Empirical reconstruction of the solar ultraviolet spectrum

Anatoliy Vuiets

UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE ENERGIE - MATERIAUX - SCIENCES DE LA TERRE ET DE L'UNIVERS

Laboratoire de Physique et Chimie de l'Environnement et de l'Espace

THÈSE présentée par :

Anatoliy VUIETS

soutenue le : **24 mars 2015**

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline/ Spécialité : **Sciences de l'Univers**

Reconstruction empirique du spectre UV solaire

THÈSE DIRIGÉE PAR :

Thierry DUDOK DE WIT	Professeur, Université d'Orléans

RAPPORTEURS :

Yvonne UNRUH	Reader, Imperial College, Londres
Jean-François HOCHEDEZ	Directeur de Recherche, LATMOS, Paris

JURY :

Alessandro SPALLICCI	Professeur, LPC2E, Univ. d'Orléans
Frédéric AUCHÈRE	Astronome-adjoint, IAS, Orsay
Thierry DUDOK DE WIT	Professeur, LPC2E, Univ. d'Orléans
Jean-François HOCHEDEZ	Directeur de Recherche, LATMOS, Paris
Matthieu KRETZSCHMAR	Maître de Conférences, LPC2E, Univ. d'Orléans
Yvonne UNRUH	Reader, Imperial College, Londres

Acknowledgments

The work comprising this thesis was carried out in the Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), Université d'Orléans, France, under the supervision of Thierry Dudok de Wit. I would like to express my gratitude to him for his kind support, guidance, and encouraging me to learn new things. Another person in LPC2E who supported me a lot is Matthieu Kretzschmar. His experience, useful discussions and friendly attitude played a huge part in accomplishing this work.

Luis Eduardo Vieira, who currently works in Espaciais Instituto Nacional de Pasquisas Espacias, Brazil, deserves special mention here, as during his post-doctoral research in LPC2E he developed several ideas that were used in this thesis.

I am very grateful to the région Centre - Val de Loire, and the Centre National d'Etudes Spatiales (CNES) for a research scholarship.

Finally I would like to thank my examiners Yvonne Unruh and Jean-François Hochedez for their thorough analysis of this thesis and giving tips for its improvement.

Résumé

L'irradiance spectrale solaire (SSI) dans la bande ultraviolette est un paramètreclé pour la spécification de la moyenne et la haute atmosphère terrestre. Elle est requise dans de nombreuses applications en météorologie de l'espace, et aussi pour l'étude du climat. Or les observations souffrent de plusieurs défauts : manque de couverture spectrale et temporelle permanente, dégradation des capteurs, désaccords entre le instruments, etc. Plusieurs modèles de reconstruction de la SSI ont été développés pour pallier à ces difficultés. Chacun souffre de défauts, et la reconstruction du spectre en-dessous de 120 nm est un réel défi.

C'est dans ce contexte que nous avons développé un modèle empirique, qui recourt au champ magnétique photosphérique pour reconstruire les variations du spectre solaire. Ce modèle décompose les magnétogrammes solaires en différentes structures qui sont classées à partir de leur aire (et non sur la base de leur intensité, comme dans la plupart des autres modèles). La signature spectrale de ces structures est déduite des observations, et non pas imposée par des modèles de l'atmosphère solaire. La qualité de la reconstruction s'avère être comparable à celle d'autres modèles. Parmi les principaux résultats, relevons que deux classes seulement de structures solaires suffisent à reproduire correctement la variabilité spectrale solaire. En outre, seule une faible résolution radiale suffit pour reproduire les variations de centre-bord. Enfin, nous montrons que l'amélioration apportée par la décomposition du modèle en deux constantes de temps peut être attribuée à l'effet des raies optiquement minces.

Abstract

The spectrally-resolved radiative output of the Sun (SSI) in the UV band, i.e. at wavelengths below 300 nm, is a key quantity for specifying the state of the middle and upper terrestrial atmosphere. This quantity is required in numerous space weather applications, and also for climate studies. Unfortunately, SSI observations suffer from several problems : they have numerous spectral and temporal gaps, instruments are prone to degradation and often disagree, etc. This has stimulated the development of various types of SSI models. Proxy-based models suffer from lack of the physical interpretation and are as good as the proxies are. Semi-empirical models do not perform well below 300 nm, where the local thermodynamic equilibrium approximation does not hold anymore.

We have developed an empirical model, which assumes that variations in the SSI are driven by solar surface magnetic flux. This model proceeds by segmenting solar magnetograms into different structures. In contrast to existing models, these features are classified by their area (and not their intensity), and their spectral signatures are derived from the observations (and not from models). The quality of the reconstruction is comparable to that of other models. More importantly, we find that two classes only of solar features are required to properly reproduce the spectral variability. Furthermore, we find that a coarse radial resolution suffices to account for geometrical line-of-sight effects. Finally, we show how the performance of the model on different time-scales is related to the optical thickness of the emission lines.

Contents

1 Solar UV irradiance **1**
- 1.1 Rationale of the thesis . . . 1
- 1.2 The Sun . . . 3
  - 1.2.1 Quantification of the solar electromagnetic radiation . . . 4
  - 1.2.2 Internal structure of the Sun . . . 4
  - 1.2.3 Solar magnetic field . . . 5
  - 1.2.4 Solar atmosphere . . . 7
  - 1.2.5 Formation of the UV spectrum . . . 14
  - 1.2.6 Variability of the UV emissions . . . 16
- 1.3 Impact of the UV irradiance on the Earth . . . 16
- 1.4 SSI observations in the UV . . . 19
- 1.5 Conclusions . . . 21
- 1.6 Résumé en Français . . . 21

2 Solar spectral irradiance models **25**
- 2.1 Proxy-based models . . . 27
- 2.2 Semi-empirical models . . . 29
  - 2.2.1 SATIRE-S . . . 30
  - 2.2.2 SRPM . . . 33
  - 2.2.3 COSI . . . 37
  - 2.2.4 SOLMOD . . . 37
- 2.3 Physical models . . . 37
- 2.4 Conclusions . . . 39
- 2.5 Résumé en Français . . . 40

3 Empirical modelling **43**
- 3.1 Data for training the model . . . 44
- 3.2 Long-term stability of the SSI observations . . . 45
- 3.3 Model description . . . 47
- 3.4 Extraction of magnetically active regions . . . 50
- 3.5 Segmentation by area . . . 53
- 3.6 Model selection . . . 55
  - 3.6.1 Optimising the number of classes and annuli . . . 56
  - 3.6.2 Optimising threshold levels . . . 57
  - 3.6.3 Single UV model . . . 59
- 3.7 Contribution of the small and large magnetic structures to the SSI variability . . . 61
- 3.8 Centre-to-limb variation . . . 62
- 3.9 The off-limb contribution . . . 64
- 3.10 Two-timescale model . . . 67

3.11 Contribution of magnetic structures to the SSI variability at different timescales . . . 70
3.12 Conclusions . . . 71
3.13 Résumé en Français . . . 72

4 Properties of the empirical model **77**
4.1 Noise in the data . . . 77
4.2 Model quality at different timescales . . . 79
4.3 Model training . . . 81
4.4 Model stability . . . 81
4.5 Intensity contrast of the magnetic structures . . . 84
4.6 Transition to operations . . . 86
4.7 Conclusions . . . 91
4.8 Résumé en Français . . . 91

5 Conclusions **95**

6 Conclusions en Français **101**

Bibliography **107**

Chapter 1

Solar UV irradiance

Contents

1.1 Rationale of the thesis 1
1.2 The Sun 3
1.2.1 Quantification of the solar electromagnetic radiation 4
1.2.2 Internal structure of the Sun 4
1.2.3 Solar magnetic field 5
1.2.4 Solar atmosphere 7
1.2.5 Formation of the UV spectrum 14
1.2.6 Variability of the UV emissions 16
1.3 Impact of the UV irradiance on the Earth 16
1.4 SSI observations in the UV 19
1.5 Conclusions 21
1.6 Résumé en Français 21

1.1 Rationale of the thesis

The spectrally-resolved solar electromagnetic radiation (SSI) in the ultra-violet (UV) band (typically between 1 and 300 nm) is a crucial quantity for specifying the state of the Earth's middle and upper atmosphere. This band is of high interest for space weather, with different societal applications such as satellite orbit prediction, and satellite telecommunications. More recently, it has also come under close scrutiny for climate studies.

Although the power emitted by the Sun in the UV accounts only for about 8% of its total (wavelength integrated) irradiance, this band has a strong impact on the Earth's environment, partly because it is highly variable: relative changes in excess of 10 to 100% can be observed (see Figure 1.1 c). UV emissions vary at timescales from seconds to decades, which makes its constant monitoring extremely important.

As most of the solar UV irradiance is heavily absorbed in the stratosphere and above, it cannot be measured from the ground. The very first measurements were obtained in 1946 with the development of sounding rockets. However, solar irradiance measurements really became available only with the beginning of the space era [Phillips 2008].

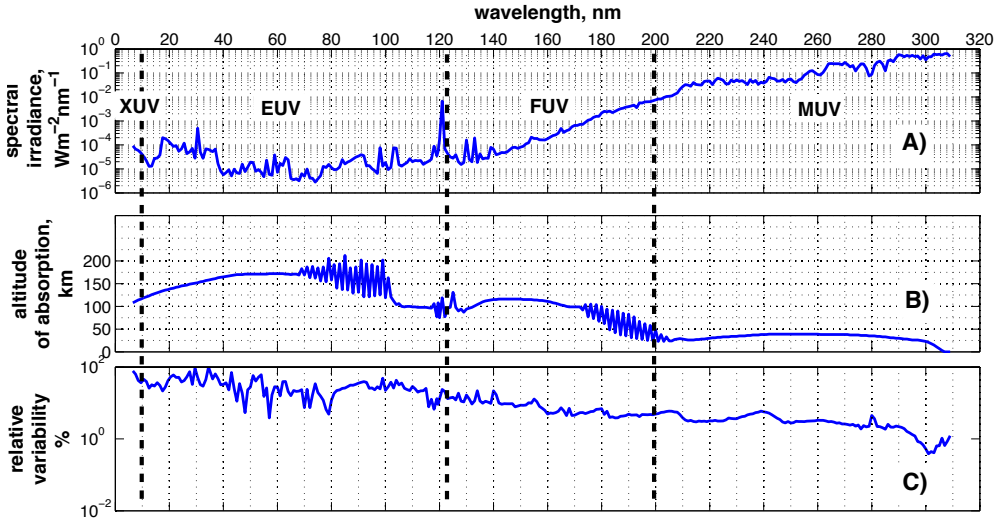

Figure 1.1: Solar spectral irradiance as observed by the SORCE and TIMED instruments from 22 April 2004 to 23 July 2010. Panel A shows the time-averaged solar spectral irradiance (spectrum). Panel B displays the characteristic altitude of absorption in the Earth's atmosphere. Panel C shows the relative variability of the SSI, for a 11-year solar cycle.

Today, these observations suffer from numerous gaps in both the spectral and temporal coverage. Moreover, observations by different instruments often disagree due to the lack of in-flight calibration. To alleviate these problems, many different SSI models have been developed over the last two decades. They can be split into two classes: proxy-based models, and semi-empirical models. The proxy-based models are based on the statistical relation between the SSI variability and different proxies of solar activity to reconstruct the latter, whereas the semi-empirical ones decompose the solar atmosphere into a set of atmospheric structures with different spectral profiles that are obtained partially or completely from physical models. Although both are able to reproduce the salient features of the SSI variability, they also suffer from weaknesses. Proxy-based models suffer from a lack of physical interpretation and rely heavily on the quality of proxies. Semi-empirical models encounter problems in reproducing the SSI variability below approximately 200 nm, where the local thermodynamic equilibrium does not hold anymore. Thus, there is a need for reconstructing the SSI in the UV, and in particular in the EUV (10 - 121 nm) band, which is crucial for specifying the thermosphere-ionosphere system.

In this context, we have decided to develop an empirical model that is based on assumption that the SSI variability is driven by the evolution of the solar surface magnetic field. Similarly to some of the existing semi-empirical models [Fligge 2000, Haberreiter 2008, Fontenla 2011], we decompose the solar atmosphere into a set of atmospheric structures with characteristic spectral intensity profiles. However, to

compensate for the incomplete physical description we derive both the classes of magnetic structures and their spectra from observations, rather than from models. As we shall see, the performance of this model is comparable to existing ones. Since it makes relatively few assumptions, we may expect it to give new insight into the nature and properties of the solar spectral variability.

Our model allows to investigate several important questions:

- To what degree can the SSI variability in the UV be described by the surface area coverage of different magnetic structures?
- What spatial resolution is needed to account for centre-to-limb variation effects?
- How many different classes of magnetic structures are needed to describe the observed SSI variability, given its uncertainty?

To proceed with the development of our model let us first discuss the structure of the Sun and the basic physical processes that lead to the formation and variability of electromagnetic emissions in the UV.

1.2 The Sun

The Sun is the brightest object in the Earth's sky, and our main source of energy. With a mass of about 2.99×10^{30} kg, it accounts for about 99.86 % of the total mass of the solar system. The solar radius is 6.96×10^8 m, leading to an average density that is 4 times smaller than that of the Earth. The distance between the Sun and the Earth changes due to the Earth's elliptic orbit, with an average distance of 1.47×10^{11} m (1 AU). The Sun rotates on its axis with an average period of 27 days. However, because the upper layers of the Sun cannot be treated as a solid body, they rotate at different rates that are latitude, longitude and altitude dependent. As we shall see later, this has a major influence on its magnetic field, which deeply impacts the variability of the SSI.

Compared to other stars in our galaxy, the Sun is a relatively large (about two times the median size) and bright star of spectral class G2V, with an effective blackbody temperature of 5777 K. Because of that, most of its energy (92%) is radiated in the visible and near infrared bands.

An important quantity that characterises the Sun as a star is its luminosity $L_\odot$, which is the total power emitted, integrated over all electromagnetic spectrum. The amount of solar radiative energy received by the Earth is called total solar irradiance (TSI), and is defined as TSI$= L_\odot/4\pi D^2$, where D is the average distance from the Sun to the Earth (1 AU). Although the TSI is now known to change with solar activity (about 0.1% during an 11-year solar cycle), in the past this quantity used to be called "solar constant". Today, its average value is estimated to be $1360.8 \pm 0.5\mathrm{W/m}^2$ [Kopp 2011].

1.2.1 Quantification of the solar electromagnetic radiation

In the analysis of the solar electromagnetic output, the most basic physical quantity is the radiative flux, which is the energy that is radiated through a closed surface per unit time. Thus, the radiative flux is measured in watts per meter squared ($W{\cdot}m^{-2}$). It is important to notice that the radiative flux does not account for the direction and the number of sources from where the radiative energy comes from. When the flux is incident on a surface it is called irradiance [Wilhelm 2010].

The most notable quantity that characterises the solar electromagnetic radiation is the above-mentioned total solar irradiance (TSI). The spectral distribution of the TSI, i.e., the solar irradiance as a function of wavelength λ, is called the solar spectral irradiance (SSI). The SSI is measured in units of $W{\cdot}m^{-2} \cdot \lambda^{-1}$. The SSI is of prime interest for solar-terrestrial sciences (e.g. atmospheric chemistry) since many physical and chemical processes require detailed knowledge of electromagnetic radiation at specific wavelengths.

In the frame of this work we focus on the ultraviolet (UV) band, which we take to correspond to wavelengths shorter than 300 nm. Indeed, as we shall see later, these wavelengths have enough energy to dissociate specific molecules, ionise atoms, and then deeply affect the physics and chemistry of all atmospheric layers. The UV band can be split into sub-bands, and the precise definition of these sub-bands depends on the context of a specific problem. However, in the field of the SSI modelling, the UV band is traditionally split into soft X-rays (XUV, 0.1 - 10 nm), extreme ultraviolet (EUV, 10 - 121 nm), far ultraviolet (FUV, 121 - 200 nm), medium ultraviolet (MUV, 200 - 300 nm) [Lilensten 2008] (see Figure 1.1).

1.2.2 Internal structure of the Sun

The Sun has a complex structure, as shown in Figure 1.2. The photosphere is the visible part of the Sun, which separates its atmosphere and the interior that is opaque to electromagnetic radiation. Although it is difficult to study the solar interior because no direct observations are possible, helioseismology provides a good picture of the internal structure [Christensen-Dalsgaard 2002]. The inner Sun consists of three parts:

- The core of the Sun is the heavier (about 50% of the total mass) and the most dense (the size is about 20-25% of the solar radius) part located in the centre. It consists mostly of hydrogen (72 %) and helium (26%). The remaining part (2%) consists of heavier elements, among which the most common are oxygen, carbon, and nitrogen. The extremely high temperature of about 27.7 $\times 10^6$ K and the pressure of about 250 $\times 10^{12}$ atm make it possible for fusion reactions to be sustained. These reactions are the main source of the energy, which propagates through other layers, and finally escapes from the Sun. The fusion of hydrogen nuclei consists of a series of reactions that are called the proton-proton chain:

$$^{1}_{1}\mathrm{H} + {}^{1}_{1}\mathrm{H} \longrightarrow {}^{2}_{1}\mathrm{H} + \mathrm{e}^{+} + \nu_{\mathrm{e}}$$

$$^{2}_{1}\mathrm{H} + {}^{1}_{1}\mathrm{H} \longrightarrow {}^{3}_{2}\mathrm{He} + \mathrm{e}^{+} + \gamma$$

$$^{3}_{2}\mathrm{He} + {}^{3}_{2}\mathrm{He} \longrightarrow {}^{4}_{2}\mathrm{H} + 2\,{}^{1}_{1}\mathrm{H}$$

where γ is a gamma photon. The energy output of this reaction chain (about 6.5×10^{14} J $\cdot$ kg^{-1}) gives about 95% of the total solar energy. The remaining part comes from the carbon cycle (CNO), which similarly produces helium from the hydrogen.

The total energy produced in the core of the Sun is transported to the photosphere by different mechanisms, and eventually, leaves the Sun in the form of energetic particles (20%), and photons (80%) [Phillips 2008].

- The radiative zone of the Sun extends from the core up to about 70% of the solar radius. In this zone, the temperature drops form about 7×10^{6} to 2×10^{6} K. Under such conditions, hydrogen and helium atoms transfer the energy emerging from the core by absorption and reemission of photons. This thermal radiation is the main energy transfer process in the radiative zone. At the top of the radiative zone there is a thin layer (about 3% of the solar radius), called tachocline, where the rigid-body rotation regime of the core and radiative zone changes to the differential rotation of the solar convection zone.

- The convection zone is the outermost layer, which comprises about 30% of the solar radius, and extends from the tachocline up to the photosphere. In this layer, because of the lower temperature and pressure, some heavy elements are not longer in an ionised state, which makes the radiative transfer less efficient. However, the relatively low gas density allows the thermal energy from the radiative zone to be transported by convection. The plasma is heated near the tachocline, which makes it expand and become less dense. Further up, it floats upwards, reaching the photosphere where the temperature drops to about 5800 K. As the hot material looses its heat, it becomes more dense and sinks back to the tachocline. This convection process forms structures called thermal cells, which are flows of heated material, similar to Rayleigh-Benard convection cells. In the photosphere, these thermal cells appear as granules and super granules [Phillips 2008].

1.2.3 Solar magnetic field

The dynamic of the Sun is heavily influenced by the a strong magnetic field that penetrates both the solar atmosphere and its interior. This magnetic field is thought to be generated in the tachocline, where the solar plasma rotation is no longer uniform but differential [Phillips 2008]. Because of this differential rotation, charged

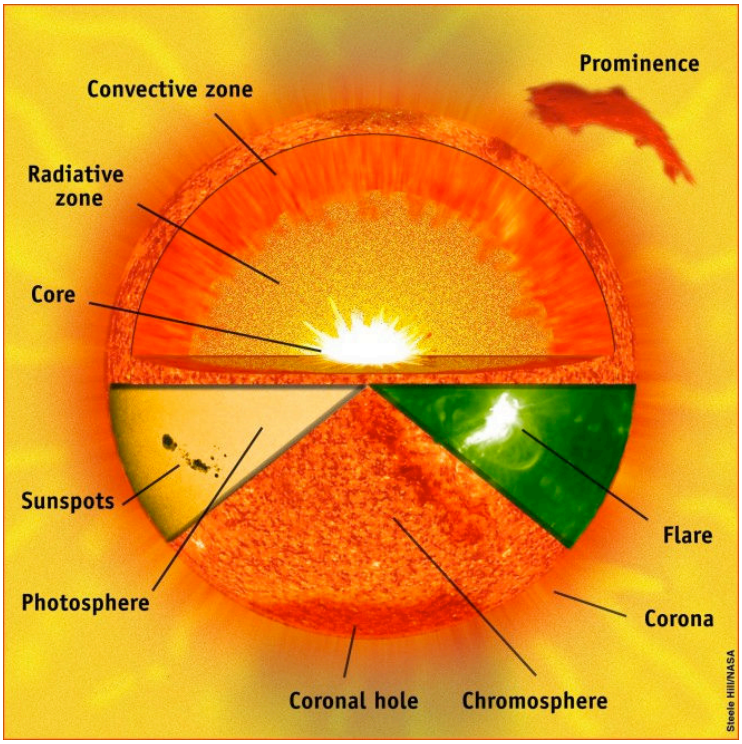

Figure 1.2: Schematic structure of the Sun. Source: NASA.

particles in different plasma layers start inducing a magnetic field through a process called solar dynamo. The solar magnetic field may then be decomposed into a poloidal and a toroidal component. The force lines of the poloidal component mostly lie in planes that include the solar rotation axis. On the contrary, the force lines of the toroidal component mostly lie in planes that are perpendicular to the axis. These two components change in time with the 11-year solar cycle, as it is illustrated in Figure 1.3.

During the period of low solar activity, when the photosphere is free of sunspots, the large scale magnetic field has mostly a bipolar structure, and the toroidal component is weak, as shown in Figure 1.3 a). Then, because of solar differential rotation the magnetic field lines carried by plasma start twisting along the direction of solar rotation (see Figure 1.3 b) and c)). Due to this twisting of the bipolar magnetic field, the toroidal component starts to develop. As this toroidal component increases, some magnetic field lines start to arise above the solar "surface", forming bipolar loops (see Figure 1.3 d)) that expand into the atmosphere. In the photosphere the footpoints of these loops are visible as groups of dark sunspots because the presence of the strong magnetic field blocks convection, leading to a temperature decrease (Figure 1.3 e)). Sunspots thus are a manifestation of an intense toroidal field component. The total sunspot area on the solar disk is largest during the maximum of the solar activity. The deformed magnetic loops start moving towards the solar

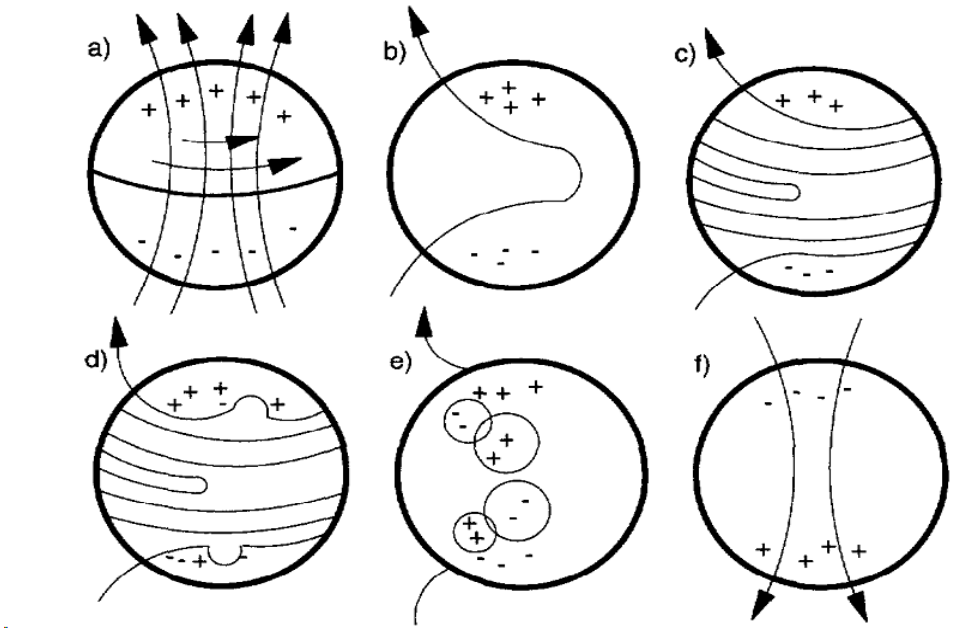

Figure 1.3: The cyclic evolution of the solar magnetic field with a period of about 11 years. Different stages of this evolution shown in panels a) - f) are described in the text below. Figure adapted from [Paterno 1998].

equator and reconnect into new, more stable loops. This leads to a decrease of the toroidal component. At the end of the reconnection process, the toroidal component disappears completely, and a new minimum of the solar cycle is reached. The newly closed magnetic field strength lines again form a bipolar magnetic field of opposite polarity. The period of this process is about 11 years, and consequently, it takes about 22 years for the solar magnetic field to reach the same polarity configuration. As we shall discuss later, this cyclic change in the solar magnetic field configuration is the main driver of the changes that affect the solar electromagnetic spectrum at different wavelengths.

1.2.4 Solar atmosphere

The first observations of the Sun were photospheric ones [Phillips 2008]. In this atmospheric layer the irradiance is dissociated from matter, in contrast to the inner (non-atmospheric) layers of the Sun. On the contrary, inside the Sun, the interaction between the plasma and the radiation it emits is complete, and the two are in equilibrium. In the solar atmosphere, radiation begins to escape into space, thus making it visible to different instruments.

The major part of the visible light is emitted in the photosphere, which makes other layers of the solar atmosphere look much dimmer. However, during total solar eclipses, when the photospheric light is blocked by the Moon, one may notice the

presence of a bright atmosphere extending far beyond the solar disk. This part of the Sun is called solar corona, and though it is barely noticeable in visible light, it is the brightest part at short wavelengths (EUV, see Figure 1.4 blue row), and is extremely hot compared to the lower layers. The vast region of space around the Sun (extending far beyond the orbit of Pluto) which is formed by the solar wind (a solar plasma that carries remnants of the solar magnetic field, and is constantly ejected from the solar atmosphere into space) is called the heliosphere.

In this work, we concentrate on the nearest layers of the solar atmosphere (from the photosphere up to corona), which are of main interest for describing the variability of the UV irradiance.

- The photosphere is the solar surface as seen in the visible part of the spectrum (see Figure 1.4 red row). It is a narrow layer, about 400 km thick, which emits about 99% of the total solar irradiance, mostly in the visible and infra-red. The solar spectrum can be remarkably well approximated be a black-body model with an effective temperature of 5777 K [Phillips 2008]. Spectroscopic analyses of the photosphere in the visible show that there is a local thermodynamic equilibrium (LTE), which means that the emitted radiation can be characterised by local values of temperature and density.

  As mentioned before, the photosphere exhibits convection cells. A typical cell is relatively short-lived (about 10 min) structure with a diameter of about 1000 km. However, in some cases, convection cells may merge, forming large structures (about 30 times larger than a regular granule) called super granules, and that may exist up to one day. The turbulent motion of plasma in the convection cells also produces oscillations with periods of the order of minutes. The dynamics of granules along with these oscillations cause changes in the SSI on short timescales.

  During solar minimum, the photosphere is homogeneous and granules are the only features present. As the solar cycle enters its rising phase, and the twisted magnetic field lines start to form magnetic loops, sunspots appear. Sunspots are regions of relatively low temperatures of about 3000 - 5500 K, with sizes ranging from 20 to 160×10^3 km. Inside of them, the convection motion is blocked by the strong magnetic field. Their lifetime varies from few days to few weeks. Typically, sunspots are pairwise structures with opposite magnetic polarity, and correspond to the photospheric footpoints of a magnetic loop where it exits and then enters again the convection zone. Because sunspots are cooler than the quiet photosphere, they are seen as dark structures in the majority of spectral bands, especially in the visible.

  In a sunspot, one may clearly distinguish two parts: the dark core, and a relatively less dark surrounding. The inner dark part, called sunspot umbra, is where the magnetic field lines are mostly normal to the solar surface, thus reducing convection heavily. In the less dark part, which is the sunspot penumbra, the magnetic field is more inclined, and the convection is less

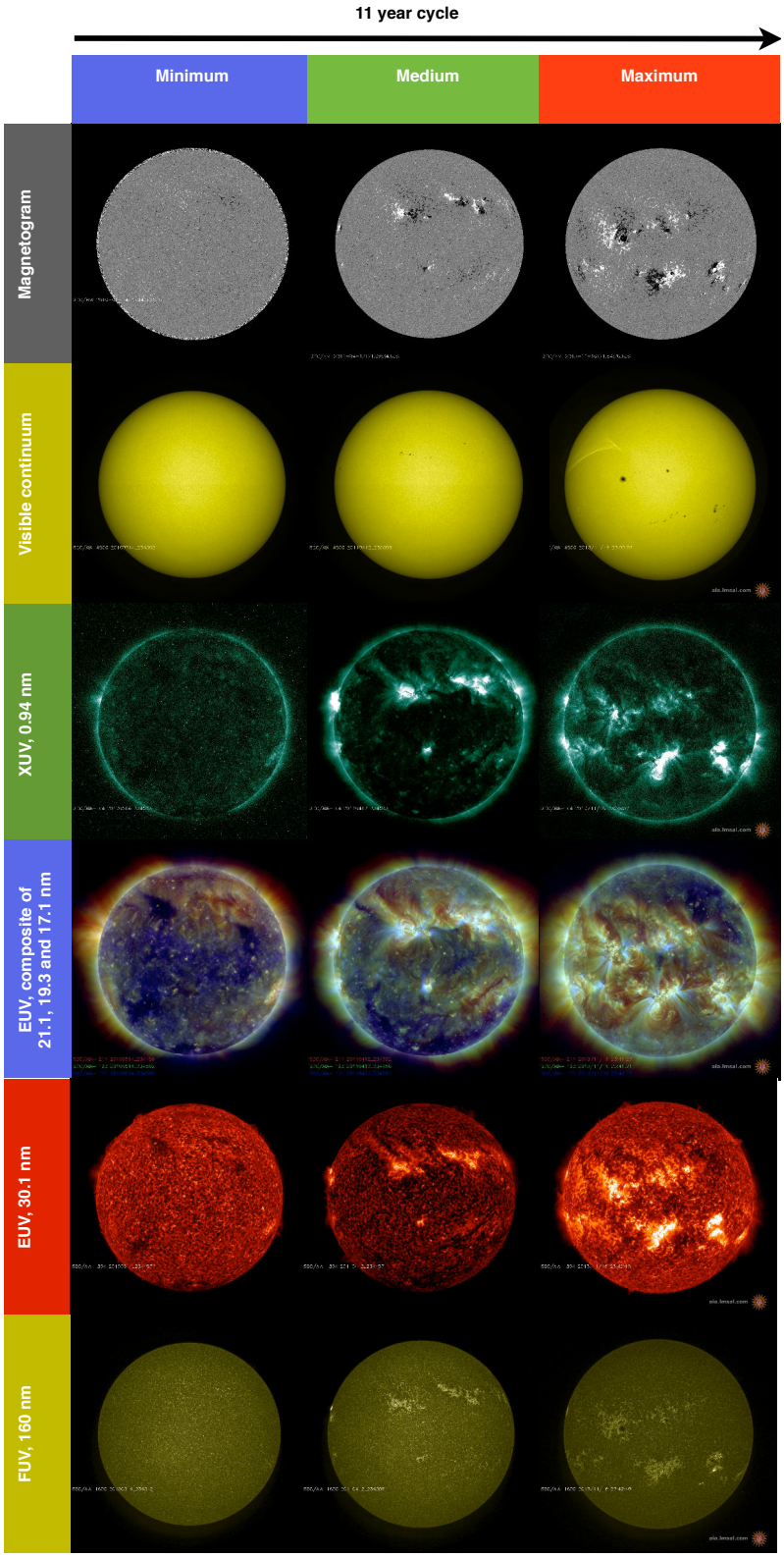

Figure 1.4: Changes in the solar atmosphere with the 11-year cyclic evolution of the magnetic field, presented here for cycle 24. Images are taken from SDO/HMI and SDO/AIA instruments on 2010/05/04 (minimum activity), 2011/04/12 (medium), and 2013/11/16 (maximum), each at 12:00 UT.

strongly reduced [Phillips 2008].

The evolution of sunspots over the 11-year solar cycle is traditionally presented by means of the "butterfly" diagram that is shown in Figure 1.5. As the solar cycle progresses, new sunspots emerge closer to the solar equator. During solar maximum the area coverage of sunspots is the largest; it subsequently decreases as the Sun enters the descending face. The maximum area sunspot coverage is cycle dependent, represent about 0.2 - 0.5 % of the surface of visible hemisphere (see Figure 1.5).

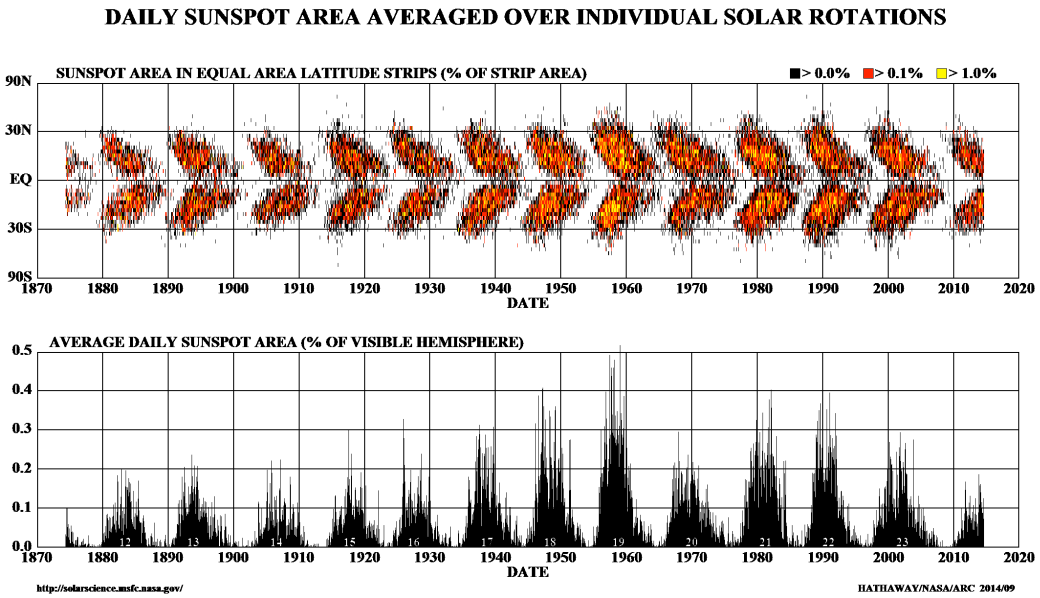

Figure 1.5: Butterfly diagram of sunspot area evolution, with the 11 year cycle. Source: http://solarscience.msfc.nasa.gov.

Another type of solar feature that is modulated by the 11-year cycle magnetic field evolution are faculae. These are small bright structures that consist of the strong magnetic flux tubes located between the edges of granules. In contrast to sunspots, faculae are brighter than the surrounding quiet photosphere. This is because the strong magnetic field makes the plasma captured in magnetic loop less dense and almost completely transparent, so that the hotter plasma located deeper in the convection zone becomes visible.

Over the 11-year cycle, sunspots and faculae are the main contributors to changes in the TSI but their effects on the visible part of the spectrum are opposite. Because both features have a very small coverage of the solar surface, the solar cycle variation of the TSI ends up being only about 0.1%. On average, facular brightening exceeds sunspot darkening, which enhances the TSI during solar maximum. As a consequence, most of the variability of the SSI in the visible and near infrared can be successfully reproduced by

combining solar proxies that reflect sunspot area and facular area. This is illustrated in Figure 1.6, in which changes in the TSI on the 11-year timescale are in phase with the MgII index [Viereck 2001] that describes facular brightening (top panel). On the contrary, on shorter timescales (i.e. solar rotation), changes in the TSI are in antiphase with the sunspot area (bottom panel). Several empirical models of the TSI are based on this property. In the UV, as we shall see later, the situation is more complex.

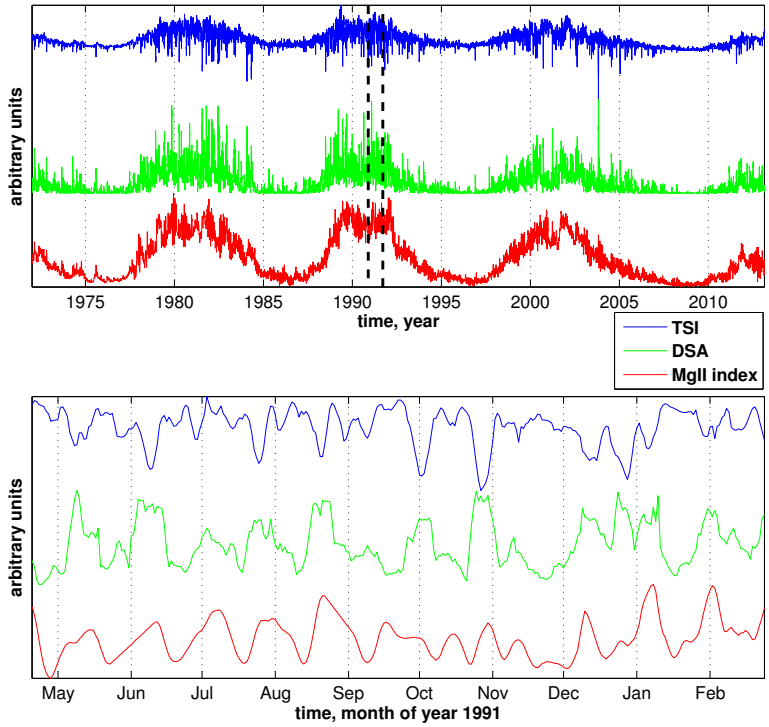

Figure 1.6: PMOD composite of the TSI versus the daily sunspot area (DSA) and MgII index. Top panel: 11-year cycle variability. Bottom panel: solar rotation timescale variability for the year 1991.

