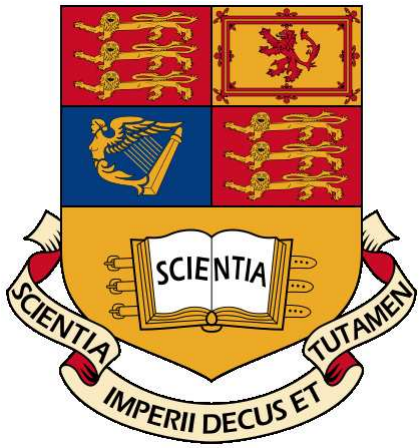
Observations and Modelling of Total and Spectral Solar Irradiance

William T. Ball
Astrophysics Group



Department of Physics
Imperial College London

Thesis submitted for the Degree of Doctor of Philosophy to
Imperial College London

· 2012 ·

Abstract

The variation of solar energy entering Earth's atmosphere, the solar irradiance, is an important influence on the Earth's climate. Total output of the Sun varies by 0.1% over an 11-year solar cycle and on longer, secular scales there still remains uncertainty. Global temperatures on Earth have increased over the last 160 years along with increasing anthropogenic impact on the environment. It is, therefore, important to fully understand how much of this increase can be attributed to changes in the Sun.

This thesis makes use of a semi-empirical version of the Spectral And Total Irradiance REconstruction model, or SATIRE-S, to reconstruct both total and spectral irradiance variations. The physical basis of SATIRE-S is that all variations in solar irradiance are caused by changes in surface magnetic flux emergence.

In this thesis, SATIRE-S is updated to accommodate the input of full-disk continuum image and magnetogram data from ground-based and satellite sources spanning three full solar cycles over the period 1974-2009. These changes are described in detail. The combination of terrestrial and space-based images allows for a reconstruction that, for the first time, has been produced independently to, and therefore provides an unbiased comparison of, the composites of direct radiometric observations of total solar irradiance (TSI), which began in 1978. The excellent agreement with, in particular, the PMOD composite supports the simple model assumptions. It also provides constraints on the influence of other mechanisms that may affect solar irradiance on these timescales.

The reconstruction of TSI over three solar cycles provides the opportunity to produce a consistent long-term spectral solar irradiance (SSI) dataset that can be put to use by the climate and atmospheric physics communities. This has been one of the goals of this work. In this thesis, comparisons are made with the new SORCE/SIM broadband spectral observations, which include for the first time the visible and infrared above 400 nm. During the declining phase of the recent solar cycle it is found that there is almost no agreement between the model and instrument on a long-term, cycle-length period, but very good agreement on short-term rotational variation. The influence of the ultra-violet region below 400 nm is important in Earth-based stratospheric chemistry and is a strong influence on both temperature and ozone concentration. Therefore, the SATIRE-S spectral dataset is employed in a simple 2D atmospheric model to evaluate its sensitivity in light of recent spectral irradiance changes suggested by data from the SORCE satellite.

For my family and friends

Contents

List of Tables

List of Figures

Declaration and Copyright

This thesis is my own work, except where explicitly indicated in the text.

William Ball
February 2012

Acknowledgments

This PhD would never have happened without the assistance and support of countless people. There are going to be those who have provided help at some stage along the way who don't get a mention here. It is not because I don't think you deserve one, but simply that I didn't remember when I wrote this!

Of course, I must first express my gratitude to my supevisor, Dr. Yvonne Unruh, who gave me this opportunity to enter into the career I had always wanted. She has provided unwavering guidance, support, attention to detail and patience and has been the essence of what it means to be a supervisor. I look forward to continuing to work with her as an RA over the next few of years and on into the future.

Special thanks goes to Natalie Krivova who has been incredibly accommodating and supportive and made my visits to Lindau both productive and something to look forward to. She has also given up a large amount of her time to read, let alone reply, to many long emails over the years. Without her input, this thesis would not be what has become today. I would also like to thank Sami Solanki and Joanna Haigh who, despite both being very busy people, have offered enthusiasm and guidance that has been helpful and very much appreciated.

There have, of course, been many others who have helped with my work, so I must thank the following for their time and input: Jerry Harder, Daniel Mortlock, Andrew Jaffe, Thomas Wenzler, Marty Snow, Greg Kopp, Rock Bush.

