecutive measurements of these fields are limited to the last few years (Kleint et al., 2010). Therefore, we assume that the existing proxies of the solar activity averaged over the two solar cycles period can also describe the activity of the quiet Sun. Averaging of proxies allows sufficient time for the magnetic components to decay to the quiet network, and then to even smaller magnetic features as the decay process can take up to several years (Solanki et al., 2000). In other words we set the small-scale activity (which defines the fractional contributions of quiet Sun components) to be proportional to the large-scale activity. We want to state clearly that this proposition is an assumption, which is however in line with a high-resolution observations of the present Sun. The large-scale structure due to strong magnetically active features is repeated on a smaller scale in less active regions, and even in the apparently quietest areas there is still a mosaic of regions of different magnetic field strengths, reminiscent of fractal structure (de Wijn et al., 2009).

Hence, the time-dependent irradiance $I_{\text{quiet}}(\lambda, t)$ of the quiet Sun in our reconstruction can be calculated as

$$
\frac{I_{\text{quiet}}(\lambda, t) - I_{\text{min.state}}(\lambda)}{<\text{Proxy}>_{22}(t)} = \frac{I_{\text{quiet}}(\lambda, t_0) - I_{\text{min.state}}(\lambda)}{<\text{Proxy}>_{22}(t_0)},
$$

where $I_{\text{quiet}}(\lambda, t)$ is the time dependent irradiance of the quiet Sun and $t_0$ denotes a reference time. $<\text{Proxy}>_{22}$ is the averaged over a 22-year period value of the proxy for the solar activity. The 22-year period was chosen because the cosmogenic isotope data used for the reconstruction (see below) are available as 22-year averaged data (Steinhilber et al., 2008). We have set 1996 as the reference year, i.e. $I_{\text{quiet}}(\lambda, t_0)$ is the irradiance of the quiet Sun as observed during the 1996 minimum. Let us notice that the quiet Sun irradiance was roughly constant for the last 3 cycles (see below). Therefore the solar spectrum during the 1996 minimum is a good representation of the present quiet Sun spectrum. $I_{\text{min.state}}(\lambda)$ is the irradiance of an absolute minimum state of the Sun, with a minimum of remaining magnetic flux emerging on the solar surface. Thus, the term in brackets can be considered as the enhancement level of the present quiet Sun with respect to the Sun in its most inactive state.

The prominent, readily observable active regions on the Sun also contribute to the variability in irradiance. Thus, the full solar variability is described by:

$$
I(\lambda, t) = I_{\text{quiet}}(\lambda, t) + I_{\text{active}}(\lambda, t),
$$

where $I_{\text{active}}(\lambda, t)$ is the contribution to solar irradiance from active regions e.g. sunspots, plages, and network. $I_{\text{active}}(\lambda, t)$ is calculated following the approach by Krivova et al. (2003) (see also the online Sect. 8.3.8) and, as calculations are done with annular resolution, is proportional to the sunspot number. The group sunspot number used in our reconstruction is taken from NOAA data center and described in Hoyt and Schatten (1998).
The term $I_{\text{active}}(\lambda, t)$ in the Eq. (8.9) describes the cyclic component of solar variability due to the 11-year activity cycle, while the slower long-term changes due to the evolution of the small-scale magnetic flux are given by the Eq. (8.8). $I_{\text{active}}(\lambda, t)$ term can only be calculated for the time when sunspot number is available and therefore our reconstruction has 22-year resolution for the Holocene and annual resolution from 1610 to present.

The choice of the model for the minimum state of the Sun is a crucial point in our technique as it defines the amplitude of the reconstructed solar irradiance variability. Observations of the Sun with relatively high spatial resolution show that the present quiet Sun is still highly inhomogeneous even though it appears spotless. Measurements obtained with the Harvard spectroheliometer aboard Skylab were used to derive brightness components of the quiet Sun and to construct corresponding semi-empirical solar atmosphere structures (Vernazza et al., 1981; Fontenla et al., 1999). The darkest regions with the least amount of magnetic flux corresponds to the faint supergranule cell interior (component A). This component comes closest to describing the most inactive state of the Sun. Therefore, in our approach we set $I_{\text{min.state}}(\lambda) \equiv I_A(\lambda)$. Thus, the basic magnitude of solar irradiance variations is given by the difference between the irradiance of the present quiet Sun (composed from a distribution of brightness components defined in supporting online material) and the irradiance from component A (see Eq. (8.8)).

For the reconstruction to the past this amplitude is scaled with proxies for solar activity. Two proxies are available for the reconstruction: Group sunspot number, which is available from the present to 1610 AD, and the solar modulation potential extending back to circa 7300 BC. The latter is a measure of the heliospheric shielding from cosmic rays derived from the analysis of cosmogenic isotope abundances in tree rings or ice cores, and is available with a time resolution of 2-3 solar cycles (Steinhilber et al., 2008). Although sunspot number dropped to zero for a long time during the Maunder minimum, the solar cycle was uninterrupted (Beer et al., 1998; Usoskin et al., 2001) and the modulation potential did not fall to zero. Hence, a reconstruction based solely on sunspot number may underestimate the solar activity during the Maunder minimum. Therefore in our reconstruction we used the solar modulation potential to calculate the long-term variations and sunspot number to superpose them with the 11-year cycle variations (see the Online Section 8.3.10).
8.3. A new approach to long-term reconstruction of TSI

**Figure 8.5:** Modulation potential and TSI reconstruction for the last 70 years. Lower panel: Yearly averaged neutron monitor data (green) and two modulation potential composites (red and cyan curves, see the discussion in the text). Upper panel: TSI reconstructions based on the two modulation potential composites (red and cyan curves). The black dashed line is the observed TSI from the PMOD composite. Error bar for 1980 corresponds to 1/4 the difference between two published TSI composites, and the error for 2008 is taken from Fröhlich (2009b). The reconstructed TSI curves are normalized to the 1996 minimum and the grey-shaded region indicates the intrinsic uncertainty due to differences in the modulation potential data.

**Figure 8.6:** Modulation potential (lower panel) and TSI reconstructions (upper panel) for the last 2500 years. Data prior to 1600 AD are based on the modulation potential derived from $^{10}\text{Be}$ records from the Greenland Ice core Project (red curves). Data since 1600 AD are based on the two composites shown in Fig. 8.5 (red and cyan curves). The grey-shaded area indicates the intrinsic uncertainty.
Figure 8.7: TSI reconstruction from 7000 BC to 500 AD. The reconstruction is based on the modulation potential derived from $^{10}$Be records from the Greenland Ice core Project. Red dotted lines indicate the estimated error bars of the reconstruction. Black dashed and dotted lines indicate TSI for the 1996 solar minimum and the lowest TSI during the Maunder minimum, respectively.
The modulation potential used in the calculations is based on the composite of data determined from the cosmogenic isotope records of $^{10}\text{Be}$ and neutron monitor. $^{10}\text{Be}$ data are available up to about 1970 (McCracken et al., 2004) and neutron monitor data, which are used to calculate the current solar modulation potential, are available since the 1950s. Three different datasets of the $^{10}\text{Be}$ data were used: measurements at DYE 3, Greenland, and the South Pole provided by McCracken et al. (2004) for the reconstruction back to the Maunder minimum, and measurements from the Greenland Ice core Project provided by Vonmoos et al. (2006) for the reconstruction back to 7300 BC. The neutron monitor data were provided by Usoskin et al. (2005). Although they are available with a monthly resolution, we calculated the 22-year mean of the neutron monitor in order to homogenize the data sets and because only averaged modulation potential can be used as a proxy for the quiet Sun activity. The transition from $^{10}\text{Be}$ to neutron monitor data is shown in the lower panel of Fig. 8.5. Prior to 1952 the composites are based on $^{10}\text{Be}$ data (cyan based on DYE 3, and red on South Pole records, both with 22-year resolution). Values between 1952 and 1998 are obtained by linearly interpolating the following consecutive data points: two blue crosses at 1952 (the last data points used from $^{10}\text{Be}$ records), green crosses at 1976 and 1998 (22-year averages of the neutron monitor data). Values after 1998 require knowledge of the next 22-year average (between 2010 and 2031) of the modulation potential. We assume this average to be 92% of the previous average (between 1988 and 2009), which allows us to reproduce the observed TSI minimum in 2008. The datasets mentioned above were derived using different assumptions of the Local Interstellar Spectra (LIS). To homogenize the data we converted the modulation potential from McCracken et al. (2004) and Usoskin et al. (2005) to LIS by Castagnoli and Lal (1980) which is used in Vonmoos et al. (2006). For conversion we applied a method suggested by Steinhilber et al. (2008). Herbst et al. (2010) analyzed the dependency of the modulation potential on the applied LIS models and show that negative values of the modulation potential in Vonmoos et al. (2006) data can be corrected with another LIS model. Let us notice that the solar forcing in our reconstruction is determined by the relative values of the modulation potential so change to the different LIS model will introduce only several percents correction, which is much less than estimated accuracy of the reconstruction. In our dataset the value of the reference modulation potential $<\text{Proxy}>_{22}(t_0)$ (see Eq. (8.8)) is equal to 631 MeV. The relative error is estimated to be less than 10% (Usoskin et al., 2005).