- The chromosphere is located directly above the photosphere. It got this name because the very fist observation was made during solar eclipses when the chromosphere has reddish colour (because of the Hα emission). The chromosphere ranges from about 500 to 2000 km, with a temperature profile increasing with height (see Figure 1.7). Chromospheric material is much less dense than the photospheric (by factor of about 10^4), and receives heat from below. Among the emissions that mainly originate in the chromosphere, the most prominent in the UV are the hydrogen (Ly-α at 121.5 nm and Ly-β at 102.5 nm), and helium lines. There are also emissions from heavier elements such as calcium

(Ca II K line) and magnesium (Mg II K line, see Figure 1.7).

Starting from the chromosphere and above there is a significant departure from the LTE as the emitted irradiance interacts less and less with the plasma.

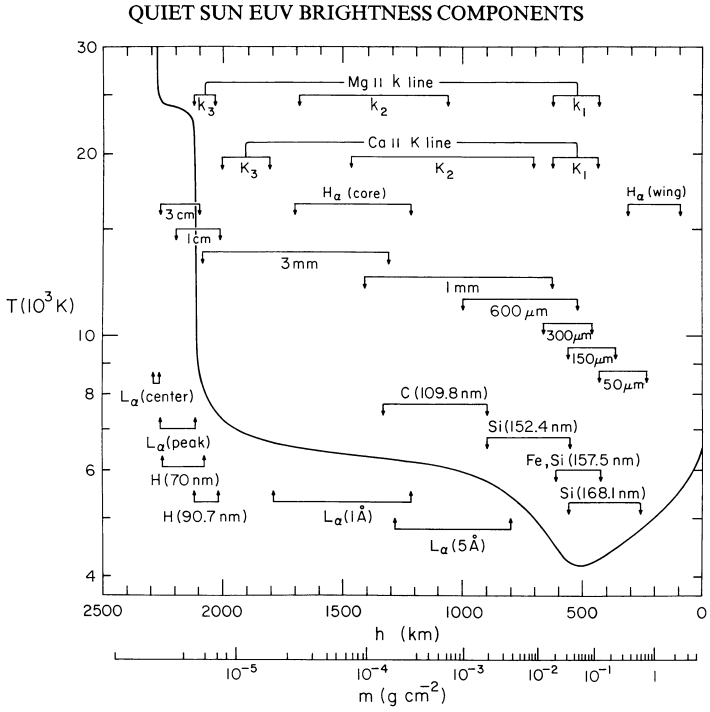

Figure 1.7: Distribution of temperature in the solar atmosphere as a function of height. Figure adapted from [Paterno 1998].

Solar UV images of the chromosphere reveal the presence of regions with enhanced brightness, covering the whole solar disk, called network. The darker part of the chromosphere that separates individual network elements is called the internetwork. The appearance of the chromospheric network is related to the heat received from the photospheric convection granulas. Large compact regions of increased intensity that develop as the 11-year solar cycle progresses are called plages. These plages are formed exactly where the emerging magnetic field loop penetrate the chromosphere. Both chromospheric plages and network appear as bright structures at wavelengths shortward of about 300 nm. Their evolution through the 11-year solar cycle is the prime driver of the SSI variability in the UV on timescales of days and above. For that reason, UV spectral reconstruction models that rely on solar images, require images taken in spectral bands that capture chromospheric features.

- The transition region is a very thin layer of about 100 km, which is located between the chromosphere and the corona. In this region, the temperature

increases extremely rapidly, rising from 60 000 to about 1.8 $\times 10^6$ K. This tremendous temperature increase in atmospheric layers that are so far away from the main heat sources is today one of the outstanding open questions in solar physics [Phillips 2008].

In the transition region, emissions are dominated by the medium-weight ions such as carbon (C IV), oxygen (O IV), silicon (Si IV). Hydrogen emissions are weak, as this element is fully ionised. All these emissions appear at short wavelengths below 121 nm.

- The solar corona is the outermost and hottest layer of the solar atmosphere, with temperatures of about 2×10^6 K, and electron densities that are about four orders of magnitude below that of the chromosphere [Aschwanden 2004]. At such high temperatures, even heavy elements such as iron are ionised many times (e.g. Fe XXIV), and most emissions occur at very short wavelengths (in the XUV and EUV). The corona is penetrated by magnetic loops that trap heated plasma, and whose number and extent evolve with the 11-year cycle. Although all loops are extremely hot, their temperature distribution is not homogeneous; the hottest loops are brighter in the soft X-rays whereas the cooler ones mostly emit in the EUV. Regions in which the magnetic loops are completely absent look dark, and consist of relatively cooler and less dense plasma called coronal holes. Coronal holes are thus associated with unipolar magnetic field regions that open up into space. During solar minimum, when the solar magnetic field configuration is mostly bipolar, coronal holes are located mostly near solar poles. On the contrary, when the 11-year cycle reaches the maximum of its activity, coronal holes can be found all over the solar disk.

  The most important sources of electromagnetic emissions in the corona are ionised heavy elements (emitting in the XUV/EUV) and particles accelerated by magnetic reconnection, emitting in centimetric and millimetric radio bands. Incidentally, because these different processes are intimately connected, radio emissions in the centimetric range end up being strongly correlated with the EUV flux [Tapping 2013], which explains the frequent use of the 10.7 cm radio flux as a proxy for the EUV flux.

  Magnetic reconnection is most active during periods of high solar activity. These are fast events occurring on timescales from milliseconds to minutes. In a reconnection event, the local configuration of one or several magnetic loops gets unstable, which leads to the subsequent reconnection to achieve a more stable state. The energy accumulated in strained loops is released swiftly. Part of the released energy pushes the solar plasma together with the remnants of the magnetic field upwards, away from the solar surface. This phenomenon is at the origin of coronal mass ejections (CMEs). The other part of the magnetic loop shrinks rapidly moving towards the solar surface. Accelerated particles (in particular electrons) as they hit the denser and cooler

chromospheric material produce hard X-ray emissions (bremsstrahlung). During such reconnection events, XUV and EUV emissions typically increase by several orders of magnitude. These are called flares, which are the main cause of variability in the UV bands on timescales below one hour.

CMEs are massive releases of the solar corona plasma into space. Similarly to flares, a CME is a phenomenon that is caused by magnetic field reconnection. The mass of ejected material is typically about 1.6 $\times 10^{12}$ kg and the velocity ranges up to 3200 km/s. The frequency of CME events is directly related to the level of solar magnetic activity, and the average number of CMEs is about 3.5 per day during periods of high activity.

Both flares and CMEs are extremely difficult to predict, making the modelling of the SSI on the timescales shorter then an hour very challenging. As a consequence, flare irradiance models cannot rely on the characteristics of the magnetic field only, and need a conceptually different approach [Chamberlin 2007].

1.2.5 Formation of the UV spectrum

From the preceding description, we can already conclude that various solar processes can lead to the emission of UV radiation, and that the latter contain a mix of line emissions and continua involving atoms and ions. Most of the UV irradiance is emitted by high-temperature regions that are located in the upper atmosphere, above the solar photosphere. These emissions mostly stem from recombination and ionisation processes of ions in the solar plasma. These processes, however, take place in different layers of the solar atmosphere, with differing temperature, pressure and chemical composition. As a consequence, the emissions vary with different layers of the solar atmosphere.

Recombination processes bring ion X^{+m+1} (i.e. ionised m + 1 times) to its ground state X^{+m}. The opposite process, that puts ions out of its ground state X^{+m} to the state X^{+m+1} is called ionisation. Since the emission from a solar plasma with locally constant physical conditions and chemical composition does not change over time the processes of recombination and ionizations are in equilibrium. Recombination and ionizations processes occur through interaction of an ion with other ions, free electrons or photons (external radiation), and can be classified pairwise as following [Phillips 2008]:

Ionization:

- Photoionisation
- Collisional ionisation
- Excitation-autoionisation

Recombination:

- Radiative recombination
- Three-body recombination
- Dielectronic recombination

Photo-ionisation and radiative recombination are both due to the interaction of an ion with an external radiation field. Collisional ionisation and three-body recom-

bination are caused by collisions between a target ion and free electrons. Finally, the excitation-autoionisation and dielectronic recombination involve excited state electrons.

The most significant processes in the solar atmosphere are radiative recombination, dielectronic recombination and collisional ionisation. In order to calculate the plasma emission over the whole spectrum one needs to incorporate all ions that are present in it along with their ionisation and recombination rate coefficients, which are temperature-dependent. The emergent irradiance intensity is then described by:

$$I = \frac{1}{4\pi} \int G(T, N_e)\xi(T)dT, \tag{1.1}$$

where $G(T, N_e)$ is the contribution function, and $\xi(T)$ is the differential emission measure (DEM). The function $G(T, N_e)$ describes all physical parameters of the electronic transitions. The second term $\xi(T)$ describes the physical parameters of the plasma (electron density, temperature).

One of the main challenges in the theoretical modelling of the SSI stems from the fact that the produced radiation has to cross the solar atmosphere before it reaches a remote observer (at the Earth). To describe the interaction between the electromagnetic radiation and the plasma of the solar atmosphere, a radiative transfer equation is required, which accounts for both emission and absorption processes.

$$\frac{dI_\lambda}{ds} = -\alpha_\lambda I_\lambda + j_\lambda, \tag{1.2}$$

where I_λ is the spectral irradiance; s is the length of the path followed by the electromagnetic beam through the plasma of the solar atmosphere; j_λ is the emission coefficient, which describes the increase of the irradiance intensity at specific wavelength λ due to the emission by the plasma being traversed; α_λ is the absorption coefficient, which is the fraction of the electromagnetic beam intensity that is lost due to absorption.

The solution of this equation allows to calculate the resulting intensity at a given wavelength that will be captured by an external observer. An important quantity that emerges in the solution of the radiative transfer equation is the optical depth $\tau_\lambda(S)$, which is an integral of the plasma absorption coefficient along the electromagnetic beam direction:

$$\tau_\lambda(s) = \int \alpha_\lambda(s')ds'. \tag{1.3}$$

This quantity describes the absorption of the light as it passes through the solar atmosphere to an observer, and is wavelength-dependent.

There are two important cases that are characteristic for UV emissions. When the optical depth is close to zero then the associated emissions are called optically thin. In this case, the electromagnetic wave hardly interacts with the solar atmosphere, and so the intensity does not depend on the location of the source of emission on the solar disk. Moreover, if along the line-of-sight there are several different emis-

sion sources, then the resulting intensity will be the sum of their intensities. This explains why the solar limb is extremely bright in the XUV band and in most of the EUV band (see Figure 1.4) whose emissions are optically thin: the resulting radiation intensity in the limb area is then a sum of contributions from many sources (due to line-of-sight projection effects). In the case of a large optical depth, emissions heavily interact with the solar atmosphere and end up being absorbed and re-emitted many times. Most emissions occurring in the MUV and FUV bands are optically thick, and thus have lower intensities at the solar limb than at the centre. For that reason, no model aiming at reconstructing the spectral irradiance in the UV can properly work without taking into account the position of the source of the emissions on the solar disk.

1.2.6 Variability of the UV emissions

From the physical point of view, the distinction of different bands in the solar spectrum is largely artificial. However, having in mind space weather applications, our interest is mainly in the UV band, so let us from now on concentrate on the latter. Although, the time evolution of the SSI in the UV is largely coherent across wavelengths (see Figure 1.8) [Lean 1982], there are also important aspects that make it wavelength-dependent.

Changes in UV emissions occur at all observed timescales, from minutes to decades. However, as we have already seen, these changes are caused by different physical processes. The 11-year cycle variability is driven by the evolution of the global solar magnetic field. On timescales of about 3-4 solar rotations (100 days) the SSI evolution is mostly governed by the emergence and decay of active regions. The 27-day timescale variability is associated with solar rotation. On shorter timescales up to several hours the variability is driven by the dynamical evolution of active regions. On extremely short timescales of minutes and below, the main drivers of variability are flares, oscillations and the motion of the granulation.

For the purpose of space weather and climate studies (discussed below in 1.3) we shall focus on timescales of 24 h and beyond, and thus shall mostly consider daily averages. Transient events such as flares, and granulation are not discussed in this study. We refer to [Chamberlin 2007] for detailed studies of flares for space weather applications.

1.3 Impact of the UV irradiance on the Earth

The SSI in the visible and near-infrared bands is mostly absorbed in the troposphere, and by continents and oceans, where is produces direct heating. On the contrary, the UV band is mostly absorbed in the higher levels of the Earth's atmosphere (see Figure 1.1b). This absorption leads to different photolysis processes [Haigh 2007] [Gray 2010] including ionisation, dissociation, excitation and heating. Photoionisation is the most effective in the ionosphere above 150 km, and disappears completely in the mesosphere, below about 70 km. Variations of the solar radiation in the UV

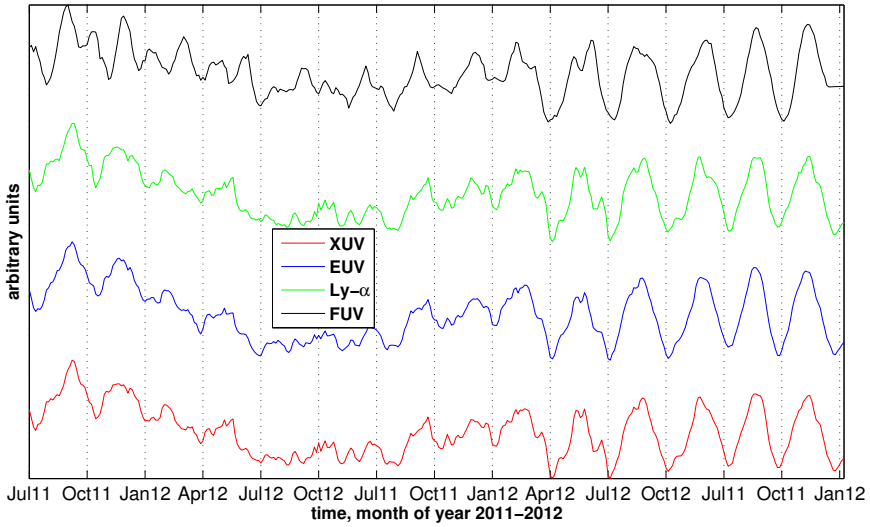

Figure 1.8: Time evolution of different UV bands, showing how coherent they are in time. The observations are from the TIMED/SEE and SORCE/SOLSTICE instruments, and are either band-integrated (XUV, EUV, FUV) or line-integrated (Ly-α).

between 120 and 300 nm cause changes in stratospheric ozone concentrations, and heating. In particular, the Herzberg continuum (185 - 242 nm) drives ozone production through photolysis. The Hartley band (200 - 300 nm) on the contrary causes photolytic destruction of ozone. A more detailed review of the physical processes in the Earth's atmosphere in response to SSI variations is given in [Schunk 2004].

Among the different consequences of the impact of the varying UV irradiance on the Earth's atmosphere, we may distinguish three principal ones. The first one involves density changes in the upper atmosphere (i.e. thermosphere), which are of major importance for satellite orbit control [Bruinsma 2012]. Here, in order to maintain a specific orbit trajectory or mitigate collisions with debris, one of the important aspects is satellite drag. The drag force depends on the density of the thermosphere, which is strongly modulated by solar XUV and EUV fluxes, and also by Joule heating (i.e. geomagnetic activity). As it is shown in Figure 1.9, during periods of high solar activity, changes in the thermosphere density may reach up to 1 order of magnitude for low Earth orbiting satellites (altitudes of about 400 km). The objective of satellite operators is the prediction of the position of a space objects (satellite or debris) with an accuracy of up to 20 km after 24 h. This is achieved by means of upper atmosphere models such as DTM2012 [Bruinsma 2012], and JB2006 [Bowman 2008]. All these models require the SSI in the UV as solar input, with a cadence of typically 24 hours. However, because of the lack of continuous and properly calibrated observations, most models rely on solar proxies such as the

radio flux at 10.7 cm [Tapping 2013].

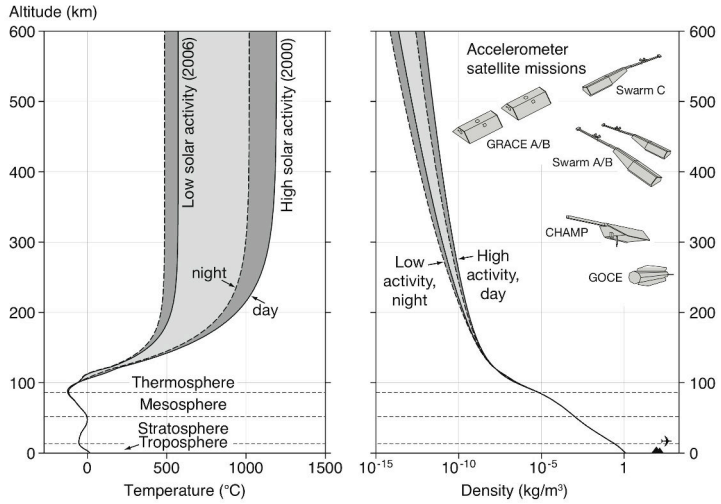

Figure 1.9: Temperature and density profiles of the Earth's upper atmosphere during minimum and maximum solar activity. Figure adapted from ESA.

Changes in the solar UV flux also impact ground-ground and ground-space telecommunications, and positioning. The propagation of electromagnetic waves through the ionosphere is affected by scattering, absorption, and Faraday rotation [Budden 1988]. The optical index depends on the electron density, and on the ionospheric composition, which, similarly to the thermosphere, are modulated by the solar XUV and EUV fluxes. However, the underlying physical processes somewhat differ: they mostly involve photoionisation, collisional ionisation, and frictional heating. One of their consequences is the emergence of small-scale disturbances such as blobs, patches and scintillations that may cause miss-positioning and perturbations, or even loss of communication. One of the most evident, and longest observed manifestations, however, is the strong solar cycle modulation of critical radio frequencies such as foF2, see Figure 1.10.

Finally, there is growing evidence that changes in the solar UV flux also affect climate. The average equilibrium temperature of the Earth is determined by the balance between the absorbed solar radiation and the loss of energy into space due to the thermal emissions. The characteristic altitudes at which the SSI is absorbed in the Earth's atmosphere are shown in Figure 1.1.

The SSI at wavelengths shorter than 121 nm is mostly absorbed at altitudes between 100 and 200 km by atomic and molecular oxygen and nitrogen that produce different ions. The Ly-α irradiance (121.5 nm) in the upper mesosphere makes a significant contribution to the heating rates and water vapor photolysis. Ozone in the middle-atmosphere modulates the outgoing infra-red irradiance from the Earth, thus eventually affecting the energy budget, and climate [Lean 1997a].

Modifications of the stratospheric temperature gradients also impact the propa-

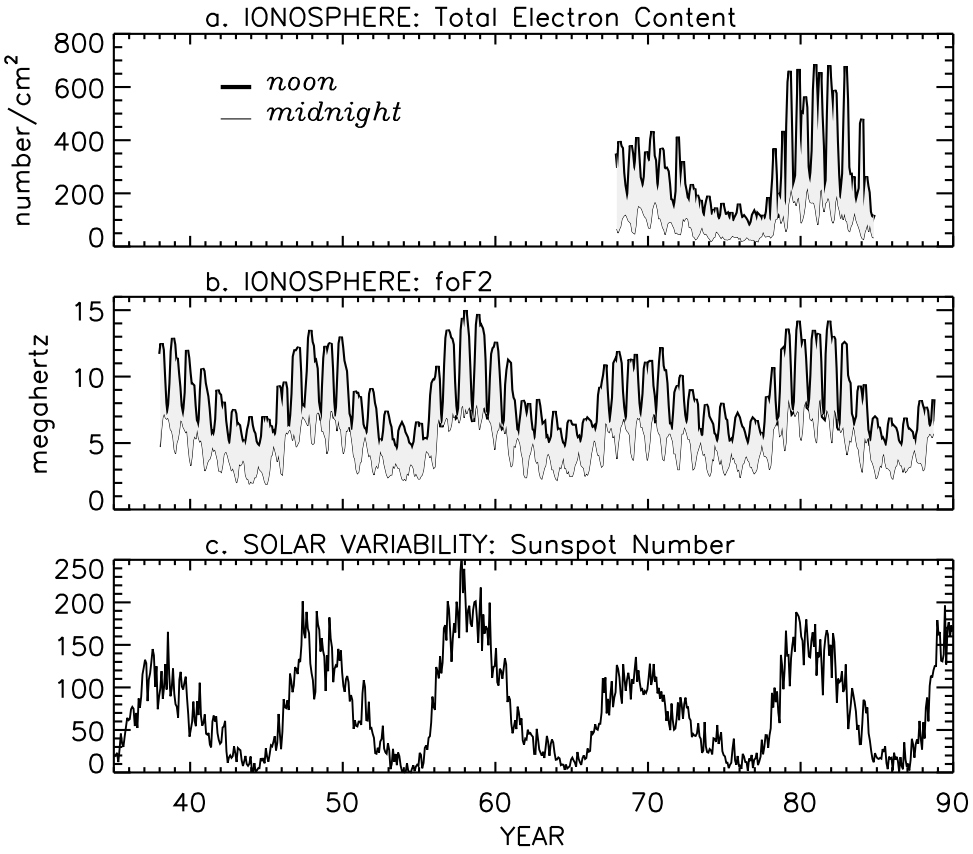

Figure 1.10: 11-year solar cycle modulates a significant changes in the ionosphere electron density (a) and the critical frequency foF2 (b), above which the ionosphere is unable to reflect radio waves back to the Earth. Figure adapted from [Lean 1997a].

gation of planetary waves, which leads to changes in the global atmospheric circulation pattern, and thus, climate too.

Currently, in climatology studies, the most advanced tools are the general circulation models that simulate the global circulation of atmosphere and ocean based on physical descriptions. These models in particular are used to reconstruct the Earth's climate in the past. There is now a growing need for specifying the solar forcing with spectral resolution, rather than with band-integrated spectra only [Ermolli 2013].

1.4 SSI observations in the UV

Because UV emissions are partially or totally absorbed in the Earth's atmosphere, it is not possible to properly measure them from the ground. The first SSI measurements were made by sounding rockets, starting in 1946. Naturally, such measurements were sporadic, and continuous observations started with the satellite era.

The very first satellite observations were made in 1962 in the frame of the Orbiting Solar Observatory (OSO I) programme by NASA. Since then, several space missions have been monitoring the UV. However, the spectral and time domain coverage of these observations has been highly variable. As a consequence, there are many gaps in the observations; there is a notable one in the EUV band, which ended only in 2002, with the launch of the TIMED mission [Woods 2005].

As shown in Figure 1.11, the full spectral coverage of the UV band was achieved for the first time with the launch of the SOlar Radiation and Climate Experiment (SORCE) in January 2003. The instruments from SORCE provide continuous measurements of the SSI in the UV from 0.1 to 400 nm, with a missing part between 35 and 115 nm [Snow 2010]. The Solar EUV Experiment (SEE) onboard the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) mission covers this part by providing measurements from 0.1 to 200 nm.

The major problem that face all SSI instruments in space is the degradation of their detectors by energetic radiation, and also by contamination. All instruments have a limited lifetime of about a decade, which makes the continuous monitoring of the SSI extremely challenging. Furthermore, because of this degradation, there is a need for constant refinement of the instrument calibration, which is rarely done in-flight. Moreover, degradation effects are strongly wavelength dependent. Because of this, the SSI measurements obtained by different instruments often disagree, and their long term stability remains an open question [Ermolli 2013].

The need to alleviate these problems has stimulated the development of different types of SSI models as complements or substitutes of these observations. This is one of the main motivations of this thesis too. The next chapter will address these models.

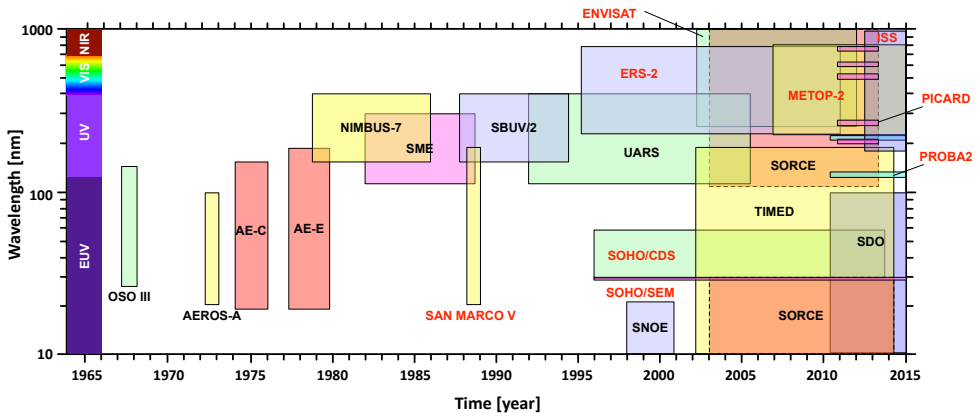

Figure 1.11: Spectral and time domain coverage of the available SSI measurements. Figure adapted from [Ermolli 2013].

1.5 Conclusions

Variations of the SSI are predominantly driven by the evolution of the surface magnetic field. The observed emission can be understood as the sum of emission coming from different atmospheric structures that are primarily driven by the solar atmospheric magnetic field.

The optical depth is an important property that characterises the observed intensity of these structures. There are two principal cases: for optically thick emissions, the same atmospheric structure may have different intensity when located at different distance from the solar disk centre. Structures that are located near the centre of the solar disk look brighter than the ones at the limb because the radiation has to penetrate a thicker layer of the absorbing atmosphere. For optically thin emissions, the same emitting region has comparable intensities regardless its position on the solar disk, as the emitted radiation does not interact with the solar atmosphere. Importantly, at the limb, several magnetic structures overlap the resulting radiation intensity is the sum of the intensities of the overlapping atmospheric structures, which makes the limb look relatively brighter. Because of this, structures that are located at the limb are important for describing optically thin emissions.

The variability of the SSI is timescale-dependent, and is driven by different physical processes. The 11-year cycle is driven by the global magnetic field evolution. Changes on the timescale of about 100 days occur due to the emergence and decay of the large magnetic regions; 27-day modulations are associated with the solar rotation. Changes on daily timescales are primarily due to the dynamical evolution of the magnetic structures. All these changes are thus well described by the photospheric magnetic field, which should therefore be a good candidate as an input to SSI models.

Short timescales (less than hours), are driven by granulation, p-mode oscillations and flares. These phenomena can not be described in terms of the surface magnetic filed, and require a conceptually different approach, which is outside of the scope of this thesis.

1.6 Résumé en Français

Le rayonnement électromagnétique solaire dans la bande ultra-violette (UV) (typiquement entre 1 et 300 nm) est une quantité-clé pour la spécification de la moyenne et haute atmosphère terrestre. Cette bande est d'un grand intérêt pour la météorologie de l'espace, et divers secteurs d'activité y ont recours: la prévision orbitographique, la communication sol-espace, et aussi le positionnement. Plus récemment, la bande UV a reçu un regain d'attention en rapport avec son rôle présumé dans le réchauffement climatique.

Bien que la puissance émise par le Soleil dans la bande UV ne représente que 8% de la puissance totale rayonnée, son impact sur l'environnement terrestre est loin d'être négligeable. En effet, son intensité est très variable, et des variations relatives de plus de 10 à 100 % sont courantes (voir Figure 1.1 c). Son intensité

varie sur des échelles de temps allant de la seconde à la décennie (et certainement au-delà). Pour cette raison, sa surveillance continue est d'un grand intérêt. On utilisera ici fréquemment le terme générique anglais SSI ("Solar Spectral Irradiance") pour désigner l'irradiance solaire résolue spectralement.

La majeure partie de l'irradiance solaire dans la bande UV est fortement absorbée dans la stratosphère et au-dessus, et ne peut donc être correctement mesurée à partir du sol. Les premières mesures remontent à 1946, avec le développement des fusées-sonde. Les premières mesures véritables d'irradiance coincident cependant avec le début de l'ère spatiale [Phillips 2008]. Il aura fallu attendre le lancement de la mission SORCE en 2003 pour avoir enfin accès au spectre solaire entier (des rayons X-mous à l'infra-rouge proche).

Or toutes ces observations souffrent aujourd'hui encore de nombreux défauts, à commencer par de nombreuses lacunes dans leur couverture spectrale et temporelle. En outre, les observations effectuées par différents instruments sont souvent en désaccord. La raison première en est la dégradation et la contamination des détecteurs, dont la durée de vie effective est limitée à une dizaine d'années. Un autre facteur handicapant est la quasi-absence d'étalonnage en vol.

Au cours des deux dernières décennies, divers modèles numériques de la SSI ont été développés dans le but de pallier ces problèmes. Ces modèles peuvent être répartis en deux catégories: ceux qui sont à base d'indices ou de traceurs d'activité solaires ("proxies" en anglais), et les modèles semi-empiriques. Les premiers sont basés sur la relation statistique qui lie la variabilité solaire dans différentes bandes spectrales et les divers indices d'activité solaire existants, tels que le nombre de taches solaires, ou encore l'indice MgII du magnésium. Les modèles semi-empiriques décomposent l'atmosphère solaire en un ensemble de structures dont les spectres d'émission caractéristiques sont obtenus partiellement ou totalement à partir de modèles physiques.

Ces deux classes de modèles parviennent à reproduire avec succès les principales caractéristiques de la variabilité de la SSI (aussi pour des bandes autres que l'UV). Cependant, ils ont aussi des défauts. Les modèles basés sur des indices n'ont pas forcément d'interprétation physique immédiate. De plus, ils sont conditionnés par la qualité des indices utilisés. Les modèles semi-empiriques ont une justification physique plus forte, mais peinent souvent à reproduire la variabilité de la SSI en-dessous de 200-300 nm, où l'hypothèse d'équilibre thermodynamique local ne s'applique plus. Il existe donc aujourd'hui un besoin important pour reconstruire la SSI dans la bande UV, et plus particulièrement dans l'Extrême-UV (EUV, 10 - 121 nm), qui est une bande essentielle pour la spécification du système thermosphère-ionosphère.

C'est dans ce contexte que nous avons décidé de développer un modèle empirique différent. A l'instar des modèles semi-empiriques, il recourt à l'évolution du champ magnétique photosphérique pour reconstruire la variabilité de la SSI. Il est donc comparable à la famille de modèles semi-empiriques actuels [Fligge 2000, Haberreiter 2008, Fontenla 2011], qui décomposent l'atmosphère solaire en un ensemble de structures atmosphériques auxquelles sont associés des profils d'émission.

En revanche, à l'inverse de ces modèles, le nôtre déduit ces profils d'émission des observations, au lieu de les imposer à partir de modèles théoriques. Comme nous le verrons, la performance de notre modèle est comparable à celle des modèles existants. Tirant profit du nombre réduit de contraintes théoriques, nous pouvons désormais aborder plusieurs questions intéressantes:

- Dans quelle mesure la variabilité de la SSI peut-elle être décrite par la décomposition du champ magnétique photosphérique en des structures différentes?
- Quelle résolution spatiale faut-il pour tenir correctement compte des effets de centre-bord (de la variation de la contribution spectrale entre le centre et le bord du disque solaire)?
- Combien de classes de structures différentes sont nécessaires pour décrire correctement la variabilité de la SSI, compte tenu de l'incertitude sur les observations?

Avant d'aborder la question de la modélisation de la SSI, décrivons rapidement les processus physiques responsables de la variabilité des émissions électromagnétiques dans la bande UV. Ces émissions proviennent de différentes couches de l'atmosphère solaire, pour lesquelles les variables thermodynamiques-clé sont la température et la densité locales du plasma. La source première de la variabilité solaire, et donc de la SSI est le magnétisme solaire. A l'échelle macroscopique, le champ magnétique solaire s'inverse périodiquement avec un cycle de 20-23 ans (cycle de Hale). Comme l'irradiance est essentiellement fonction de l'intensité du champ magnétique, et non du champ signé, c'est en réalité un cycle de 10-12 ans qui est observé (cycle de Schwabe) dans la SSI.

Pendant un cycle de Schwabe, les changements dans la topologie du champ magnétique solaire conduisent à l'apparition de différentes structures atmosphériques telles que les taches solaires, les plages et le réseau actif, qui vont moduler la SSI. Les changements d'intensité qui en résultent dans la bande UV varient de quelques % à plus de 100% pour les plus petites longueurs d'onde. Certaines des ces variations s'observent à des échelles de temps spécifiques. Notons en particulier la modulation de 27 jours (une période moyenne de rotation solaire), une constante de temps d'environ 100 jours (durée de vie moyenne de grandes régions magnétiquement actives), et le cycle de Schwabe susmentionné. Des variations s'observent en outre sur des durées plus courtes, notamment en raison de phénomènes de reconnexion magnétique, d'éruptions et d'oscillations acoustiques de la photosphère.

Les processus d'émissions dans l'UV varient en fonction de leur localisation dans l'atmosphère solaire. On est habituellement amené à distinguer ceux agissant dans de la couronne, dans la région de transition, dans la chromosphère et dans la photosphère. Même si le champ magnétique qui structure ces régions engendre une forte cohérence spatiale et temporelle, il existe aussi des différences notoires. C'est ainsi que la plupart des émissions coronales sont optiquement minces en raison de la très haute température de la couronne (de l'ordre de 10^6 K). D'autres émissions coronales

proviennent du plasma piégé dans les boucles magnétiques qui s'élèvent au-dessus de la photosphère. Lorsque ces boucles se situent au-dessus du disque, leur émission est intégrée selon la ligne de visée. Pour celles situées au-delà du limbe (bord du disque), une part substantielle de l'émission peut provenir de boucles dont les pieds ne sont pas visibles car situés sur la face cachée du disque. Dans les bandes EUV et XUV (rayons X-mous, 0.1 - 10 nm) cette contribution peut représenter jusqu'à 30 % de l'éclairement total du disque. Il est donc particulièrement important de connaître l'emplacement exact des boucles.

Les émissions issues des plus basses couches de l'atmosphère solaire (chromosphère et photosphère) sont pour la plupart optiquement épaisses. Ces émissions sont issues d'une région nettement moins chaude et plus mince que la couronne. Les structures caractéristiques de la chromosphère sont les plages et le réseau actif, qui correspondent à l'ancrage des boucles magnétiques coronales. Toutes ont donc une origine magnétique. A cause des effets de projection, celles situées plus près du bord solaire apparaissent relativement plus contrastées que celles situées au centre du disque solaire. Dans la photosphère, les structures les plus apparentes sont les taches solaires, dont le déficit d'émission est dû à une réduction locale du flux de chaleur provenant de l'intérieur du Soleil; cette baisse de la conductivité thermique est liée à une concentration locale du champ magnétique. Nous en déduisons que la variation des émissions chromosphériques et photosphériques (qui s'observent essentiellement dans les bandes FUV 122 - 200 nm et MUV 200 - 300 nm) est principalement dictée par la structuration magnétique au niveau de la photosphère. Ce fort lien entre magnétisme photosphérique et variabilité de la SSI sert actuellement de base à une famille de modèles de la SSI, que nous aborderons dans le chapitre suivant.

Ce rayonnement UV (pour des longueurs d'onde inférieures à 300 nm) est principalement absorbé dans la haute et moyenne atmosphère terrestre, via divers processus physico-chimiques tels que la photolyse, la dissociation, l'excitation, etc. En particulier, l'EUV est la principale source d'ionisation de l'ionosphère; tout changement dans son intensité se répercute donc instantanément dans les caractéristiques de la haute et moyenne atmosphère. Nous avons déjà évoqué plus haut les conséquences sociétales, qui font du suivi de la bande UV une des priorités en météorologie de l'espace.

Chapter 2

Solar spectral irradiance models

Contents

2.1 Proxy-based models 27
2.2 Semi-empirical models 29
2.2.1 SATIRE-S 30
2.2.2 SRPM 33
2.2.3 COSI 37
2.2.4 SOLMOD 37
2.3 Physical models 37
2.4 Conclusions 39
2.5 Résumé en Français 40

The preceding chapter revealed the need for having models that can overcome some of the limitations of SSI observations by reconstructing the SSI and its variations by alternative means. However, because of the large number of different physical processes involved, and the varying spatio-temporal distribution of the emitting regions, the design of such models may seem like a hopeless task.

Present SSI models describe irradiance changes by means of changes in the area coverage of solar atmospheric structures at the solar disk that have different brightness, or, more precisely, emission spectra (sunspots, faculae, network etc.). These structures are formed because of the interaction between the solar surface magnetic field with solar atmosphere, and have different physical characteristics. Their temporal evolution is driven by the evolution of the magnetic field and can be extracted from solar images, or described globally by different disk-integrated proxies of solar activity. Therefore, two different types of solar inputs can be used to derive SSI models: images and disk-integrated proxies. Because emissions at different wavelengths are formed in different layers of the solar atmosphere, with different corresponding atmospheric structures, SSI models use specific solar images and proxies for each spectral band. For example, coronal images (such as SOHO/EIT, SDO/AIA) are generally used to reconstruct the EUV/XUV band, whereas in the FUV/MUV band chromospheric images (e.g. CaK line) or magnetograms (e.g. SOHO/MDI, or SDO/HMI) are used.