At Imperial, I have had the pleasure of making friends with some awesome people. I'd like to mention Harsit Patel and Marisa March, who started at the same time and together we all managed to get through it successfully. Well done guys.

The road to academia began back in Durham where many people helped me get to Imperial. I just wanted to thank one individual in particular from there, John Bolton. Without

his support, advice and friendship I would never have go here, so thank you.

Closer to home, I must thank my family for their support, and also the same to my friends, especially Paddy, who spent many hours in the pub listening, with apparent genuine interest, to what I was saying about my work. I must also thank Barry and Catherine, who have been like parents to me. I am indebited to them for their encouragement and love.

Finally, thank you Alyson, for your love and for keeping me sane and grounded throughout the years, especially in the last few months.

Publications

- “Solar irradiance variability: a six-year comparison between SORCE observations and the SATIRE model”
  Ball, W. T., Unruh, Y. C., Krivova, N. A., Solanki, S., Harder, J. W., 2011, A&A, 530, 71

- “Solar irradiance models and measurements: a comparison in the 220 nm to 240 nm wavelength band”
  Yvonne C. Unruh, **Will T. Ball** & Natalie A. Krivova, 2011, Surv. Geophys.

- “Reconstruction of total solar irradiance 1974-2009”
  Ball, W. T., Unruh, Y. C., Krivova, N. A., Solanki, S., Wenzler, T., Mortlock, D. J., Jaffe, A. H., 2012, A&A, accepted.

- “(Working title) Spectral solar irradiance reconstruction for solar cycles 21-23 using SATIRE-S”
  Ball, W. T., Krivova, N. A., Unruh, Y. C., Haigh, J. D., Solanki, S., in preparation.


SATIRE-S
TSI
irradiance
magnetic
spectral
variability
data
change
model
flux
solar
MDI
Fig
period
cycle
time
nm
PMOD
Chapter
et
al
Bsat
SATIRE-SInd
magnetograms
wavelengths
instrument
using
reconstruction.
regions
observations
changes
shown
lines
Sun
SIM
difference
plot
sunspot
section
region
agreement
uncertainty
correlation
regression
rotational
variation
chapter
trends
NRLSSI
lower
also
smoothed
results
minimum
range
SORCE/SOLSTICE
SORCE/TIM
comparison
mean
seen
higher
observed
Comparison
three
shows
original
ozone
correction
degradation
December
KP/SPM
KP/512
TIM
used
two
level
value
see
surface
top
number
considered
fit
temperature
dataset
images
minima
reconstructions
found
long-term
umbral
intensity
prior
facular
within
plots
result
Wm
sunspots
series
red
composites
increase
function
levels
though
larger
facular
UV
03
use
first
relative
pixels
line
wavelength
models
ISS
observations
focus
detrended
blue
different
figure
trend
factor
may
well
limb
right
composite
km
left
free
field
continuum
disk
black
periods
compared
contrast
activity
datasets
large
important
parameter
therefore
lift-SIM

Chapter 1

Introduction: background physics and context

Overview

This chapter is intended as an introduction to the field of solar irradiance. It begins by putting the Sun in the context of other solar-type stars and then considers the effect the Sun has on Earth's climate, providing a justification for gaining a deeper understanding of changes in the Sun. An explanation for the Sun's underlying variability begins with the solar dynamo before discussing the emergent surface features, such as sunspots and faculae, that cause irradiance variability. An understanding of variability depends on the time period considered, so a brief overview of solar variation on various time scales is undertaken. The chapter ends with a discussion of some of the irradiance missions that have contributed to our current understanding of solar variability and a short summary of irradiance modelling techniques in preparation for the next chapter, which focuses on the model used in this thesis.

1.1 The Sun in a stellar context

Stellar variability occurs on all time scales, from small-scale surface convection over periods of minutes, pulsation and starspot transits over days and weeks, and the mega- to giga-year gradual increase in energy output as a star ages. In this thesis, the focus is on changes in the Sun that occur on scales from days to centuries. This is the period most important for studies of global warming and present day climate change. It excludes surface turbulence and pulsations, both of which are insignificant on periods longer than a few minutes. The increase in solar luminosity owing to aging processes is less than 0.01% Myr^{-1} (Ribas, 2010), or a millionth of the Sun's variation over the 11-year solar cycle. The remaining known cause of solar variability is from changes in the distribution of surface magnetic fields in the form of sunspots and faculae. Observations of the Sun's variability in brightness have only been directly and continuously measured since 1978. Indirect proxy data hint at a longer-term larger variability, but there are severe limitations in determining the full range of solar variability unless it is understood in a wider, stellar context.