Synthetic solar spectra are calculated with a state-of-the-art radiative transfer code (Haberreiter et al., 2008; Shapiro et al., 2010). We discuss the calculations in more detail in online Sect. 8.3.8, where in addition we show that the concept of quiet Sun activity scaling can also be expressed in terms of varying fractional contributions from different components of the quiet Sun.

### 8.3.5 Results and Discussion

In the upper panels of Fig. 8.5 and Fig. 8.6 we present the TSI reconstructions, which are obtained after the integration of the Eqs. (8.8) and (8.9) over the wavelengths and normalization of the quiet Sun value for the reference year 1996 to 1365.5 W/m$^2$. As the sunspot number is only available since 1610 AD the reconstruction of the full solar cycle variability with an annual resolution extends back only 400 years. Both reconstructions in the right-hand panels of Fig. 8.6 are based on the $^{10}\text{Be}$ data sets mentioned above. The difference in the reconstructions allows the error originating from the uncertainties in the proxy data to be estimated (20–50 % in the solar forcing value, depending on the year). This is large but still significantly less than the change in irradiance between the present and the Maunder minimum. Both reconstructions suggest a significant increase in TSI during the first half of the twentieth-century as well as low solar irradiance during the Maunder and Dalton minima. The difference between the current and reconstructed TSI during the Maunder minimum is about $6 \pm 3$ W/m$^2$ (equivalent to a solar forcing of $\Delta F_{P-M} \sim 1.0 \pm 0.5$ W/m$^2$) which is substantially larger than recent estimates (see Sect. 8.3.2). Note that as our technique uses 22-year means of the solar modulation potential our approach cannot be tested with the last, unusual solar minimum in 2008. In order to reproduce the current minimum as shown in Fig. 8.5 we have adopted a value of 584 MeV for the future 22-year average in 2020 (which is 92% of the 22-year average for 1988–2009).
The reconstruction before 1500 AD in the left-hand panel of Fig. 8.6 (which stops for clarity at 500 BC) is based on $^{10}$Be records from the Greenland Ice core Project (Vonmoos et al., 2006). The modulation potential during the Maunder minimum is about 3–4 times less than at present but not zero. However, it has decreased to zero several times in the past and the corresponding TSI was even smaller than during the Maunder minimum. There were also several periods when the modulation potential, and hence TSI, were higher than the present value. The reconstruction back to 7000 BC is presented in Fig. 8.7. The choice of model A introduces an uncertainty of the order of 30%, which is estimated by comparing model A to other possible candidates for the minimum state of the quiet Sun, e.g., model B from Vernazza et al. (1981). Combining this with the uncertainties of the proxy data outlined in Fig. 8.6 we can roughly estimate the uncertainty of our solar forcing value to be 50%. In addition to the $^{10}$Be-based data of solar activity there are several $^{14}$C-based datasets (e.g., Solanki et al., 2004; Vonmoos et al., 2006; Muscheler et al., 2007; Usoskin, 2008). Employment of these datasets will lead to a somewhat different values of the solar variability, which is, however, covered by our rough order of magnitude estimate of the overall uncertainty of the reconstruction.

Our TSI reconstructions give a value of $\sim 1 \text{ W/m}^2$ per decade for the period 1900–1950. The Smithsonian Astrophysical Observatory (SAO) has a 32-year record of ground-based observations for 1920–1952. Although the SAO data are disputed in reliability (Solomon et al., 2007) and clearly contain a non-solar signal they are the only available long-term measurements of TSI in the first half of the twentieth-century. The data show an increase of 1±0.5 (1.5±0.5) W/m$^2$ per decade for the period 1928–1947 (1920–1952). As we are aware that maintaining a stable calibration to better than 0.1 % over 30 years is very demanding we cannot claim that the historical data confirm our reconstruction. Nevertheless, it is intriguing to note how well the SAO trend agrees with our TSI reconstruction.

Our reconstructed solar spectral irradiance comprises spectra from 130 nm to 10 µm. Fig. 8.8 presents a reconstruction of the integrated flux for several, selected spectral regions. The contrast between different brightness components of the quiet Sun is especially high in the UV, which results in a large historical variability of the UV spectral irradiance. The irradiance in the Schumann-Runge bands and Herzberg continuum increases from the Maunder minimum to the present by about 26.6% and 10.9% respectively, which is much larger than 0.4% for TSI and the visible region. The variability is also relatively high around the CN violet system whose strength is very sensitive to even small temperature differences due to the high value of the dissociation potential. The large UV variability reported here is especially of importance to the climate community because it influences climate via an indirect, non-linearly amplified forcing (Haigh, 1994; Egorova et al., 2004).

We are aware that the choice of model A is responsible for a relatively large fraction of the uncertainty in our results. Higher resolution observations have recently become available, and model A could possibly be improved in future studies. We emphasize that model A is not the coldest possible quiet Sun model and therefore our estimate is not a lower limit of the Sun’s energy output. The coldest model would be a non-magnetic atmospheric structure without a chromosphere and corona. As model A contains some remaining magnetic activity our approach does not imply that the solar dynamo stops during the periods when the modulation potential is equal to zero. Let us also notice that the modulation potential never reached zero for the last 400 years (see Fig. 8.6).

8.3.6 Conclusions

We present a new technique to reconstruct total and spectral solar irradiance over the Holocene. We obtained a large historical solar forcing between the Maunder minimum and the present, as well as a significant increase in solar irradiance in the first half of the twentieth-century. Our value of the historical solar forcing is remarkably larger than other estimations published in the recent literature.

We note that our conclusions can not be tested on the basis of the last 30 years of solar observations because, according to the proxy data, the Sun was in a maximum plato state in its long-term evolution. All recently published reconstructions agree well during the satellite observational period and diverge only in the past. This implies that observational data do not allow to select and favor one of the proposed reconstructions. Therefore, until new evidence become available we are in a situation that
8.3. A new approach to long-term reconstruction of TSI

Figure 8.8: Reconstruction of the integrated spectral irradiance in selected wavelength bands. This reconstruction is based on the composite of the $^{10}$Be South Pole record and neutron monitor data (red curve in Fig. 8.6). Panels from bottom to top: The 500–600 nm band is representative of the time evolution of the visible irradiance. The CN violet system is one of the most variable bands accessible from the ground. The Herzberg continuum is crucial for ozone production in the stratosphere. The Schumann-Runge band is important for heating processes in the middle atmosphere.
different approaches and hypothesis yield different solar forcing values. Our result allows the climate community to evaluate the full range of present uncertainty in solar forcing.

The full dataset of the solar spectral irradiance back to 7000 BC is available upon request.

8.3.7 Acknowledgements

We are grateful to Friedhelm Steinhiiber and Jürg Beer for useful discussions and help with data. The research leading to this paper was supported by the Swiss National Science Foundation under grant CRSI122-130642 (FUPSOL) and also received funding from European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement N 218816 (SOTERIA). M.H. appreciates funding from the Swiss Holcim Foundation.