The brightness of the atmospheric structures at different wavelengths can be either calculated based on a physical model, or derived empirically from observations. In this sense, SSI models can be classified according to the amount of the physical

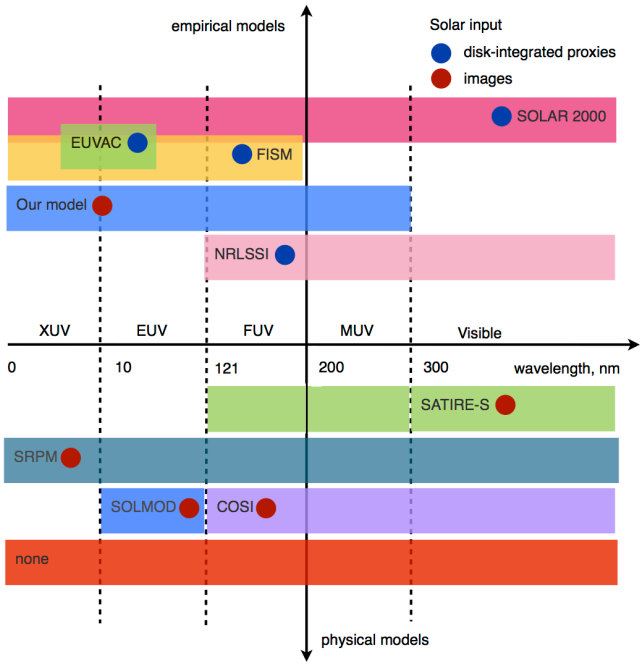

Figure 2.1: Spectral coverage and the amount of physical information of some of the routinely used SSI models.

information that is incorporated in the calculation of spectra of the atmospheric structures.

Several different SSI models that are classified according to the amount of the physical information, and the type of solar input, are shown in Figure 2.1, along with their spectral coverage. Most physical models (bottom part) such as COSI [Haberreiter 2008] calculate spectra of atmospheric structures accounting for large number of ions and molecules, and do not assume LTE. The SATIRE-S model [Fligge 2000] in this sense is more restrictive because it calculates spectra of atmospheric structures under LTE assumption, which limits it spectral coverage down to about 300 nm. Irradiance at shorter wavelengths (down to 115 nm) is modelled by rescaling the model outcome to match SORCE/SOLSTICE spectra (115-180 nm) and the Whole Heliospheric Interval (WHI) reference spectra (180-300 nm). These models are also commonly referred to as semi-empirical because they combine physics-based spectra with the empirical connection relating them to solar proxies. All of these models use solar images because the spatially resolved information about the atmospheric structures is needed to account for the centre-to-limb variation effects.

In the top part of Figure 2.1 we see models that use less physical information and

in this sense are more empirical. The NRLSSI model [Lean 1997b, Lean 2005], although it uses physical models for calculating spectra of atmospheric structures, uses disk-integrated proxies to estimate the area coverage of atmospheric structures, and consequently, cannot accurately reproduce some of the centre-to-limb variation effects. Other models, such as SOLAR2000 [Tobiska 2000] or EUVAC [Richards 1994] do not use physical models at all to calculate the spectra, and merely establish a linear connection between the spectral variability and some proxies such as the F10.7 index. Most of these proxies are disk-integrated quantities.

Let us now briefly review the main characteristics of some of these models.

2.1 Proxy-based models

Proxy-based models use the statistical relation between observed variations in the SSI at different wavelengths, and different proxies of solar activity to reconstruct the former. Their large success owes to the remarkably strong correlation between the two [Schmidtke 1976, Dudok de Wit 2009].

As discussed in Chapter 1, variations in the SSI are driven by processes that take place at different layers of the solar atmosphere. For that reason, a set of different proxies is needed to properly capture the dynamics of these processes and provide inputs to the SSI models.

The proxies that are most frequently used for the UV are presented in Table 2.1.

Table 2.1: Proxies of solar activity for the UV band.

Name	Description	spectral band	observable since
ISN	sunspot number		1608
F10.7	radio flux at 10.7 cm	EUV/XUV	1947
F30	radio flux at 30 cm	EUV/XUV	1957
Mg II	the core-to-wing ratio of the Mg II line at 280 nm	FUV	1978
CaK	the mormalised intensity of the Ca II K-line at 393 nm	FUV/MUV	1974
MPSI	magnetic plage strength index (the relative fraction of the solar surface that is covered by magnetic fields with $10 \leq \|B\| \leq 100$ G)	FUV/MUV	1968
MWSI	sunspot index, defined as the MPSI, but for intense magnetic field ($\|B\| \geq 100$ G)	EUV	1968

Most of these proxies benefit from long time records, which makes them attractive for the SSI reconstruction purposes. However, since each one tends to capture a particular process, or describe a specific layer, no single one can reconstruct irradiance variability in the UV at all timescales. The performance of different indices at different timescales was studied by [Dudok de Wit 2009], and is summarised in Figure 2.2.

Here, the distance between different proxies and spectral bands in the UV corresponds to the degree of linear dependency between them (the closer to points are to each other the closer is the corresponding linear scaling coefficient to 1). We notice that indeed, different proxies have different performance in different UV bands. Moreover, their performance varies at different timescales.

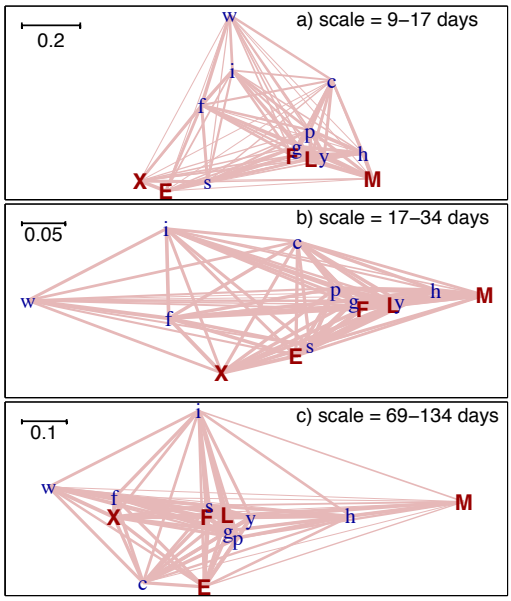

Figure 2.2: Performance of different indices of solar activity with respect to different UV bands at three charscteristic timescales. Five spectral bends are coded with the capital letters: (X)UV, (E)UV, HI (L)yman-α, (F)UV and (M)UV. Solar indices are coded as following: (i)SN, (f) 10.7, M(g)II, (c)aK, M(p)SI, M(w)SI, L(y)aman-α channel and (h)erzberg channel. Picture adapted from [Dudok de Wit 2009].

The Mg II index reproduces solar rotation variability of the FUV emissions and Ly-α line quite well on all timescales, however it diverges with the MUV and EUV/XUV variability. The ISN is the worst index to reconstruct the UV irradiance variations at all timescales [Dudok de Wit 2009]. Naturally, proxies that are derived directly from the SSI measurements (such as Ly-α and Mg II) are the best ones to describe their corresponding spectral bands.

Typically, proxy-based EUV models incorporate the F10.7 as a proxy for the coronal emissions and MgII for describing the chromosphere, and most have two-timescales that are separated by a cutoff period of 81 days (three solar rotations).

The two timescales allow to fit observational data better and indicate different scaling of the SSI variability with proxies at short and long timescales. A simplistic equation of such a model is the following:

$$I(\lambda, t) = S_0 + (F_{10.7}(t) - F^*_{10.7}(t)) \cdot S_1(\lambda) + F^*_{10.7}(t) \cdot S_2(\lambda), \tag{2.1}$$

where $I(\lambda, t)$ is the estimated spectral irradiance, $F^*_{10.7}(t)$ is the value of the $F_{10.7}$ index averaged over 90 days (the long timescale). S_0 is the quiet Sun spectrum, and S_1, S_2 are the short and long timescale scaling coefficients. All three parameters (S_0, S_1 and S_2) can be estimated through a linear regression. However, in some cases reference spectra are obtained directly from observations. Examples of proxy-based EUV models are: SERF1 [Hinteregger 1981], EUVAC [Richards 1994], EUV97 [Tobiska 1998].

In the FUV spectral region, where the SSI originates in the solar chromosphere and photosphere, proxy based models typically use the MgII and CaK proxies to model facular activity. Typical examples are the models: NRLSSI by [Lean 1982], [Cook 1980, Worden 1996].

More recently, the SOLAR2000 model [Tobiska 2000] combined a set of band-specific models to reconstruct the whole solar spectrum. In the UV band the model reconstructs individually XUV, EUV, FUV and UV(MUV) bands using different proxy models. In particular the SSI reconstruction in the XUV band is partly based on F10.7 coronal index. The EUV band has a contribution from both solar corona and chromosphere, thus F10.7 and Ly-α indices are used. The FUV and MUV bands are reconstructed using chromospheric Ly-α and MgII, and photospheric MPSI proxies.

The major limitation for the majority of the proxy-based models is that they are trained to reproduce the SSI observations, and consequently rely heavily on the quality of the data. Overall, most proxy based model offer a good quality of the SSI reconstructions no the 11-year cycle timescales, however, the fine details of the short timescale (27-day solar rotation and shorter) variability can not be described without taking into account the centre-to-limb variation of different atmospheric structures, which is not always possible with the use of proxies as they are disk-integrated quantities. Incorporation of the spatially resolved inputs of solar activity is used in the next generation of the SSI models which are the semi-empirical models.

2.2 Semi-empirical models

This class of models considers contributions to the SSI from individual atmospheric structures such as sunspots, plages, network, etc. separately. The SSI evolution is described as a superposition of contributions from all atmospheric structures that cover the solar disk, weighted with their area coverage. The number and types of structures depend on the model. Different types of atmospheric structures have different characteristic emission spectra. The area coverage is the fraction of the solar disk area covered by the atmospheric structures of the same type; they are

typically estimated from intensity images and magnetograms. As the radiation at different wavelengths originates at different heights in the solar atmosphere, images taken at different wavelengths are needed to reconstruct a specific UV band. Thus, the bright atmospheric structures in FUV and MUV bands (plages, network) can be extracted from images that capture the solar photosphere and chromosphere, such as images taken in the CaK line. Also, plage regions can be extracted from solar magnetograms (from ground-based KP/NSO or space-based SOHO/HMI, and SDO/MDI instruments). In the EUV and XUV bands, where the radiation is mostly emitted from the coronal loops, i.e. from three-dimensional structures that expand high in the solar atmosphere, the identification of individual atmospheric structures is more complex, and requires images that capture the solar corona. The emission intensity profiles (spectra) of the atmospheric structures are calculated using semi-empirical models that make use of atmospheric radiative transfer codes.

Thus the solar atmosphere is approximated as a two-dimensional surface which consists of the set of different vertical atmospheric structures, which are distributed over the plane surface. The empirical atmospheric structures are defined to correspond to different atmospheric structures that are identified observationally [Fontenla 1999]. For each type of structure, the temperature and the density are given as functions of height, and typically cover the photosphere and transition region (which leads to inapplicability in the EUV/XUV bands).

Starting from the empirical temperature and density profiles, radiative transfer codes provide a spectrum of each atmospheric structure. As we already discussed in Chapter 1, the radiative transfer code describes the interaction of the radiation with the solar atmosphere. There are three types of the interactions: emission, absorption, and scattering. Radiance emission and absorption come from different atoms and ions in the solar atmospheric plasma. These radiance processes can be classified as bound - bound, bound - free and free-free processes, depending on the initial and the final state of electrons. Changes in the state of atoms (as well as ionization and recombination) occur due to collisional processes (collisional excitation and de-excitation) with other atoms, electrons or protons. Radiative transfer codes account for all these radiative and collisional processes to calculate the resulting emergent spectrum. Details about the radiative transfer codes are presented in [Clough 2005].

2.2.1 SATIRE-S

SATIRE-S (Spectral And Total Irradiance REconstructions for the Satellite era) is an often-cited semi-empirical model [Fligge 2000, Krivova 2005a, Krivova 2005b]. This model has three atmospheric components: quiet Sun, faculae and sunspots. Their corresponding spectra are the following: the quiet Sun spectrum is calculated with the Kurucz (1991) ATLAS9 code, and has an effective temperature of 5777 K; a similar model is used for sunspots (umbra and penumbra), and a modified model P by [Fontenla 1999] has been used for faculae. The ATLAS9 code uses an LTE approximation, which means that the model is likely to provide less good estimates

for the SSI variability below 300 nm. The code considers the solar photosphere and chromosphere only, which means that it is inappropriate for the EUV/XUV emissions originating from the corona. The atmospheric structures are identified in magnetograms and continuum images. Magnetograms from the SDO/HMI (since May 2010), SOHO/MDI (1999 - 2010), and Kitt Peak Vacuum Telescope (KPVT) (1974-2003) instruments are employed to identify faculae, which partly covers solar cycles 22 and 24.

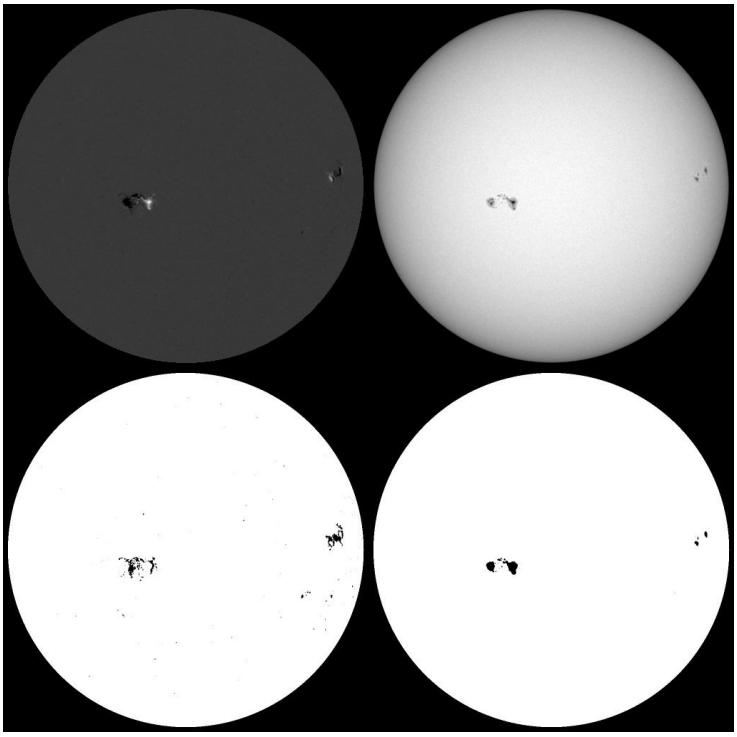

Figure 2.3: Example of sunspot (bottom right) and facular regions (bottom left) extracted on 24 November 1996. The corresponding magnetogram (top left) and intensity image (top right) are taken by SOHO/MDI instrument. Image adapted from [Fligge 2000].

The spectral irradiance $I(\lambda, t)$ at wavelength λ and time t is calculated as the sum of contributions from all solar disk elements (pixels) (i, j). The contribution of each individual solar disk element $I_{i,j}(\lambda, t)$ depends on the type of the associated atmospheric component, and its position (to account for the centre-to-limb variation).

$$\begin{aligned} I(\lambda,t) &= \sum_{i,j} I_{i,j}(\lambda,t), \text{ where} \\ I_{i,j}(\lambda,t) &= (1-\alpha^s_{i,j}(\Phi,t)-\alpha^f_{i,j}(\Phi,t))\cdot I^q(\mu(i,j),\lambda) \\ &+\alpha^s_{i,j}(\Phi,t)\cdot I^s(\mu(i,j),\lambda)+\alpha^f_{i,j}(\Phi,t)\cdot I^f(\mu(i,j),\lambda). \end{aligned} \tag{2.2}$$

Here $I^s(\mu(i,j),\lambda)$, $I^f(\mu(i,j),\lambda)$, and $I^q(\mu(i,j),\lambda)$ are spectra of sunspots, faculae and the quiet Sun with respect to the cosine of the corresponding heliocentric angle $\mu(i,j)$. The angles $\mu(i,j)$ are binned into 11 values from $\mu = 1$ (Sun disk centre) to $\mu = 0.05$ (limb). The solar disk area where $\mu \geq 0.05$ is not included into the model because of the extremely high noise level in magnetograms, which is caused by the projection effects and the line-of-sight integration. $\alpha^s_{i,j}(\Phi,t)$, $\alpha^f_{i,j}(\Phi,t)$, and $\alpha^q_{i,j}(\Phi,t)$ are the corresponding filling factors. The filling factors are introduced to account for the finite resolution of the magnetograms and allow to describe the sub-pixel size elements. Φ is the threshold value that is used to separate facular regions form the quiet Sun.

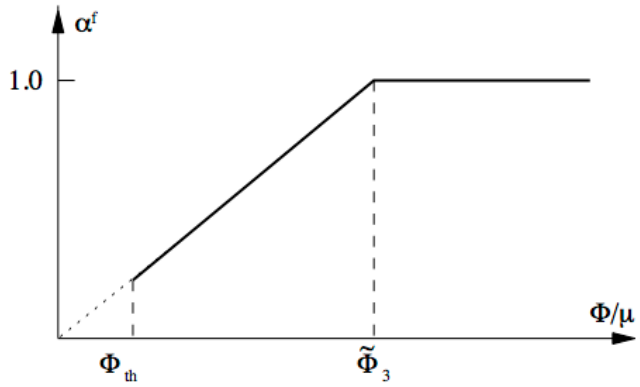

Figure 2.4: Facular filling factor versus the magnetic flux intensity measured within a single magnetogram pixel. Φ_{th} is the threshold value which separates facular regions from the surrounding quiet Sun. Image adapted from [Fligge 2000].

This model offers several degrees of freedom. In particular the spectra of atmospheric structures can be refined, and the filling factors can be adjusted to match better the SSI observations. For the quiet Sun and sunspot solar disk elements the respective filling factors equal one. The facular filling factor $\alpha^f_{i,j}(\Phi,t)$ accounts for the small magnetic flux tubes whose size is less than one pixel. For such small structures the spectral irradiance intensity associated with such a pixel has to be reduced. In the SATIRE-S model this reduction coefficient (which is the filling factor) is assumed to scale linearly with the magnetic field intensity for those pixels for which the magnetic field intensity has values lower then a saturation level $\tilde{\Phi}_3$. This saturation level $\tilde{\Phi}_3$ is adjusted to achieve better agreement with the SSI observations, and have values of about 75 - 300 G, depending on the instrument which measures the magnetograms.

The SATIRE-S model provides SSI reconstruction from the UV up to the IR band. However, at the shortest wavelengths, typically below 300 nm, the model fails to reproduce the observed variation in the SSI due to the LTE approximation, which is used in the calculation of the spectra of atmospheric components. The ro-

tational variability of the UV is found to be in good agreement with observations by SORCE/SOLSTICE and SORCE/SIM. On the longer timescales (more than a few solar rotations), the reconstructions (along with other models and UARS/SUSIM observations) depart more strongly from the SORCE/SOLSTICE and SORCE/SIM observations [Unruh 2012] (see Figure 2.5). This raises a question: is this discrepancy due to a problem with the long-term stability of the SORCE/SOLSTICE and SIM instruments or because of the model assumptions? One possible explanation may be insufficient number of contributing atmospheric components or an incomplete description of their spectra. Another possibility is that the basic assumption which implies that the surface magnetic field evolution is the only driver of the SSI variability is wrong.

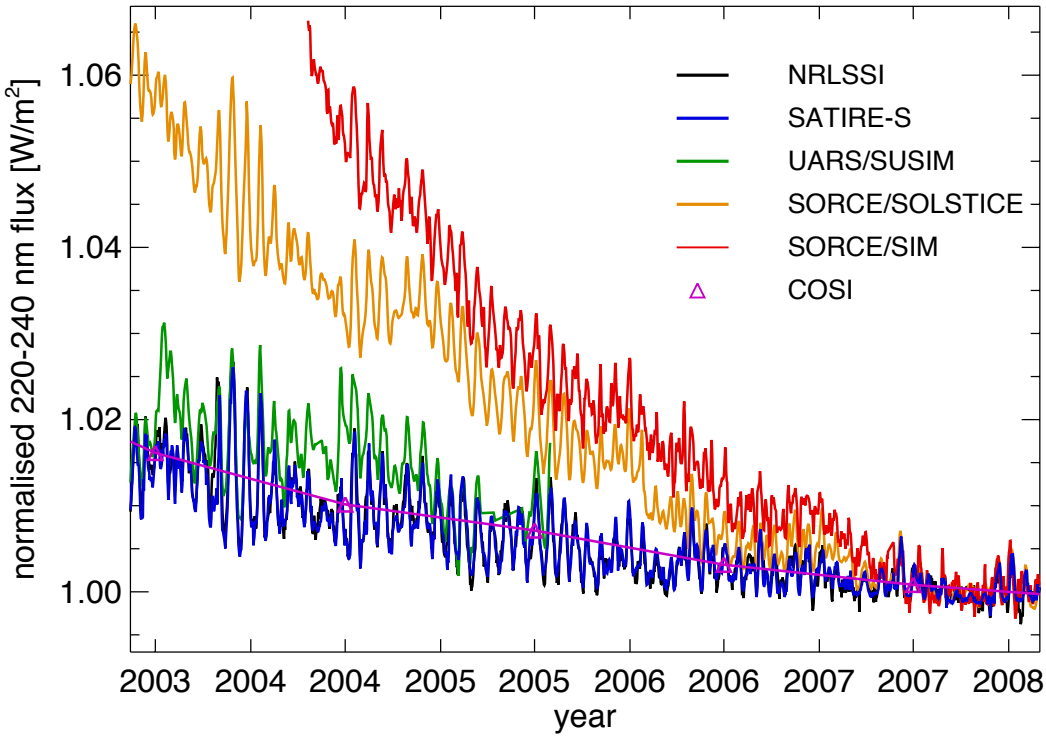

Figure 2.5: Discrepancy between the long-term SSI variability in the UV (integrated over the 220-240 nm band) and three models: NRLSSI, SATIRE-S, and COSI. Observations are from the SORCE/SOLSTICE, SORCE/SIM, and UARS/SUSIM instruments. Image adapted from [Ermolli 2013]

2.2.2 SRPM

Solar Irradiance Physical Modelling (SRPM) is a model that provides high-resolution SSI reconstructions. A detailed description of the model is provided in [Fontenla 2011].

Here we discuss its main aspects, that are relevant for the modelling of the UV band. The spectral coverage of the model runs from 0.12 nm to 100μm, thus covering the entire UV band.

The SRPM model decomposes the solar atmosphere into a set of empirical components that are assumed to have fine spatial and temporal structure. In contrast to the early versions of atmospheric models, the ones used in the SRPM include high layers of solar atmosphere, including the solar corona. This important aspect makes the SRPM model favorable for reconstructing the UV band. Moreover, the calculations used to derive the associated spectra do not assume LTE.

The full set of the atmospheric models that are incorporated in the SRPM framework is as follow:

- A - Dark quiet-Sun inter-network
- B - Quiet-Sun inter-network
- D - Quiet-Sun network lane
- F - Enhanced network
- H - Plage
- P - Facula (bright plage)
- S - Sunspot umbra
- R - Sunspot penumbra
- Q - Hot facula

Spectra are derived from the atmospheric models and are reviewed in details in [Fontenla 2009, Fontenla 2006, Fontenla 2007]. Currently, however, only seven atmospheric structures are used for the SSI reconstruction: quiet Sun, quiet Sun network-lane, enhanced network, plage, faculae (bright plage), sunspot umbra and sunspot penumbra. For the SSI modelling the SRPM model uses the spatial distribution of the atmospheric structures derived form full-disk filtergrams provided by the Precision Solar Photometric Telescopes (PSPT), and solar images from the Catania Astrophysical Observatory. Examples of segmentation maps for different levels of solar activity are shown in Figure 2.6.

The modelled rotational variability is found to be in good agreement with the observed one (see Figure 2.7). However, the observed variability (from SORCE/SIM and SORCE/SOLSTICE instruments) on the longer timescales is significantly larger in the FUV and especially in the EUV bands (see Figure 2.5). The possible reason for this mismatch might be some missing atmospheric components [Fontenla 2011].

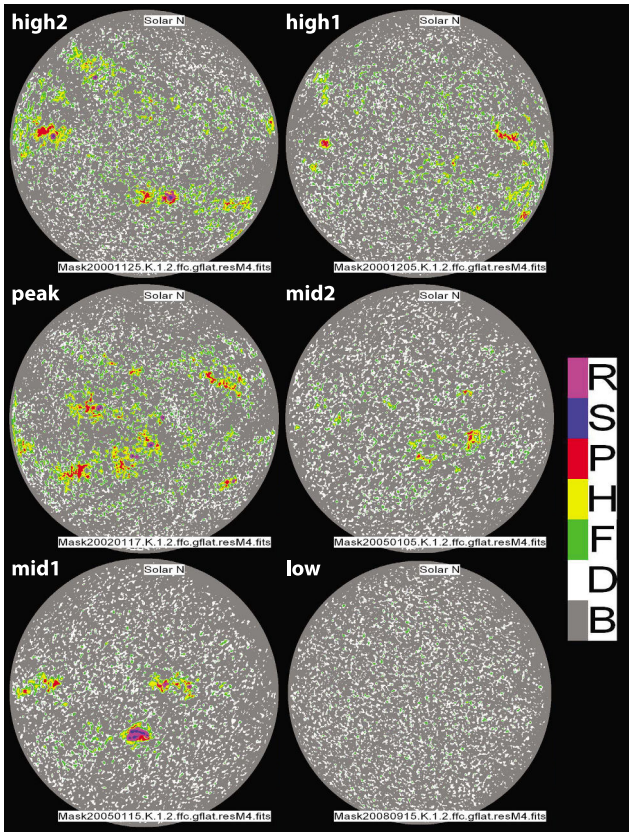

Figure 2.6: Segmentation map of chromospheric features for several days that correspond to different levels of solar activity during solar cycle 23. Figure adapted from [Fontenla 2011].

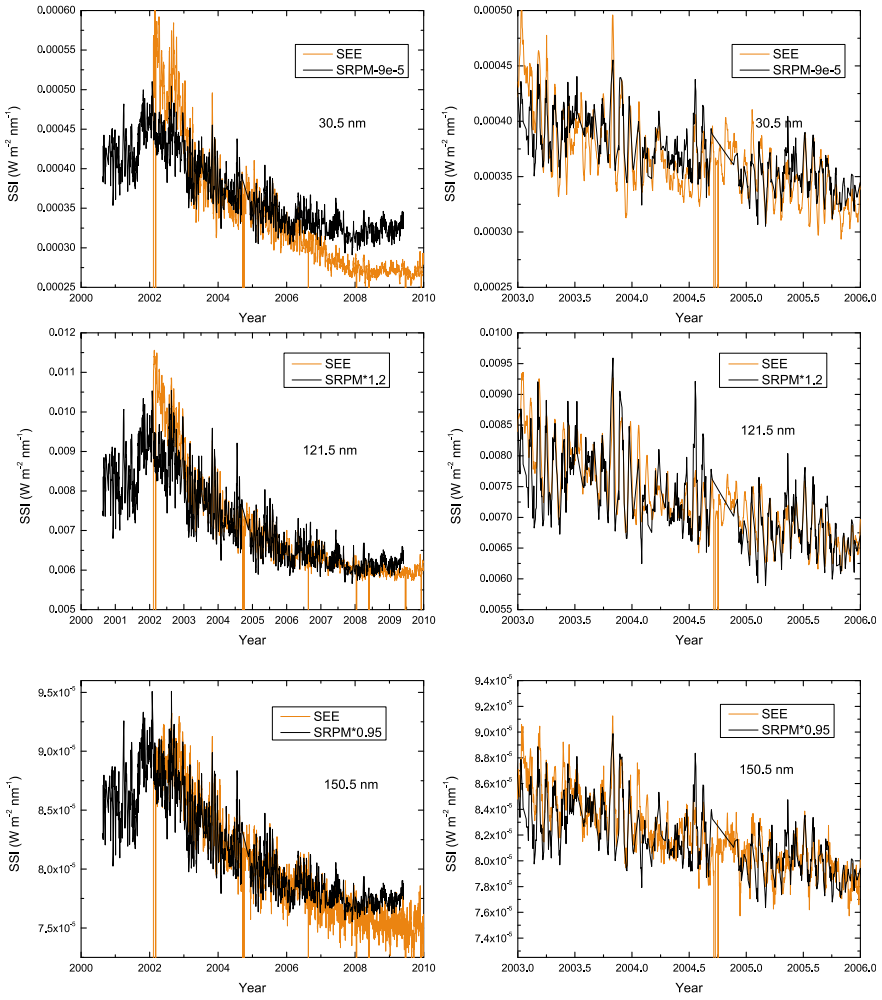

Figure 2.7: Comparison of the SSI reconstruction obtained by the SRPM model and SORCE/SIM observations in the UV. The solar rotation variability (left panels) agrees between both modeled and observed SSI. On the longer timescales there is a significant discrepancy which is the largest in the EUV band. Figure adapted from [Fontenla 2011].

2.2.3 COSI

COde for Solar Irradiance (COSI) stands out as one of the very few models that calculate the spectra of different atmospheric components under non-LTE conditions. NLTE effects are indirectly taken into account by using the NLTE opacity distribution functions [Hubeny 1995], which were incorporated in COSI by [Haberreiter 2008]. The quiet Sun spectrum obtained with COSI is in good agreement with both SORCE/SOLSTICE and SORCE/TIM observations down to 120 nm during the period of solar minimum 2008 [Haberreiter 2008].

The SSI variability can be derived using spectra obtained with the COSI code, weighting them with the corresponding area coverage of the atmospheric structures, similarly to the approach applied in SATIRE-S. Also, similarly to SATIRE-S, COSI decomposes the solar atmosphere into 3 components by using full-disk magnetograms: quiet Sun, faculae and sunspots. The facular saturation magnetic field intensity value B_{sat} is adjusted to get the better agreement with the observations.

2.2.4 SOLMOD

The Solar Modeling (SOLMOD) code [Haberreiter 2011] produces synthetic spectra of five atmospheric components of the solar corona: quiet corona, coronal quiet network , coronal active network, active regions, and bright active regions. One of the key features of this code is that it calculates the emergent spectrum in spherical symmetry, which allows the emission beyond the solar limb to be calculated [Thuillier 2011]. This is important for reconstructing the optically thin emissions in the EUV/XUV, which have a strong contribution from the limb and off-limb regions due to the line-of-sight integration effects.

For the SSI reconstruction the distribution of the atmospheric structures (coronal structures) is derived from SOHO/EIT coronal images taken at 17.1 nm, and 19.5 nm. The segmentation of the coronal images extends up to 1.4 solar radii, and is based on a fuzzy-logic algorithm described in [Verbeeck 2014].

The reconstructed SSI covers the UV band from 10 to 100 nm, and is in good agreement with SOHO/SEM measurements for the time period of about one solar cycle (see Figure 2.8). The question here, however, is the link to the photospheric magnetic field.

2.3 Physical models

Today no single physical model is able yet to properly reconstruct the SSI variability over the full solar spectrum, and in particular in the UV band. Such a model would require a full description of all physical processes involved in the formation of the SSI, including both the evolution of the solar magnetic field and its interaction with the solar 3D atmosphere. Because of our incomplete understanding of the physical processes (e.g. chromospheric and coronal heating mechanisms, magnetic field emergence activity) and high computational complexity there is no reason to

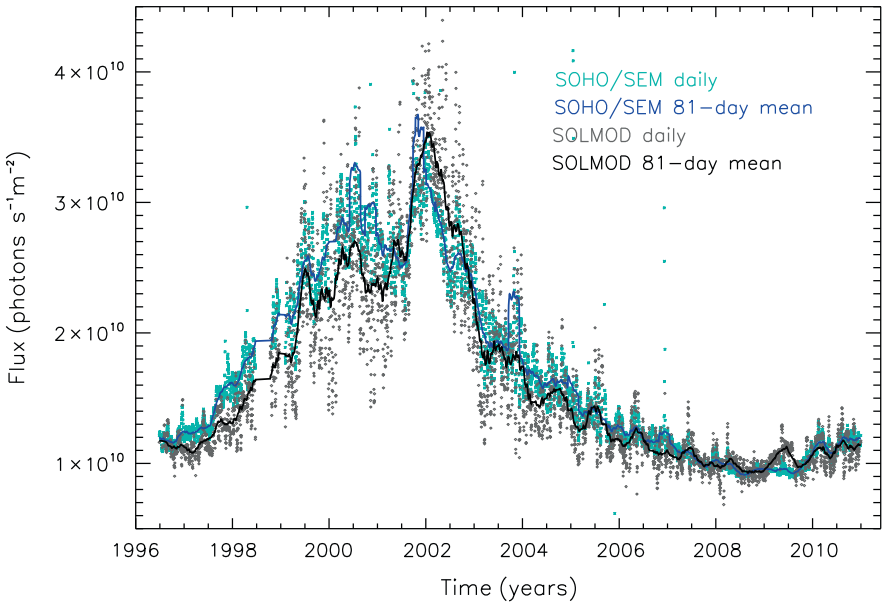

Figure 2.8: Comparison of the SOLMOD flux reconstruction (plotted in gray) with the SOHO/SEM measurements (plotted in blue) between 26 and 34 nm. Figure adapted from [Haberreiter 2014].

expect an operational physical model to be developed in the near future. Thus, for the time being semi-empirical models should be considered as the best tool available for the SSI reconstructions. However, 3D MHD codes that describe the solar magnetic field evolution are under constant improvement e.g. [Vögler 2005], and high-resolution 3D simulations of the solar atmosphere are also becoming available [Socas-Navarro 2011].

2.4 Conclusions

Many different models have been developed for reproducing SSI variations in the UV. All are based on the premise that SSI changes are driven by the evolution of different bright atmospheric structures (active regions, faculae, plages, network etc.) on the "visible" solar disk. These structures are in turn governed by the solar magnetic field that permeates the solar atmosphere. These SSI models can be matched to one of two principal classes: proxy-based models, and semi-empirical models. Physical models are still beyond reach.

Most proxy-based models use different proxies of solar magnetic activity such as the F10.7 index, the MgII core-to-wing ratio, etc., that globally describe the evolution of specific atmospheric structures. The variability of the SSI is fitted through a multiple regression of these proxies. These models are conceptually the simplest ones, and are still widely used for historical reasons. The availability of long time records of proxies makes these proxy-based models capable of covering long time periods. In particular, they are highly relevant for making reconstructions before the space era.

However, the physical connection between these different solar proxies and the SSI evolution is often at best indirect, and so the relation between the two is primarily statistical. Proxies are disk-integrated quantities, and represent activity at different layers of the solar atmosphere, whose connection with the SSI variability can be very indirect (e.g. for the solar radio flux).

For this reason, there are concerns about the ability of proxy models to reproduce small details in the SSI, and also to properly capture secular variations, especially at a time where we are witnessing an unusually weak cycle 24. Moreover, no single proxy can reproduce the SSI variability across the whole UV spectrum.

Most semi-empirical models use empirical inputs either by proxies or segmented solar images (or magnetograms) in combination with physical solar atmosphere models to reconstruct the SSI. They typically decompose solar images into a set of empirical atmospheric structures with characteristic temperature and density profiles. The SSI is the sum of the emissions that come from all the structures that cover the solar disk weighted with their area coverage. Spectral profiles of the atmospheric structures express their contribution to the spectral irradiance at a given wavelength as function of distance from the Sun disk centre. The semi-empirical models offer better accuracy on short timescales (solar rotation and less), and some of them (SRPM) offer extremely high spectral resolution. A question here is the

number of the empirical atmospheric structures that have to be used for the solar images decomposition. This number varies from 3 to 7 depending on the model. Large numbers of atmospheric structures may give a good SSI reconstruction but may make model less stable and more sensitive to the noise in the solar images and SSI time series that are used for the model calibration. Some semi-empirical models do not perform well in the EUV/XUV band, which is of importance for the space weather purposes, because of the departure from LTE.

Disagreements on the long timescales between SSI observations from SORCE/SOLSTICE and semi-empirical models may indicate that the properties (temperature and density), and, consequently the intensity spectral profiles of the atmospheric structures change over time. It is not fully understood yet if the area coverage of atmospheric structures properly capture the long-term evolution of the SSI. Thus, the reliability of long-term SSI reconstructions remains an open question.

Because of this, we decided to develop a data-driven model that derives intensity spectral profiles of different magnetic structures from observations, in this way alleviating some of the issues associated with the incomplete physical knowledge. Moreover, in our model we want number of classes of the magnetic structures to be inferred from the observations, thus making the model more data-adaptive.

2.5 Résumé en Français

Le chapitre précédent a révélé la nécessité de disposer des modèles capables de reconstruire la SSI afin de pallier les limitations des observations, qui sont souvent lacunaires. Or le grand nombre de processus physiques impliqués dans l'émission UV, et l'extrême complexité spatiale et temporelle des régions émettrices pourraient d'emblée nous faire croire que cet exercice est voué à l'échec.

Les modèles actuels de la SSI cherchent tous à reproduire la variabilité de l'irradiance à partir de la superposition linéaire de contributions qui décrivent partiellement ou entièrement la contribution individuelle des différentes structures de l'atmosphère solaire (taches, facules, plages, réseau actif, boucles, etc.) à cette variation. Toutes ces structures sont in fine liées à l'interaction du champ magnétique photosphérique avec l'atmosphère solaire, d'où l'idée de partir du premier pour chercher à reproduire cette variabilité de la SSI.

On rencontre ici deux catégories de modèles: ceux qui se basent sur des images 2D évoluant au cours du temps (prises dans l'UV, dans le visible, ou encore, des magnétogrammes) et ceux qui recourent à des indices, qui des séries temporelles 1D. Notons bien que ces derniers ne contiennent pas d'information spatiale puisqu'ils sont intégrés sur le disque solaire.