To identify luminosity changes related to magnetic activity in stars, the magnetic field or some physically related variable needs to be observed. For the resolved Sun this can be directly mapped, both spatially and temporally, while, with rare exceptions, only temporal resolution can be obtained for point-source stars. From an analysis of solar magnetograms and the relative core strengths of Ca II H & K lines at 396.9 nm and 393.3 nm in solar plages, a direct correlation between magnetic field strength and heating of the chromosphere was found (Skumanich et al., 1975). The chromosphere is a layer of the Sun's atmosphere approximately 2000 km thick lying on top of the photosphere, which is the effective 'surface' where the Sun becomes opaque. The ratio of the HK line cores to nearby continuum bandpasses is known as the S-index and is a measure of stellar activity. Combining the S-index with measurements of stellar magnitude over time reveals the effect that the surface magnetic field has on the stellar luminosity.

Such a program of observations, covering a range of main-sequence stars, was initiated by Olin Wilson at the Mount Wilson Observatory. He found variations in Ca II line strengths over days and months (Wilson, 1978), showing that magnetic activity on stars caused variations in their spectra. Indeed, 60% of F to K[1] type stars have clearly identifiable S-index periodicity ranging from 7 to 25 years, more commonly occuring in late-type stars. Stars

[1]F, G and K type stars cover an approximate range of temperatures of 8000-4500 K with white, yellow and reddish colours, respectively (Zeilik and Gregory, 1998).

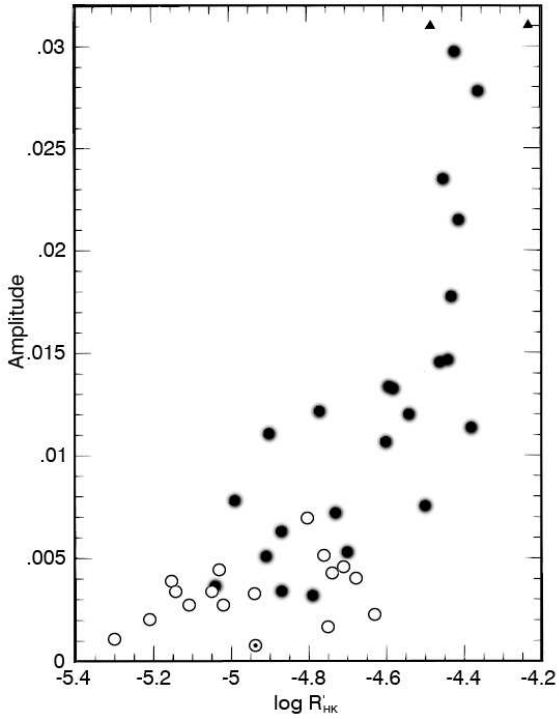


Figure 1.1: Plot from Lockwood et al. (1997) of the amplitude of variation, measured with b and y Strömgren filters, (b+y)/2 against the mean normalised Ca II H & K emission (R'$_{HK}$) for solar analogs from 12 years of observations. Filled circles represent stars with clear variation in brightness, open circles represent those with constant brightness within errors and triangles are points off the scale. The dotted circle is the Sun.

with a higher S-index have more irregular cycles and possibly more active chromospheres (Baliunas and Jastrow, 1990; Radick et al., 1998). The Sun, as a G-type star, goes through cycles of activity approximately 11 years in length.

Studying a range of stellar types is useful to first order, but to determine a more precise range of variation for the Sun, solar analogs need to be observed. Solar analogs are stars that have similar properties to the Sun (by rotation rate, age, metallicity) and so are more informative probes of the dynamic range of the Sun's magnetic activity and brightness. The Sun's magnitude varies by 0.001, or 0.1% in luminosity, over an 11-year activity cycle. However, a study by Lockwood et al. (1997) into the variation of stars with Sun-like Ca II emission levels found some stars with cycle amplitudes 30 times greater than the Sun. In Fig. 1.1 the HK emission of Sun-like stars, normalised by the bolometric luminosity to produce the chromospheric emission ratio R'$_{HK}$, is plotted against the amplitude of variability as measured by (b+y)/2, where b and y are the Strömgren filters at 467 and 547 nm, wavelengths of light formed in the photosphere. In this plot, the Sun is the dotted circle, open circles are stars showing no photospheric variability (within errors) and