8.3.8 Online Material

Spectral synthesis

8.3.9 Quiet Sun

The quiet Sun is a combination of different brightness components and the evolution of their fractional contributions drives its activity and long-term irradiance variability. The four main components of the quiet Sun (Vernazza et al., 1981; Fontenla et al., 1999) are: component $A$ (faint supergranule cell interior), component $C$ (average supergranule cell interior), component $E$ (average network or quiet network), and component $F$ (bright network).

We calculated the synthetic spectra $I_A$, $I_C$, $I_E$, $I_P$ of all these components employing the NLTE (non-local thermodynamic equilibrium) COde for Solar Irradiance (Haberreiter et al., 2008; Shapiro et al., 2010) (COSI). Recently, Shapiro et al. (2010) showed that COSI calculations with the atmosphere model for component $C$ reproduces spectral irradiance measurements from the last two solar minima with good accuracy. This is used in Eq. (8.9) from the main text, where $I_{\text{present}}(\lambda)$ is substituted by the irradiance $I_C(\lambda)$ for component $C$.

The evolution of the magnetic activity of the quiet Sun can be represented by the time-dependent fractional contributions (i.e. filling factors) of different components of the quiet Sun. The quiet Sun can be described by model $A$ with varying contributions from the brighter components. We compose the quiet Sun using model $A$, and model $E$ which is an adequate representation of the brighter contributions.

Using COSI we demonstrate (see Fig. 8.9) that the solar irradiance for model $C$ and, therefore, measured solar irradiance for the last two minima, can be successfully reproduced with a combination of $\alpha_A^{\text{present}} = 43\%$ model $A$ and $\alpha_E^{\text{present}} = 57\%$ model $E$ ($\alpha_A^{\text{present}} + \alpha_E^{\text{present}} = 1$):

$$I_{\text{present}}(\lambda) = I_C(\lambda) = \alpha_A^{\text{present}} I_A(\lambda) + \alpha_E^{\text{present}} I_E(\lambda).$$  \hspace{1cm} (8.10)

The linear scaling of the magnetic activity of the quiet Sun with proxies, as described in the main text, is equivalent to setting the model $E$ filling factor to be proportional to a chosen proxy. So that the time-dependent irradiance of the quiet Sun can be calculated:

$$I_{\text{quiet}}(\lambda, t) = (1 - \alpha^E(t)) I_A(\lambda) + \alpha^E(t) I_E(\lambda),$$  \hspace{1cm} (8.11)

where

$$\alpha^E(t) = \frac{\mathcal{P}_{\text{proxy}}(t)}{\mathcal{P}_{\text{proxy}}^{\text{present}}} \alpha_E^{\text{present}}.$$  \hspace{1cm} (8.12)

The set of Eqs. (8.10-8.12) is equivalent to Eq. (8.9).

There is an ongoing discussion of whether the trend in the filling factor of the quiet network can be detected. It was suggested (Foukal and Milano, 2001) that the analysis of the historical Mt. Wilson observations exclude the existence of the trend between 1914 and 1996. However, this analysis was
Discussion

8.4. Discussion

8.4.1 Uncertainties

In the work above we assumed that the quiet sun is only composed of model A and model E. This assumption can be changed to e.g. the current day quiet sun consisting of 10% A, 77% C, 10% E and 3% F, as proposed by Fontenla et al. (1999). This will decrease the variability depending on how the long-term trend of the model E and model F are treated, given that model A of the current day quiet sun stays constant.

Figure 8.9: Representation of the quiet Sun by a combination of models. Ratios of synthetic solar spectra of models A (blue) and E (red) to model C. The black line results by combining 57% of the model E spectrum with 43% of the model A spectrum. The spectral signatures are almost perfectly cancelled, implying that the combination of both spectra is equal to the model C spectrum. The latter has been demonstrated to accurately reproduce the observed solar spectra during the last two minima (Shapiro et al., 2010). Hence, a combination of spectra for models A and E reproduces the observed quiet Sun spectrum.

Figure 8.9: Representation of the quiet Sun by a combination of models. Ratios of synthetic solar spectra of models A (blue) and E (red) to model C. The black line results by combining 57% of the model E spectrum with 43% of the model A spectrum. The spectral signatures are almost perfectly cancelled, implying that the combination of both spectra is equal to the model C spectrum. The latter has been demonstrated to accurately reproduce the observed solar spectra during the last two minima (Shapiro et al., 2010). Hence, a combination of spectra for models A and E reproduces the observed quiet Sun spectrum.

criticized for using uncalibrated data (Solanki and Krivova, 2004) and also contradicts other studies (Lockwood, 2009). Thus only the future long-term monitoring of the quiet Sun with high resolution or recently proposed (Kleint et al., 2010) monitoring of weak turbulent magnetic fields can help to clarify this question.

8.3.10 Active Sun

From 1610 onward we have additional information from sunspot number, which allows the calculation of the active regions contribution to the solar irradiance ($I_{\text{active}}(\lambda, t)$ in Eq. (8.9)). For this we follow the approach by Krivova et al. (2003). Because our main goal is to reproduce centennial solar variability and because magnetograms are unavailable for historical time periods, we scale the faculae and the active network filling factors with the sunspot number instead of using filling factors derived from available magnetogram data. The synthetic spectra are then added according to their filling factors, and TSI is then determined by integrating over all wavelengths.

End of Paper
Another assumption is the one of the quiet network. The model of the quiet network may be improved, however replacing the quiet network with another model will not change the result unless the magnetic activity gets above $1/e_m$, i.e. it saturates the filling factor of the quiet network. However, since the magnetic activity of the last 400 years was always below the current level of activity, this problem has no application for the content of this work. The driver of the variability is the C–A difference, not the E–A difference.

The third assumption is the one of a linear combination of A and E. This assumption is made for practical purposes, for there are only two points to calibrate the model, the zero-activity, assumed to be an “all model A”, and the current magnetic activity. The upcoming solar cycle 24 may be used as an falsification, if the long-term magnetic activity changes significantly. Also see Section 11.1 for another approach to calculate the long-term evolution of the filling factors.

### 8.4.2 Extrapolation – Interpolation

It should be noted that this approach does not use the the measured TSI variability to extrapolate the long-term trend, as used by Schöll et al. (2007); Steinhilber et al. (2009); Fröhlich (2009b). Especially it is not an extrapolation of neither the short-term (i.e. 11-year cyclic) TSI variation to the long-term variation nor an extrapolation of the observed (small) variation of the long-term trend as done by Steinhilber et al. (2009); Fröhlich (2009b)where the uncertainties of the extrapolation is multiplied by the error of the measurements. Instead, it is an interpolation between the most quiet observed region and the current activity of the observed quiet sun. The uncertainties of this model is in the definition of what constitutes a “most quiet” or, rather zero-$\Phi$ sun, and wether zero-$\Phi$ really can describe this theoretical most quiet sun.

Another method that can be used to reconstruct the quiet-sun filling factors is described in Section 11.1. Applying the same method as described above, but using these new filling factors results in a weaker long-term trend, with an non-linear dependence on $\Phi$, as shown in Figure 11.6. Also compare to Figure 9.9, utilizing the A/E model for the quiet sun.
Chapter 9

Spectral Reconstruction

The same method can be used to reconstruct spectral variability. However, as discussed in the following chapter 9.1, the filling factors for the active regions have to be corrected for the formation height of the specific frequency. This is most notable in the UV where the line formation is in the transition region.

9.1 Active Area Expansion

As the magnetic field expands from the photosphere towards the transition region it follows that the area covered by active regions also depends on height, called the Active Area Expansion (AAE).