Les images prises dans des longueurs d'onde différentes décrivent des structures qui sont majoritairement localisées dans des couches atmosphériques spécifiques. Il est donc nécessaire de recourir à plusieurs longueurs d'onde convenablement choisies pour décrire une partie du spectre solaire. Par exemple, les structures observées dans les images EUV de SOHO/EIT et SDO/AIA sont généralement utilisées pour

reconstruire la bande EUV/XUV, alors que pour la bande FUV/MUV on aura davantage recours à des images chromosphériques, comme par exemple la raie du Ca K. Pour la photosphère, on aura essentiellement recours à des images prises dans la bande visible du spectre.

Une autre approche consiste à mesurer le champ magnétique photosphérique, dont nous avons déjà vu que sa connaissance permet de reconstruire la contribution associée à la plupart des structures solaires. Cette contribution peut être estimée soit à partir de modèles semi-empiriques, soit dérivée de manière plus empirique à partir des observations. En ce sens, les modèles de la SSI peuvent être classés en fonction de la quantité d'information physique qui est incorporée dans le calcul des spectres caractéristiques des structures atmosphériques.

Certains modèles semi-empiriques tels que COSI [Haberreiter 2008] calculent les spectres des structures atmosphériques en incorporant un grand nombre d'ions et de molécules, et n'imposent pas d'équilibre thermodynamique local. Le modèle SATIRE-S [Fligge 2000] requiert un équilibre thermodynamique local, ce qui limite sa couverture spectrale à environ 300 nm. Ces modèles sont dits semi-empiriques parce qu'ils ont des spectres issus de modèles théoriques d'une part, et un lien empirique qui les lie aux traceurs d'activité solaire (le magnétisme dans le cas de SATIRE-S) d'autre part. Tous ces modèles ont besoin de l'information spatiale contenue dans les images pour décrire correctement les variations intensité associées au défilement des structures du centre au bord du disque solaire.

Les modèles dits empiriques s'appuient davantage sur des indices d'activité solaire et des spectres majoritairement issus des observations. Le modèle NRLSSI [Lean 1997b, Lean 2005], par exemple, se base uniquement sur des indices solaires pour reproduire la variabilité du spectre, et par conséquent parvient moins bien à décrire les variations de type centre-bord. Néanmoins, sa performance est excellente, et ce modèle est un des plus utilisés. D'autres modèles, tels que SOLAR2000 [Tobiska 2000] ou EUVAC [Richards 1994] se bornent à établir un lien statistique entre la variabilité spectrale et des indices comme le flux radio à 10.7 cm (ou indice F10.7).

Les indices solaires et les modèles qui en découlent ont le grand avantage de reposer sur des longues séries d'observations, dont certaines remontent au XIXe siècle, et même au-delà. Ces modèles sont donc intéressants pour reconstruire la SSI sur de longues durées. Toutefois, comme le lien entre la SSI et les indices est essentiellement statistique, et parfois dénué d'interprétation physique directe, une grande prudence est de mise lorsqu'aucune observation directe ne permet de les étayer. En particulier, les variations à long terme ne sont pas forcément reproduites de manière correcte par ces indices. La récent cycle solaire 24, et le minimum anormalement long qui l'a précédé en 2007-2010, nous ont montré que des indices pourtant fortement corrélés pouvaient subir des évolutions légèrement différentes, avec des conséquences sur notre interprétation de la variation à long terme de la SSI. Par ailleurs, aucun indice ne peut reproduire à lui seul la variabilité de la SSI dans l'intégralité de la bande UV.

Les modèles semi-empiriques offrent en comparaison de meilleures performances:

leur précision est meilleure sur les variations de l'ordre du jour au mois. Certains d'entre eux, tels que SRPM, offrent en outre une très haute résolution spectrale. Leur principal défaut est le recours aux images solaires et aux magnétogrammes, ce qui les rend quasiment inutilisables seulement avant 1980.

Dans tous ces modèles semi-empiriques se pose la question du nombre de structures atmosphériques nécessaires pour reproduire correctement la variabilité du spectre UV. En d'autres termes: combien de degrés de liberté faut-il pour la reproduire? Selon les modèles, ce nombre de structures varie de 3 à 7. La présence d'un grand nombre des structures peut donner l'illusion d'obtenir une bonne reconstruction de la SSI; or le modèle sera alors moins stable et plus sensible aux défauts omniprésents dans les observations. Rappelons aussi que la plupart des modèles actuels ne permettent pas de descendre en dessous de 200-300 nm en raison des la non prise en compte des conditions hors équilibre thermodynamique local.

Les récents désaccords entre ces modèles d'une part et les observations de SORCE/SIM et de SORCE/SOLSTICE d'autre part ont ravivé le débat autour de la précision des modèles, et soulevé de nouvelles questions. Ainsi, comme le Soleil traverse actuellement une phase de faible activité, certains ont émis l'hypothèse selon laquelle la structure atmosphérique (et donc les profils d'émission sur lesquels sont basés ces modèles) aurait changé. La fiabilité des reconstructions à long terme de la SSI reste donc une question ouverte.

C'est dans ce contexte que nous avons décidé de développer un modèle différent et davantage piloté par les observations ("data driven"). Ce modèle s'apparente aux modèles semi-empiriques dans la mesure où il recourt lui aussi au champ magnétique photosphérique pour reconstruire la variabilité de la SSI. Il s'en différencie toutefois par la déduction des spectres caractéristiques des observations, et non de modèles. Nous espérons ainsi rendre ce modèles moins contraignant, et explorer des questions que les autres modèles ne permettent pas. L'objectif in fine est de poser un cadre pour un modèle simple mais fiable, qui puisse reconstruire en temps réel la variabilité dans l'UV.

Chapter 3

Empirical modelling

Contents

3.1 Data for training the model . . . 44
3.2 Long-term stability of the SSI observations . . . 45
3.3 Model description . . . 47
3.4 Extraction of magnetically active regions . . . 50
3.5 Segmentation by area . . . 53
3.6 Model selection . . . 55
3.6.1 Optimising the number of classes and annuli . . . 56
3.6.2 Optimising threshold levels . . . 57
3.6.3 Single UV model . . . 59
3.7 Contribution of the small and large magnetic structures to the SSI variability . . . 61
3.8 Centre-to-limb variation . . . 62
3.9 The off-limb contribution . . . 64
3.10 Two-timescale model . . . 67
3.11 Contribution of magnetic structures to the SSI variability at different timescales . . . 70
3.12 Conclusions . . . 71
3.13 Résumé en Français . . . 72

So far we have reviewed the main approaches for reconstructing the SSI. We saw that models combining information from segmented solar images with characteristic emission spectra associated with various classes often provide superior performance. Two important issues with such models, however, are whether they can easily be run in near-realtime, and whether they can be better tuned to the observations by inferring their characteristic spectra from the observations (instead of deriving them from atmospheric models). For that reason, we investigate here a different approach that relies on segmented solar images, and empirically derive characteristic spectra. In this chapter we describe our empirical approach and discuss the physical information that can be inferred from it.

Our working hypothesis, which is similar to models such as SATIRE-S, is that the evolution of the solar surface magnetic field is the primary driver of the changes in the SSI [Krivova 2003].

The principal difference of our work is that we do not assume beforehand what the relevant magnetic structures and their associated spectra are. Both are inferred from the observations. In particular, we want to retrieve the most compact set of atmospheric structures that allows to describe the observed SSI variability in the UV. And, in contrast to the majority of semi-empirical models, we determine empirically the spectra that are associated with these structures.

To extract the distribution of atmospheric structures on the solar disk we rely on solar magnetograms only. Some atmospheric structures, such as sunspots and coronal loops, cannot be accurately determined in magnetograms. As we shall see later (see 3.5) the performance of the model indeed does not significantly improve by including the area coverage of umbra and penumbra, as deduced from SDO/HMI continuum images. Coronal loops, which emit mostly in the XUV/EUV bands, are complex three-dimensional structures (in contrast to plages and network that have negligible height, and thus can be consider as 2D structures), and cannot be fully described by the two-dimensional magnetograms without extrapolating the magnetic field. Thus, with our model we also aim to check to what degree the photospheric magnetic field only can describe the variability of the SSI in the whole UV band.

In this chapter we describe in detail how our empirical model is designed, calibrated, and what its physical outcomes are.

3.1 Data for training the model

In this work, the two prime solar inputs are magnetograms and SSI measurements. Magnetograms are used to extract the distribution of magnetic structures, whereas the SSI measurements serve for model selection and calibration.

The magnetograms that we use are obtained by the Helioseismic and Magnetic Imager (HMI) instrument onboard the Solar Dynamics Observatory (SDO) mission [Scherrer 2012]. For our study we employ only the more reliable line-of-sight magnetograms. HMI observes the full solar disk with a resolution of 0.5 arcsec. The cadence of the observations is 45 seconds. However, for the reasons mentioned in the introduction, we use one observation per day only, which is the one nearest 12:00 UT.

The full-disk integrated SSI measurements, which serve for model calibration, are normalised to 1 AU. Here, we concentrate on the full ultraviolet irradiance range from 6.5 nm up to 300 nm, which includes soft X-rays, extreme ultraviolet, the Ly-α line, far ultraviolet, and medium ultraviolet. For this purpose the data from different instruments are merged into a single spectrum. We employ the SSI measurements from the Extreme ultraviolet Variability Experiment (EVE) [Woods 2012], which is part of SDO mission, the Solar EUV Experiment (SEE) of the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) mission [Woods 2005], and the SOlar Stellar Irradiance Comparison Experiment (SOLSTICE) of the Solar Radiation and Climate Experiment (SORCE) [Rottman 2005] (see Table 3.1).

The SSI data are daily-averaged values without flares, expressed in 1 nm bins.

Ten small data gaps (1-3 days) are filled by a linear interpolation. The time period considered in this study runs from 24 April 2010 (i.e., the beginning of SDO observations) until 1 July 2013, which covers more than 3 years of SSI observations during the ascending phase of solar cycle 24.

Table 3.1: The SSI observations that are used for the model calibration.

Instrument	Wavelength range (nm)	Spectral band	Data version
SDO/EVE	$6.5 - 10.5$	XUV	4
SDO/EVE	$11.5 - 36.5$	EUV	4
TIMED/SEE	$37.5 - 119.5$	EUV	11
SORCE/SOLSTICE	121.5	Ly-α	11
SORCE/SOLSTICE	$120.5 - 200.5$	FUV	11
SORCE/SOLSTICE	$201.5 - 300.5$	MUV	11

3.2 Long-term stability of the SSI observations

As we have discussed in Section 1.4, SSI instruments that measure UV radiation suffer from different effects such as detector degradation, and contamination of optical surfaces, which lead to a decrease of their response over time. Typically their response decreases exponentially, and is the strongest at the beginning of instrument operation. Because of this, the response from SSI instruments has to be properly corrected in order to deliver calibrated SSI values.

Although instrument teams typically provide data to which all known corrections are applied, and the quality of the data improves iteratively with different data versions, in some cases a significant uncorrected long-term trend of the instrument might remain. Among the instruments that we use to calibrate our model, such uncorrected trends are suspected to occur in the TIMED/SEE and SORCE/SOLSTICE instruments. Several studies that compared data from SORCE/SOLSTICE (version 10) with data from other missions, and proxy models (MgII), reveal a significant disagreement in the long term variability for the end of solar cycle 23 [DeLand 2012b, Morrill 2014]. In the TIMED/SEE instrument, many intense lines in the EUV suffer from a strong degradation of the detectors, which reduces the photon count by orders of magnitude as compared to the beginning of the mission. Moreover, typically the photon count of the TIMED/SEE detectors does not decay exponentially, but behaves in a more complex way. Thus it is very likely that long-term instrumental trends are present in the TIMED/SEE SSI data products. Also, because of the low photon count, the short-term variability is more strongly affected by instrumental noise.

More recently, data from SORCE/SOLSTICE (version 13) and TIMED/SEE (version 11) has been tested for the presence of the uncorrected long-term trends by [Dudok de Wit, 2014 personal communication]. This author estimated the long-term

trend based on equal activity levels that are derived from different proxies such as F10.7, F30, MgII etc. All proxies revealed significant trends in both TIMED/SEE and SORCE/SOLSTICE data, which suggests that these trends are most likely not physical but instrumental. This analysis is illustrated in Figure 3.1.

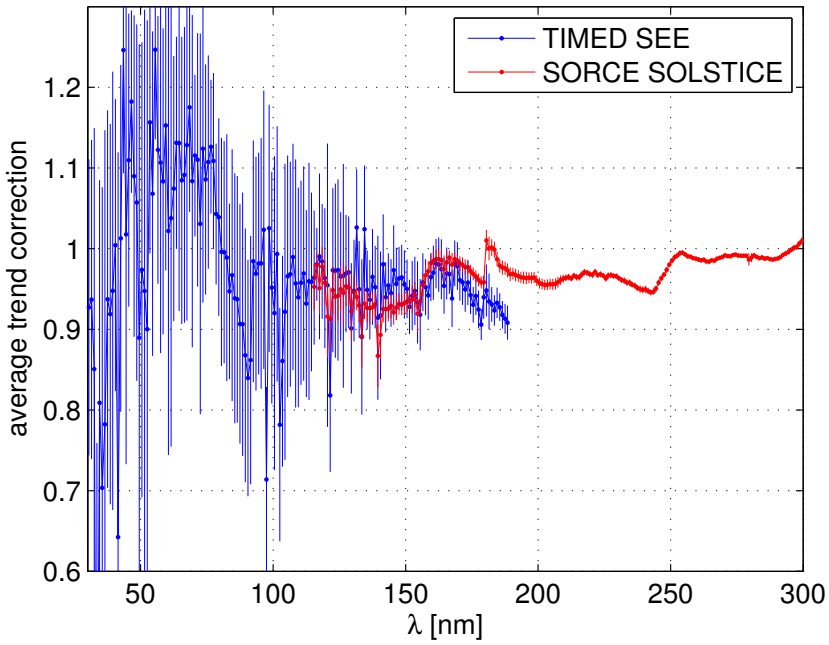

Figure 3.1: Average trend correction estimated from the TIMED/SEE and SORCE/SOLSTICE instruments for a time period ranging from 2000 to 3000 days after launch [Dudok de Wit, 2014 personal communication].

In Figure 3.1 the solid red curve represents the average trend correction for SORCE/SOLSTICE observations versus different wavelengths for a time period ranging from 2000 to 3000 days after launch. The error bars reflect the scatter that results from the use of different proxies. We notice that for the most wavelengths the average trend is significantly different from 1. It is typically underestimated for wavelengths shorter than about 250 nm. Also, a pronounced discontinuity at 180 nm corresponds to the change between two different detectors of the SORCE/SOLSTICE instrument [McClintock 2000], which gives evidence that this trend correction curve is at least partly capturing instrumental effects. Thus, we notice that in the version 13 data, there are still significant uncorrected trends present, and conclude that in the data version 11 that we use for our analysis, long-term instrumental trends should be present as well.

The solid blue curve in Figure 3.1 corresponds to the average trend correction for the TIMED/SEE instrument. We notice that the uncorrected long-term trends are larger than for SORCE/SOLSTICE for comparable wavelengths, and that their corresponding uncertainties are larger. Therefore, we conclude that the long-term

instrumental effects in the TIMED/SEE data are likely to be even stronger than those of SORCE/SOLSTICE. They are extremely pronounced for the intense lines in the EUV.

As we shall see later, these long-term instrumental trends in the SSI data have a strong impact on the reliability of our model on the 11-year timescale.

3.3 Model description

The basic assumption of our model is that the SSI variability is primarily driven by the evolution of the photosperic magnetic field. The same assumption is made in many semi-empirical models, such as SATIRE-S, which has already been described in Chapter 2. Similarly to these models, we first decompose the magnetic field into a set of magnetic structures, and, secondly, fit the spectral irradiance as the superposition of their corresponding spectral profiles (contributions). In contrast to other models, however, we make no assumptions on the number of magnetic structures, nor do we impose any spectral profile on them. Our main objective is to derive empirically how many different magnetic structures are needed, and how they should be defined in order to properly reproduce the observed SSI variability, given its uncertainties.

By the latter we mean that a more sophisticated model with more structures may provide a better description of the observations for the interval the model is trained on, but not necessarily for an out-of-sample interval.

Let the spectral irradiance $I(\lambda, t)$ at wavelength λ and time t be the sum of the spectra associated with different magnetic structures S, weighted by their corresponding area coverage coefficients F

$$I(\lambda, t) = \sum_f \sum_r S_{f,r}(\lambda) F_{f,r}(t) + S_{QS}(\lambda) F_{QS}(t) + \xi(\lambda, t), \tag{3.1}$$

where $\xi(\lambda, t)$ denotes the residual error. Each area coverage coefficient $F_{f,r}(t)$ expresses the fraction of the solar disk area that is covered by magnetic structures of different classes (labeled f) with respect to their radial location (labeled r) on the solar disk. $S_{QS}(\lambda)$ is the spectral profile of the quite Sun, and $F_{QS}(t)$ is the associated area coverage coefficient. The quiet Sun area coverage coefficient is basically what is left over after subtracting all the other features,

$$F_{QS}(t) = 1 - \sum_f \sum_r F_{f,r}(t) \ . \tag{3.2}$$

Different classes of magnetic structures are defined according to the area of the magnetic structures (we discuss the motivation for this later). Each class f is associated with area threshold levels AT_{f-1} and AT_f. A magnetic structure belongs to class f if its area A (corrected for the foreshortening) falls in the interval $[AT_{f-1}, AT_f)$.

Magnetic structures that belong to the same class but are located at different positions may have different contributions to the irradiance $I(\lambda, t)$ due to the centre-

to-limb variation. In our model we assume that the centre-to-limb profile does not depend on the direction, which means that the magnetic structures of the same class f that are located at the same distance from the solar disk centre, give equal contribution to the spectral irradiance $I(\lambda, t)$. The position of the magnetic structures is determined by splitting the solar disk into a set of heliocentric annuli r. Magnetic structures located in the same annulus are assumed to be at equal distance from the solar disk centre (equal position). All annuli are designed to have the same area, in order to have comparable sample sizes for calculating the area coverage s. Thus, the area coverage coefficients $F_{f,r}(t)$ are calculated from magnetograms, based on the parameters f, r and the set of area thresholds $\{AT_0, AT_1, \ldots, AT_{N_f}\}$. The number of classes is N_f. We detail this step in section 3.6.

To estimate the model coefficients $S_{f,r}(\lambda)$, which represent the spectrum of the different magnetic structures, we proceed in several steps. First, we estimate the area coverage coefficients F for predefined values of f, r, and $\{AT_0, AT_1, \ldots, AT_{N_f}\}$.

Secondly, we estimate the coefficients $S_{f,r}$ and S_{QS} as a solution of a linear regression equation to reproduce at best the SSI observations. Using equations 3.1 and 3.2 we have:

$$I(\lambda, t) = \sum_f \sum_r S_{f,r}(\lambda) F_{f,r}(t) + S_{QS}(\lambda) \left(1 - \sum_f \sum_r F_{f,r}(t)\right) + \xi(\lambda, t), \quad (3.3)$$

and, consequently

$$I(\lambda, t) = \sum_f \sum_r (S_{f,r}(\lambda) - S_{QS}) F_{f,r}(t) + S_{QS}(\lambda) + \xi(\lambda, t). \quad (3.4)$$

Introducing $S'_{f,r}(\lambda) = S_{f,r}(\lambda) - S_{QS}$, the former equation can be interpreted as a linear regression problem:

$$I(\lambda, t) = \sum_f \sum_r S'_{f,r}(\lambda) F_{f,r}(t) + S_{QS}(\lambda) + \xi(\lambda, t), \quad (3.5)$$

where $I(\lambda, t)$ is a dependent variable, and $F_{f,r}(t)$ are regressors. $S'_{f,r}$ and S_{QS} are coefficients to be estimated.

Since our model has many variables, which are spectral irradiance time series at different wavelengths, it is more convenient to consider a matrix form of the problem:

$$\mathbb{I} = \mathbb{F} \cdot \mathbb{S}', \quad (3.6)$$

where

$$\mathbb{S}' = \begin{pmatrix} S'_{1,1} & S'_{1,2} & \cdots & S'_{1,N_\lambda} \\ S'_{2,1} & S'_{2,2} & \cdots & S'_{2,N_\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ S'_{N_f,1} & S'_{N_f,2} & \cdots & S'_{N_f,N_\lambda} \\ S_{QS,1} & S_{QS,2} & \cdots & S_{QS,N_\lambda} \end{pmatrix} \quad (3.7)$$

is a $[N_f + 1, N_\lambda]$ matrix of the spectral contrast profiles,

$$\mathbb{I} = \begin{pmatrix} I_{1,1} & I_{1,2} & \cdots & I_{1,N_\lambda} \\ I_{2,1} & I_{2,2} & \cdots & I_{2,N_\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ I_{N_t,1} & I_{N_t,2} & \cdots & I_{N_t,N_\lambda} \end{pmatrix} \tag{3.8}$$

is a $[N_t, N_\lambda]$ matrix with the SSI observations, and

$$\mathbb{F} = \begin{pmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,N_f} & 1 \\ F_{2,1} & F_{2,2} & \cdots & F_{2,N_f} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{N_t,1} & F_{N_t,2} & \cdots & F_{N_t,N_f} & 1 \end{pmatrix} \tag{3.9}$$

is a $[N_t, N_f+1]$ matrix with the area coverage coefficients. Here N_λ is the number of the wavelengths, and N_t is the number of samples for each λ. In our model we have $N_\lambda = 305$, and $N_t = 1200$. The solution of this linear problem is formally given by: $\mathbb{S}' = \mathbb{F}^*\mathbb{I}$, where $\mathbb{F}^* = (\mathbb{F}^T\mathbb{F})^{-1}\mathbb{F}^T$. Once the matrix $\mathbb{S}'$ is estimated we calculate the spectra:

$$\mathbb{S} = \begin{pmatrix} S'_{1,1} + S_{QS} & S'_{1,2} + S_{QS} & \cdots & S_{1,N_\lambda} + S_{QS} \\ S'_{2,1} + S_{QS} & S'_{2,2} + S_{QS} & \cdots & S'_{2,N_\lambda} + S_{QS} \\ \vdots & \vdots & \ddots & \vdots \\ S'_{N_f,1} + S_{QS} & S'_{N_f,2} + S_{QS} & \cdots & S'_{N_f,N_\lambda} + S_{QS}. \end{pmatrix} \tag{3.10}$$

Numerically, however, it is more efficient to solve the linear overdetermined system $\mathbb{I} = \mathbb{F}\mathbb{S}'$ with a specific algorithm (Gaussian elimination).

Based on the known physics and observations we may expect the solutions of this linear regression problem to be positive defined. As we shall discover later on during model selection and calibration, the solutions are indeed positive for those wavelengths where the SSI observations do not suffer from strong instrument degradation effects.

The matrix $\mathbb{F}$ could potentially become ill-conditioned because of the high degree of correlation between the area coverage coefficients of different magnetic structures, which makes the solution of the linear problem unstable. This may arise when the number of classes N_f or annuli N_r is too large, or when different area coverage coefficients contain magnetic structures that exhibit similar time evolution. To quantify the ill-conditioning of a linear system we use the condition number [Cheney 2007]. The area coverage coefficients, that allow the model to reach its best performance, typically have a condition number of about 10 - 20. With this we can roughly estimate that the uncertainty of the model coefficients due to the statistical redundancy of classes f is about one order of magnitude higher than the uncertainty in the SSI observations. Ill-conditioning thus is not an issue.

We quantify the model quality by using the normalised root mean square devi-

ation (NRMSD):

$$\mathrm{NRMSD}(\lambda) = \frac{\sqrt{\frac{1}{N}\sum_t (I_t(\lambda) - \hat{I}_t(\lambda))^2}}{\sigma_I}, \tag{3.11}$$

where $I_t(\lambda)$ is the observed SSI, $\hat{I}_t(\lambda)$ stands for the model estimate, and σ_I is the standard deviation of the observed SSI. Because of the normalisation with respect to the standard deviation, the NRMSD(λ) can be interpreted as the amount of unexplained error versus the variability of the observed SSI; its value equals zero for a perfect match, and NRMSD(λ) $\geqslant 1$ means that the residual error is comparable or larger than the SSI variability. In many studies, the error is normalised with respect to the average value of the SSI, which leads to much smaller values of the NRMSD. Our definition is more relevant insofar we aim to model the variability of the SSI.

The next step consists in optimising the number of classes N_f, annuli N_r, and the size thresholds $\{AT_0, AT_1 \dots AT_{N_f}\}$. To do so, we vary these parameters and look for the model that offers the best reconstruction capacity, which is discussed in section 3.6.

3.4 Extraction of magnetically active regions

Let us now concentrate on the extraction of the magnetic structures that are required for estimating the area coverage coefficients, before we optimise the number of classes of these structures. The solar surface is traditionally split into magnetically active regions, and the quiet Sun region. Magnetically active regions can be further classified into a set of different structures (sunspots, faculae, network) based on their properties such as contrast at different wavelength in intensity images, magnetic field value, and morphology [Zwaan 1987].

We proceed with the identification of magnetically active regions. Quiet solar regions in magnetograms are assumed to have magnetic field intensity around 0 G, which is mixed with instrumental noise. Magnetically active regions, on the contrary, have magnetic field intensity values that differ significantly from 0 and are significantly greater than the noise level.

SDO/HMI magnetograms (see Figure 3.2) capture the line-of-sight magnetic field B_{LOS} at photospheric level, which can be considered as a projection of a two-dimensional cut through magnetic flux tubes in the solar atmosphere. This means that even strong active regions can have low line-of-sight magnetic field intensity values when they are located near the limb. However, due to the magnetic buoyancy, the magnetic field lines can be assumed to be perpendicular to the solar "surface". Thus, in order to treat in an equal way magnetic regions that are located at different distances from the solar disk centre, we correct the magnetic flux for the projection effect with

$$B_N(i,j) = \frac{|B_{LOS}(i,j)|}{\sqrt{1 - (d(i,j)/R_0)^2}},$$

where i and j are coordinates of pixel in magnetograms with respect to the solar

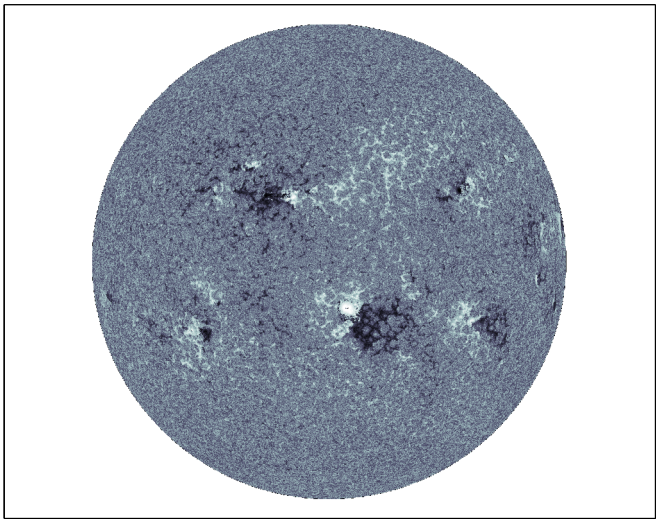

Figure 3.2: SDO/HMI magnetogram taken on 5 April 2013 at 12:00 UT. Here white color indicates magnetic regions where the magnetic field force lines are oriented in the direction towards an observer, whereas the black colour corresponds to the opposite orientation. Gray areas correspond to the quiet Sun.

disk centre, R_0 is the solar disk radius in pixels, and $d(i, j)$ is the distance from the pixel to the solar disk centre; $B_N(i, j)$ is the radial magnetic field intensity. The magnetic field in the magnetic flux tubes has different signs depending on its orientation. We take the absolute value of the magnetic field intensity to incorporate both orientations.

To determine the threshold between magnetically active and quiet regions that is valid for the whole time period considered, we need a representative distribution of intensities which covers different levels of solar activity. For this purpose we take 116 randomly selected magnetograms (which cover 10% of the whole dataset) that are corrected for the projection effect, and calculate the histogram of the magnetic field intensity. Pixels that are located near the limb, where $d(i, j) \geq 0.95R_0$, are not included because of their high noise level, which is amplified by projection effects. Note that by doing so, we implicitly assume that the thresholds are not solar cycle-dependent. The same assumption is made for semi empirical models such as SATIRE-S.

The histogram of the intensities is presented in Figure 3.3; it exhibits several features, with a transition point near 20 G-70 G, separating two regimes of magnetically quiet and active pixels. Several authors similarly separate magnetically quiet and active regions. [Yeo 2013], for example, finds the threshold to be 40 G - 70 G with SDO/HMI, depending on the distance from the solar disk centre. Other stud-

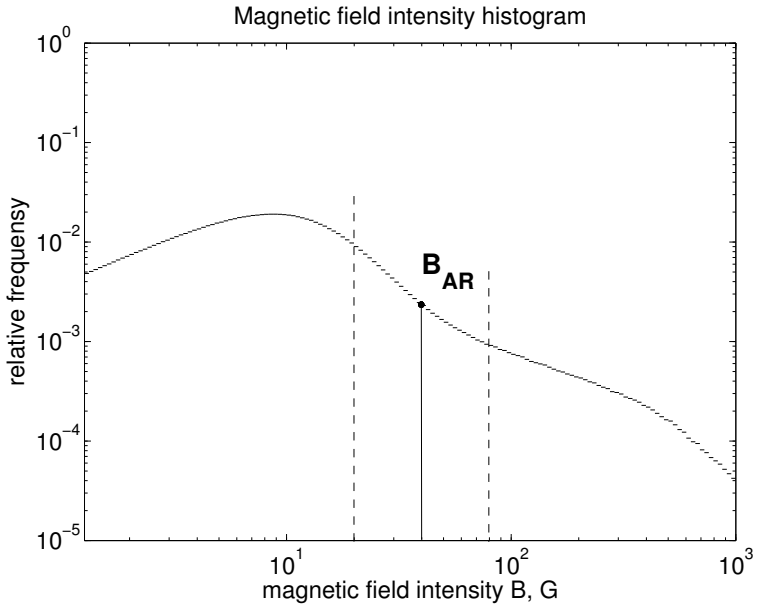

Figure 3.3: Histogram of the absolute values of the sample of magnetic field intensities, calculated from 116 randomly selected magnetograms that are corrected for the line-of-sight projection effect. The transition point B_{AR} at about 40 G separates magnetically active regions from the quiet Sun.

ies [Fligge 2000, Higgins 2011], based on SOHO/MDI magnetograms, locate this threshold at about 50 G-75 G. All these values are located within the transition interval shown in Figure 3.3). Note that thresholds obtained from SOHO/MDI are generally found to be larger, but pixel intensities from SOHO/MDI are also known to exceed those from SDO/HMI by a factor of about 1.4 [Liu 2012].

In accordance with these results, we also associate small intensities with the quiet Sun, and large ones ($>$ 20 G) with magnetically active regions. However, for the purpose of our work, we wish to retain a single threshold value 20 G$\leq B_{AR} \leq$70 G. Excessively large values are likely to underestimate the size of magnetically active regions, and useful signal may be lost. Conversely, if the threshold is too low, then magnetically active regions will be more strongly contaminated by noise. To find the best compromise, we determine the model performance for a variety of threshold values. The optimum threshold is wavelength-dependent, but typically occurs between 20 G and 40 G. Based on this, we choose the largest value in order to minimise the impact of noise. From now on, our threshold will be B_{AR} =40 G.

Magnetically active regions can now be extracted with a simple mask:

$$M_{AR}(i,j) = \begin{cases} 1 \text{ if } B_N(i,j) \geq B_{AR} \\ 0 \text{ if } B_N(i,j) < B_{AR} \end{cases} . \tag{3.12}$$

The binary mask M_{AR} (see Figure 3.4) flags magnetically active pixels with 1, and

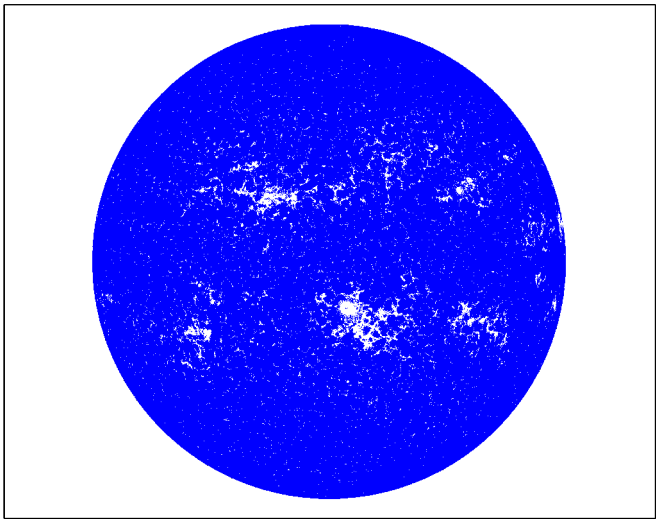

Figure 3.4: Binary mask of magnetically active regions M_{AR} for a magnetogram taken by SDO/HMI on 5 April, 2013.

quiet Sun pixels with 0. Further classification of the magnetically active regions is done according to their area (see section 3.5). To perform such a classification, we need to identify individual regions in the binary mask M_{AR}. To this end, we identify 4-connected pixels in the binary mask and apply the Dulmage-Mendelsohn decomposition to find connected components [Haralick 1992]. The definition of the connectivity in the binary mask affects the outcome of the active region identification procedure. Two default options would be to use 4 or 8-component connectivity criterion. We choose the 4-connected connectivity criterion because it produces much less stitching of active regions that are located close to each other. However, the drawback is that those active regions that have a complex structure in some cases might be split into several parts. The final result of the extraction procedure is a mask which labels all pixels by the identification number of their corresponding magnetic region.

3.5 Segmentation by area

The next step is the classification of the magnetically active regions according to their spectral profile. In semi-empirical models, the classes of magnetic structures are pre-defined (sunspots, plages, network); these structures are identified in magnetic and intensity images [Fligge 2000, Haberreiter 2012, Haberreiter 2014]. In the case where the area coverage coefficients of atmospheric structures (typically plages) are derived from magnetic images, the latter are segmented according to the mag-

netic field intensity.

One of the disadvantages of this approach is its sensitivity to instrumental artifacts that may affect the magnitude of the magnetic field (in particular absolute calibration errors). Such artifacts then inevitably affect the result of classification procedure. Another problem is the processing of magnetic structures that are located near the limb, where small uncertainties in the line-of-sight magnetic field intensity are amplified by projection effects.

To alleviate these problems, we consider a different approach. Several studies have found monotonic statistical relations between the size of magnetic structures and their magnetic field intensity [Ringnes 1960, Kiess 2014]. This suggests that magnetic structures may be classified based on their area. The area metrics is much less sensitive to absolute calibration errors of the magnetograms. However, in the case where two different regions are located close to each other so that their contours touch, these regions will merge into a single region, which will affect the classification procedure. In particular, this means that it is not possible to distinguish between a sunspot and its surrounding plage. This problem however, does not affect the SSI reconstruction quality since the sunspots are bright features for the EUV and XUV emissions. In the MUV and FUV, where sunspots gradually become dark features, their area coverage coefficients (estimated form the SDO/HMI visible continuum images) are about 2 orders of magnitude weaker than the ones for bright features. By comparing the model reconstruction quality between two cases: with, and without including umbra and penumbra area coverage coefficients, we find that the effect of sunspots in the FUV/MUV bands is negligible within the uncertainty of the SSI observations (see Figure 3.5).

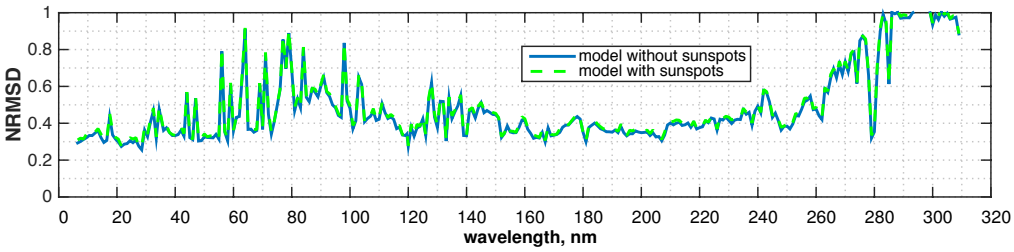

Figure 3.5: Model error at different wavelengths for two cases: with (green) and without (blue) including sunspot areas. The two curves overlap, which suggests a negligible effect of sunspots in the UV in our model.

Therefore, to proceed with the classification of different magnetic regions we calculate only two parameters: area and distance from the solar disk centre. Both parameters are estimated by using the binary mask M_{AR}.

3.6 Model selection

Now, we want to find the different classes of magnetic structures that are needed to properly describe the observed SSI variations in the UV.

The area coverage coefficients $F_{f,r}(t)$ depend on the number of different classes of magnetically active regions N_f, the number of annuli N_r, and the threshold levels $\{AT_0, AT_1 \dots AT_{N_f}\}$ used for segmenting the magnetograms. Our objective is to estimate the values of these parameters that minimise the model error for each wavelength, without making the model overdetermined.

We solve this optimisation problem in two steps. First, for each wavelength we consider different numbers of classes from 1 to 5, as well as different numbers of annuli from 1 to 5. Thus, there are a total of 25 different sets of area coverage coefficients. Within each set we vary the threshold levels $\{AT_0, AT_1 \dots AT_f\}$. This produces all possible variations in the determination of the area coverage coefficients. At each wavelength λ, and for each value of f and r,we then look for the threshold levels $\{AT_0, AT_1 \dots AT_f\}$ that minimise the model error.