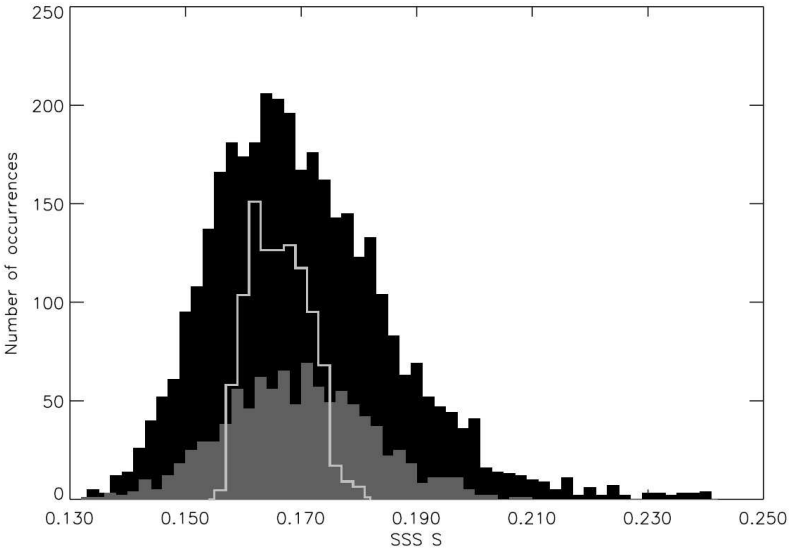


Figure 1.2: From Hall and Lockwood (2004), a histogram of S-index observations (black) from 57 solar-type stars from the Solar-Stellar Spectrograph (SSS S). The grey histogram are observations from a subsample of 10 low activity stars. The white outlined histogram are 3038 solar observations, with the occurance scale reduced by a factor of 3.

filled circles show clear photospheric variability. It shows that there is a relationship of increasing photospheric variability with an increase in chromospheric activity related to variations in the surface magnetic fields.

Stars with the largest observed photometric variation appear to show an anti-correlation with magnetic activity, but these stars are very much younger than the Sun (Radick et al., 1998; Lockwood et al., 1997). This anti-correlation is observed, to a lesser extent, on the Sun over rotational time scales: during active periods, regions of high magnetic concentration are associated with sunspots, causing a brightness decrease as it crosses the disk. Sunspots are associated with bright regions, known as faculae. On the Sun faculae cover a larger surface area than spots, but as sunspots become very large (e.g. 1000 millionths of a hemisphere) they significantly counteract the brightening effect of facular regions (Foukal, 1994). This provides an explanation as to why very active solar-type stars have an inverse relation between activity and brightness. During periods of high activity the stronger surface magnetic flux density results in a dominance of sunspots over faculae and hence a decrease in brightness.

Other studies (e.g. Hall and Lockwood (2004); also see the discussion and references within Haigh et al., 2005) have suggested that around 10% of solar twins are in a much less active Maunder Minimum-type phase (see section 1.9) while 20% exceed currently

documented solar maximum levels. In Fig. 1.2, taken from Hall and Lockwood (2004), the S-index from 3709 observations of 57 Sun-like stars are plotted as the black histogram. Measurements of the Sun are plotted as the white histogram outline, with the y-scale reduced by a factor of 3; the grey histogram are a subset of stars with essentially flat time series, as discussed in Hall and Lockwood (2004). The implication from this plot is that the Sun may vary more than current observations suggest, as the S-index from the stars in this plot lie outside the Sun's observed range approximately a third of the time (Haigh et al., 2005). Since magnetic activity is related to changes in luminosity, this has implications for the amount of energy received from the Sun at Earth. Therefore, solar analogs and their properties need to be accurately determined over long periods of time to precisely place the Sun among them. The conclusion from solar analog studies is that the range of possible fluctuations in the Sun may be very much greater than that currently observed. However, cosmogenic isotopes suggest that the Sun has been unusually active in recent decades (Solanki et al., 2004; Steinhilber et al., 2008). These results lead to the question, how significant an effect on terrestrial climate can the Sun have, and over what time scales?