The photospheric filling factors from Wenzler (2005) used for this work are compared with Barra et al. (2009) EUV filling factors extracted from SOHO Extreme ultraviolet Imaging Telescope (EIT) 17.1 nm and 19.5 nm images. Since those are coronal images, the formation height is above 5000 km and the active regions display a loop structure. Hence the increase in active regions can only be used as an indicator of AAE, not as a quantifier. The corona filling factors are shown in Figure 9.2, upper panel. The lower panel displays the ratio of Wenzler active area filling factor to Barra’s, showing an increase of about eleven in coverage for the Fe-XII 19.5 nm line and about half this increase for the Fe-IX 17.1 nm line. For a discussion of the form of those magnetostatic flux tube models see also Zwaan (1978).

For a more detailed analysis, the Solar Ultraviolet Monitor of Emitted Radiation (SUMER) Ly-ε images are compared to Michelson Doppler Imager (MDI) images, where the Ly-ε formation height is about the same as the Ly-α formation height, while the MDI images are of photospheric origin. MDI active regions are defined as the regions that are 3σ above average. The same method has been applied to the SUMER images, but with a different threshold. This threshold is selected to fit the form, but not the area of the MDI active regions which holds for thresholds in the range of 3σ to 4σ, forcing the active area expansion factor between 1.8 and 4.7. Using a factor of $f_{\text{AAE}} = 2.4$, which translates to a threshold of 3.6σ results in the best agreement with Ly-α measurements. It should be noted that this is a constrained free parameter.

McIntosh et al. (2001) extract a three dimensional potential field based from high-resolution co-aligned MDI and SUMER images. A sample of a SUMER image aligned to MDI is given in Figure 9.3. Based on this magnetic field extrapolation, the AAE factor is with 2.6 slightly higher. However, the
Chapter 9. Spectral Reconstruction

Figure 9.2: Top panel shows filling factors for both photospheric (solid) and coronal active regions. The coronal AR are calculated from two different datasets, the EIT-19.5 nm and the EIT-17.1 nm images. The first is the Fe-XII 19.5 nm which is formed above the Fe-IX 17.1 nm line. Bottom panel shows the ratio (dotted) of the aforementioned regions and the mean ratio (solid) for the whole time (mean value = 11.8), during solar maximum (11.7) and during solar minimum (mean value of 14.8), remaining relatively stable over a whole solar cycle. The mean for the Wenzler-Barra Fe-IX 17.1 nm line is half the mean-values above. Given that those are coronal filling factors it should be noted that the AAE formula does not hold any more since it does not take the loop-structure into account. This is only an indication for the existence of AAE, not a quantification.

atmospheric heights of the models used may differ from the one used to extract the formation height of Lyman-α. Figure 9.4 visualizes the modelled magnetic widening by interpolating the direction of the magnetic field over height. Figure 9.5 shows the extracted AAE from the previous figure. For this, photospheric active regions are selected and the boundary of those active regions are followed.

Based on the field extrapolation and the MDI and SUMER images a tube growth model is derived by assuming a linear increase in diameter over height, where the boundary condition is such that the AAE factor at the photosphere is unity and the factor at the formation height of Ly α is 2.4. Figure 9.1 shows the schematics.

The photospheric active filling factors are multiplied by the AAE, resulting in wavelength dependent filling factors (Eq. 9.2), the quiet sun is defined as usual as the remaining area (Eq. 9.3). The "active area enhancement" parameter $f_{AAE}$ is the enhancement of Lyman-α compared to the photosphere. It is set to 2.4 for this work.

$$AAE(\lambda) = 1 + \begin{cases} \text{if } h_\lambda > h_{500 \text{ nm}} & \left( f_{AAE} - 1 \right) \left( \frac{h_\lambda - h_{500 \text{ nm}}}{h_{Ly\alpha} - h_{500 \text{ nm}}} \right)^2 \left( \frac{1}{0} \right) \\ \text{otherwise} \end{cases} \quad (9.1)$$

$$\alpha'_{active}(t, \lambda) = \alpha(t) \times AAE(\lambda) \quad (9.2)$$

$$\alpha'_{quiet}(t, \lambda) = 1 - \sum_{x=\text{quiet}} \alpha'_x(t, \lambda) \quad (9.3)$$
9.1. Active Area Expansion

Figure 9.3: Aligned SUMER (left) and MDI (right) images with an active region highlighted. The active SUMER region has a larger area compared to the MDI active area. For large active regions as shown here the AAE effect is smaller compared to small active regions.

Extending the SATIRE model with AAE for active components results in

\[ I(\lambda, t) = \sum_x a'_x(t, \lambda) \times I_x(\lambda) \]  

(9.4)

where the quiet sun component is separated into several components as discussed in Subsection 8.2.3. Also see Section 11.1 for another method of calculating the long-term quiet sun contribution.

Figure 9.6 shows the reconstructed spectrum from 90 nm to 1000 nm. Figure 9.7 displays the relative change to the average irradiance for each wavelength for the same wavelength range. In the following sections I will discuss several wavelength ranges, from the UV to the IR and, separately the Lyman-\(\alpha\) line.
Chapter 9. Spectral Reconstruction

Figure 9.4: Modelled magnetic widening using the data by McIntosh et al. (2001). The X and Y axis are arbitrary data. Colour describes the final relative magnetic field strength.

Figure 9.5: Extracted AAE from Figure 9.4 by following the magnetic field lines of active regions over height.
9.1. Active Area Expansion

Figure 9.6: Reconstructed spectral solar irradiance from 90 nm to 1000 nm. The Lyman-α line at 121.5 nm is clearly visible. The eleven year solar cycle is noticeable at 300 nm and below. Figure 9.7 shows the relative change of the SSI. The AAE is applied to the whole spectrum.
Figure 9.7: Relative spectral reconstruction from 1609 to 2009. For each wavelength the relative change to its average is shown.
9.2 Lyman-\(\alpha\)

Several Lyman-\(\alpha\) reconstruction exist, with different resulting long-term trends, from a zero-long-term trend, as done by Woods et al. (2000) to 14% (Lean et al., 2011). Lean et al. (2011) also presents several reconstruction of the EUV with long-term variabilities from 0 to 100%, the latter case assuming no chromospheric faculae during Maunder minimum, utilizing a three component model, utilizing the MgII and F\(10,7\) indices with both daily values for Mg-II and an 81 day running mean of Mg-II to account for the long-term behaviour.

9.2.1 Results

Figure 9.8 shows a reconstruction of the Ly-\(\alpha\) with AAE, together with a comparison to both SOLSTICE and SUSIM measurements with several smoothing and time intervals to highlight different parts of the reconstruction. Figure 9.8(a) presents the long-term trend of the Ly-\(\alpha\) reconstruction with two different \(^{10}\)Be sources, displaying the error of the reconstruction due to different sources of radionuclide isotopes. Panel 9.8(c) and 9.8(d) compares the same data with a one year smoothing to the cyclic variability of SUSIM, SOLSTICE on UARS (1991–2002) and SORCE (2002–) and the Ly-\(\alpha\) reconstruction by Woods et al. (2000). For the 1-year intra-cycle and inter-cycle variation it is in good agreement with Woods, but for the current minimum. This disagreement is possibly due to missing data for the running mean of the quiet network. Panel 9.8(c) extends the previous data into the past while panel 9.8(b) compares the reconstruction to the weekly variation of SOLSTICE and SUSIM measurements. It shows an increased variability of the reconstruction, which is due to the possible overestimation of the sunspot model. Also this panel shows that the relationship of sunspots to plage break down on those time scales.

9.3 Total Solar Irradiance

The TSI is calculated as the integrated spectral solar irradiance, with a correction value applied to the integrated SSI to account for missing irradiance of the models in the infra-red and EUV. The integrated spectral irradiance of the five models used is given in table 8.1. The mixing ratio for the current time is also fixed by this values, i.e. \(\alpha_E = \frac{I_C - I_A}{I_E - I_A} \alpha_C\).

The reconstruction of the TSI is the same as in Shapiro et al. (2010) for the yearly data. The monthly data is obtained from the sunspot group number and the international sunspot number with monthly resolution. For a shorter resolution, one must take into account the \(\mu\) dependency of the filling factors and the spectrum. This would require a reconstruction of the \(\mu\) dependency by e.g. reconstructing the butterfly diagram over time. While this has been done by Jiang et al. (2011), this work does not take into account the short-term variability of the sun over the long-term trend.