When fitting the model to the data, we must ensure that it properly matches the observed variability of the SSI rather than spurious noise patterns in it. We assume here that the observations are corrupted by additive noise that does not correlate with the SSI signal, and is identically distributed in time. We split the available dataset, which has 1160 samples per wavelength bin of 1 nm, into two subsets: the first one is used to train the model only, and therefore has to be representative of the different levels of solar activity that occur during the three years of observations. For that purpose we randomly select 80% (928 data points) of the full dataset. The second subset is used to validate the model; it does not overlap the former one and contains the remaining 20% of data. We use this second subset to check how well the model reproduces the out-of-sample data, which is essential for preventing it from overfitting.

Since our dataset merely covers 3 years of solar activity, with the ascending phase only of solar cycle 24, which moreover is a weak one, we should stress that it may not be representative (yet) of the full solar cycle. Furthermore, since we randomly select the training sample, model coefficients may slightly vary from one run to another. To overcome such variations, we repeat the model training and validation procedures 30 times, and only keep the ensemble-averaged values for further interpretation. Reshuffling the samples more than 30 times does not change the results significantly.

The estimated optimum values of the parameters ($N^o_{f\lambda}$, $N^o_{r\lambda}$, and $\{AT^{\lambda o}_0, AT^{\lambda o}_1 \dots AT^{\lambda o}_f\}$) have an uncertainty, which give some freedom in defining their value. For the sake of convenience, we check whether a common solution (N^o_f, N^o_r, and $\{AT^o_0, AT^o_1 \dots AT^o_f\}$) is possible for all wavelengths, in order to have a single model for the whole UV spectrum.

3.6.1 Optimising the number of classes and annuli

The minimum of the model error with respect to the threshold levels $\{AT_0, AT_1 \dots AT_f\}$ and for different values of N_f and N_r at particular wavelengths is illustrated in Figure 3.6. For each spectral band we choose the optimum values of N_f and N_r. Because the model is trained, and then tested on different samples, the residual error does not decrease monotonically as the number of degrees of freedom increases; it usually reaches a minimum, and then starts increasing slowly again. This important (and often overlooked) aspect implies that models with large number of structures or annuli, actually perform worst because they try to describe the instrumental noise in addition to the variability of the SSI.

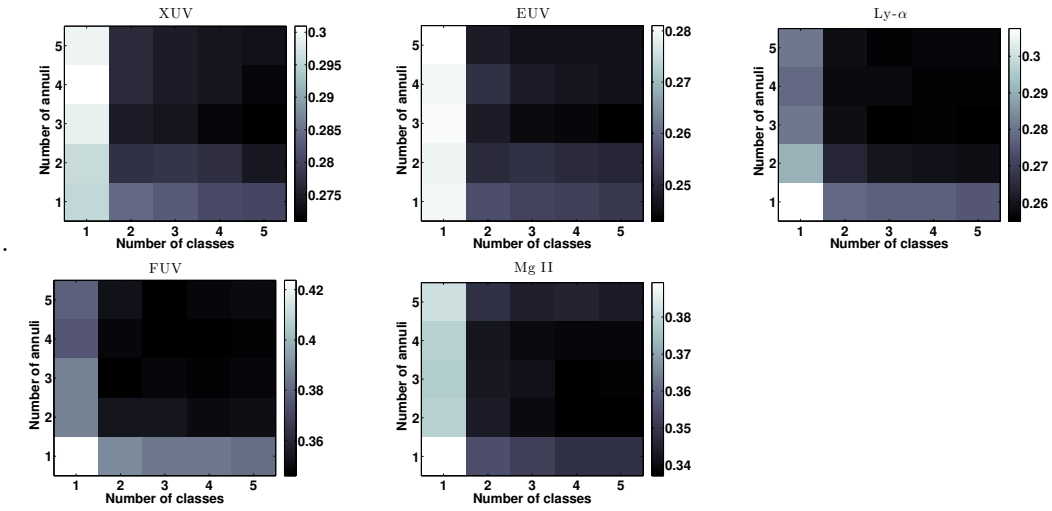

Figure 3.6: The model error (calculated for the validation set) with respect to the number of annuli and number of classes for emissions from different spectral bands. Top left: 8 nm (XUV), middle: 20 nm (EUV) right: 121.5 nm Ly-α. Bottom left: 190 nm (FUV), centre: 280 nm MgII. The value of the uncertainty of the model error, based on different runs, is typically 0.01, and most solutions that appear in dark grey are thus comparable in performance.

The uncertainty of the model error based on different runs is typically 0.01, which means that we do not distinguish between error values that are within 0.01of each other.

From Figure 3.6 we see that the absolute minimum of the error in the XUV band is reached for a model with 5 classes and 3 annuli. However, all the models with ≥ 2 classes and ≥ 2 annuli have errors that are not significantly greater than the absolute minimum value as compared to the uncertainty of the model error. In this case we select the model which has the least number of parameters. Therefore, we conclude that two classes and two annuli are the optimum values for the XUV. We inspect other spectral bands in the same manner. In the EUV, the optimum model has 2 classes, and 2 annuli. The Ly-α line, the FUV, and the MUV require three annuli and two classes. We notice, that the XUV and EUV bands are less

sensitive to the number of annuli, and more sensitive to the number of classes. The FUV and MUV bands have equal sensitivity to classes and annuli. This was expected, because most part of the XUV/EUV emissions are optically thin, whereas the FUV/MUV emissions are optically thick. We expect the model to require more annuli for optically thick lines because of the centre-to-limb variation effects.

3.6.2 Optimising threshold levels

The range of values of the threshold levels $\{AT_0, AT_1 \dots AT_f\}$ is split into 14 bins $\{AB_0, AB_1 \dots AB_{13}\}$. The size of the bins increases on a dyadic scale, i.e. 2^n; each value is two times larger than the previous one. Magnetic structures which correspond to bin AB_k have a size between 2^{k-1} and 2^k pixels. The very first bin corresponds to magnetic structures, that have a size of 1 pixel when located in the centre of the solar disk; if thresholds for a class correspond to bins number m and n then the class contains all magnetic structures which have area between 2^{m-1} and 2^n pixels. We note that the original HMI magnetograms were decimated by factor of 2 for the sake of faster processing. For each fixed number of classes N_f we test all possible values of thresholds in order to find the optimum ones that minimise the model error (NRMSD).

Here we review how the optimum threshold is found for the case of two classes ($N_f = 2$). Figure 3.7 plots the model error versus different thresholds levels $AT_1 = AB_1 \dots AB_{13}$ in the XUV/EUV band. We notice that the error gradually decreases, then reaches its minimum values, and starts to increase again as the threshold goes from its lowest to highest value. The difference between the lowest and highest values of the model error reaches about 0.02. We consider this difference to be significant because the uncertainty in the error based on different runs is about 0.01. The model reaches its best performance when the threshold AT_1 equals $AB_8 \dots AB_{12}$. This shows that it is important to separate the large and small magnetic regions.

Similar analysis of the impact of the threshold on the model error in the FUV/MUV bands is illustrated in Figure 3.8. Here the model has the lowest error values with correspond to the threshold levels of $AB_8 \dots AB_{11}$ which corresponds to the ones found for the XUV/EUV. We choose the single optimum value of AT_1 to be $AT_1^o = AB_{10}$ and stress that variations of this parameter within $AB_8 \dots AB_{11}$ range do not change further calculations significantly. This separation in the size of magnetic regions may be caused by sunspot darkening, which decreases the average intensity of the large regions. Small magnetic regions typically are not associated with sunspots. Reconstructions from 121 nm to 185 nm are not sensitive to the threshold, which suggests that magnetic structures of all sizes give equal contribution to the variability captured by the observations.

The threshold AT_0 defines the size of the smallest structures which are of importance for the describing of the SSI variability. Figure 3.9 shows how the model error behaves with respect to this parameter for a fixed value of $AT_1 = AB_{10}$. We note that the error does not increase significantly in the whole UV range when the threshold AT_0 is not greater then AB_7. With this we find the optimum value of

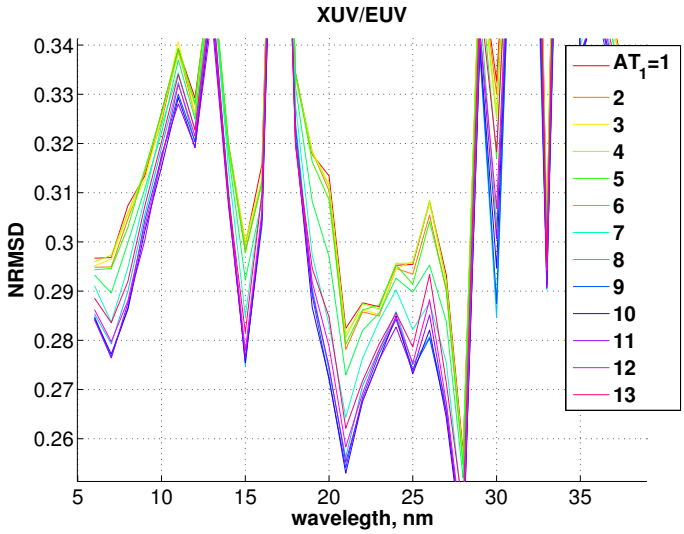

Figure 3.7: Model error with respect to the threshold between classes ($N_f = 2$) in the XUV/EUV.

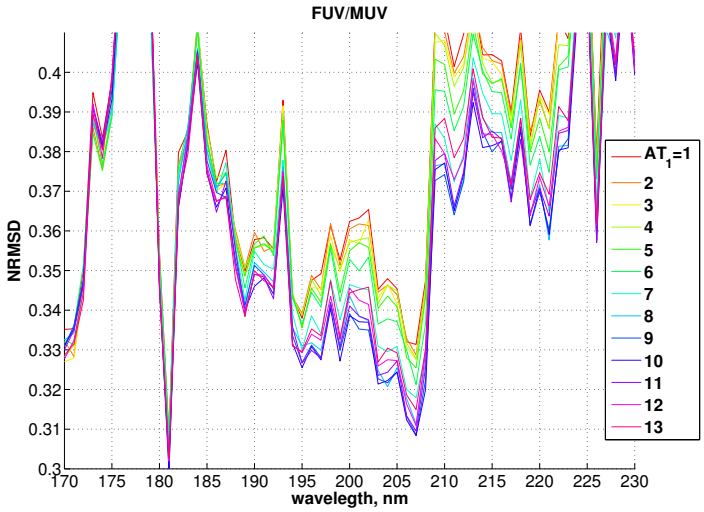

Figure 3.8: Model error with respect to the threshold between classes ($N_f = 2$) in the FUV/MUV.

AT_0 to be $AT_0^o = AB_6$. This means that magnetic structures with sizes below this value do not bring new information about the SSI variability within the uncertainty of the reconstruction.

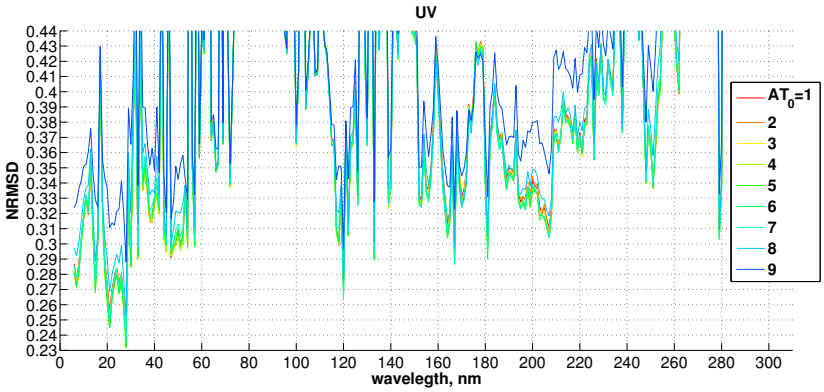

Figure 3.9: Model error with respect to the threshold of the small structures ($N_f = 2$).

3.6.3 Single UV model

For the sake of convenience, we select 3 annuli and 2 classes as the optimal numbers for all spectral bands. Similarly, we apply the same threshold levels $\{AT_0, AT_1 \dots AT_f\}$ since they do not change significantly with wavelength. Thus, we find the optimum values to be: $N_f^o = 2$; $N_r^o = 3$; $AT_0^o = 32$ pxls; $AT_1^o = 512$ pxls (which corresponds to areas of 32 arcsec^2-512 arcsec^2 and > 512 arcsec^2 in physical units, given the size of pixel of the decimated HMI magnetogram 1 arcsec). Two classes of magnetic features suffice for reconstructing the observed SSI variations in the UV. The model error does not significantly decrease, and sometimes even increases when more than two classes are considered. Most likely, this number would have been larger if we had extended our model to the visible part of the spectrum, where other solar features contribute to the SSI variability. Meanwhile, it means that the variability in the UV, when expressed in terms of surface solar magnetic field, only exhibits two degrees of freedom, which are associated with large and small structures (see Figure 3.10). We tentatively associate the large magnetic structures with plages and the small structures structures with network.

Additionally, we find that three principal positions (annuli) of these structures suffice for describing the observed variability of the SSI (see Figure 3.10). The number of annuli expresses how the contrast of a magnetic structure at a particular wavelength changes as it moves across the solar disk. This number is, thus, related to the centre-to-limb variation.

The quality of the optimal model, quantified with the NRMSD is displayed in Figure 3.11 as a function of wavelength. The NRMSD varies strongly, ranging from 0.25 to 1. However, the larger values (typically above 0.45) are associated with the

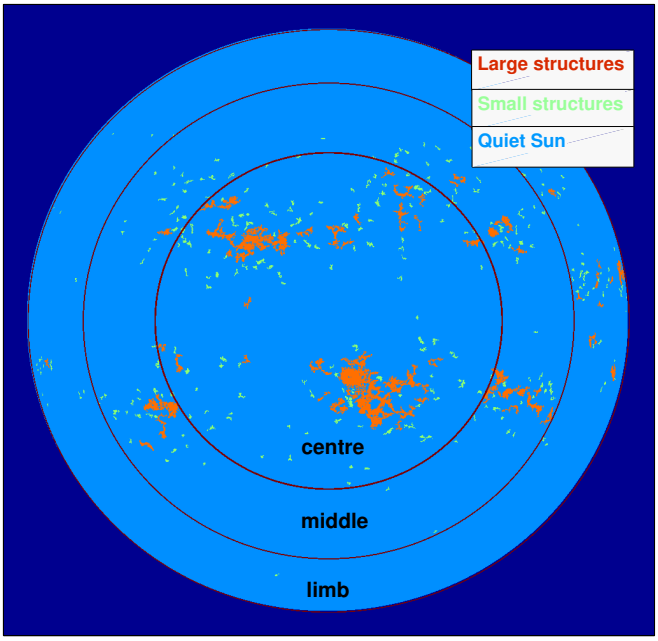

Figure 3.10: Two principal classes of magnetic structures (large in orange, and small in green) and their three principal position on the solar disk (centre, middle, and limb). The segmentation map corresponds to a SDO/HMI magnetogram of 5 April 2013.

SSI records that have significant instrumental noise patterns or problems with the long-term stability. The remaining NRMSD values (not greater than 0.45) have a mean value of 0.35. Thus, the model typically explains 65% of the observed SSI variability in most bands of the UV. The EUV band below 35 nm has an average NRMSD of 0.29 which is the lowest one. Such a good performance may be explained due to the good quality of the SDO/EVE measurements. In the remainder of the EUV band the SSI is not as well reproduced by the model, most probably because of the aforementioned long-term stability issues with the TIMED/SEE observations. The FUV and MUV bands, where the SSI records come from SORCE/SOLSTICE, have average NRMSD values of 0.35, and 0.37 respectively. The NRMSD error increases dramatically in the MUV band above about 260 nm. This is most probably due to the low signal-to-noise ratio.

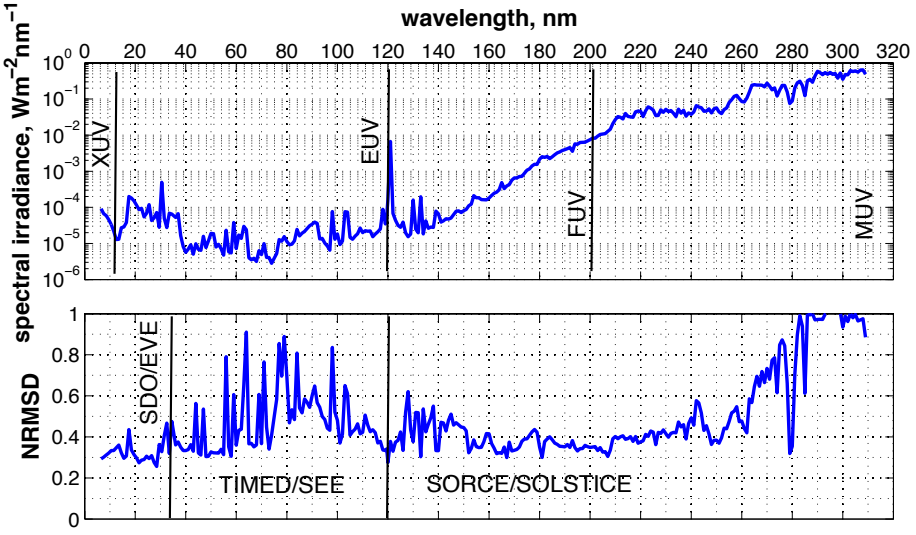

Figure 3.11: The optimal model quality at different wavelengths on the UV. Top panel: time-averaged solar spectrum. Bottom panel: NRMSD of the optimal model. Spectral intervals which belong to different UV bands and covered with observations from the different instruments are shown.

3.7 Contribution of the small and large magnetic structures to the SSI variability

In Section 3.6.3, we have derived two classes of magnetic structures that are sufficient for describing the observed SSI variability within its uncertainty. Now we want to examine what is the contribution of these two classes to emissions at different wavelengths. Typically, in literature authors compare contrasts of different structures (plages, sunspots) at different wavelengths to check the agreement of brightening and darkening effects with available observations. However, in our study the classes of magnetic structures are derived empirically and consequently they cannot be directly matched to different atmospheric structures but has to be understood as proxies that represent the variability. This means in particular that the values of the area coverage coefficients $F_{f,r}\,(t)$ are not necessary the same as the area coverage coefficients of the corresponding structures. Because of this, interpretation of the absolute values of contrasts $S_{f,r}(\lambda)$ (that one may expect to be equal to the contrasts of atmospheric structures from semi-empirical models) is meaningless. Instead we study how the SSI variability is shared between classes $F_{f,r}\,(t)$ by inspecting their relative contribution.

To calculate the relative contribution of class f to wavelength λ we simply integrate its spectrum over all annuli r, and then normalise it to the sum of spectra of

all classes:

$$C_f(\lambda) = \frac{\sum_r S_{f,r}(\lambda)}{\sum_{f,r'} S_{f,r'}(\lambda)}. \tag{3.13}$$

The quantity is shown in Figure 3.12 as function of wavelength.

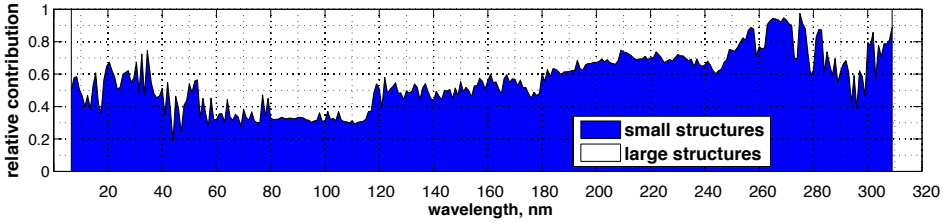

Figure 3.12: Relative contribution of different magnetic structures to the SSI variability at different wavelengths.

The wavelength-average uncertainty of the model coefficients is about 13 % of their magnitude. The corresponding uncertainty of the relative contribution is about 0.14. However, the possible problems with the long term stability of the measurements and the correlated noise still affect the absolute values of coefficients. For this reason, tiny details (below 0.14) in the behavior of the relative contribution curve should not be considered as significant. Let us therefore inspect the relative changes in the relative contribution curve in different spectral regions.

In Figure 3.12 we see that the small magnetic structures contribute more to the long-wavelength irradiance, whereas the large structures contribute more to the short-wavelength emissions. We associate the decreasing contribution of the large structures to wavelengths above 180 nm with the darkening of sunspots because sunspots decrease the average plage intensity.

We now investigate the larger contribution of the large magnetic structures to the EUV/XUV variability (in comparison to the FUV/MUV). This suggests that large magnetic structures are more representative of intense emissions which originate in the corona; the smaller structures, on the contrary, are representative of emissions occurring below in the chromosphere or in the photosphere. These results are expected, because large magnetic structures represent footpoints of the large magnetic loops that reach the solar corona, while small magnetic structures are associated with the footpoints of the small size loops that only reach up to the transition region.

3.8 Centre-to-limb variation

One of the assumptions behind our model is that the contribution of magnetic structures of the same class depends on their distance from the Sun disk centre. Let us see how the contribution of the small and large structures behaves as a function

of their position at solar disk at different wavelengths. The centre-to-limb variation of magnetic structures at wavelength λ can be defined as:

$$CLV_r(\lambda) = \frac{\sum\limits_f S_{f,r}(\lambda)}{\sum\limits_f S_{f,1}(\lambda)}, \tag{3.14}$$

where $S_{f,1}$ is the contribution of magnetic structures of class f located in the central annuli of the solar disk. In order to compare the $CLV_r(\lambda)$ functions for different wavelengths λ we have to normalise them:

$$CLV_r^N(\lambda) = \frac{1}{\sum\limits_{r'} CLV_{r'}(\lambda)} CLV_r(\lambda) = \frac{\sum\limits_f S_{f,r}(\lambda)}{\sum\limits_{f',r'} S_{f',r'}(\lambda)}. \tag{3.15}$$

The normalised centre-to-limb variation of the small and large magnetic structures is plotted as function of wavelength in Figure 3.13.

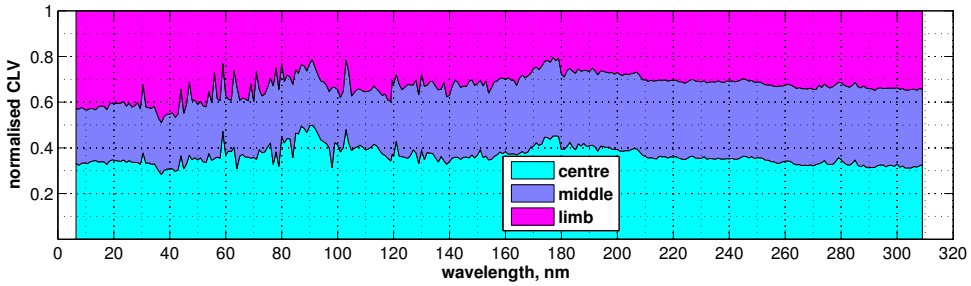

Figure 3.13: Normalised centre-to-limb variation of emissions at different wavelengths in the UV.

We stress that the absolute values of the model coefficients $S_{f,r}$ are sensitive to the SSI data that is used for calibration of the model. Small variations in the CLV behavior (below 0.07) that are present in the plot are, therefore, not significant and should be ignored. However, the relative changes in the CLV from one wavelength to another are usually meaningful. The spectral range from 37 nm to 120 nm is strongly affected by instrumental artifacts in the TIMED/SEE data. In particular, the 11-year cycle sometimes is not properly captured in the observations. Not surprisingly, the model NRMSD sometimes reaches values as high as 0.9. Another problematic spectral band is 265 nm - 310 nm, where artifacts in the SSI data similarly require a great care.

We conclude at this stage that information about the CLV can be inferred from emissions in 8 - 36 nm and 121 — 235 nm bands, and for the Mg ll line at 280 nm. We then find that the XUV/EUV emissions from 8 nm to 36 nm get the largest contribution from the limb annulus, whereas the contribution from this same annulus to the FUV/MUV emissions is smaller. On the contrary, the contribution from the

central and middle annuli tends to increase with wavelengths. In the XUV/EUV region, most emissions are optically thin and originate high in solar atmosphere, which leads to limb brightening. Because of this, the contribution for the limb annulus is much higher. Emissions coming from the FUV/MUV bands are optically thick, and originate in a narrow layer of the solar atmosphere (chomosphere, transition region). Magnetic structures that are located close to the limb have effective emitting areas that are smaller than for central annuli because of projection effects. Moreover, the light emitted near the limb has to penetrate a thicker level of the solar atmosphere, which leads to significant intensity reduction. These effects make magnetic structures look "brighter" near the centre of the disk, and cause the intensity to drop at the limb. We thus conclude that the properties of our model are consistent with the salient characteristics of the different spectral bands.

3.9 The off-limb contribution

The solar limb looks bright in the XUV in most of EUV images, and accounts for a significant part of the total (on-disk and off-limb) emission. This is well illustrated by the intensity images taken by SDO/AIA instrument at different wavelengths in XUV/EUV (see Figure 3.14). These emissions originate in magnetic loops that expand high in solar atmosphere. Due to their spread in altitude, these magnetic loops remain partly visible on the Sun limb even when their footpoints have already passed behind it and are no longer visible. Typically, a coronal loop is visible until the associated footpoints of the loop make an angle of about 120° with the central median; however, this limit is wavelength-dependent. Because of this geometrical effect, and the optical thinness of the majority of coronal emissions, there is a substantial integration of the contributions from different coronal loops when looking at the limb and off-limb area. This means that the magnetic structures which are not further than about 30° behind the limb are important to properly reproduce the XUV band, and part of the EUV band as well.

In magnetograms however, magnetic structures that are located behind the limb are not visible. Moreover, magnetic structures that are located near the limb cannot be estimated precisely due to projection effects.

To examine how these geometrical effects affect our model, we consider a schematic centre-to-limb profile of an optically thin emission line (see Figure 3.15). This profile shows how the intensity of a single magnetic structure changes as this structure passes the solar disk. The whole area below the profile, including the off-limb contribution, shows the contribution of the structure to the solar irradiance on the long-timescale (the 11-year cycle). The amplitude of the curve corresponds to the amplitude of the solar rotation variability. The yellow area corresponds to the irradiance that is emitted when the magnetic structure is behind the resolution limit of the model (95% of the solar disk radius). Thus we see that if the solar rotation variability is properly estimated then the 11-year cycle variability will be underestimated due to the missing off-limb contribution. On the contrary, in the case where

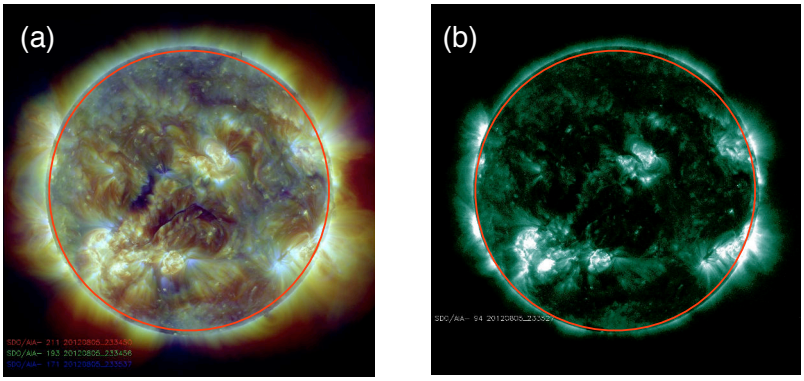

Figure 3.14: Intensity images of the Sun from SDO/AIA instrument taken on 05 August 2012. Panel a: a composite of 17.1 nm, 19.3 nm, and 21.1 nm images (EUV). Panel b: 9.4 nm (XUV). The red ring corresponds to the limb resolution limit of the area coverage coefficients, which is 95% of the solar disk radius.

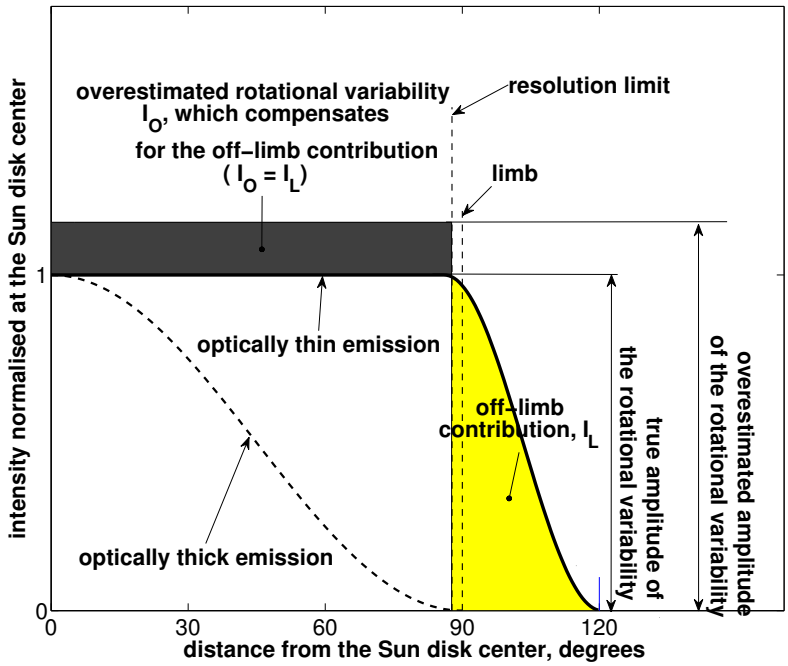

Figure 3.15: Schematic centre-to-limb profile of an optically thin coronal emissions.

the 11-year cycle variability is properly estimated the amplitude of the rotational variability has to be overestimated in order to compensate for the lack of the off-limb contribution. Since the NRMSD is dominated by the largest timescale, the model

tends to reproduce the 11-year cycle, and, thus we expect the solar rotational variability to be to overestimated for optically thin coronal emissions. Because of this disagreement between the 11-year cycle and solar rotation, a two-timescale model may be more appropriate. In such a model the solar cycle, and solar rotational timescales are modelled separately.

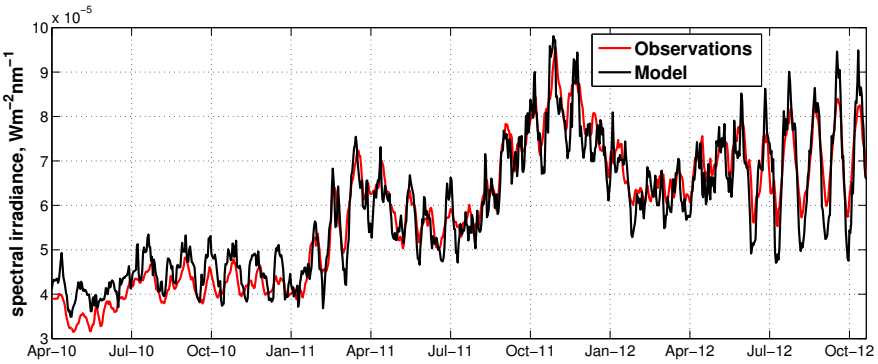

Figure 3.16: Model fit of the spectral irradiance at 8 nm. The solar rotation variability is overestimated, see in particular October - December 2010, and July - October 2012, whereas the 11-year cycle is reproduced more accurately.

Figure 3.16 displays an example of the model fit of the SSI at 8 nm. There are two time intervals in which the irradiance variability is mostly driven by the solar rotation, and the 11-year cycle trend does not have big variability; these are October - December 2010, and July - August 2012. Visually, we notice that the solar rotation timescale of the model fit is overestimated. To further inspect and quantify this overestimation we plot the model fit versus observations (see Figure 3.17). Blue points correspond to the full time record. A linear fit of these points is plotted by the solid line. The slope of this line expresses the linear scaling coefficient between the model fit and observations. In our case this coefficient equals 0.92, which means that the 11-year cycle variability is properly reproduced.

Data points that correspond to the 3 solar rotations during the period of high solar activity in July - August 2012 are shown as circles. Data points from the time period of 3 solar rotations with low solar activity in October - December 2010 are plotted as stars. These two subsets are representative for the solar rotational variability only; their linear fits are plotted respectively as dot-dashed and dotted lines. The slopes of these two lines show the linear scaling coefficients of the model fit with respect to observations at the solar rotation timescale. The value of these coefficients are 1.1 and 1.4 respectively, which tells us by how much the rotation variability is overestimated on average for the corresponding time intervals. Assuming that this overestimation is mostly due to the off-limb contribution we can now estimate by how much the off-limb emission contributes. We find that it represents 10% - 40% of the on-disk emission, which corresponds to 9% - 29% of the total (on-disk and off-limb) irradiance. The difference between the off limb contributions

can be explained by the different levels of solar activity.

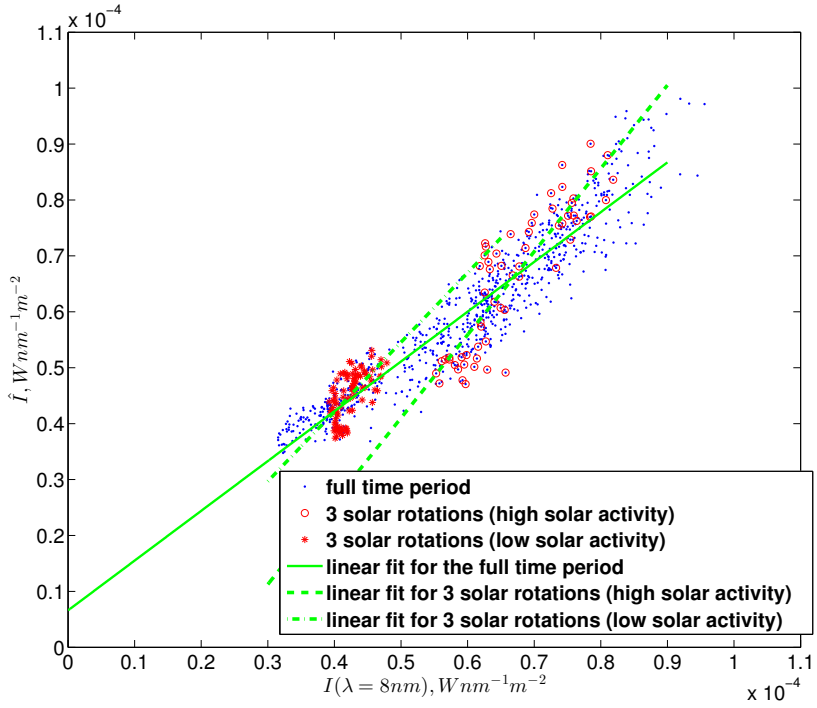

Figure 3.17: Model fit of the spectral irradiance at 8 nm is plotted versus observations. Time intervals that have a small 11-year cycle variability highlight the overestimated rotational variability.

3.10 Two-timescale model

So far we have considered a single-timescale model, which implies a one-to-one relation between the surface magnetic field and the SSI. The same holds for models such as SRPM and SATIRE-S. However, when investigating the linear relation between the model fit and SSI observations in Section 3.9, we have found that this relation may not be the same for different timescales because of the off-limb contribution, which is important for optically thin emissions. Let us, thus, check whether the reconstruction of the SSI can be improved by using a multi-timescale model. In our dataset we identify at least three characteristic timescales: solar rotation (27 days), decay time of large active regions, which is typically of the order of several solar rotations, and the 11-year solar cycle. However, the second one is the weakest of the three. For this reason, we test a two-timescale model which separates only solar rotation, and the 11-year cycle. We split the initial area coverage coefficients $F_{f,r}(t)$ into short and long-timescale components $F_{Sf,r}(t)$ and $F_{Lf,r}(t)$ respectively. We select a cutoff period of 90 days, which corresponds to three solar rotations and

thus include the dynamical evolution of sunspots in what we call the solar rotation timescale.

The multiscale SSI model now becomes:

$$I(\lambda,t) = \sum_f \sum_r S_{Lf,r}(\lambda) F_{Lf,r}(t) + \sum_f \sum_r S_{Sf,r}(\lambda) F_{Sf,r}(t) + S_{QS}(\lambda) F_{QS}(t) + \xi(\lambda,t),$$

with

$$F_{f,r}(t) = F_{Lf,r}(t) + F_{Sf,r}(t). \tag{3.16}$$

Figure 3.18 (top panel) plots the model error NRMSD(λ) for a single-timescale model (green) and for a two-timescale model one (blue). The ratio between the two quantities is shown in the bottom panel. The two-timescale model exhibits significantly lower error for most of the emissions below 121 nm, with some exceptions.

The two-timescale model is capable of describing different contribution of the magnetic structures to the short and long timescales. Thus, it allows to better describe the SSI variability in the XUV/EUV bands, in which the off-limb contribution is significant. However, the model is also capable of fitting the degradation trend of the SSI instruments.

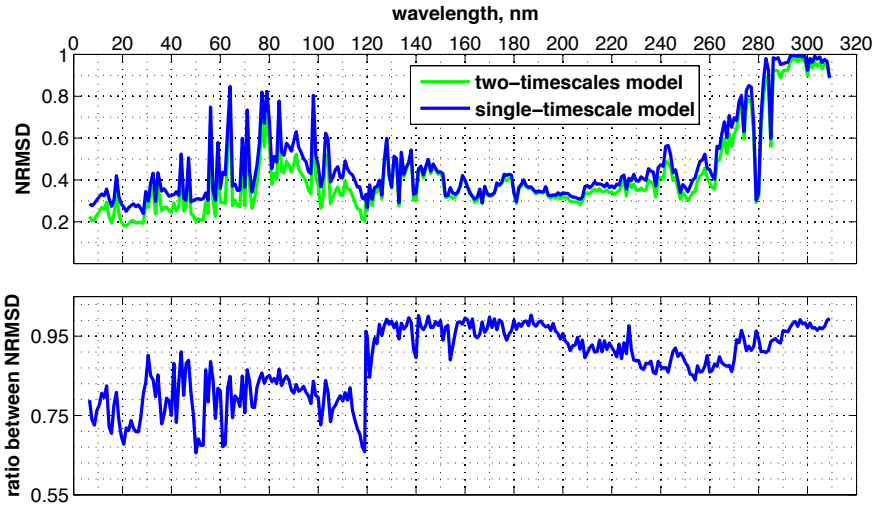

Figure 3.18: Comparison of quality of the single-timescale and two-timescale models. Top panel: NRMSD of the single-timescale and two-timescales models respectively. Bottom panel: Ratio between NRMSD of the single and two-timescales models.