1.2 The Sun's impact on our climate

The Sun, as Earth's primary source of energy, has a direct impact on the lives of Earth's inhabitants. Changes in the total solar irradiance (TSI), the radiant energy emitted over all wavelengths (see section 1.3), will modify the radiative forcing of the climate, although the actual response will depend on terrestrial based factors. It is important to understand that one of the reasons for producing an accurate reconstruction of total and spectral solar variability in this thesis is to help isolate its signal within the climate record, among the many contributors to climatic variation. Other influences include: the effects of Milankovitch cycles resulting from changes in the eccentricity, precession and axial tilt of the Earth; changes in albedo due to cloud, aerosol and polar ice coverage; the influence of atmospheric composition and gases released by volcanoes; feedback from the biosphere; and any amplifications of these effects as a result of their interaction.

Measurements of solar irradiance have been conducted almost continually since 1978, but this is a short period compared with the time scale over which recent anthropogenic activity has likely caused a change in global temperatures, around 160 years (IPCC, 2007). To quantify the Sun's influence on changes in global temperature change, the record of direct

solar observations needs to be extended back in time. Indirect, proxy variables that correlate well with direct observations, coupled with models, are required. One such proxy is the group sunspot number (SSN, discussed in section 1.5) which correlates well with TSI on time scales above the 27-day solar rotation. The sunspot dataset extends back to 1610 and when coupled to recent solar irradiance observations, a long term variation in TSI emerges. However, there are wide range of estimates that have been put forward for the change in TSI over the last 400 years (see section 1.9). Prior to 1610, there are no direct solar observations and other proxies are needed.

The solar signal can be inferred over longer time scales by considering the Sun's influence on the galactic cosmic ray (GCR) flux. The GCR flux is a measure of the number of high energy charged particles originating from outside the solar system. It has been shown to vary in anti-phase with solar activity on 27-day and the 11-year cycles (Olemskoy and Mordvinov, 2009). This happens because the open solar flux (see section 1.8), carried by charged particles ejected from the Sun as the inter-planetary magnetic field (IMF), alters the trajectories of GCRs. The IMF is strongest at solar maximum and restricts GCR penetration into the core of the solar system (Wang et al., 2006), while the opposite is the case at solar minimum. Carbon-14 and Beryllium-10 are produced by GCR interaction in the Earth's atmosphere and deposited in organic remains and ice cores. As such, IMF modulation provides a record of solar activity on millenial scales. An example of this, for a shorter time scale, is shown in in Fig. 1.3, from Beer et al. (1990). In this plot, a correlation of increasing solar activity (sunspot number, with inverted right axis), with decreasing ^{10}Be deposition can be seen. This result shows that the isotope concentration varies on time scales of centuries to millenia and can therefore provide a proxy record for solar influence on global climate and temperature variation (Bond et al., 2001).

There is significant evidence to suggest that solar irradiance correlates with variations in the Earth's atmosphere over the 11-year solar cycle. The Earth's atmosphere is categorised into layers: the troposphere extends up to ~15 km above the surface, the stratosphere continues to a height of 50 km, the mesosphere to 85 km and the thermosphere terminates at 600 km. Using the 10.7 cm radio flux (another solar activity proxy that again correlates well with cycle length variation in TSI), correlations suggest that over a solar cycle, regions of the sub-tropical Pacific Ocean below 24km vary by 0.5-1.0 K (Labitzke and van Loon, 1995) and the upper troposphere varies by 0.15-0.20 K, averaged over the northern hemisphere (van Loon and Shea, 1999). These examples highlight the possible direct effect on terrestrial temperatures as a result of solar irradiance changes, but one must note that correlation does not ensure identification of cause and, indeed, these correlations may