9.4 Spectral Variability

The spectral variability for the most active day to the least active is shown in Figure 9.10. The variability changes from a factor of 1.7 in the EUV to a small negative factor of -1\% in the infra-red. This negative factor is due to the high sunspot quiet-sun contrast in comparison to the contrast of the bright regions with respect to the quiet sun.

9.5 Conclusion

Using the COSI code to calculate the spectral solar irradiance together with the SATIRE model, the observed spectrum can be reproduced. Furthermore, assuming a linear relationship between the
long-term trend of the quiet sun and \( \Phi \), the long-term variability of the TSI is with 6.5 W/m\(^2\) significantly greater than the eleven year intra-cycle variability of 1 W/m\(^2\). Furthermore, when extending the SATIRE model with the active area expansion, the UV spectrum can be reconstructed, with a high accuracy for the Ly-\( \alpha \) when compared to space-time and other reconstructions. The long-term variability of the Ly-\( \alpha \) using the same model results in 3 W/m\(^2\), which is about the same as the intra-cycle variability (See Figure 9.8). However, the model of the long-term trend is based on assumptions that can not yet be verified by observations due to the constant quiet-sun activity of the space time. Other models are possible, one is presented in the Outlook section of Part III, which results in a different trend.
Figure 9.8: All four figures display the reconstructed Lyman-α using different smoothing intervals and different timescales. The solid line is the reconstructed Ly-α. They are compared to the reconstruction by Woods et al. (2000) dating back to 1947 (dotted), UARS-SOLSTICE measurements (dot-dashed, 1991–2001) and SORCE-SOLSTICE data (dashed, 2003–2011). Panel 9.8(a) shows the reconstructed Ly-α, where only the long-term trend remains visible, while panel 9.8(b) compares the short-term variation during the beginning of solar cycle 23 to the various instruments and Woods reconstructions. Panels 9.8(c) and 9.8(d) are using a 1-year smoothing to be able to compare the long-term trend.
Figure 9.9: Figures a-c display the reconstructed integrated TSI using different smoothing intervals and different timescales. Figure (d) indicates the Beryllium 10 used, data from DYE 3, Greenland (blue) and the South Pole (red) (McCracken et al., 2004).
Figure 9.10: SSI variability from the Maunder minimum to today. The solid line represents the variability when using South Pole data for Φ and the dashed line uses the Dye-3 Data.
Part III

Conclusion & Outlook
Chapter 10

Discussion of Thesis Results

A new model to reconstruct both the Spectral Solar Irradiance (SSI) and Total Solar Irradiance (TSI) back to the Maunder Minimum has been introduced, with special emphasis on the EUV, in particular the Lyman-\(\alpha\) (121.5 nm) line. It shows a good agreement with both measurements and other reconstructions. To achieve this, the spectral radiative transfer code used had to be improved (Part I) and new models had to be developed to account for the long-term trend in solar activity as done in Part II, both in the temporal (Section 8.2) and spatial (Section 9.1) domain.

COde for Solar Irradiance (COSI) has two aims, understanding the physics and modelling the solar spectral irradiance. Ideally, it models the irradiance based only on the physical implementation. However, at some parts, when the physics is not yet understood, it is necessary to introduce synthetic parameters to account for unknown effects, as done in the case of missing opacity in the UV as described below in Section 10.2.

10.1 COSI Improvements

The improvements of the radiative transfer code consists of several parts, consisting of updating the code from Fortran 77 to Fortran 95, also including improvements in stability and implementing MPI capabilities. As part of the updates, an optional dependence on the linear algebra package lapack95 (Barker et al., 2001) has been introduced for performance optimization.

In Section 3.2.2 an accelerated lambda iterator has been implemented successfully for the lines, but not the continuum. The continuum could also be implemented, but requires some reformulation of the algorithm used.

The third improvement is the implementation of ambipolar diffusion of hydrogen described in Section 3.2.3. Its most pronounced effect is in the transition region where the Lyman-\(\alpha\) line is formed. It decrease the ionization gradient of hydrogen in the lower transition region due to diffusion of neutral hydrogen into the ionized hydrogen region. It also partly stabilizes the numerics and partly introduces numerical problems that has to be taken into account by carefully adjusting the parameters of convergence.

The turbulent velocities of the Fontenla atmospheric model are multiplied by four to compensate for the missing corona. This results in an improved Lyman-\(\alpha\) profile by widening the wings due to the amplified Doppler broadening (Section 3.3.2).

10.2 UV Opacities

The computed spectrum is compared to the ATLAS3 reference spectrum, showing an increased flux in the UV region, possibly due to missing lines. This is a known problem in NLTE model calculations, c.f. Busá et al. (2001); Short and Hauschildt (2009). Haberreiter (2006) used an frequency dependent
artificial line broadening to account for the missing opacities. While it does account for the missing opacities, it does not reproduce the line profiles. Hence the method has been changed by assuming background lines whose opacity add to the continuum opacity, improving the continuum irradiance, while preserving the observed line profile of all lines calculated explicitly as described in Shapiro et al. (2010) and Chapter 4. The amount of the missing background lines are calculated iteratively to account for NLTE effects.

10.3 1-D Modelling

COSI, as described in part I calculates the spectral irradiance of different solar features like sunspots, faculae, etc. in a one dimensional model atmosphere. This does not account for loop structure or enhancement of e.g. the magnetic field over sunspots, which implies different models as the line of sight goes through the different solar structures, or lateral heating/cooling effects from other models.

10.4 Variability over Time

COSI calculates a time-invariant spectrum for the different parts of the sun. In order to reconstruct the spectral solar irradiance over time, the changes of the solar parts have to be reconstructed, following the Spectral and Total Irradiance REconstruction (SATIRE) model by Krivova et al. (2003). For this work, international sunspot number is used to reconstruct all active solar components, i.e. sunspots, plage and the active network.

The long-term change in solar activity is modelled by use of the solar modulation potential $\Phi$ based on radionuclide isotopes extracted from ice-cores in Greenland and the South Pole by McCracken et al. (2004). This data is available as long-term (22 year) averages, where long-term is defined as more than eleven years, i.e. one solar cycle, with short term being the solar cycle variability. For this work it is assumed that $\Phi$ varies linearly with the long-term trend from the very quiet sun model A and the current day quiet sun, as observed by ATLAS3 and modelled by COSI (c.f. Section 8.3, theory part). Another method is to assume that the active sun can be described by the average network model $E$, with parameters $\alpha$ and $\epsilon$ chosen such that $I_{\text{ATLAS3}} = \alpha I_A + \epsilon I_E$. This is the actual method of implementation of the above part.

The advantage of this method is the interpolation nature. It does neither assume an extrapolation of the long-term trend nor any connection between the short-term and the long-term trend. The disadvantage for the first method is the mixture of a calculated model and observation. and for the second method, the discrepancies between observations and the mixture used in the UV (c.f. Section 8.2.3). In both cases, the main assumption is that all steps between the theoretical minimal sun and the current-day quiet sun can be interpolated, which may not be the case. E.g. it would be very difficult to interpolate between the two most extreme cases, the very quiet sun, and sunspots in a meaningful way. Both are darker than the quiet sun.

In Chapter 11 an alternative method is presented which is based on the magnetic distribution on the solar surface as observed by Stenflo and Holzreuter (2003). This method enables us to reconstruct the full spectrum using physics based models. This approach does not assume that any one solar feature can be described by a combination of other solar features, however, to obtain the necessary parameters for the distribution, either an extrapolation from current day minima or some assumptions of the connection between the short-term and long-term trend is necessary.