We can also associate the decrease in the NRMSD ratio above 210 nm to the long-term instability of the SORCE/SOLSTICE instrument. There is a sharp transition in the NRMSD ratio at 119 nm, which is exactly where we switch data from the

SORCE/SOLSTICE to the TIMED/SEE instrument. This transition might be caused by degradation of TIMED/SEE channels, which affects the data quality at solar cycle timescale. The transition between the TIMED/SEE and the SDO/EVE data occurs at 35 nm. However we cannot spot any abrupt behavior in error ratio; this is most likely due to the presence of numerous spectral lines that mask the sharp transition. Since the measurements from the SDO/EVE instrument do not suffer from the degradation trend in the considered time interval, we conclude that the improvement of the two-timescale model is most likely due to a better description of the off-limb contribution.

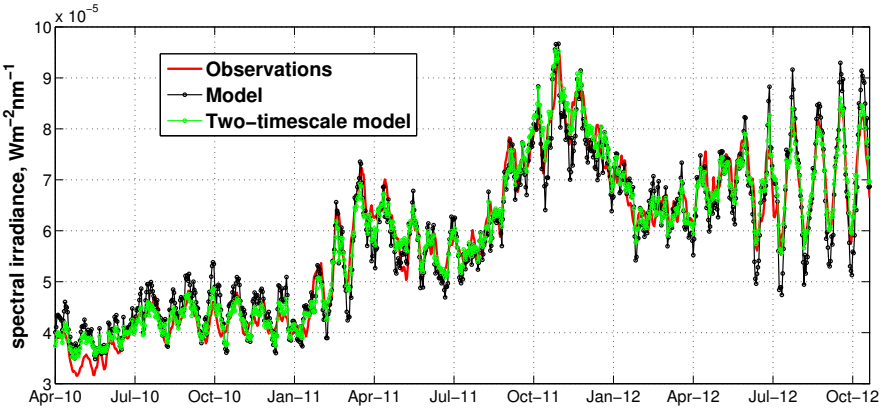

Figure 3.19: Comparison of the single timescale and two-timescales model fits of a wavelength at which the emission is optically thin (8 nm).

To inspect the improvement of the two-timescale model for optically thin emissions we check the model fit of the spectral irradiance at 8 nm (see Figure 3.19) (green dot-dashed line). Visual inspection shows that the solar rotation variability is reproduced more accurately in comparison with the single-timescale model (black dot-dashed line). The scatter plot of the model fit versus observations at this particular wavelength is displayed in Figure 3.20. Here, similarly to the analysis done in Section 3.9, we do a linear fit for the whole time interval and for two time intervals which do not have a significant variability on the 11-year cycle timescale. The corresponding linear scaling coefficients are respectively 0.97, 0.92, and 0.8. These values are closer to 1 than for the single-timescale model (where we had respectively 0.92, 1.10, and 1.40) which indicates that the agreement with the solar rotation and 11-year cycle is better. The 11-year solar cycle is also better reproduced since the linear scaling coefficient of the two-timescale model for the whole time period is closer to 1.

Thus we conclude that the two-timescale model leads to an improved fit of the SSI because it reproduces more accurately the solar rotation timescale of optically thin emissions in the XUV/EUV bands.

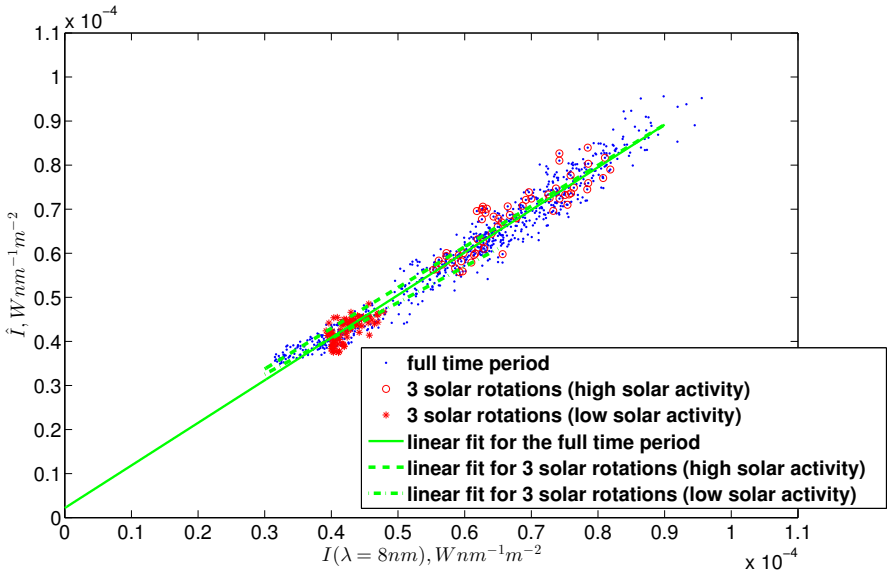

Figure 3.20: Two-timescale model fit of the spectral irradiance at 8 nm is plotted versus observations. The time intervals which have small 11-year cycle variability are in good agreement with the rotational variability.

3.11 Contribution of magnetic structures to the SSI variability at different timescales

Similarly to section 3.7 let us now inspect the contribution of the small and large magnetic structures to the short and long timescales, based on the coefficients $S_{Lf,r}(\lambda)$ and $S_{Sf,r}(\lambda)$ of the two-timescale model. The relative contribution of magnetic structures of class f to the variability at long and short timescales at wavelength λ is given respectively by

$$C_{Lf}(\lambda) = \frac{\sum\limits_{r} S_{Lf,r}(\lambda)}{\sum\limits_{f,r} S_{Lf,r}(\lambda)}, \tag{3.17}$$

and

$$C_{Sf}(\lambda) = \frac{\sum\limits_{r} S_{Sf,r}(\lambda)}{\sum\limits_{f,r} S_{Sf,r}(\lambda)}. \tag{3.18}$$

These two quantities are displayed in Figure 3.21. Interestingly, the contributions from the large and small structures are completely different for the two timescales, which suggests that the one-to-one relation between surface magnetism and spectral irradiance may be more complex than expected.

The long-timescale component (which is associated with the 11-year cycle) is

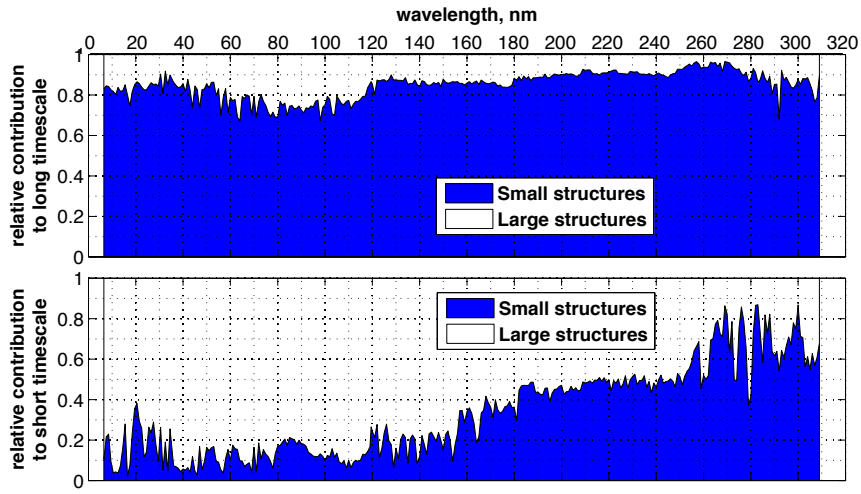

Figure 3.21: Relative contribution of the small and large magnetic structures to the SSI variability at different timescales. Top panel: long timescale ($\geq$90 days). Bottom panel: short timescale ($<$90 days).

mostly driven by the dynamics of the small magnetic structures. We note that the TIMED/SEE data (35 nm to 119 nm) has a different contribution profile, which again suggests a contamination of the results by instrument degradation. The short-timescale component has a stronger contribution from large magnetic structures with a tendency to decrease at the long wavelengths. Large coronal loops are indeed the main contributors to short timescale variations of the coronal emissions in XUV/EUV and are mostly associated with large magnetic structures. The short timescale variations of the emissions, which are formed in the chromosphere and above, are driven by both the large and the small loops simultaneously.

3.12 Conclusions

We have developed an empirical model that describes the SSI variability based on the evolution of the solar surface magnetic field. Our model derives both the classes of magnetic structures and their spectral intensity profiles from the SSI observations. Although the parameters of the magnetic structures that offer the best model performance are wavelength-dependent, we find that it is possible to use the same classes of magnetic structures to describe the SSI variability in the UV band from 6 nm up to 310 nm within the uncertainty of the observations.

Thus, our model suggests that two classes of magnetic structures only (or 3 atmospheric structures, including the quiet Sun) suffice for describing the observed SSI variability. These magnetic structures are distinguished according to their area;

we call them the large and small structures. The large ones are representative of the footpoints of the coronal loops, and, thus, contribute more to the SSI variability in the XUV/EUV bands. On the contrary, small structures represent footpoints of the smaller loops that expand not higher than the transition region, and thus, are of importance for the FUV/MUV emissions. Centre-to-limb effects can be described with a coarse spatial resolution by introducing only three heliocentric annuli. They are of greater importance in the EUV/XUV band, where the large contribution from the outermost (limb) annulus is related to the limb brightening. They also matter for the optically thick FUV emissions in the 160-180 nm range, where limb darkening is strong.

A two-timescale model offers a better reconstruction quality (with a reduction up to about 10% in the model error) in the EUV/XUV bands, and in some of FUV/MUV emissions (typically above 170 nm). In the EUV/XUV bands, where most of emissions are optically thin, this improvement can be explained by the ability of the two-timescale model to account for the off-limb contribution. The slightly better performance of the two-timescale model in the FUV/MUV is not justified physically, and an alternative explanation lies in the plausible presence of a long-term instrument degradation trend in the SORCE/SOLSTICE observations, which is better fitted by a model that has more degrees of freedom.

Since our model is data-driven, it is only as good as the observations are. The low performance of our model (typically when it explains less than 60% of the solar cycle variability) in some bands is related to either the plausible existence of a long term trend or the low signal-to noise ratio in the SSI observations. Thus, the model may be used as well to diagnose the SSI observations.

The presence of the eclipses, instrumental noise, and the instrument degradation trends in the SSI data induces uncertainties in the estimation of the parameters of our model. These various sources of SSI data pollution affect different timescales: eclipses and instrumental noise affect mostly the short timescales, whereas the long timescales are mostly affected by the instrument stability. Moreover, the magnetogram segmentation procedure generates a high-frequency discretisation noise in the area coverage time series that also affects our model. Therefore we need to estimate more accurately how our model performs on different timescales and how stable the estimates arethat it provides, which brings us to the next chapter.

3.13 Résumé en Français

Nous avons jusqu'ici passé en revue les principales approches qui permettent de reconstruire la variabilité de la SSI. Nous avons vu que les modèles combinant des informations spatiales issues d'images solaires segmentées en différentes classes à chacune desquelles est associé un spectre d'émission, offrent généralement une performance supérieure. Ces modèles constituent aujourd'hui la voie la plus prometteuse pour reconstruire la SSI.

Ceci soulève deux questions: 1) ces modèles se prêtent-ils à un fonctionnement

en temps réel, voire à une utilisation à des fins opérationnelles? et 2) peuvent-ils être améliorés en déduisant leurs spectres d'émission des observations, au lieu de les imposer? Pour répondre à ces questions, nous suivrons une voie nouvelle, qui consiste à segmenter des magnétogrammes, mais à dériver les spectres d'émission de manière empirique. Notre hypothèse de base reste ce donc ce lien profond entre la variabilité de la SSI et l'information contenue dans l'évolution du champ magnétique photosphérique.

Notre second objectif consiste à déterminer le plus petit ensemble se structures nécessaires à la reproduction de la SSI. Comme nous l'avons déjà expliqué dans le chapitre précédent, les modèles recourent habituellement à un ensemble de 3-7 structures (ou classes), qui ont de fortes chances d'être en partie redondantes.

Pour extraire ces structures nous recourons au champ magnétique solaire projeté selon la ligne de visée, qui est le produit par défaut des magnétogrammes. Certaines structures atmosphériques, telles que les taches solaires et les boucles coronales ne peuvent être déterminées correctement à partir de tels magnétogrammes. Pour les premières il faudrait ajouter l'analyse des images prises en lumière visible. Or nous avons trouvé que la performance du modèle dans la bande UV ne s'améliore pas de façon significative en présence de telles images prises dans le visible (qui donnent accès à l'ombre et la pénombre des taches). Nous nous contenterons donc dans ce qui suit de magnétogrammes comme seul produit d'entrée de notre modèle.

Les boucles coronales, qui émettent principalement dans l'XUV et l'EUV sont elles aussi difficiles à modéliser à partir de magnétogrammes. Ce sont en effet des structures tridimensionnelles complexes (contrairement aux plages et au réseau qui ont une hauteur négligeable, et peuvent donc être considérées comme des structures 2D), dont la description correcte requiert une extrapolation du champ magnétique. Ce problème dépasse le cadre de notre étude et fait actuellement l'objet de quelques travaux; nous ne l'avons donc pas abordé. Ainsi, avec notre étude, nous visons également à vérifier dans quelle mesure le champ magnétique photosphérique peut décrire la variabilité de la SSI dans toute la bande UV.

Pour extraire les régions magnétiquement actives des magnétogrammes, la méthode classique consiste à identifier des classes de régions en fonction de l'intensité du champ magnétique. Ceci revient à effectuer un seuillage par intensité. Son principal défaut est sa sensibilité à l'étalonnage de l'instrument. Or divers travaux ont révélé l'existence d'un lien statistique entre intensité et taille d'une structure. Nous procédons donc différemment et identifions des classes dont les régions se distinguent par leur taille. Cette approche est appliquée aux magnétogrammes de SDO/HMI pour fournir au final l'évolution temporelle de la fraction du disque solaire occupée par chaque classe. Par régression linéaire, nous cherchons ensuite à identifier les spectres associés qui reproduisent au mieux la SSI entre 6.5 et 310 nm, avec une résolution de 1 nm. Les observations qui servent d'apprentissage et de test du modèle sont issues du des instruments: SDO/EVE (6 - 36 nm), TIMED/SEE (37 - 120 nm), ainsi que SORCE/SOLSTICE (121 - 310 nm).

Pour trouver les paramètres optimaux des classes, nous avons fait varier progressivement leur nombre ainsi que les seuils qui séparent des régions de taille différente.

Au final, nous avons retenu les paramètres qui minimisent l'erreur entre les observations et le résultat du modèle. Cette erreur est évidemment estimée sur un intervalle autre que celui sur lequel se fait l'apprentissage. Le paramétrage dépend un peu de la longueur d'onde. Néanmoins, il est facile d'imposer le même choix pour toutes les longueurs d'onde sans perdre en performance. Les coefficients issus du modèle sont dès lors l'évolution temporelle de la couverture spatiale de chaque classe, ainsi que les profils d'intensité spectrale (ou spectres d'émission) des différentes structures magnétiques. Notre étude suggère que deux classes seulement de structures magnétiques (ou seulement trois structures atmosphériques, si on y inclut le Soleil calme) suffisent pour décrire de manière satisfaisante la variabilité observée SSI. Autrement dit, l'ajout d'une classe supplémentaire ne permet pas de réduire l'erreur de manière significative, compte tenu de l'incertitude sur les observations.

Ces deux classes de structures magnétiques se distinguent pas leur taille: nous les appellerons grandes et petites structures. Les grandes ont le même emplacement que les pieds des boucles coronales. Par conséquent, elles contribuent surtout à la variabilité de la SSI dans les bandes EUV/XUV. Les petites structures représentent au contraire le point d'ancrage de petites boucles qui ne dépassent guère la région de transition et émettent donc davantage dans la bande FUV/MUV.

Pour la prise en compte des effets de centre-bord, nous trouvons qu'une résolution spatiale grossière avec seulement trois anneaux héliocentriques de même aire suffit. Les classes sont dont traitées séparément dans chaque anneau. Ceci est surtout important dans la bande EUV/XUV, en raison de la faible épaisseur optique des émissions.

Nous nous sommes aussi aperçus que la performance du modèle s'améliore si les données sont au préalable décomposées en haute et basse fréquence, avec une coupure aux alentours de 80-100 jours. Nous obtenons donc deux jeux d'observations, avec une variabilité dominée respectivement par la rotation solaire à 27 jours, et le cycle de Schwabe. La réduction de l'erreur qui en résulte est de l'ordre de 10%, mais s'observe surtout dans l'EUV et dans l'XUV. Nous l'expliquons par la faible épaisseur optique des émissions concernées, pour lesquelles le modèle à deux constantes de temps parvient mieux à décrire les part des émissions qui se produisent au-dessus du limbe.

Notre modèle étant entièrement défini à partir d'observations, il ne peut évidemment prétendre à faire mieux que celle-ci. Il se peut aussi qu'il soit affecté par la dérive à long terme observée dans certaines des longueurs d'onde de SORCE/SOLSTICE et dont l'origine a déjà fait l'objet de nombreux débats. Un autre défaut de notre modèle est le nombre limité d'observations (3 ans) liées au lancement de la mission SDO en 2010. Pour cette même raison, nous ne pouvons évidemment prétendre reproduire un cycle solaire entier.

Dans la plupart des bandes, notre modèle explique 30 à 45 % de la variation observée. En cela, il fait pareil que le modèle NRLSSI. Dans certaines bandes, toutefois, l'erreur grimpe à 60%, voire plus. Ceci semble être lié à des dérives instrumentales ou à un mauvais rapport signal/bruit dans certaines longueurs d'onde dégradées de l'instrument TIMED/SEE. Différents facteurs contribuent en effet à

augmenter l'incertitude des coefficients du modèle. En première place figurent le bruit instrumental dans les observations, et la dégradation des instruments utilisés. La procédure de segmentation des magnétogrammes génère lui aussi un bruit de discrétisation de haute fréquence. Il donc important de pouvoir estimer ces effets qui affectent le comportement du modèle, ceci fait l'objet du chapitre suivant.

Chapter 4

Properties of the empirical model

Contents

4.1 Noise in the data 77
4.2 Model quality at different timescales 79
4.3 Model training 81
4.4 Model stability 81
4.5 Intensity contrast of the magnetic structures 84
4.6 Transition to operations 86
4.7 Conclusions 91
4.8 Résumé en Français 91

In this chapter we discuss some of the more specific aspects of our model, including its stability, and issues related to the data availability, which are important for turning the model into a more robust and operational service. In particular, we study how the different sources of uncertainty affect the model reconstruction stability and quality on different timescales. The key question we are addressing here is whether our model can meaningfully reconstruct the SSI in the UV band, given the uncertainties in the SSI observations and in solar magnetograms.

4.1 Noise in the data

Our model has been trained with data that are affected by various types of noise. The area coverage time series are affected by:

- various types of noise in SDO/HMI magnetograms. Foremost among these is the impact of stray light, which is discussed in detail in [Yeo 2014].
- degradation of the SDO/HMI instrument, which may cause non-physical long-term trends in the area coverage coefficients.
- discretisation noise induced by the segmentation procedure that is used to separate the magnetically active regions

The SSI time series are predominantly affected by:

- instrumental noise, which is partly physical (photon count statistics), and partly instrumental only (temperature variations, amplifier noise, etc.).

- instrumental drifts that may have various origins: uncorrected stray light, detector degradation etc. Although the versions we use included all corrections known to date, some uncorrected artifacts are still present. Several recent studies, for example, have questioned the unusually large amplitude in the UV as observed by the SORCE instruments [DeLand 2012b, Morrill 2014].

All these different sources of uncertainty affect our model in different ways, and on different timescales. Random (and spatially incoherent) fluctuations in the SDO/HMI magnetograms affect the area of the magnetically active regions and as such naturally impact the estimation of their area. Small regions, in particular, may easily merge into larger ones, which may thereby end up belonging to a different class. However, since the size of the smallest structures that are included in our model is between 64 and 512 pixels, and the original HMI magnetograms downsampled by factor of 2 (which improves their signal-to-noise ratio by 2) we do not expect this kind of noise to affect our model outcome significantly. And yet, the sharp threshold between our classes will inevitably lead to high frequency discretisation noise. This noise is prevalent on timescales $\leq$1 day. Although the degradation of the SDO/HMI instrument is potentially an important issue, we do not see evidence of its presence in the area coverage time series because the 11-year timescale variability of the SSI in the EUV band observed by SDO/EVE instrument agrees with one estimated by the linear model.

The instrumental noise in the SSI observations is instrument-specific and its characteristics are provided by the instrumental teams, see for example [Woods 2005, Rottman 2006, Woods 2012]. In practice, however, the theoretical specifications strongly vary from one instrument to another, and the noise estimates are often found to be underestimated comparing to the estimates obtained by other means (see section 3.2). In some cases the signal-to-noise ratio may get low either because the solar signal exhibits little variability as compared to the instrumental noise (typically SORCE/SOLSTICE above 270 nm) or because of degradation of the sensor, which is the case for most of the intense lines in TIMED/SEE, whose photon count is now orders of magnitude below what it used to be at launch. The instrumental noise affects our model on the short timescales, and to some degree can be assumed to be independent and identically distributed in time. The latter assumption, however, is not realistic for the instrumental drifts that have a pronounced temporal structure.

The impact of long term degradation on our model is evident. In the case when the degradation trend is linear this results in a decrease of our single-timescale model performance at short timescales for optically thick emissions. Indeed, the model tends to fit the longest timescale which has the largest variance. When the degradation trend is non-linear it causes a discrepancy between the SSI modelled by the two-timescale model and corresponding observations on the long timescales (see Section 4.2).

Our SSI records are affected by 10 eclipses that produce data gaps of about 1 - 3 days. For such short timescales, the gaps can be effectively filled using the linear interpolation. This has no major incidence on the results.

4.2 Model quality at different timescales

The cadence of our model is 1 day, and the total duration of our observations is 1160 days (about 3 years). We thus capture at least three characteristic solar timescales: the 27-day solar rotation, the average decay time of large magnetic regions (below about 100 days), and part of the rising 11-year cycle. As we pointed out before, the noise characteristics of the SSI are timescale dependent. Here we want to test the model reconstruction at two characteristic timescales that account for the largest amount of the total variance, namely, the solar rotation, and the 11-year cycle.

To this aim we split both the SSI observations and the model fit into "low" and "high" frequency components by using a filter with cutoff period of 100 days. The low frequency component is obtained by convolving both the SSI observations and the model fit with a 100-day Gaussian window. The difference between the low frequency component and the initial time series then yields the so-called high frequency component. Next, for each component, we estimate the NRMSD.

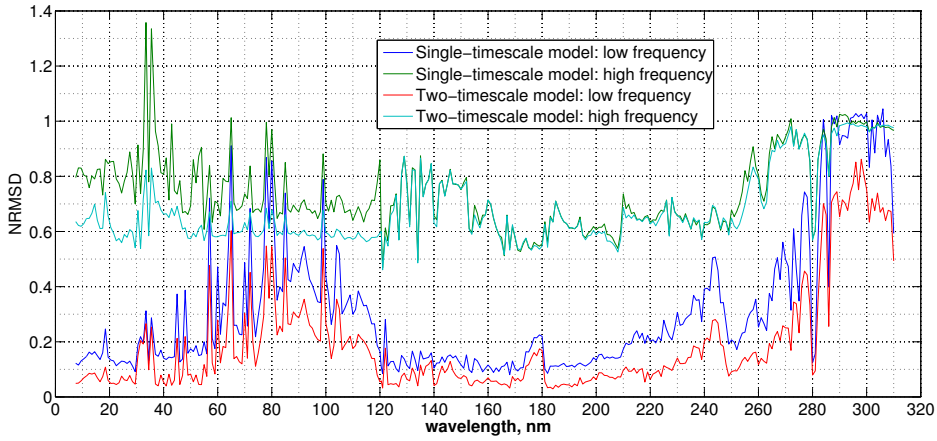

Figure 4.1: Model quality, quantified by the NRMSD on different timescales.

In Figure 4.1 the two values of the NRMSD which correspond to the high and low frequency error of the single-timescale model are plotted respectively in green and blue. We notice that the NRMSD is completely different for the two frequencies: it is much lower for the low frequency (dominated by the 11-year cycle), where it reaches up to 0.1. Much higher values are observed for the high frequency, whose lowest value of 0.45 is reached for the Ly-α line at 121.5 nm. This difference arises for two different reasons. Firstly, our model is heavily biased towards fitting slow variations, and, as a consequence, lower values of the NRMSD are expected for the low frequency component. Secondly, as we have already discussed, our model cannot fit high frequency noise that comes from both the SSI observations and the magnetogram segmentation, both of which lead to higher values of the NRMSD for the high frequency component.

Thus, our model fit quality is biased in favour of the 11-year cycle variability. Moreover, the short timescale SSI variability is affected by fine details in the evolution of the active regions, which are not sufficiently captured by the evolution of their corresponding footpoints. And finally, since we use only one magnetogram per day, we cannot fully describe the daily-averaged SSI value by using one magnetogram per day only. Long-timescale errors can be due either to the lack of long term stability of the SSI observations, or the inappropriate description of the model for the long timescales.

In the FUV band, the single-timescale model explains about 44% of the observed high frequency SSI variability, whereas in the XUV/EUV this amount drops to about 19%. This decrease of the model reconstruction capacity for high frequencies is expected since the single-timescale model is not able to properly reproduce the solar rotation SSI variability in the EUV/XUV band due to the off-limb contribution (see Chapter 3.9).

For the low-frequency component, the single-timescale model explains about 90% of the observed SSI variability in the FUV up to about 170 nm. This indicates relatively good long-term stability of the SORCE/SOLSTICE measurements in that wavelength range. Note however a rapid increase from 170 to 180 nm, and again from 220 to 248 nm. Interestingly, these are precisely the bands in which alternative methods suggest that an uncorrected instrumental drift may be at play (see chapter 3.2). Remarkably, for the Mgll line at 280 nm the single-timescale model has good high and low frequency reconstructions, that are comparable with the ones of FUV emissions. Also we note that in the 55-119 nm spectral range, where the observations are from TIMED/SEE, the low frequency error is large. This is not surprising, as the observations from TIMED/SEE are suffering from strong detector degradation.

For high frequencies, there is no significant improvement in the FUV and MUV bands when moving from a single-timescale to a two-timescale model. Indeed, in these bands a satisfactory reconstruction is achieved with a single timescale only, hence the absence of further improvement. However, in the EUV/XUV band due to the off-limb contribution the two-timescale model explains about 20% more of the short-timescale SSI variability than the single timescale model. On the long timescale the two-timescale model yields better reconstructions for the EUV/XUV and MUV bands. In the MUV, this improvement is likely because the two-timescale model is capable of fitting the instrument degradation trend which may be significant for the SORCE/SOLSTICE instrument in this spectral band. Thus, in particular, we note that the discrepancy between two models increases with the wavelength starting from 180 nm to 250 - 270 nm, indicating that the SORCE/SOLSTICE degradation might be the strongest around 240 -250 nm.

The improvement in the XUV/EUV band occurs because the two-timescale model offers a better trade-off between the short and the long term trend irradiance variability of the optically thin emissions. In the FUV band the two-timescale model exhibits the weakest increase in the reconstruction capacity. The same is true for the Mgll line. This suggests that the 11-year cycle variability is more consistent here with the rotational variability, and the relation between the surface magnetic field

and the SSI is equal for both timescales.

4.3 Model training

To build our model, we must ensure that it properly matches the observed variability of the SSI rather than spurious noise patterns. By this we mean that the model should rely on the relationship between the magnetograms and the SSI to fit the latter, and avoid using its degrees of freedom to fit in addition SSI variations that cannot be described by the segmented magnetograms. To this end, we use cross-validation, and split the available dataset, which has 1160 data points, into two subsets: the first subset (the "in sample" set) is used to train the model only, and therefore has to be representative of the different levels of solar activity that occur during the three years of observations. For that purpose we randomly select 80% (928 data points) from the full dataset. The second subset (the "out of sample" set) is used to validate the model; it does not overlap the former one and contains the remaining 20% of data. We use the latter subset to test how well the model reproduces the out-of-sample data. This cross-validation is essential for preventing the model from overfitting the SSI. The sample size we used for model training was chosen as to achieve the most stable estimates for the model coefficients. Note that for timescales in excess of a few months the training and the validation subsets are not truly independent anymore, and so our model may be slightly biased for the long timescales.

Since our dataset covers about 3 years only of solar activity, with the ascending phase of solar cycle 24, which moreover is a peculiar one, we should stress that it may not be representative (yet) of the full solar cycle. Furthermore, since we randomly select the training sample, the estimate of the model coefficients may slightly vary from one run to another. To overcome the latter effect, and estimate the variation of the model coefficients, we use a bootstrapping-like approach [Efron 1981]: we repeat the model training and validation procedures with randomly-selected values for the in-sample set 30 times, and only keep the ensemble-averaged values for further interpretation. Reshuffling the samples more than 30 times does not improve the results significantly.

Bootstrapping is powerful for assessing the variance of the model coefficients with such a small datasets. Note, however, that it cannot overcome the sparse coverage of solar cycle 24, and consequently, model coefficient estimates for solar cycle time scales should be considered with great care until the model has been trained with a full cycle of observations. This will require another 4-5 years of observations, assuming that SORCE will remain operational.

4.4 Model stability

Our empirical model is affected by various types of noise in both the SSI observations, and the magnetograms. We already found that the presence of noise in the

data brings uncertainties in the model coefficients and in the SSI reconstructions.

Indeed, our model regression can be viewed as an inverse problem, for which it is well known that the variance of the solution gets amplified by the variance of the noise, multiplied by a factor >1 that expresses the lack of conditioning of the regression matrix in equation 3.6, see [Groß 2003].

We just mentioned that the model coefficients are estimated by using a bootstrapping-like technique. This technique provides a distribution of the model coefficients estimates, from which we can infer information about the model stability. To asses the stability of the model coefficients we use the following ratio:

$$T_{S_{f,r}}(\lambda) = \frac{\sigma_{S_{f,r}(\lambda)}}{\mu_{S_{f,r}(\lambda)}}, \tag{4.1}$$

where $\sigma_{S_{f,r}(\lambda)}$ and $\mu_{S_{f,r}(\lambda)}$ are respectively the standard deviation, and the mean value of the model coefficients $S_{f,r}$, as estimated from different runs. The ratio $T_{S_{f,r}}(\lambda)$ quantifies the relative degree of variability of the estimates of the model coefficients. The smaller it is, the more stable the estimates are; its value ought to be $\ll 1$ in order to be acceptable.

We estimate the $T_{S_{f,r}}(\lambda)$ ratio for both large and small magnetic structures individually. This quantity is plotted for different spectral bands in Figure 4.2 to Figure 4.5. The stability of model coefficients for small structures in the XUV/EUV (see Figure 4.3) is about 0.04 to 0.16 whereas for large structures (Figure 4.2) it is about 0.04 to 0.08. While these values are satisfactory in absolute terms, we note that the stability of the coefficients estimates is slightly better for large structures. Our interpretation for this is the relatively better signal-to-noise ratio of the large magnetic structures. Regarding the effect of the position on the solar disk, we find that coefficients associated with the solar disk centre and the middle annulus are more stable than the ones of the limb area. This is consistent with the degradation of the model performance near the limb, because of geometrical effects.

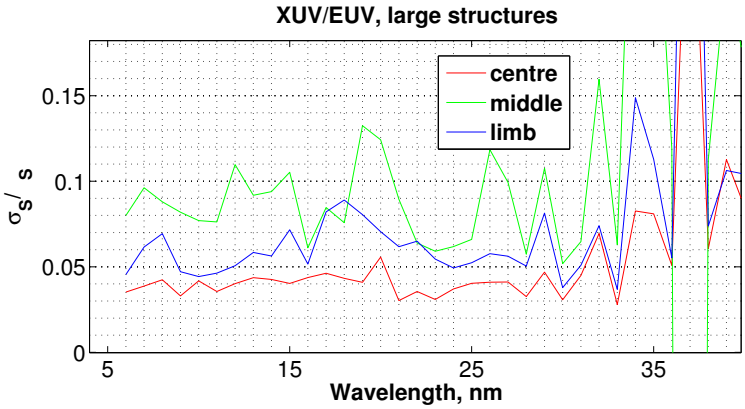

Figure 4.2: Stability of the model coefficients, for large magnetic structures, expressed by the dimensionless ratio $T_{S_{f,r}}(\lambda)$.

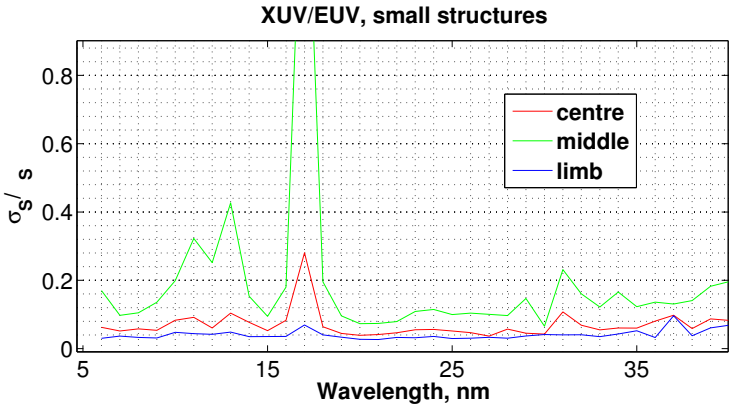

Figure 4.3: Same as Figure 4.2, but for small-size magnetic structures.

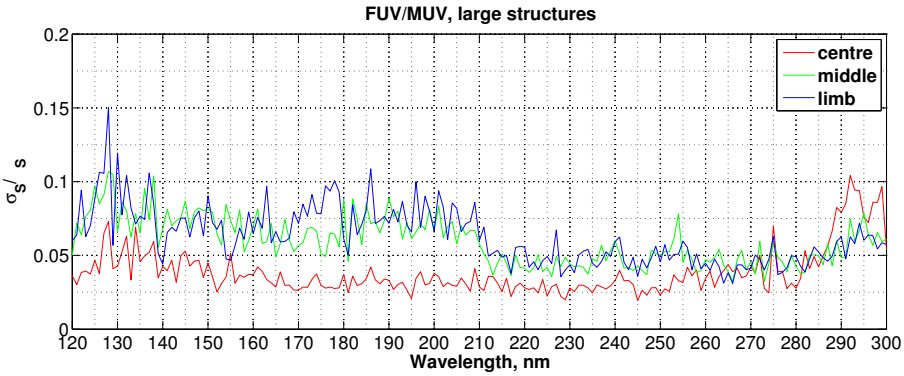

Figure 4.4: Same as Figure 4.2, but for the FUV-MUV band.

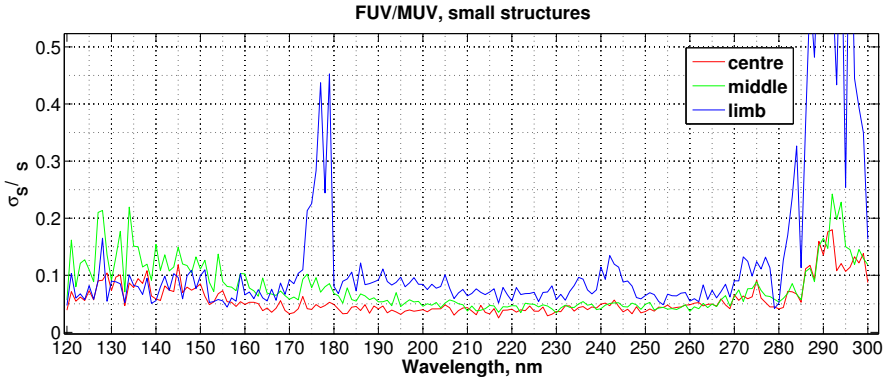

Figure 4.5: Same as Figure 4.3, but for the FUV-MUV band.

In the FUV/MUV band, the stability criterion gives about 0.06 to 0.09 respectively for the large and small structures (see Figure 4.5 and Figure 4.4). As for the EUV/XUV band, the uncertainty in the model coefficients is largest for magnetic structures that are located in the limb area. Overall, the stability of the model coefficients is always small enough to confirm the good stability of the model. There are two exceptions to this: one is the 175 - 180 nm range, which may of have uncorrected trends, and wavelengths above 280 nm for which the model performance is low anyway.

Let us now examine how stable the reconstructed SSI is at different wavelengths. Although this could be estimated directly through error propagation of the above mentioned stability ratio $T_{S_{f,r}}(\lambda)$, we consider here a somewhat different definition, and quantify the reconstruction stability with the dimensionless ratio:

$$T_{\hat{I}}(\lambda) = \frac{M(\sigma_{\hat{I}(\lambda)})}{\sigma_{I_{Obs}}(\lambda)}, \tag{4.2}$$

where $M(\sigma_{\hat{I}(\lambda)})$ is the median value of the standard deviation of the SSI estimates at wavelength λ, and $\sigma_{I_{Obs}}(\lambda)$ is the standard deviation of the SSI observations at wavelength λ. Lower ratios correspond to more stable model reconstructions. $T_{\hat{I}}(\lambda)$ is plotted for different wavelengths in Figure 4.6. The model reconstruction is remarkably stable, and has typical values of about 2%. The reconstruction stability is better for those wavelengths for which the observations are the least affected by noise.