An implicit assumption in both approaches is that the long-term trend is directly coupled to $\Phi$ and independent of the short-term trend. However, $\Phi$ also has a partial contribution from the active features, whose strength vary from cycle to cycle, with a 50% contribution for the current solar cycle and a non-existent contribution during the mauerden minimum. Hence, it is reasonable to assume that the evolution of active features also contribute to the long term trend of $\Phi$ and that the long-term trend is overestimated for times with an increase of the active solar features.
10.5 TSI Reconstruction

Based on the time dependent reconstructed regions and corresponding irradiance values for each region, the TSI can be calculated by following the SATIRE model, that is taking the sum over the product of the irradiance value with the relative area of its corresponding region as done in Part II. This results in both a good agreement with observations for cycles 21-23 and a radiative forcing of 6 W/m\(^2\) for the Maunder minimum when using the linear interpolation as described above. The radiative forcing decreases to 1 W/m\(^2\) when applying the “magnetic distribution” method presented in Chapter 11.

10.6 Spectral Reconstruction

Replacing the irradiance scalar of each region of the TSI reconstruction by the corresponding spectral irradiance model as calculated by COSI, the spectrum can be calculated. The resulting spectra how a good agreement with measurement from 200 nm to 2 \(\mu\)m. However, it does break down in the UV, including Lyman-\(\alpha\), resulting in a 50\% lower irradiance variability compared to SUMER (Solar Ultraviolet Monitor of Emitted Radiation) and SUSIM (Solar Ultraviolet Spectral Irradiance Monitor) measurements.

10.6.1 Active Area Enhancement

One possible explanation of the decreased variability of the UV lies in the formation region of the UV. The UV originates in the chromosphere and transition region, which are the two regions right above, and \(\sim 2000\) km above the photosphere. Hence the size of the active regions may be different for different wavelength. It is also well-known that active areas enhance due to the magnetic pressure (Wedemeyer-Böhm and Wöger, 2008). Additional, Barra et al. (2009) extracted filling factors from coronal images, also displaying an enhancement of the active regions by a factor of \(\sim 5\) and \(\sim 11\) for the Fe-IX 17.1 and the Fe-XII 19.5 lines, respectively. Based on this, a simple one-dimensional (height-dependent) model is developed in Section 9.1 to account for this effect. It is based on the assumption of quadratic growth over height and calibrated against Lyman-\(\alpha\) images and MDI photospheric images. They restrain the Active Area Expansion (AAE) factor between 1.8 and 4.7 for the formation height of Lyman-\(\alpha\). This result agrees well with a best-fit of Lyman-\(\alpha\) irradiance measurement, that requires an AAE factor of 2.4.

This enhancement factor is an average over different sized active regions, that is the AAE factor is inverse proportional to the size of the active region. However, once again, as the goal is to reconstruct SSI back to the Maunder minimum, for which the data-source consists of sunspots only, the average AAE factor of 2.4 is used.

McIntosh et al. (2001) extract a three dimensional potential field based from high-resolution co-aligned MDI (Michelson Doppler Imager) and SUMER images. Based on this magnetic field extrapolation, the AAE factor is with 2.6 slightly higher than the one chosen, but within the range of error from the SUMER images. Also, this number depends on the atmosphere model used.

10.6.2 Lyman-\(\alpha\)

Applying the one-dimensional AAE model to the active regions of SATIRE, the spectral reconstruction improves, showing a good agreement with measurements of the Lyman-\(\alpha\) line as done in Section 9.2.

The long-term trend of Lyman-\(\alpha\) results in a 3 mW/m\(^2\) decrease compared to a modern time minimum of 6 mW/m\(^2\), with a 4 mW/m\(^2\) short-term variability. This is based on the linear interpolation model of the long-term trend reconstruction of the filling factors as described above. Using the distribution model, the long-term trend decreases to 0.5 mW/m\(^2\).
Chapter 11

Outlook

This reconstruction does not contain very short-term (below monthly) due to the limb-darkening effect and the need to reconstruct the \( \mu \) dependency of the filling factors, as done by Wenzler (2005) for the space-time.

It is also based on sunspots and cosmogenic radionuclides, hence the reconstruction can be taken back further in time when limiting oneself only to the long-term variability as done in Section 8.3.

As already mentioned in Section 3.2.2, it would be useful to also implement the Accelerated Lambda Iterator (ALI) for lines.

11.1 Three Component Quiet Sun – Stretched Distribution Model

Next to the short-term variability modelling for which the \( \mu \) dependency is needed, it is also interesting to look at the assumption of the linear relation of \( \Phi \) to the quiet network. Since the event of satellite measurements, there was no long-term trend, this assumption can not yet be verified, and other methods are possible to use. In the following section, I will devise a model which is based on the magnetic distribution of SOHO MDI measurements. The advantage of this model is the ability to fully reproduce the current-time observed spectrum by using as many atmosphere models as necessary.

However, this distribution is somewhat artificial in the sense that it is caused by the instruments resolution, and actually describes the aligned magnetic fields within a resolution element, where the turbulent magnetic fields chancel out. Stenflo and Holzreuter (2003) show that the average magnetic flux field strength of the quiet sun is in the kG range.

The two-component model described in Part II works because of the linearity of the temperature in the photosphere. This is not the case in the chromosphere, and hence the divergence from unity in Figure 8.4.

Also the model atmosphere correspond to an area of a specific magnetic strength, with the strength increasing from model A to model S (compare for Figure 11.1 for the first five models. Fontenla et al. (1999) improved on those models by Vernazza et al. (1981).) Stenflo and Holzreuter (2003) showed that the distribution of magnetic activity follows a Voigt distribution, shown in Figure 11.2. Hence, assuming that the Interplanetary Magnetic Field (IMF) is coupled to the integrated magnetic intensity, and that the nature of the magnetic distribution does not change over time, it is possible to calculate the filling ratio for each component according to the distribution, by coupling \( \Phi \) to e.g. the peak of Figure 11.1 or the integrated magnetic field strength of the Voigt distribution, or the Doppler width \( \Delta B \) (c.f. Stenflo and Holzreuter, 2003). This last method is used here to reconstruct the filling factors.

In this model, it is still assumed that the quiet network and the faint super-granule cells follow \( \Phi \) while the active regions are coupled to the sunspots. The latter assumption is due to availability of
Figure 11.1: Histogram of intensity measurements at $\lambda = 90\text{nm}$ of the quiet sun plotted as a function of the count rate (Vernazza et al., 1981, Figure 7, data extracted from Paper).

high-frequency data. For the current-time, Fontenla estimates the quiet-sun to be composed of 10% model A, 77% model C, again 10% of model E and 3% for model F.  

The 10% for each of model A and E and 3% for model F is with good agreement to the value we obtained for the A+E mixture in the previous section.

The Voigt profile has two parameters, the half-width of the Doppler broadening and the half-width of the Lorenz-profile, the ratio is denoted as the dampening factor $a$, and the Doppler broadening described by $\Delta B$. I assume that $\Delta B$ is linearly correlated to the long-term solar activity.

This leaves $a$ as an unknown parameter. The dependency of $a$ to $\Delta B$ are investigated with the help of MDI Magnetograms. For this analysis a total of 380 Magnetogram from different dates over the whole mission-time of MDI are used. From this dataset, 42 are removed due to problems with the magnetograms. For each Magnetogram a best-fit for both $\Delta B$ and $a$ are computed with a linear least-square, by use of lookup tables for fixed $\Delta B$, going linearly from 1 to 25 G with a step-size of 0.1 G and $a$ going logarithmically from $10^{-8}$ to 100, with 199 steps. This can be done successfully for 75% of all magnetograms. Figure 11.2 shows the best fit for different MDI images with varying $\Delta B$ and $a$. The dampening factor $a$ goes from $3.8 \cdot 10^{-4}$ to 0.6. $\Delta B$ ranges from 10 G to 14 G. The data displays a linear correlation ($\sigma = 0.75$) between $\Delta B$ and $a$, displaying a weak coupling between the dampening factor $a$ and $\Delta B$, as shown in Figure 11.1 and Equation 11.6.

Figure 11.3: Scatterplot of $\Delta B$ vs the dampening factor $a$ and a linear best-fit, see Equation 11.6.
11.1. Three Component Quiet Sun – Stretched Distribution Model

When assuming that the long-term quiet-sun features exhibit the same behaviour as the active sun, and that the long-term filling factors depend only on one value, $\Phi$, it can be coupled to both $\Delta B$ and $a$.

See Figure 11.5 for a comparison of the reconstructed filling factors using both methods described. For the three component model, model C has been partitioned into two parts and added to A and E, using the same parameter $c$ as in the previous section for the two component model. Both the current time solar modulation and the minimum solar modulation are marked to show that the difference of the two methods are marginal.

This method is independent of the absolute value $\Delta B$, but on the relative change of $\Delta B$ w.r.t. $\Phi$.

\[
\alpha_A = \alpha_{\text{quiet}} \int_{B_A} v(\Delta B, a(\Delta B)) \, d\lambda 
\]
\[
\alpha_E = \alpha_{\text{quiet}} \int_{B_E} v(\Delta B, a(\Delta B)) \, d\lambda 
\]
\[
\alpha_C = \alpha_{\text{quiet}} - (\alpha_A + \alpha_E) 
\]
with
\[
\alpha_{\text{quiet}} = 1 - \alpha_{\text{active}} = 1 - (\alpha_F + \alpha_P + \alpha_S) 
\]
\[
\Delta B = 17 \frac{\Phi}{\Phi_{1996}} 
\]
\[
a := \Delta B \rightarrow 2.24 - 0.16 \Delta B 
\]
and the Model A and E magnetic activity range,
\[
B_A = [0, 2.14] \, \text{G}, \quad B_E = [23, 35.8] \, \text{G} 
\]

where Equation 11.6 is a linear-logarithmic best-fit for the $\Delta B - a$ relation and the intervals in Equation 11.7 are chosen such that the integral for the current time is consistent with Fontenla et al. (1999).

The long-term difference of both the spectral and total solar irradiance differs significantly when compared to the two component model from Section 8.2.3. Compare e.g. Figure 11.6 using the three component model with Figure 9.9 which utilizes the two component model. The difference in TSI from Maunder minimum to today using the distribution model is $1 \, \text{W/m}^2$, compared to a $6 \, \text{W/m}^2$ long-term trend that results from the two component model.
Figure 11.2: Best-fit Voigt distributions for different times of full disk MDI measurements. The $\Delta B$ is usually between 12 G and 30 G, with the dampening parameter $a$ between 0.018 and 0.68. Figures (a)–(c) show a typical Voigt distribution, while Figure (d) displays atypical behaviour (seen in 5 from 47 images).
11.1. Three Component Quiet Sun – Stretched Distribution Model

Figure 11.4: Reconstructed Filling Factors utilizing the model of the “Stretched Voigt distribution”. The solid and dotted lines are the same as in Figure 8.3, that is model A and model E. The new component is model C, shown as a dash-dotted line. Again, thick lines correspond to Dye-3 based Φ and the thin lines to the South Pole archive.

Figure 11.5: Quiet sun components A (dashed) and E (solid) versus the solar modulation potential Φ using the two component model described in Subsection 8.2.3 and the three component model described above. For the three component model, the quiet sun model C is split into two parts, as described in the two component model subsection, with A+(1-c)C (dashed) and E+cC (solid) shown. This is done to ease comparison of the two filling factors. The horizontal lines marked with A and B is the minimal solar modulation during Maunder minimum and the current-time solar modulation, respectively. The difference between the two models is small for the time between Maunder minimum and the current time. However, this difference increases for an increase in Φ.
Figure 11.6: Figures a–c display the reconstructed integrated TSI using different smoothing intervals and different timescales using the three component model described in Section 11.1. Compare to Figure 9.9 for a reconstruction using the two component model. Figure (d) shows the \( \Phi \) based on the two Beryllium 10 archives used, data from DYE 3, Greenland (blue) and the South Pole (red) from McCracken et al. (2004).
Part IV

Appendix
Appendix A

H-Minus Manual

Since there is no written manual of the COde for Solar Irradiance (COSI) code, I decided to include one here. The source code for all the programs is at svn://corona.pmodwrc.ch/cosi/ using the svn version control system. In section A.2.4 I also included the manuals for some small programs I wrote to make it easier to handle COSI runs. This appendix is only useful for persons wanting to use COSI.

A.1 H-Minus

A.1.1 Input

The input files are the atomic data, solar composition and the frequency grid.

Elx----n opacities for the element El at ionization level x and level n, e.g. MgI------2.

DATOM Atomic Data file, links the opacity files. See datom.for.

TABLE|FAL_VD solar atmosphere. Either an Asplund atmosphere (TABLE) or an Fontenla atmosphere with included doppler broadening (FAL_VD).

fort.99 Control file. To start, this file must contain the line wrstart. Other options are wrcont, repeat or stop. See hminus.for for more information

correction.xlbkg The multiplication for the lineblanketing. If the correction is below 1 and the convergence coefficient is below 0.01, this value gets multiplied by 10 up to a maximum value of 1.

A.1.2 Output

POPNUM Population Numbers of all elements given in DATOM.

BROYDEN Broyden Matrix

RADIOC Continuum radiation field

RADIOL Line radiation field

*.out Informational output files

1http://subversion.tigris.org
A.2 FIOSS

The program fioss8.exe needs several input files and the control file CARDS. The input files are the hminus output files, POPNUM, MODFILE, RADIOC and RADIOL, the CARDS control file and the following files:

A.2.1 Input

fort.19 The line database. It must exist, but may be empty.

fort.55 Control file for lines. Depending on the format of the file, additional input files are needed. The file contains keys, with each entry a key; the names of the keys are IMODE, IDSTD, IHEMC, INLTE, INLIST, IFHE2, IIVCS, IBVCS, IHE1, IHE144, IHE2UV, IHE2VI, IHE2RE, ALAMO, ALAST, CUTOFO, RELOP, SPACE.

The third line contains the file units for the stark profile for different elements. E.g. the entry IBVCS is the Index Balmer Vidal, Cooper, Smith (now extended to all hydrogen lines, as given by Lemke (1997))

Elx-----n opacities for the element El at ionization level x and level n, e.g. MgI------2.

VELO or TABLE The temperature structure. If the key FAL in CARDS is set, the Fontenla structure in VELO is read, otherwise the file TABLE is used.

A.2.2 Control file

The CARDS file consists of KEY [VALUE...] entries, with the following keys:

**ABEMLIN WRITE|READ|AUTO (path).** If it is set to WRITE, write the abemlin into (path)/\(\lambda\).abemlin. If it is set to READ, read from the given path. AUTO reads the absorption and emission profile, if it exists, otherwise it writes it out.

**NFOBS** number of frequency points for the output file, evenly spaced. This setting is independent of the internal frequency grid.

**MAXITER** set to 1.

**PFIRST** First impact point, impact points 1 to 9 correspond to \(\mu\) values between 1 and 0. Impact points above 9 are for off the limp calculations. Defaults to 1.

**PLAST** Last impact point. Defaults to number of depth points as given in MODFILE.

**VDOP** The minimum turbulent velocity. If a Fontenla atmosphere is given, and the turbulent velocity for a depthpoint is above VDOP, this turbulent velocity is used. Variables that influence this behavior are the logics ADDVELO and ADDVAR in fioss8.for

**INTERVAL** Last line in the CARDS. This line causes the fioss program to run. It has two arguments, given the lower and upper limit of the frequencies to calculate. This must be the last line in the CARDS file.

A.2.3 Output

The output files are given below. \(\lambda\) is the center frequency,

\[
\langle \lambda \rangle = \text{round}((\text{INTERVAL}_1 + \text{INTERVAL}_2)/2)
\]

\(\langle \lambda \rangle.mdisp\) The irradiance for the given frequencies in erg/s/cm\(^2\)/Hz.
A.2. FIOSS

\(\langle \lambda \rangle .\)lopa The line opacities for the frequency the given frequencies.

\(\langle \lambda \rangle .\)abemlin The line opacities and emissions for the given frequencies. This file is only written if the option ABEMLIN in CARDS is set to WRITE or AUTO.

\(\langle \lambda \rangle .\)tau The line formation height of the given frequencies. The format of the file is frequency(Å), formation height (km), interpolation error (km), mode (0,1,2) and nearest depthpoint (#).