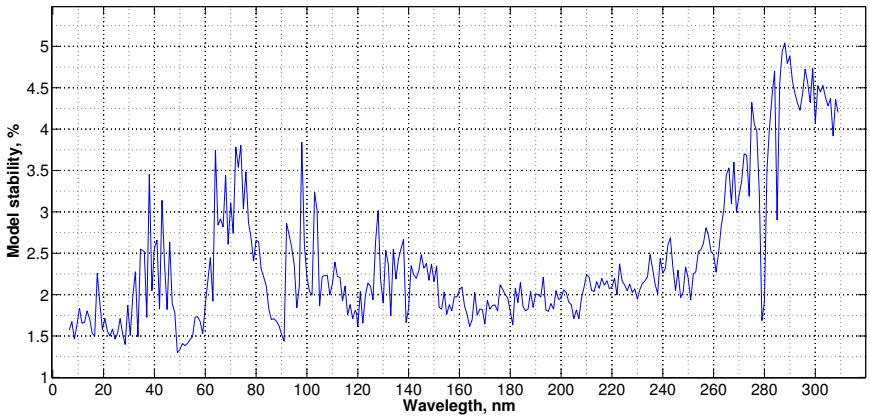

Figure 4.6: SSI reconstruction stability ratio $T_{\hat{I}}(\lambda)$ at different wavelengths.

4.5 Intensity contrast of the magnetic structures

Here we want to study how the small and the large magnetic structures that we derive with our model are related to the intensity of these structures in UV images.

Such a comparison may give an additional insight about the reliability of the classes of magnetic structures that are derived empirically from our model based on the disk-integrated SSI observations only. Such a connection can be established by comparing the magnetic structures with the UV images taken at wavelengths that correspond to optically thick emissions. Indeed atmospheric structures seen in these images are almost flat, and therefore can be directly matched to magnetic structures. [Roger 2014] studied how the contrast of pixels in the UV images from SDO/AIA instrument is related to the magnetic field intensity (obtained from the SDO/HMI instrument). Among different images available from the SDO/AIA instrument only 2 are associated with optically thick emission: 160 nm and 170 nm. Here we consider the UV intensity contrast of atmospheric structures as seen in 160 nm images. To compare our classes of magnetic structures (that are defined according to their area) with these results, we first need to stablish a correspondence between magnetic field intensity and area for each class.

Figure 4.7 shows the relation between the average magnetic field intensity (corrected for the line-of-sight projection effect) of the magnetically active regions that we extract from magnetograms and their corresponding area. Because of the broad scattering of the results, to study the average tendency in the relation between the two quantities we calculate the median area as plotted in green. We find a monotonic increase, which confirms the connection between area and magnetic field intensity, which is at the heart of our model. Note, however, that this connection is statistical, with a considerable scatter that also contributes to the uncertainty of the model reconstructions. Also, we note that there are two regimes, which are separated by a knee at about 2^9 pixels: this is exactly where we placed the threshold between large and small structures AT_1^o (see Section 3.6.2 for details).

White dashed lines at $A_1 = 2^5$ and $A_2 = 2^9$ correspond to the threshold levels that respectively define the small and the large structures in our model. The magnetic field intensity values M_1 and M_2 where these white lines intersect the green line can be tentatively considered as typical levels of the magnetic field intensity that are associated with the small and large magnetic structures. Thus, small structures have magnetic field intensity values of typically 50 - 103 G, whereas large ones have an intensity $\geq$103 G.

The intensity contrast at 160 nm versus the magnetic field intensity is shown in Figure 4.8. Here we see two cases: in panel a) pixels are located at the solar disk centre, in panel b) pixels are located near the limb.

Because the points are broadly scattered, the median value is calculated to show the average tendency in the relation between the intensity contrast and the magnetic field intensity. Thus, we see that for the low magnetic filed intensity values the contrast is about 1, which corresponds to the quiet Sun. Then, starting from about 20 G (panel a) solar disk centre) and 35 G (panel b) limb) the contrast starts to increase almost linearly with the magnetic field intensity. Here, values of magnetic field intensity above which we start see the response in the intensity contrast can be understood as the threshold between the quiet Sun and magnetically active regions.

Also we note that this threshold is higher for pixels that are located near the

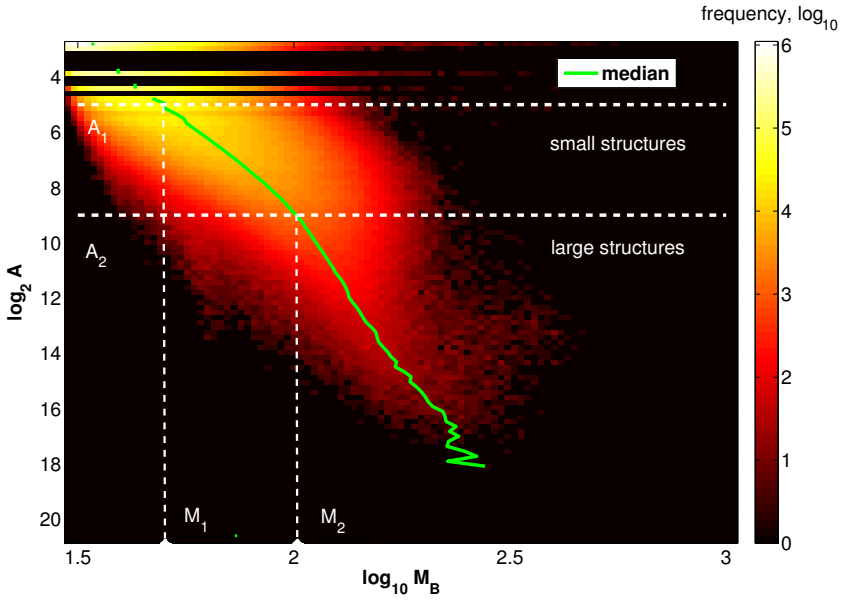

Figure 4.7: 2D histogram showing the average magnetic field intensity of active regions versus their corresponding area.

limb. This happens because pixels in solar magnetograms are increasingly affected by noise when near the limb. Incidentally, that suggests that our model may be improved by applying a threshold for extracting magnetically active regions, which depends on their distance from the solar disk centre.

Next, we see that above about 200 G (panel a) solar disk centre) and 130 G (panel b) limb) the intensity contrast ceases to increase with the magnetic field intensity, reaching saturation and starts to drop. This effect may be related to the magnetic field saturation described in [Fligge 2000]. For our model it means that the spectral intensity profile of the small magnetic structures has to be scaled linearly with the area.

Above about 400 G (panel a, solar disk centre) and 500 G (panel b, limb) the average intensity contrast drops rapidly, reaching values below 1. Pixels in this region mostly belong to sunspots. However, as we have already discussed the effect of sunspot darkening is not directly taken into account in our model.

4.6 Transition to operations

Although several models have been developed for reproducing the variability of the SSI, very few of them are operational. By operational, we mean with continuous and near-realtime operation (with a latency of few hours at most for daily values), and with backup solutions in the event of a data outage. One of the few exceptions is the Solar2000 model developed by the commercial Space Environment Technologies

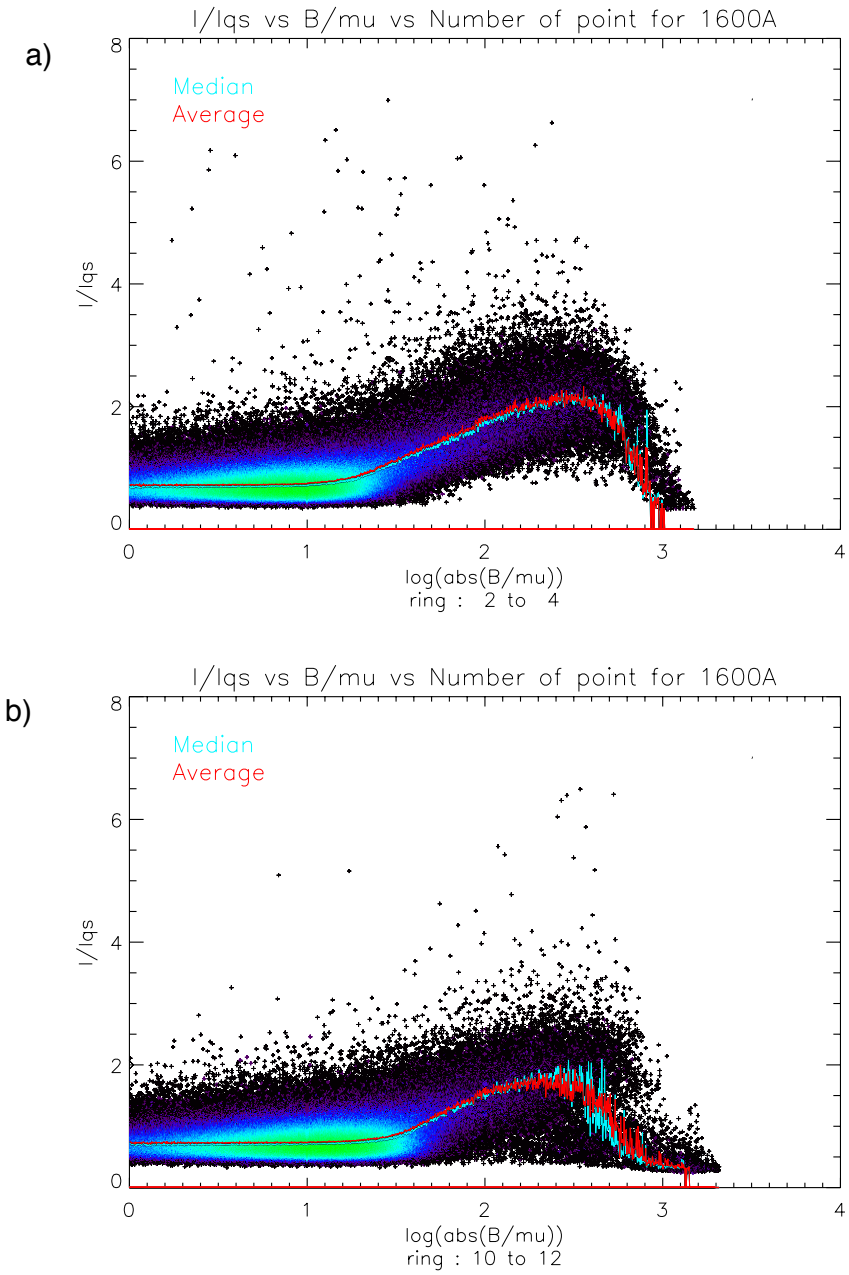

Figure 4.8: Magnetic field intensity (from SDO/HMI magnetograms, corrected for the line-of-sight projection effect) versus the irradiance intensity contrast (from HMI/AIA images) at 160 nm. Panel a: disk centre. Panel b: near the limb. Picture taken from [Roger 2014]

company [Tobiska 2000].

Our single timescale model can be converted into an operation tool provided some additions are made. The model provides estimates of the daily average SSI expressed in 1 nm bins, in the 6 - 300 nm range (with exception of the intense in the UV), and the MgII line at 280 nm based on SDO/HMI magnetograms. In the EUV band, for several of the intense lines between 36 nm and 120 nm, the model calibration is too strongly affected by the low quality of the observations from TIMED/SEE. To estimate the average SSI on a given day a single magnetogram is needed, which is taken in the middle (at about 12:00 UT) of the corresponding day. The model cadence can be easily increased up to several reconstructions per day. Higher cadences however are not recommended, as our model is not capable of reproducing the short timescale variability driven by flares, granulation, etc.

The main requirement for an operational service is robustness, namely, resilience to outages or corruptions in the input images from SDO/HMI. Another important aspect is providing confidence intervals for the estimates of the SSI.

This involves the following issues:

- Changes in instrument calibration. Here we focus on the simple case wherein changes in the calibration of instruments are known, and are introduced by the instrument teams. In this case the model has to be re-calibrated by using the latest version of the data, and the information about the data version has to be stored. When different instruments that provide the SSI measurements overlap in time or spectral coverage (or in both), the model has to be trained independently for each dataset, because there is no clear answer about the reliability of the data from different instruments. An alternative solution would be to use composite SSI timeseries such as provided by the SOLID project [http://projects.pmodwrc.ch/solid/] which aims at delivering a single continuous SSI record along with its uncertainties by combining data from different instruments and models.

  Our model can also be used to spot anomalous behaviour in the SSI measurements. In the case when the calibration of the magnetograms changes the model also has to be re-calibrated by using the procedure for refining magnetic field intensity threshold value B_{AR} described in Section 3.4.

- Migration from one instrument to another. An example of such a case might be a change from SDO/HMI to SOHO/MDI magnetograms. This is a more complex case, as the noise properties, calibration and instrument resolution change simultaneously. To handle this situation there is a need to repeat the steps described for the previous case. However, changes in the spatial resolution of magnetograms may affect the definition of the magnetic structures that are used in the model. Thus, the whole procedure of the model selection that derives the optimal classes for the SSI reconstruction has to be repeated.

- No observations available. In the case when no magnetograms are available (for example, during one of the few solar eclipses encountered by SDO twice

per year for three weeks near the equinox and which last for a period of time each day) then the area coverage of magnetically active regions cannot be evaluated and consequently, the model cannot produce the estimation of the SSI. One possible solution here is to switch to an empirical proxy-based model. For this purpose some solar UV indices might be incorporated, such as the Mg II index, or the the radio flux at 10.7 cm, or at 30 cm. Many empirical models use these proxies and their performance have been discussed in [Schmidtke 1976, Judge 2012, Dudok de Wit 2014]. Another possibility consists in using a flux transport model [Wang 1991, Dikpati 2004] to extrapolate the magnetic field evolution by producing the synthetic magnetograms that cover the time gap. However, this solution can give a reliable SSI estimates no longer then 1 month without magnetograms.

Since the time interval used for the model training can not be yet considered as representative for a full solar cycle, our model calibration will need to be improved progressively in time as new SSI data come in. Here we assume that both the SSI measurements and the magnetograms will come from the same source, with no calibration issues for the years to come. If an undocumented change occurs in the data, then we need to be able to detect this.

To determine if a model re-calibration is needed we suggest a following workflow: firstly, the estimate of the SSI for the time period that was not used for the model calibration has to be obtained based on the current model state, and the corresponding quality estimate (NRMSD_N) has to be calculated. Now we want to check if this particular value of the NRMSD_N is likely to be observed, based on the information that has already been ingested in our model. If the answer is positive, then there is no need to refine the model with new SSI data, since these do not bring new information. In the opposite case, the model has to be re-calibrated by using the extended SSI dataset (which includes new SSI data).

To decide if the particular NRMSD_N is likely to be observed, we study how the NRMSD is expected to be distributed, based on the data that has been used for the model calibration. For this purpose we compute the NRMSD values for randomly located time windows of different size. Thus, for each window size we obtain a corresponding distribution of the NRMSD values. The mean values along with the standard deviations are plotted as function of the window size for the typical case of the Lyα line at 121 nm, see Figure 4.9.

By inspecting this plot we note that the NRMSD value is large for very short windows but rapidly decrease to about 0.44 as the window size reach values about 60 days. The NRMSD then continues to decrease at a much slower rate. The corresponding standard deviation of the NRMSD also decrease with a growing window size, but more slowly. With this we see that only about two months of observations are needed to train the model to match the observations (although this does not imply that the model is representative of longer timescales).

To decide if the model recalibration is needed for a particular observed NRMSD_N value we check if it is within the uncertainty of the corresponding expected value.

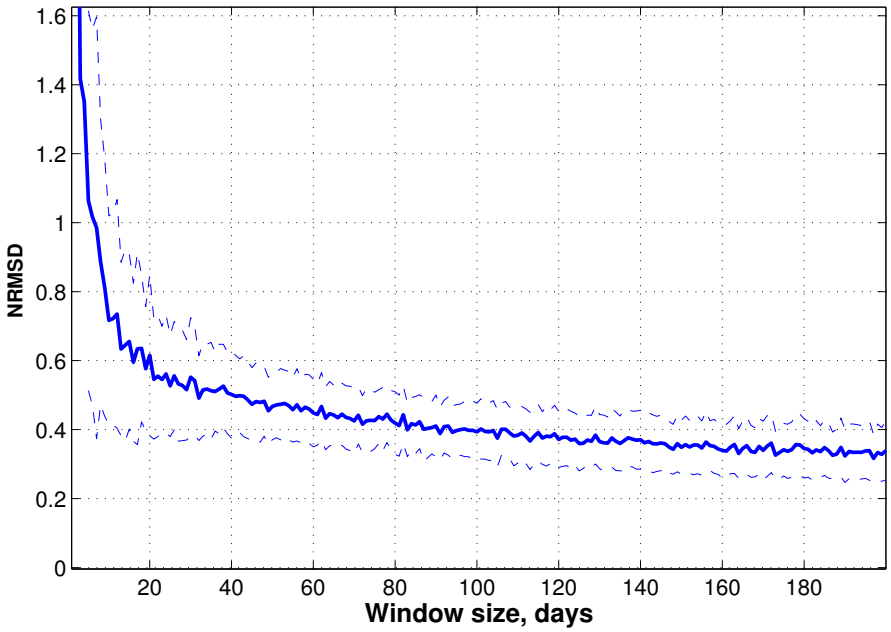

Figure 4.9: The NRMSD value for the SSI at 121 nm, evaluated based on the training dataset with respect to the window size. The mean value is plotted as a solid blue line, the dashed blue lines correspond to the $\pm 1\sigma$ level.

For example, if the observed NRMSD_N value is 0.6 and the size of the new data chunk is not larger then 44 days we do not have enough evidence to conclude that the re-calibration is needed. However, if the same NRMSD_N value is observed with the larger data chunk (>44 days) then the model recalibration is needed.

The two-timescale model, which is more relevant for the XUV/EUV bands, has one additional problem regarding operability. A causal filter is mandatory for separating the long and short timescales. The one we used so far requires past and future values, and thus is non-causal. There are two possibilities to overcome this problem; one consists in using a causal filtering technique [Press 2007] which is specifically developed to handle such a case.

The other possibility is to use a surface flux transport model to predict the evolution of the magnetic field, generate synthetic magnetograms for the future time period (about 30 days ahead), and then use a non-causal filtering technique to separate the short and the long timescale components of the time evolution of the area coverage of magnetic structures. The second solution is more favorable, as it accounts for the evolution of the magnetically active regions.

4.7 Conclusions

Despite of the partial statistical redundancy of the input data, that are the area coverage timeseries of the magnetic structures, our model can be considered to be unique within the uncertainty of the SSI observations. The uncertainty of the SSI estimates provided by our model is wavelength-dependent and is typically about 1.5-2.2% of the solar cycle variability. However, the model stability on the long term remains an open question, and can be fully understood only once an extended time interval of at least 1 solar cycle is available. The latter can be also achieved by going backward in time, and incorporating magnetograms from the SOHO/MDI instrument, thus extending the model back to 1996. However, the full coverage of the UV band can be achieved only starting from 2003 (start of the SORCE mission).

The performance of the model is timescale-dependent. Thus on the short timescales (solar rotation), the model explains about 40% of the observed variability, whereas for the long timescales this number reaches about 90-95%. However, as pointed out above, the model performance on long timescales is most likely to be overestimated.

The two-timescale model improves both the short and the long timescale performance of the model in the EUV/XUV band accounting for the off-limb contribution. In the MUV/FUV band at wavelengths above 180 nm, the two-timescale model improves only on the long timescales, thus pointing to the possible problems with the long term stability of the SORCE/SOLSTICE observations.

The model performance on the short timescales (solar rotation and shorter) might be improved by increasing the model cadence to several observations per day.

A comparison of the classes of magnetic structures derived from our model, and EUV intensity images gives a physical justification that the small and the large magnetic structures indeed have different spectral profiles.xs Also, this comparison suggests that the spectral profile of the small magnetic structures scales with their size (linearly as the first order approximation), and that the extraction of the magnetically active regions can be improved by applying threshold a value that depends on the distance from the solar disk centre.

4.8 Résumé en Français

Dans ce chapitre, nous discutons plus en détails certains aspects de notre modèle, comme sa stabilité et les données sur lesquelles ils reposent, qui sont importants pour faire évoluer le modèle vers un service opérationnel de prédiction de l'irradiance solaire. En particulier, nous étudions comment les différentes sources d'incertitude affectent la stabilité et la qualité de la reconstruction sur différentes échelles de temps. La question clé que nous abordons ici est de savoir dans quelles mesures notre modèle reproduit les vraies variations de l'irradiance solaire dans la bande UV, étant données les incertitudes sur les mesures d'irradiance d'une part et sur le champ magnétique d'autre part.

Comme l'évolution temporelle des différentes structures magnétiques est en très grande partie cohérente, nous pourrions nous attendre à ce que la matrice contenant les observations d'irradiance soit potentiellement mal conditionnée et conduise ainsi à une estimation instable des coefficients du modèle, et/ou à une sensibilité de ces coefficients aux bruits instrumentaux figurant dans les données. Cependant, nous avons calculé le nombre de condition de la matrice et trouvé une valeur typiquement autour de 10-20 qui indique que le mauvais conditionnement n'est pas un problème.

L'incertitude sur les séries d'irradiance calculées par notre modèle dépend de la longueur d'onde et est typiquement d'environ 1.5 à 2.2% de la variation au cours du cycle solaire. La stabilité du modèle sur le long terme reste une question ouverte, et ne pourrait être pleinement décrite qu'avec des observations couvrant au moins 1 cycle solaire. Une telle étude pourrait éventuellement être effectuée en incorporant les magnétogrammes de l'instrument SOHO/MDI, étendant ainsi le modèle jusque en 1996, mais les observations d'irradiance dans l'UV ne remonte qu'à 2003 (lancement de la mission SORCE). Pour la période 1996-2003, les seules observations d'irradiance disponibles viennent de la mission (alors vieillissante) UARS et de SOHO/CDS, qui ne couvrent pas entièrement le domaine spectral qui nous intéresse.

Parce que notre modèle est piloté par les données, il est par ailleurs sensible aux différents artefacts présents dans les magnétogrammes et dans les observations d'irradiance. Les mesures d'irradiance sont connues pour souffrir d'une forte dégradation de la réponse instrumentale, qui doit donc être correctement corrigée avant de fournir les données à la communauté. Même avec cette correction, et dans le cas de très forte dégradation, le rapport signal-sur-bruit peut devenir très faible, de sorte que la variabilité présente dans les données soit dominée par le bruit instrumental (ce qui est par exemple le cas pour les raies EUV intenses observées par TIMED/SEE après plusieurs années d'opération). Dans de tels cas, notre modèle reste relativement stable mais ne peut pas reproduire le bruit de l'instrument à toutes les longueurs d'onde, et nous trouvons ainsi un désaccord significatif entre l'ajustement du modèle et les observations.

Lorsque la dégradation est moins forte, la dérive instrumentale résiduelle (non corrigée) affecte la performance du modèle sur les longues échelles de temps (cycle de 11 ans), comme cela est probablement le cas pour SORCE/SOLSTICE. En particulier, cela peut entraîner une reconstruction de l'irradiance par le modèle qui est biaisée sur le long terme et qui rend donc notre modèle pas encore pleinement représentatif pour le cycle de 11 ans.

La répartition du bruit dans les magnétogrammes de SDO/HMI est étudiée en détail par [Yeo 2014] et ne devrait pas avoir un impact significatif sur notre modèle. Toutefois, nous avons effectué une comparaison entre les images prises par SDO à différentes longueurs d'onde dans l'ultraviolet lointain et les magnétogrammes qui suggère que l'extraction des régions actives dans les magnétogrammes doit être fait en utilisant un seuil qui dépend de la distance au centre du disque solaire, en accord avec les travaux cités ci-dessus.

La performance du modèle dépend de l'échelle de temps considérée. Ainsi, à l'échelle de la rotation solaire, le modèle explique environ 40% de la variabilité

observée, alors que pour les longues échelles de temps ce nombre atteint environ 90-95%. Toutefois, comme indiqué ci-dessus, la performance du modèle sur les longues durées est plus susceptible d'être surestimée en raison des dérives instrumentales.

La version du modèle qui distingue deux échelles de temps améliore à la fois la performance du modèle à l'échelles de la rotation et du cycle dans la bande EUV/XUV, car elle capture la contribution des régions actives au dessus du limbe. Dans la bande MUV/FUV, aux longueurs d'onde supérieures à 180 nm, le modèle à deux échelle de temps présente une amélioration seulement pour les longues échelles de temps, indiquant ainsi de potentiels problèmes de stabilité instrumental à long terme pour SORCE/SOLSTICE.

La performance du modèle sur les petites échelles de temps (rotation solaire et inférieures) pourrait être améliorée en augmentant la cadence du modèle à plusieurs observations par jour.

Enfin, la comparaison des classes de structures magnétiques définies par notre modèle avec les images du disque solaire dans l'EUV donne une justification physique au fait que les petites et grandes structures magnétiques aient des profils spectraux différents. En outre, cette comparaison suggère que le profil spectral des petites structures magnétiques évolue linéairement (au premier ordre) avec leur taille, et à nouveau que l'extraction des régions actives pourraient être améliorée pen appliquant une valeur de seuil qui dépend de la distance au centre du disque solaire.

Chapter 5

Conclusions

In this thesis we have developed an empirical model that allows the reconstruction of SSI variations in the UV band, based on solar magnetograms only. Similarly to semi-empirical models such as SATIRE-S [Fligge 2000], our model expresses this variability in terms of a sum of contributions coming from different classes of photospheric magnetic structures, each of which has a time-independent spectral intensity profile. In contrast to semi-empirical models, however, both the number and the characteristics of these classes, and also the spectral intensity profiles, are derived empirically from the observations. In this sense, our model is data-driven.

We trained and calibrated our model by using 3 years of SSI observations from the TIMED/SEE, SORCE/SOLSTICE and SDO/EVE instruments. We configured the model to deliver daily-averaged values of the SSI with 1 nm spectral resolution, covering the 6-300 nm range. This range is particularly relevant for space weather applications that require a specification of the thermosphere-ionosphere system: orbit control of low Earth orbiting satellites, ground-ground and ground-space telecommunications, positioning, and also climate studies.

Because our model is data-driven, it is less affected by some of the limitations of semi-empirical models such as the departure from non-local thermodynamic equilibrium, which makes the theoretical modelling of the spectral irradiance below 300 nm very challenging. Furthermore, recent observations from SORCE have revealed unusually large variations in the UV, which are not reproduced by most models [Ermolli 2013]. In this context, there is a clear incentive for modelling the SSI based on observations only, until the cause of the discrepancy between the observations and the models is settled.

The performance of our model is comparable to that of NRLSSI, which is the most widely model used in the UV band (see Figure 5.1). Our metric for model performance is the root mean square of the out-of-sample error, normalised to the standard deviation of the observations (i.e. the solar cycle variability). For most wavelengths, this quantity lies between 0.3 and 0.41, which means that the model explains about 59 - 70% of the solar cycle variability.

Let us now consider some more specific issues and results:

- Model limitations.

  Let us first stress again that our model is data driven, and consequently it is only as good as the observations are. The decision to train it with observations only was a deliberate one, and our limitation to observations made

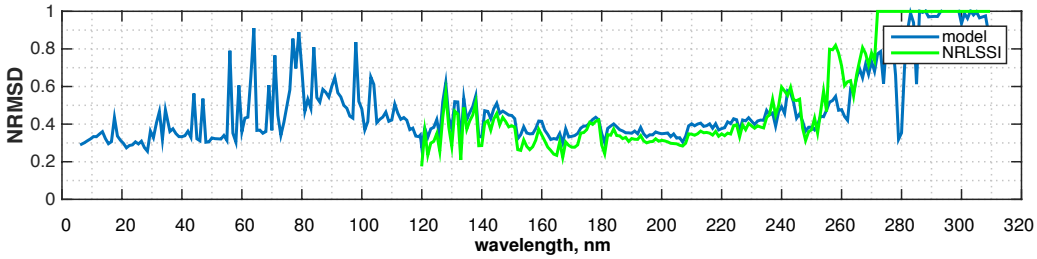

Figure 5.1: Performance of our model (blue) compared to the performance of the NRLSSI model (green) at different wavelengths. Our model on average performs 3-5% worse, however it should be noted that we transformed NRLSSI reconstructions by linear rescaling in order to get better agreement with SORCE observations.

after 2010 was dictated by the availability of magnetograms from SDO which became available after April 2010 only. The controversy regarding the possible existence of an uncorrected instrumental trend in the recent observations from SORCE/SOLSTICE [Deland 2012a] may indeed require future revision of our model parameters in the MUV band. The short duration of our training set also questions the representativity of our conclusions, and whether our model remains applicable to a full 11-year cycle. Only longer observations may enable us to answer that question in a decisive way. Meanwhile, since our model uses solar inputs that are similar to models such as SATIRE-S, which have been duly validated over several solar cycles [Yeo 2014], we may a priori expect its performance to remain good on longer time spans.

Although our model can easily deliver spectra with a sub-daily resolution, cadences faster than 3-6 hours are not recommended. Indeed, at such short timescales, small changes in the photospheric magnetic field tend to be hidden by random fluctuations that are caused by the discretisation of the magnetogram segmentation procedure. Shorter transients such as flares are definitely beyond reach because their irradiance cannot be properly described by the properties of the surface magnetic field. For such events, which do matter for many space weather applications, dedicated models exist, such as FISM [Chamberlin 2008].

- What does our model tell us about the underlying physics?

  Our model suggests that two classes only of magnetic structures are needed to properly describe the salient spectral variations in the UV, given the uncertainties in the observations. Additional classes are likely to be needed, because there are definitely more than two degrees of freedom in the spectral variability. However, our statistical analysis clearly shows that additional model complexity does not lead to better reconstructions.

  We distinguish the magnetic structures according to their size, and call them large (size $>$ 512 arc sec^2) and small (size $32 - 512$ arc sec^2) structures.

Magnetic structures smaller than 32 arc sec^2 are not found to be informative enough for being necessary for reconstructing the irradiance. This conclusion may of course change when longer records become available, as small-scale magnetic features might contribute to secular trends in the SSI [Fontenla 2011].

Although our classes of magnetic structures cannot be rigorously compared to the structures that are used in semi-empirical models (i.e. sunspots, plages, and network) we find evidence that the large structures correspond to plages, and the small structures to faculae. To asses the importance of the sunspots for describing SSI variations in the UV we also extracted the area coverage of sunspots from continuum images from the SDO/HMI instrument. By comparing the performance of our model with and without the sunspots we conclude that the latter do not significantly improve the SSI observations in the UV.

In addition, we find that three principal positions of the magnetic structures on the solar disk suffice for describing the centre-to-limb variation of UV emissions. Most semi-empirical models in comparison use between 7 and 11 positions.

The one-to-one relation between the solar surface magnetic field and SSI variations suggests that this relation should hold at all timescales. Surprisingly, we find that a two-timescale model offers a better performance, with a reduction of up to 0.1 in the NRMSD value. These two timescales typically correspond to solar rotational variations, and to the 11-year solar cycle variations. The gain in performance is most marked in the XUV/EUV bands, which are characterised by optically thin emissions that have a significant off-limb contribution. Our model thereby offers a way to estimate this off-limb contribution. We find that it represents about 12-40% of the spectral irradiance, depending on the wavelength.

- Can the model be used to diagnose SSI observations?

  Although it seems difficult to test with our model the SSI data it has been trained with, because of inbreeding, some sanity checks are nevertheless possible. In particular, we find that some bands with optically thick emissions are significantly better reproduced with a two-timescale model. Since there is no sound physical justification for this improvement, we ascribe it to the presence of an uncorrected trend in the observations. Interestingly, this improvement is most evident in the 170-190 nm, and 230-250 nm bands, in which SORCE/SOLSTICE is likely to suffer from anomalous trends [DeLand 2012b, Morrill 2014].

- Making SSI forecasts.

  Our model works in near-real time, because the processing of a magnetogram takes a few seconds only, once the solar data have been downloaded. Many

space weather users, however, also require forecasts. For satellite orbit prediction, for example, the prediction horizons typically range from 3 to 30 days.

Spectral variations can be predicted in two different ways with our model. For short term predictions (typically 1-3 days ahead), empirical forecasts of the SSI give reasonable results with a modest computational investment [Lean 2009]. For longer prediction horizons, however, physical insight becomes mandatory. The most promising approach to date consists in modelling the evolution of the surface magnetic field with flux transport models [Jiang 2014]. Prediction horizons of about one month have been achieved that way. Of course, the forecast of flares is beyond reach. The dynamics of rapid flux emergence can also become a challenge without access to the solar interior by means of helioseismology.

- Roads for improvement.

  Our model can be improved in several ways. As we have repeatedly mentioned, the model is not yet representative of a 11-year cycle. A first step would be to extend it backwards in time until 1996 by using SOHO/MDI data as input, and SSI observations from various missions such as UARS (excluding the EUV range). In doing so, we would also cover solar cycle 23.

  However, because SOHO/MDI has different noise characteristics, and is calibrated differently, our model would need to be retrained. The spatial resolution of SOHO/MDI instrument is 2 arc sec^2. Although this is 4 times less than for SDO/HMI, it remains largely sufficient for our purposes because the size of the smallest magnetic structures we need to extract exceeds 32 arc sec^2.

  The spectral coverage of our model can be extended to longer wavelengths, to encompass both the visible and near-infrared bands. Conceptually, there is no difference. Doing so, however, requires information about sunspot area and location in order to include sunspot darkening effects. This not only requires a third class of structures, but also access to solar continuum images [Lean 1988].

  Because our model can work in near real-time, it is also well suited for operational services. The two main issues are the provisioning and the filtering of the data. By provisioning, we mean the existence of backup solutions in the case when magnetograms are not accessible. This may happen during periodic eclipses of the SDO spacecraft, or during severe solar proton events, when the data become too corrupted. One solution consists in relying on alternate observations such as magnetograms from the ground-based Kit Peak Observatory. A more pragmatic solution, however, consists in using ground-based solar proxies such as the f10.7 index for a temporary backup. A Kalman filter [Kalman 1960] could then be used to define a good strategy to combine data from proxies with past SSI reconstructions (i.e. persistence scenarios).

A second but minor problem is the filtering of the observations, in which case causal filters need to be applied.

In saying this, we assume that SSI observations are not interrupted, so that the model training can continue over a larger time interval. Although this is not mandatory, it is highly recommended for enabling the model to be further tested and validated. Unfortunately, the SORCE spacecraft, after a decade of successful operations, is now nearing the end of its lifetime. No backup observations are expected for several years, except in the MUV, and in some bands of the EUV. This means that our model may soon need to rely only on composite data that consist of a mix of proxies, direct SSI observations, and even SSI model data.

An important issue regarding the operation of our model at time scales below one day is the high frequency noise that is generated by the segmentation procedure. To improve the reconstruction quality of the short timescales a fuzzy-clustering technique is recommended because it provides a more gradual transition between classes. The parameters and the shape of the membership functions near the edges of the classes has to be chosen to offer the best model performance.

The model performance in the XUV/EUV can be further improved. The three-dimensional coronal loops cannot be fully described by the photospheric magnetic field, and moreover the description of the limb region cannot be obtained at all. We repeatedly found that the description of solar magnetograms near the solar limb is critical for reconstructing the the short-term variability in the UV, especially for optically thin lines in the XUV-EUV bands. These issues can be alleviated by making use of the coronal EUV images from SDO/AIA and SOHO/EIT instruments. However, this requires a different image segmentation procedure, and raises the question about the link to the solar surface magnetic field.

A comparison between our model and other semi-empirical models (e.g. SATIRE-S) has to be done in order to explore deeper the weaknesses and advantages of our approach. In particular the connection between empirical classes of magnetic structures that we derive in our model and atmospheric structures that are used in different semi-empirical models can be established by training our model to reproduce SSI reconstructions obtained by correspond models.

Connection between our empirical classes of magnetic structures and atmospheric structures that are observed in the UV images can be established by comparing segmentation masks that we use in our model with the UV images from SDO/AIA instrument that correspond to optically thick emissions (160 nm and 170 nm).

Chapter 6

Conclusions en Français

Dans ce mémoire, nous avons présenté un modèle empirique nouveau dédié à la reconstruction des variations d'irradiance spectrale solaire (SSI) dans la bande UV. Ce modèle recourt uniquement à des magnétogrammes solaires; comme dans des modèles semi-empiriques tels que SATIRE-S [Fligge 2000], il exprime cette variabilité en termes de contributions associées à différentes classes de structures magnétiques de la photosphère, dont chacune possède un profil spectral invariant au cours du temps. Toutefois, contrairement aux modèles semi-empiriques, le nombre et les caractéristiques de ces classes, ainsi que les profils spectraux, sont ici déduits des observations. En ce sens, notre modèle est empirique, et piloté par les observations.

L'apprentissage et la calibration de notre modèle ont été faits sur 3 années d'observations issues des missions TIMES/SEE, SORCE/SOLSTICE et SDO/EVE. Notre modèle fournit actuellement une moyenne journalière de la SSI avec une résolution spectrale d'un nm, couvrant la gamme 6-310 nm. Cette gamme est pertinente pour les applications en météorologie spatiale qui requièrent une spécification du système thermosphère-ionosphère: contrôle d'orbite pour des satellites situés en orbite basse, communications sol-sol et sol-espace, positionnement, ainsi que l'étude du climat.

Comme notre modèle est dérivé des observations, il est relativement peu affecté par certaines des contraintes propres aux modèles semi-empiriques. Il s'agit notamment de l'absence d'équilibre thermodynamique local en dessous de 121 nm, ce qui rend la modélisation théorique ces longueurs d'onde plus délicate. Par ailleurs, des observations récentes de SORCE ont révélé des variations anormalement fortes dans l'UV, qui ne sont pas reproduites par la plupart des modèles [Ermolli 2013]. Dans ce contexte, la modélisation empirique devient particulièrement pertinente, en attendant que la l'origine de cette variation anormale soit comprise.

La performance du modèle est comparable à celle du modèle empirique NRLSSI, qui est couramment utilisé pour décrire la bande UV. En effet, NRLSSI est une des rares modèles à couvrir l'intégralité de la bande UV, contrairement aux modèles tels que SATIRE-S, qui ne décrivent pas la composante EUV.

Nous utilisons comme métrique de la performance du modèle la moyenne quadratique de l'erreur sur l'ensemble de test, normalisée par rapport à l'écart-type des observations (c'est-à-dire la variabilité du cycle solaire). Pour la plupart des longueurs d'onde, cette quantité est comprise entre 0.3 et 0.41; cela signifie que notre modèle explique environ 59 à 70 % de la variabilité du cycle solaire.

Concentrons-nous maintenant sur quelques aspects précis:

- Les limites de notre modèle.

Rappelons d'abord que l'apprentissage de notre modèle est basé sur des observations seules. Dans ce sens, sa qualité reflète celle des observations.

La recours à des observations postérieures à 2010 est une contrainte forte, qui est hélas imposée par la disponibilité des magnétogrammes de SDO. Or la recentre controverse sur l'existence éventuelle d'une dérive instrumentale non corrigée dans les observations de SORCE/SOLSTICE [Deland 2012a] pourrait avoir pour conséquence la nécessité de devoir un jour reprendre à zéro l'apprentissage de notre modèle dans la bande MUV.

Les trois années de notre échantillon d'apprentissage pourraient mettre en doute la capacité de notre modèle à décrire un cycle complet de 11 ans. Seule la mise à disposition future d'observations plus longues nous permettra de répondre à cette question de manière définitive. En attendant, d'autres modèles, qui sont eux aussi basés sur des magnétogrammes, ont été dûment testés et validés sur plusieurs cycles solaires [Yeo et al. 2014]. Ceci conforte l'idée selon laquelle les performances de notre modèle resteraient valables sur des durées plus longues. Notre modèle peut aisément fournir des spectres avec une résolution temporelle inférieure au jour. Néanmoins, les cadences inférieures à 3-6 heures ne sont pas recommandées. En effet, à ces courtes échelles de temps, les petites variations du champ magnétique photosphérique sont noyées par les fluctuations aléatoires engendrées par le procédure de discrétisation des magnétogrammes. Les transitoires plus brefs tels que les éruptions sont évidemment hors de portée car leur irradiance spectrale ne peut être décrite à partir des propriétés des magnétogrammes. Il existe pour les éruptions des modèles dédiés tels que FISM [Chamberlin 2007].

- Que dit notre modèle sur la physique sous-jacente? Notre étude montre que deux classes de structures magnétiques suffisent à décrire correctement les variations spectrales dans l'UV, compte tenu des incertitudes dans les observations. Le nombre de degrés de liberté de la variabilité spectrale est très probablement supérieur à deux; néanmoins, un modèle plus complexe avec des classes supplémentaires ne conduit pas à une amélioration significative du spectre.

  Nous distinguons les structures en fonction de leur taille, que nous appelons grande (taille $>$ 512 arc sec^2) et petite (taille de 32 à 512 arc sec^2). Les structures magnétiques de taille inférieure à 64 secondes d'arc ^{2} ne contiennent pas assez d'information pertinente pour justifier leur inclusion dans le modèle. Nous pourrions être amenés à revoir cette conclusion lorsque des enregistrements plus longs deviendront disponibles, d'autant plus que les structures de petite échelle pourraient contribuer aux variations séculaires de la SSI [Fontenla 2011].

  Même si nos classes de structures magnétiques ne peuvent pas être directement comparées aux structures extraites dans les modèles semi-empiriques (régions

actives, taches solaires, plages, etc.), nous trouvons les grandes structures correspondent davantage à des plages et les petites structures aux facules. Nous avons aussi évalué l'importance des taches solaires en recourant à des images de SDO/HMI prises dans le domaine visible pour extraire correctement l'aire des taches. En comparant la performance de notre modèle avec et sans taches solaires, nous pouvons conclure que la contribution des taches n'est pas essentielle et n'améliore pas de façon significative la reconstruction dans l'UV.

En outre, nous constatons que trois postes principaux des structures magnétiques sur le disque solaire suffisent pour décrire la variation des émissions UV centre à membre. La plupart des modèles semi-empiriques de comparaison utilisent entre 7 et 11 positions.

La relation un-à-un entre le champ magnétique de la surface solaire et les variations SSI suggère que cette relation devrait tenir à toutes les échelles temporelles. Étonnamment, nous constatons qu'un modèle à deux calendrier offre une meilleure performance, avec une réduction de 0,1 dans l'erreur de modèle normalisé. Ces deux échelles de temps correspondent généralement à des variations de rotation solaire, et aux variations du cycle solaire. Le gain de performance est plus marquée dans les groupes/EUV XUV, qui sont caractérisés par des émissions optiquement minces qui ont une contribution significative hors-membre. Notre modèle offre ainsi un moyen d'estimer ce hors-branche contribution. Nous trouvons qu'il représente environ 12-40 % de l'éclairement spectral, en fonction de la longueur d'onde.

- Notre modèle peut-il servir à diagnostiquer les observations ?

  Bien qu'il semble difficile de tester notre modèle les données SSI il a été formé avec, à cause de la consanguinité, certains contrôles sont néanmoins possibles. En particulier, nous constatons que certains groupes dont les émissions optiquement épaisses sont nettement mieux reproduits avec un modèle à deux délais. Comme il n'y a aucune justification physique pour décrire cette amélioration, nous attribuons à la présence d'une tendance non corrigée dans les observations. Fait intéressant, cette amélioration est la plus ev identin le 170-190 nm, et 230-250 nm bandes, dans lequel SORCE/SOLSTICE expose précisément les tendances anormalement fortes [DeLand 2012b, Morrill 2014].

- Vers des prévisions de la SSI. Notre modèle fonctionne en temps quasi-réel, parce que le traitement d'une magnétogramme prend que quelques secondes, une fois que les données solaires ont été téléchargés. Beaucoup d'utilisateurs de la météo spatiale, cependant, nécessitent également des prévisions. Pour la prédiction de l'orbite du satellite, par exemple, les horizons de prédiction varient généralement de 3 à 30 jours.

  Variations spectrales peuvent être prédits de deux manières différentes avec notre modèle. Pour prévisions à court terme (généralement 1-3 jours à l'avance),

les prévisions empiriques de la SSI donnent des résultats raisonnables avec un modeste investissement de calcul [Lean 2009]. Pour les longs horizons de prédiction, cependant, la perspicacité physique devient obligatoire. L'approche la plus prometteuse à ce jour consiste à modéliser l'évolution du champ magnétique de surface avec des modèles de transport de flux [Jiang 2014]. Horizons de prévision de un mois ont été réalisés de cette façon. Bien sûr, la prévision des éruptions est hors de portée. La dynamique d'émergence des flux rapide est également un défi majeur sans accès à l'intérieur du Soleil au moyen d'héliosismologie.

- Les améliorations possibles. Le modèle peut être amélioré de diverses façons. Comme nous l'avons mentionné à plusieurs reprises, ce modèle n'est pas encore représentative d'un cycle de 11 ans. Une première étape serait d'étendre vers l'arrière dans le temps jusqu'en 1996 en utilisant des données SOHO/MDI en entrée, et les observations SSI de diverses missions telles que UARS (à l'exclusion du domaine EUV). Ce faisant, nous souhaitons également couvrir cycle solaire 23. Toutefois, en raison SOHO/MDI présente des caractéristiques différentes de bruit, et est calibré différemment, notre modèle aurait besoin d'être recyclés. La résolution spatiale de l'instrument SOHO/MDI est de 2 arc sec^2. Bien que ce soit 4 fois moins que pour SDO/HMI, il reste largement suffisant pour nos besoins parce que la taille des structures magnétiques plus petits nous avons besoin d'extraire plus de 32 arc sec^2.

  La couverture spectrale de notre modèle peut être étendu à des longueurs d'onde, englobant à la fois les bandes visibles et proche infrarouge. Conceptuellement, il n'y a aucune différence. Ce faisant, toutefois, nécessite des informations sur la zone de taches solaires et l'emplacement afin d'y inclure les effets de taches solaires assombrissement. Ce ne requiert pas seulement une troisième classe de structures, mais aussi l'accès à des images du continuum solaires [Lean 1988].

  Parce que notre modèle peut fonctionner en temps quasi-réel, il est aussi bien adapté pour les services opérationnels. Les deux questions principales sont l'approvisionnement et le filtrage des données. En approvisionnement, on entend l'existence de solutions de sauvegarde dans le cas où magnétogrammes ne sont pas accessibles. Cela peut se produire pendant les éclipses périodiques de l'engin spatial SDO, ou au cours de rompre les événements de protons solaires, lorsque les données deviennent trop endommagés. Une solution consiste à s'appuyer sur les observations de rechange tels que magnétogrammes de l'Observatoire du Mont Wilson au sol. Une solution plus pragmatique, cependant, consiste à utiliser des proxys solaires au sol tel que l'indice de F10.7 pour une sauvegarde temporaire. Un filtre de Kalman pourrait alors être utilisée pour définir une bonne stratégie pour combiner les données de procurations avec dernières reconstructions SSI (ie scénarios de persistance). Un deuxième problème est mineur, mais le filtrage des observations, dans lequel les filtres cas de causalité doivent être appliquées.

En disant cela, nous supposons que les observations SSI ne sont pas interrompues, de sorte que la formation de modèle peut continuer sur un plus grand intervalle de temps. Bien que ce ne soit pas obligatoire, il est fortement recommandé pour permettre au modèle d'être encore testé et validé. Malheureusement, le vaisseau spatial SORCE, après une décennie de succès des opérations, est maintenant presque la fin de sa durée de vie. Aucune observation de sauvegarde sont prévues pour plusieurs années, à l'exception de la VUM, et dans certaines bandes de l'EUV. Cela signifie que notre modèle pourrait bientôt besoin d'être formés en outre des données composite qui se compose d'un mélange de procurations, des observations directes SSI, et même les données du modèle SSI.

Une question importante en ce qui concerne le fonctionnement de notre modèle à des échelles de temps inférieures à un jour est le bruit haute fréquence générés par la procédure de segmentation. Pour améliorer la qualité des délais courts de reconstruction d'une technique floue regroupement est recommandée car elle permet une transition plus graduelle entre les classes. Les paramètres et la forme des fonctions d'appartenance à proximité des bords des cours doit être choisi d'offrir la meilleure performance du modèle.

La performance du modèle dans l'XUV/EUV peut être encore améliorée. Les boucles coronales en trois dimensions ne peuvent pas être décrits en détail par le champ magnétique photosphérique, et d'ailleurs de la description de la région de la branche ne peut être obtenu du tout. Ces problèmes peuvent être atténués par l'utilisation des images EUV coronales de SDO/AIA et instruments EIT sur SOHO. Toutefois, cela nécessiterait une procédure de segmentation d'image complètement différente, et s'élève à la question sur le lien avec le champ magnétique de la surface solaire.

Bibliography

[Aschwanden 2004] M.J. Aschwanden. Physics of the solar corona: An introduction. Springer Praxis Books. Springer, 2004. (Cited on page 13.)

[Bowman 2008] B. R. Bowman, W. Kent Tobiska, F. A. Marcos and C. Valladares. *The JB2006 empirical thermospheric density model*. Journal of Atmospheric and Solar-Terrestrial Physics, vol. 70, pages 774–793, March 2008. (Cited on page 17.)

[Bruinsma 2012] S. L. Bruinsma, N. Sánchez-Ortiz, E. Olmedo and N. Guijarro. *Evaluation of the DTM-2009 thermosphere model for benchmarking purposes*. Journal of Space Weather and Space Climate, vol. 2, no. 27, page A4, June 2012. (Cited on page 17.)

[Budden 1988] K.G. Budden. The propagation of radio waves: The theory of radio waves of low power in the ionosphere and magnetosphere. Cambridge University Press, 1988. (Cited on page 18.)

[Chamberlin 2007] P. C. Chamberlin, T. N. Woods and F. G. Eparvier. *Flare Irradiance Spectral Model (FISM): Daily component algorithms and results*. Space Weather, vol. 5, page 7005, July 2007. (Cited on pages 14, 16 and 102.)

[Chamberlin 2008] P. C. Chamberlin, T. N. Woods and F. G. Eparvier. *Flare Irradiance Spectral Model (FISM): Flare component algorithms and results*. Space Weather, vol. 6, page 5001, May 2008. (Cited on page 96.)

[Cheney 2007] E. Cheney and D. Kincaid. Numerical mathematics and computing. International student edition. Cengage Learning, 2007. (Cited on page 49.)

[Christensen-Dalsgaard 2002] J. Christensen-Dalsgaard. *Helioseismology*. Reviews of Modern Physics, vol. 74, pages 1073–1129, November 2002. (Cited on page 4.)

[Clough 2005] S. A. Clough, M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara and P. D. Brown. *Atmospheric radiative transfer modeling: a summary of the AER codes*. J. Quant. Spec. Radiat. Transf., vol. 91, pages 233–244, March 2005. (Cited on page 30.)

[Cook 1980] J. W. Cook, G. E. Brueckner and M. E. Vanhoosier. *Variability of the solar flux in the far ultraviolet 1175-2100 A*. J. Geophys. Res., vol. 85, pages 2257–2268, May 1980. (Cited on page 29.)

[Deland 2012a] M. T. Deland and R. P. Cebula. *Solar UV variations during the decline of Cycle 23*. Journal of Atmospheric and Solar-Terrestrial Physics, vol. 77, pages 225–234, March 2012. (Cited on pages 96 and 102.)

[DeLand 2012b] M. T. DeLand, S. L. Taylor, L. K. Huang and B. L. Fisher. *Calibration of the SBUV version 8.6 ozone data product.* Atmospheric Measurement Techniques, vol. 5, pages 2951–2967, November 2012. (Cited on pages 45, 78, 97 and 103.)

[Dikpati 2004] M. Dikpati, G. de Toma, P. A. Gilman, C. N. Arge and O. R. White. *Diagnostics of Polar Field Reversal in Solar Cycle 23 Using a Flux Transport Dynamo Model.* ApJ, vol. 601, pages 1136–1151, February 2004. (Cited on page 89.)

[Dudok de Wit 2009] T. Dudok de Wit, M. Kretzschmar, J. Lilensten and T. Woods. *Finding the best proxies for the solar UV irradiance.* Geophys. Res. Lett., vol. 36, page 10107, May 2009. (Cited on pages 27 and 28.)

[Dudok de Wit 2014] T. Dudok de Wit, S. Bruinsma and K. Shibasaki. *Synoptic radio observations as proxies for upper atmosphere modelling.* Journal of Space Weather and Space Climate, vol. 4, no. 27, page A6, February 2014. (Cited on page 89.)

[Efron 1981] BRADLEY Efron. *Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods.* Biometrika, vol. 68, no. 3, pages 589–599, 1981. (Cited on page 81.)

[Ermolli 2013] I. Ermolli, K. Matthes, T. Dudok de Wit, N. A. Krivova, K. Tourpali, M. Weber, Y. C. Unruh, L. Gray, U. Langematz, P. Pilewskie, E. Rozanov, W. Schmutz, A. Shapiro, S. K. Solanki and T. N. Woods. *Recent variability of the solar spectral irradiance and its impact on climate modelling.* Atmospheric Chemistry & Physics, vol. 13, pages 3945–3977, April 2013. (Cited on pages 19, 20, 33, 95 and 101.)

[Fligge 2000] M. Fligge, S. K. Solanki and Y. C. Unruh. *Modelling irradiance variations from the surface distribution of the solar magnetic field.* A&A., vol. 353, pages 380–388, January 2000. (Cited on pages 2, 22, 26, 30, 31, 32, 41, 52, 53, 86, 95 and 101.)

[Fontenla 1999] J. Fontenla, O. R. White, P. A. Fox, E. H. Avrett and R. L. Kurucz. *Calculation of Solar Irradiances. I. Synthesis of the Solar Spectrum.* ApJ, vol. 518, pages 480–499, June 1999. (Cited on page 30.)

[Fontenla 2006] J. M. Fontenla, E. Avrett, G. Thuillier and J. Harder. *Semiempirical Models of the Solar Atmosphere. I. The Quiet- and Active Sun Photosphere at Moderate Resolution.* ApJ, vol. 639, pages 441–458, March 2006. (Cited on page 34.)

[Fontenla 2007] J. M. Fontenla, K. S. Balasubramaniam and J. Harder. *Semiempirical Models of the Solar Atmosphere. II. The Quiet-Sun Low Chromosphere at Moderate Resolution.* ApJ, vol. 667, pages 1243–1257, October 2007. (Cited on page 34.)

[Fontenla 2009] J. M. Fontenla, W. Curdt, M. Haberreiter, J. Harder and H. Tian. *Semiempirical Models of the Solar Atmosphere. III. Set of Non-LTE Models for Far-Ultraviolet/Extreme-Ultraviolet Irradiance Computation.* ApJ, vol. 707, pages 482–502, December 2009. (Cited on page 34.)

[Fontenla 2011] J. M. Fontenla, J. Harder, W. Livingston, M. Snow and T. Woods. *High-resolution solar spectral irradiance from extreme ultraviolet to far infrared.* Journal of Geophysical Research (Atmospheres), vol. 116, page 20108, October 2011. (Cited on pages 2, 22, 33, 34, 35, 36, 97 and 102.)

[Gray 2010] L. J. Gray, J. Beer, M. Geller, J. D. Haigh, M. Lockwood, K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, L. Hood, J. Luterbacher, G. A. Meehl, D. Shindell, B. van Geel and W. White. *Solar Influences on Climate.* Reviews of Geophysics, vol. 48, page 4001, October 2010. (Cited on page 16.)

[Groß 2003] J. Groß. Linear regression. Lecture Notes in Statistics. Springer Berlin Heidelberg, 2003. (Cited on page 82.)

[Haberreiter 2008] M. Haberreiter, W. Schmutz and I. Hubeny. *NLTE model calculations for the solar atmosphere with an iterative treatment of opacity distribution functions.* A&A., vol. 492, pages 833–840, December 2008. (Cited on pages 2, 22, 26, 37 and 41.)

[Haberreiter 2011] M. Haberreiter. *Solar EUV Spectrum Calculated for Quiet Sun Conditions.* Solar Phys., vol. 274, pages 473–479, December 2011. (Cited on page 37.)

[Haberreiter 2012] M. Haberreiter. *Towards the reconstruction of the EUV irradiance for solar Cycle 23.* In C. H. Mandrini and D. F. Webb, editeurs, IAU Symposium, volume 286 of *IAU Symposium*, pages 97–100, July 2012. (Cited on page 53.)

[Haberreiter 2014] M. Haberreiter, V. Delouille, B. Mampaey, C. Verbeeck, G. Del Zanna and S. Wieman. *Reconstruction of the solar EUV irradiance from 1996 to 2010 based on SOHO/EIT images.* Journal of Space Weather and Space Climate, vol. 4, no. 27, page A30, October 2014. (Cited on pages 38 and 53.)

[Haigh 2007] J. D. Haigh. *The Sun and the Earth's Climate.* Living Reviews in Solar Physics, vol. 4, page 2, October 2007. (Cited on page 16.)

[Haralick 1992] Robert M. Haralick and Linda G. Shapiro. Computer and robot vision. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st édition, 1992. (Cited on page 53.)

[Higgins 2011] P. A. Higgins, P. T. Gallagher, R. T. J. McAteer and D. S. Bloomfield. *Solar magnetic feature detection and tracking for space weather monitoring.* Advances in Space Research, vol. 47, pages 2105–2117, June 2011. (Cited on page 52.)

[Hinteregger 1981] H. E. Hinteregger, K. Fukui and B. R. Gilson. *Observational, reference and model data on solar EUV, from measurements on AE-E.* Geophys. Res. Lett., vol. 8, pages 1147–1150, November 1981. (Cited on page 29.)

[Hubeny 1995] I. Hubeny and T. Lanz. *Non-LTE line-blanketed model atmospheres of hot stars. 1: Hybrid complete linearization/accelerated lambda iteration method.* ApJ, vol. 439, pages 875–904, February 1995. (Cited on page 37.)

[Jiang 2014] J. Jiang, D. H. Hathaway, R. H. Cameron, S. K. Solanki, L. Gizon and L. Upton. *Magnetic Flux Transport at the Solar Surface.* Space Sci. Rev., September 2014. (Cited on pages 98 and 104.)

[Judge 2012] P. G. Judge and M. J. Thompson. *Solar and stellar activity: diagnostics and indices.* In C. H. Mandrini and D. F. Webb, editeurs, IAU Symposium, volume 286 of *IAU Symposium*, pages 15–26, July 2012. (Cited on page 89.)

[Kalman 1960] R. E. Kalman. *A New Approach to Linear Filtering And Prediction Problems.* ASME Journal of Basic Engineering, 1960. (Cited on page 98.)

[Kiess 2014] C. Kiess, R. Rezaei and W. Schmidt. *Properties of sunspot umbrae observed in cycle 24.* A&A., vol. 565, page A52, May 2014. (Cited on page 54.)

[Kopp 2011] G. Kopp and J. L. Lean. *A new, lower value of total solar irradiance: Evidence and climate significance.* Geophys. Res. Lett., vol. 38, page 1706, January 2011. (Cited on page 3.)

[Krivova 2003] N. A. Krivova, S. K. Solanki, M. Fligge and Y. C. Unruh. *Reconstruction of solar irradiance variations in cycle 23: Is solar surface magnetism the cause?* A&A., vol. 399, pages L1–L4, February 2003. (Cited on page 43.)

[Krivova 2005a] N. A. Krivova and S. K. Solanki. *Modelling of irradiance variations through atmosphere models .* Mem. Soc. Astron. Italiana, vol. 76, page 834, 2005. (Cited on page 30.)

[Krivova 2005b] N. A. Krivova and S. K. Solanki. *Reconstruction of solar UV irradiance.* Advances in Space Research, vol. 35, pages 361–364, 2005. (Cited on page 30.)

[Lean 1982] J. L. Lean, W. C. Livingston, D. F. Heath, R. F. Donnelly, A. Skumanich and O. R. White. *A three-component model of the variability of the solar ultraviolet flux 145-200 nM.* J. Geophys. Res., vol. 87, pages 10307–10317, December 1982. (Cited on pages 16 and 29.)

[Lean 1988] J. Lean and P. Foukal. *A Model of Solar Luminosity Modulation by Magnetic Activity between 1954 and 1984.* Science, vol. 240, pages 906–908, May 1988. (Cited on pages 98 and 104.)

[Lean 1997a] J. Lean. *The Sun's Variable Radiation and Its Relevance For Earth.* Ann. Rev. Astron. Astrophys., vol. 35, pages 33–67, 1997. (Cited on pages 18 and 19.)

[Lean 1997b] J. L. Lean, G. J. Rottman, H. L. Kyle, T. N. Woods, J. R. Hickey and L. C. Puga. *Detection and parameterization of variations in solar mid- and near-ultraviolet radiation (200-400 nm).* J. Geophys. Res., vol. 102, pages 29939–29956, 1997. (Cited on pages 27 and 41.)

[Lean 2005] J. Lean, G. Rottman, J. Harder and G. Kopp. *SORCE Contributions to New Understanding of Global Change and Solar Variability.* Solar Phys., vol. 230, pages 27–53, August 2005. (Cited on pages 27 and 41.)

[Lean 2009] J. L. Lean, J. M. Picone and J. T. Emmert. *Quantitative forecasting of near-term solar activity and upper atmospheric density.* Journal of Geophysical Research (Space Physics), vol. 114, page 7301, July 2009. (Cited on pages 98 and 104.)

[Lilensten 2008] J. Lilensten, T. Dudok de Wit, M. Kretzschmar, P.-O. Amblard, S. Moussaoui, J. Aboudarham and F. Auchère. *Review on the solar spectral variability in the EUV for space weather purposes.* Annales Geophysicae, vol. 26, pages 269–279, February 2008. (Cited on page 4.)

[Liu 2012] Y. Liu, J. T. Hoeksema, P. H. Scherrer, J. Schou, S. Couvidat, R. I. Bush, T. L. Duvall, K. Hayashi, X. Sun and X. Zhao. *Comparison of Line-of-Sight Magnetograms Taken by the Solar Dynamics Observatory/Helioseismic and Magnetic Imager and Solar and Heliospheric Observatory/Michelson Doppler Imager.* Solar Phys., vol. 279, pages 295–316, July 2012. (Cited on page 52.)

[McClintock 2000] W. E. McClintock, G. J. Rottman and T. N. Woods. *Solar Stellar Irradiance Comparison Experiment II (SOLSTICE II) for the NASA Earth Observing System's Solar Radiation and Climate Experiment mission.* In W. L. Barnes, editeur, Earth Observing Systems V, volume 4135 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 225, November 2000. (Cited on page 46.)

[Morrill 2014] J. S. Morrill, L. Floyd and D. McMullin. *Comparison of Solar UV Spectral Irradiance from SUSIM and SORCE.* Solar Phys., vol. 289, pages 3641–3661, October 2014. (Cited on pages 45, 78, 97 and 103.)

[Paterno 1998] L. Paterno. *Do we understand the 22-year solar activity cycle?* Academie des Sciences Paris Comptes Rendus Serie B Sciences Physiques, vol. 326, pages 393–405, June 1998. (Cited on pages 7 and 12.)

[Phillips 2008] Kenneth J. H. Phillips, Uri Feldman and Enrico Landi. Ultraviolet and x-ray spectroscopy of the solar atmosphere. Cambridge University Press, 2008. Cambridge Books Online. (Cited on pages 1, 5, 7, 8, 10, 13, 14 and 22.)

[Press 2007] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery. Numerical recipes 3rd edition: The art of scientific computing. Cambridge University Press, New York, NY, USA, 3 édition, 2007. (Cited on page 90.)

[Richards 1994] P. G. Richards, J. A. Fennelly and D. G. Torr. *EUVAC: A solar EUV flux model for aeronomic calculations.* J. Geophys. Res., vol. 99, pages 8981–8992, May 1994. (Cited on pages 27, 29 and 41.)

[Ringnes 1960] T. S. Ringnes and E. Jensen. *On the relation between magnetic fields and areas of sunspots in the interval 1917-56.* Astrophysica Norvegica, vol. 7, page 99, 1960. (Cited on page 54.)

[Roger 2014] Thibaut Roger. Etude de la variabilite solaire a l'aide des donnees du satellite sdo. Master's thesis, University of Orleans, August 2014. (Cited on page 85.)

[Rottman 2005] G. Rottman. *The SORCE Mission.* Solar Phys., vol. 230, pages 7–25, August 2005. (Cited on page 44.)

[Rottman 2006] G. J. Rottman, T. N. Woods and W. McClintock. *SORCE solar UV irradiance results.* Advances in Space Research, vol. 37, pages 201–208, 2006. (Cited on page 78.)

[Scherrer 2012] P. H. Scherrer, J. Schou, R. I. Bush, A. G. Kosovichev, R. S. Bogart, J. T. Hoeksema, Y. Liu, T. L. Duvall, J. Zhao, A. M. Title, C. J. Schrijver, T. D. Tarbell and S. Tomczyk. *The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO).* Solar Phys., vol. 275, pages 207–227, January 2012. (Cited on page 44.)

[Schmidtke 1976] G. Schmidtke. *EUV indices for solar-terrestrial relations.* Geophys. Res. Lett., vol. 3, pages 573–576, October 1976. (Cited on pages 27 and 89.)

[Schunk 2004] R. W. Schunk and A. F. Nagy. Ionospheres. November 2004. (Cited on page 17.)

[Snow 2010] M. Snow, W. E. McClintock and T. N. Woods. *Solar spectral irradiance variability in the ultraviolet from SORCE and UARS SOLSTICE.* Advances in Space Research, vol. 46, pages 296–302, August 2010. (Cited on page 20.)

[Socas-Navarro 2011] H. Socas-Navarro. *A high-resolution three-dimensional model of the solar photosphere derived from Hinode observations.* A&A., vol. 529, page A37, May 2011. (Cited on page 39.)

[Tapping 2013] K. F. Tapping. *The 10.7 cm solar radio flux ($F_{10.7}$).* Space Weather, vol. 11, pages 394–406, July 2013. (Cited on pages 13 and 18.)

[Thuillier 2011] G. Thuillier, J. Claudel, D. Djafer, M. Haberreiter, N. Mein, S. M. L. Melo, W. Schmutz, A. Shapiro, C. I. Short and S. Sofia. *The Shape of the Solar Limb: Models and Observations.* Solar Phys., vol. 268, pages 125–149, January 2011. (Cited on page 37.)

[Tobiska 1998] W. K. Tobiska and F. G. Eparvier. *EUV97: Improvements to EUV Irradiance Modeling in the Soft X-Rays and FUV.* Solar Phys., vol. 177, pages 147–159, 1998. (Cited on page 29.)

[Tobiska 2000] W. K. Tobiska, T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottman and O. R. White. *The SOLAR2000 empirical solar irradiance model and forecast tool.* Journal of Atmospheric and Solar-Terrestrial Physics, vol. 62, pages 1233–1250, September 2000. (Cited on pages 27, 29, 41 and 88.)

[Unruh 2012] Y. C. Unruh, W. T. Ball and N. A. Krivova. *Solar Irradiance Models and Measurements: A Comparison in the 220-240 nm wavelength band.* Surveys in Geophysics, vol. 33, pages 475–481, July 2012. (Cited on page 33.)

[Verbeeck 2014] C. Verbeeck, V. Delouille, B. Mampaey and R. De Visscher. *The SPoCA-suite: Software for extraction, characterization, and tracking of active regions and coronal holes on EUV images.* A&A., vol. 561, page A29, January 2014. (Cited on page 37.)

[Viereck 2001] R. Viereck, L. Puga, D. McMullin, D. Judge, M. Weber and W. K. Tobiska. *The Mg II index: A proxy for solar EUV.* Geophys. Res. Lett., vol. 28, pages 1343–1346, April 2001. (Cited on page 11.)

[Vögler 2005] A. Vögler, S. Shelyag, M. Schüssler, F. Cattaneo, T. Emonet and T. Linde. *Simulations of magneto-convection in the solar photosphere. Equations, methods, and results of the MURaM code.* A&A., vol. 429, pages 335–351, January 2005. (Cited on page 39.)

[Wang 1991] Y.-M. Wang and N. R. Sheeley Jr. *Magnetic flux transport and the sun's dipole moment - New twists to the Babcock-Leighton model.* ApJ, vol. 375, pages 761–770, July 1991. (Cited on page 89.)

[Wilhelm 2010] K. Wilhelm. *Quantitative solar spectroscopy.* Astronomische Nachrichten, vol. 331, page 502, May 2010. (Cited on page 4.)

[Woods 2005] T. N. Woods, F. G. Eparvier, S. M. Bailey, P. C. Chamberlin, J. Lean, G. J. Rottman, S. C. Solomon, W. K. Tobiska and D. L. Woodraska. *Solar EUV Experiment (SEE): Mission overview and first results.* Journal of Geophysical Research (Space Physics), vol. 110, page 1312, January 2005. (Cited on pages 20, 44 and 78.)

[Woods 2012] T. N. Woods, F. G. Eparvier, R. Hock, A. R. Jones, D. Woodraska, D. Judge, L. Didkovsky, J. Lean, J. Mariska, H. Warren, D. McMullin,

P. Chamberlin, G. Berthiaume, S. Bailey, T. Fuller-Rowell, J. Sojka, W. K. Tobiska and R. Viereck. *Extreme Ultraviolet Variability Experiment (EVE) on the Solar Dynamics Observatory (SDO): Overview of Science Objectives, Instrument Design, Data Products, and Model Developments.* Solar Phys., vol. 275, pages 115–143, January 2012. (Cited on pages 44 and 78.)

[Worden 1996] J. R. Worden. *A Three Component Proxy Model for the Solar Far Ultraviolet Irradiance.* PhD thesis, Colorado University, January 1996. (Cited on page 29.)

[Yeo 2013] K. L. Yeo, S. K. Solanki and N. A. Krivova. *Intensity contrast of solar network and faculae.* A&A., vol. 550, page A95, February 2013. (Cited on page 51.)

[Yeo 2014] K. L. Yeo, N. A. Krivova, S. K. Solanki and K. H. Glassmeier. *Reconstruction of total and spectral solar irradiance from 1974 to 2013 based on KPVT, SoHO/MDI, and SDO/HMI observations.* A&A., vol. 570, page A85, October 2014. (Cited on pages 77, 92 and 96.)

[Zwaan 1987] C. Zwaan. *Elements and patterns in the solar magnetic field.* Ann. Rev. Astron. Astrophys., vol. 25, pages 83–111, 1987. (Cited on page 50.)

Curriculum Vite

Personal data

Anatoliy Vuiets
Geroyev Dnepra. 30, 51
04214 Kiev, Ukraine

Tel.: +38(044) 414 82 68
E-Mail: anatoliy.v.vuets@gmail.com

Education

2004–2008 Bachelor graduate student of Institute for Applied System Analysis, National technical university of Ukraine Kiev Polytechnic Institute.

2008–2009 Master graduate student of Institute for Applied System Analysis, National technical university of Ukraine Kiev Polytechnic Institute.

2009–2010 Magister(Master+1) graduate student of Institute for Applied System Analysis, National technical university of Ukraine Kiev Polytechnic Institute, Thesis: Algorithms and software development for detecting large-scale events on the Sun.

2011–2014 PhD, LPC2E, University of Orleans, France.

Articles

2014 A. Vuiets, T. Dudok de Wit M. Kretzschmar, L. E. A. Vieira, and M, Scholl: Empirical modeling of the solar spectral irradiance in the UV, Astronomy and Astrophysics under review.

Oral presentations

08–10–2014 Solar metrology Symposium: An empirical approach to the modelling of the solar spectral irradiance in the UV.

31–01–2014 SORCE scientific meeting: What can we learn from SORCE about the contribution of different magnetic structures to the solar spectral irradiance variability in the UV, Cocoa Beach USA.

22–11–2013 European Space Weather Week: What is the contribution of different magnetic structures to the solar spectral irradiance variability in the UV?, Antwerp.

22–02–2013 FP7 SOLID meeting: Empirical modelling of SSI, Davos.

08–05–2013 PROBA2 science days meeting: Connectivity in oscillation processes, Brussels.

Posters

27–11–2014 SIP: Automated nowcast of the solar spectral irradiance in the UV, La Roche en Ardenne.

10–04–2013	European Geosciences Union conference: Near real-time reconstruction of the solar spectral irradiance in the UV, Vienna.
27–11–2012	European Space Weather Week: What do causal relations between flare irradainces at various wavelengths tell us?, Brussels.

28. November 2014

Anatoliy VUIETS

Reconstruction empirique du spectre solaire UV

Résumé : (1700 caractères max.)

L'irradiance spectrale solaire (SSI) dans la bande ultraviolette est un paramètre-clé pour la spécification de la moyenne et la haute atmosphère terrestre. Elle est requise dans de nombreuses applications en météorologie de l'espace, et aussi pour l'étude du climat. Or les observations souffrent de plusieurs défauts : manque de couverture spectrale et temporelle permanente, dégradation des capteurs, désaccords entre le instruments, etc. Plusieurs modèles de reconstruction de la SSI ont été développés pour pallier à ces difficultés. Chacun souffre de défauts, et la reconstruction du spectre en-dessous de 120 nm est un réel défi.
C'est dans ce contexte que nous avons développé un modèle empirique, qui recourt au champ magnétique photosphérique pour reconstruire les variations du spectre solaire. Ce modèle décompose les magnétogrammes solaires en différentes structures qui sont classées à partir de leur aire (et non sur la base de leur intensité, comme dans la plupart des autres modèles). La signature spectrale de ces structures est déduite des observations, et non pas imposée par des modèles de l'atmosphère solaire. La qualité de la reconstruction s'avère être comparable à celle d'autres modèles.
Parmi les principaux résultats, relevons que deux classes seulement de structures solaires suffisent à reproduire correctement la variabilité spectrale solaire. En outre, seule une faible résolution radiale suffit pour reproduire les variations de centre-bord. Enfin, nous montrons que l'amélioration apportée par la décomposition du modèle en deux constantes de temps peut être attribuée à l'effet des raies optiquement minces.

Mots clés : irradiance spectrale solaire, météorologie de l'espace

Empirical reconstruction of the solar UV spectrum

Résumé : (1700 caractères max.)

The spectrally-resolved radiative output of the Sun (SSI) in the UV band, i.e. at wavelengths below 300 nm, is a key quantity for specifying the state of the middle and upper terrestrial atmosphere. This quantity is required in numerous space weather applications, and also for climate studies. Unfortunately, SSI observations suffer from several problems : they have numerous spectral and temporal gaps, instruments are prone to degradation and often disagree, etc. This has stimulated the development of various types of SSI models. Proxy-based models suffer from lack of the physical interpretation and are as good as the proxies are. Semi-empirical models do not perform well below 300 nm, where the local thermodynamic equilibrium approximation does not hold anymore. We have developed an empirical model, which assumes that variations in the SSI are driven by solar surface magnetic flux. This model proceeds by segmenting solar magnetograms into different structures. In contrast to existing models, these features are classified by their area (and not their intensity), and their spectral signatures are derived from the observations (and not from models). The quality of the reconstruction is comparable to that of other models. More importantly, we find that two classes only of solar features are required to properly reproduce the spectral variability. Furthermore, we find that a coarse radial resolution suffices to account for geometrical line-of-sight effects. Finally, we show how the performance of the model on different time-scales is related to the optical thickness of the emission lines.

Keywords : solar spectral irradiance, space weather

Laboratoire de Physique et Chimie de l'Environnement (LPC2E), 3A Avenue de la Recherche Scientifique, 45071 Orléans cedex 2, France