\(\langle \lambda \rangle .\)title one line containing the TITEL variable containing NFOBS from CARDS, center frequency and the date.

The mode is set to 1 if \(\tau_\lambda = 1\) is already at the top of the atmosphere, and two if the atmosphere is transparent. Otherwise it is 0.

A.2.4 Helper programs

Fioss Prepare

fioss_prepare must be run in the main project directory. It uses the file fioss.settings, created by cosi_project_prepare and creates the fioss/IT* directory with all the files needed to run fioss8.exe or fioss_do

Fioss Do

The helper program fioss_do needs the variable \$COSI_BIN to be set to the location of the script. The bash script fioss_do [from step to] runs \$MAX_CPU number of parallel sessions of fioss8.exe. The system variable \$MAX_CPU sets the maximum number of programs running in parallel. If run for the first time in a project directory, it creates the file CARDS.TEMPLATE, which is the CARDS file without the INTERVAL setting.

fioss_do operates in three possible modes, the default mode is to run fioss8.exe for the intervals from 900Å to 10000Å in 10Å steps. The arguments are from, setting the first frequency step in Ångstrom (default 900), step, the step width (default 10) and the maximum frequency to (default 10000). fioss_do creates a control directory in the working directory and in there it creates a file for each run with the INTERVAL line. fioss_do can be run with the argument --ctrl <ctrl directory> where it uses this directory to run an additional \$MAX_PROC processes. If the program is run with the argument --single <step> <list of frequencies> it runs fioss8.exe for each entry in list of frequencies with a window width of step.

For example, the following runs fioss on four parallel processes from 900 Å to 1000 Å with a step-width of 10 Å.

\[$\text{export MAX\_PROC=4}$\]
\[$\text{fioss\_do}$ The first run creates the template file, while the second run uses this file
\[\text{> CARDS.TEMPLATE created. Please check.}$\]
\[$\text{fioss\_do 900 10 1000}$ Run fioss8.exe for the wavelength intervals 900–910, 910–920, . . . , 990–1000

A.2.5 Opacity

The opacity program sorts the line opacities in 1nm bins and writes them out as the line blanketing files. See Haberreiter et al. (2008) for a detailed description of the physics. The necessary input file is CARDS.LOPA consisting of key value pairs with

- XLAM1 start frequency
- XLAM2 end frequency
- DIM number of depthpoints
DLAM width of bins in Ångstrom. For this work 10 Å are used
INPUT PATH path to lopa files
OUTPUT PATH path to write files to

A.3 Module Dependencies

The dependencies of the modules are shown for each sub-program, \texttt{wrstart}, \texttt{wrcont}, \texttt{como}, \texttt{etl}, \texttt{steal} and the program \texttt{fioss}. Modules that have a dashed box around them are shown separately in another figure. Utility modules and modules containing constants etc. are not shown.

![Diagram of COSI dependencies](image1)

\textbf{Figure A.1:} COSI dependencies. The dependencies of highlighted modules are shown in subsequent plots.

![Diagram of GREYM dependencies](image2)

\textbf{Figure A.2:} Dependencies of GREYM, including COOP\_M
A.3. Module Dependencies

Figure A.3: Dependencies of WRSTART. For module dependencies of GREYM see Figure A.2
Figure A.4: Dependencies of WRCONT. Module dependencies of ELIMIN and COOP_M are shown in Figures A.10 and A.2 respectively.
Figure A.5: Dependencies of COMO
Figure A.6: Dependencies of ETL
A.3. Module Dependencies

Figure A.7: Dependencies of STEAL
Figure A.8: Dependencies of FIOSS
A.3. Module Dependencies

Figure A.9: Dependencies of LINPOP

Figure A.10: Dependencies of ELIMIN
Appendix B

Acronyms

AAE  Active Area Expansion ................................................................. 93
ALI  Accelerated Lambda Iterator ......................................................... 95
COSI  COde for Solar Irradiance ............................................................. 103
EIT  SOHO Extreme ultraviolet Imaging Telescope ..................................... 77
GSS  Group Sunspot Number ................................................................. 54
IMF  Interplanetary Magnetic Field ........................................................ 95
ISN  International Sunspot Number, see SSA and GSS ............................. 51
LISIRD  LASP Interactive Solar Irradiance Datacenter ............................ 55
LTE  Local Thermodynamic Equilibrium ................................................ 9
MDI  Michelson Doppler Imager .............................................................. 93
NLTE  Non-Local Thermodynamic Equilibrium ....................................... 9
NM  Neutron Monitor ............................................................................. 59
SATIRE  Spectral and Total Irradiance REconstruction ............................ 92
SOLSTICE  Solar Stellar Irradiance Comparison Experiment ..................... 55
SORCE  Solar Radiation and Climate Experiment .................................... 55
SSA  Sunspot Area
SSI  Spectral Solar Irradiance ................................................................. 91
SUMER  Solar Ultraviolet Monitor of Emitted Radiation ............................ 93
SUSIM  Solar Ultraviolet Spectral Irradiance Monitor .............................. 93
TE  Thermodynamic Equilibrium ............................................................ 9
TSI  Total Solar Irradiance .................................................................... 91
UARS  Upper Atmosphere Research Satellite ........................................... 54
VIRGO  Variability of Solar Irradiance and Gravity Oscillations ............... 54
Bibliography


Harder, J. (2009). personal communication. 33


Penza, V., Caccin, B., and Del Moro, D. (2004a). The sensitivity of the C I 538.0 nm Fe I 537.9 nm and Ti II 538.1 nm lines to solar active regions. A&A, 427:345–351. 43


Acknowledgements

Primarily I would like to thank Werner Schmutz, who let me participate in the Polyproject and introduced me to solar physics. His door was always open and he let me take advantage of his great knowledge and overview of the field astrophysics, next to his willingness to answer not only complicated questions, but also very simple ones. He also let me go to plenty of conferences.

I also would like to thank Marcella Carollo, who agreed to became my supervisor at the astrophysical department at ETH, after the solar physics department closed down. Thanks also go to her predecessors, Svetlana Berdyugina, who became my supervisor after my first supervisor at ETH, Jan O. Stenflo retired.

I am grateful to Sami Solanki that he accepted to become my second examiner. I was lucky enough to listen to several talks and always found his discussion and feedback to my work useful.

Further thanks go to Juan Fontenla, who let me stay at LASP in Boulder and taught me about his solar atmospheres and ambipolar diffusion. I learned a lot from his insight and sharp thinking. Werner Curdt provided me with the SUMER data and went into the archives to obtain those and answered all my questions about the data usage. Many thanks for that.

With regards to my colleagues at PMOD/WRC, my warmest thanks go to Margit Haberreiter, who was extremely helpful especially during my first part of the PhD, where she showed me how to use the COSI code and taught me a lot about solar physics. It was always a pleasure to discuss solar physics with her.

I would like to thank all the members of the polyproject VSCG-II, which showed me the broader impact of solar science and also lead to the publication of my first paper together with Friedhelm Steinhilber and his supervisor, Jürg Beer.

Further thanks go to my office-colleagues Alexander Shapiro, with whom I had plenty of interesting discussions about pretty much anything solar and some sword-fights (applying a very liberal definition of the word sword), Daniel Lachat, Anna Shapiro, Edgar Schmucki and Uwe Schilffowitz with whom I also shared an office at one time or another. It was always a nice atmosphere in the office, and not only in the office, but at the whole institute. So thanks also goes to the whole of the PMOD/WRC. I also would like to thank the it support, Marco Senft. He fixed all the problems that I, as the first-level technical support guy wasn’t able any time, including weekends. I am especially grateful for his fixes of problems I created.
Curriculum Vitae

Name Micha Schöll
Date of Birth April 4, 1983
Place of Birth Bensheim, Germany
Nationality German

CSYS in Mathematics I and II, Physics, Computational Sciences and German.

Aug. 2000 – Jun. 2004 Student at the University of St-Andrews,
2004 MSc. Mathematics, “Measure Theory”


Thesis title “Reconstruction of spectral solar irradiance back to the Maunder minimum.”,