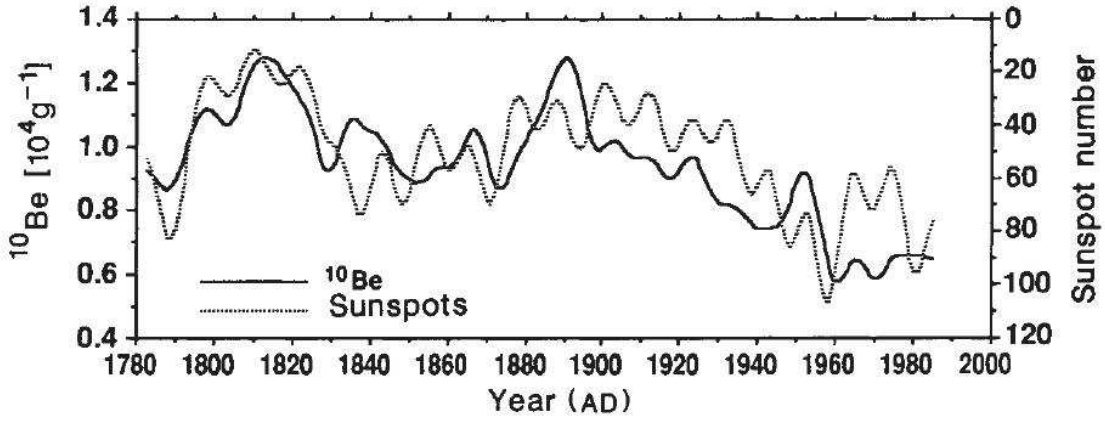


Figure 1.3: Smoothed sunspot group number, R_G (see section 1.5), (*right-axis*, inverted) and Beryllium-10 (*left-axis*) isotopes extracted from ice cores covering the past three centuries. From Beer et al. (1990).

be more related to feedback and enhancement of the Earth's atmosphere from variations in only some parts of the solar spectrum.

Therefore, the spectral solar irradiance (SSI), the irradiance as a function of wavelength (see section 1.3), is possibly even more important. Due to the presence in the atmosphere of molecules such as ozone (O_3), water, NO_2, CO_2, CH_4 and CFCs, the amount of radiation absorbed varies by wavelength, height and latitude. SSI variability is very much greater in the ultraviolet (UV) than in the visible or infra-red (IR). Wavelengths in the solar spectrum shorter than 400 nm contribute 30-60% to TSI variability (Krivova et al., 2006) but account for less than 10% of the absolute value of TSI. The impact of the UV region is very important in many processes in the atmosphere, for example in the production of ozone in the stratosphere.

Ozone production and destruction is greatly modulated by the Sun (Labitzke et al., 2002; Haigh, 2007). The ozone column density at solar maximum is up to 2% greater than at minimum (Huang and Brasseur, 1993). The presence of ozone leads to enhanced absorption of UV light resulting in a shielding of the atmosphere below. This absorption also has a heating effect (as do other atmospheric constituents) and results in a complex temperature redistribution as a function of height and latitude (e.g. Matthes et al., 2004). In the thermosphere the temperature response can range over 100 K. 1 K responses have been detected in the upper stratosphere and, below this, variation on the order of 0.1 K may occur, though datasets do vary (Haigh, 1994). Atmospheric constituents affect visible and IR absorption, but are poorly understood areas of atmospheric physics, not least due to

the fact that above-atmosphere SSI for wavelengths longer than 400 nm have not been observed on a long enough period of time, or with sufficient reliably, prior to 2003.

Predicting climate change requires the use of general circulation models (GCMs). These models use SSI and TSI data to predict past and future possibilities and therefore accuracy is paramount. They require a full knowledge of solar variability across the entire spectrum and over time scales of a decade and longer. In order to properly understand the consequences of physical processes involved in GCMs, simpler models are used. In Chapter 7, a 2D atmospheric model by Harwood and Pyle (1975) is used to investigate the sensitivity of ozone concentration in the stratosphere when modelled and observed spectral irradiance variations are used as inputs. Discussion of the interaction between SSI and ozone in the Earth's atmosphere will be expanded in that chapter.

1.3 Definition of solar irradiance

The luminosity of the Sun is measured to be $\sim 3.85 \times 10^{26}$ W and is the radiative power emitted over the entire surface. Intensity is the power emitted per unit area per steradian, in $\mathrm{Wm^{-2}sr^{-1}}$, whereas flux, or irradiance, is the power received per unit area, measured in $\mathrm{Wm^{-2}}$. In the context of this thesis, solar irradiance is defined as the radiant energy passing through a surface area of 1 $\mathrm{m^2}$ parallel to the surface of the Sun, at a distance of 1 AU (1.496×10^{11} m) in the plane of Earth's orbit.

The absolute level of TSI was previously accepted by the solar community to be approximately 1365 $\mathrm{Wm^{-2}}$ ±0.15% (Crommelynck et al., 1995). However, the SOlar Radiation and Climate Experiment (SORCE) satellite with the on board Total Irradiance Monitor (TIM), with supporting laboratory tests, has now seen a revision of this estimate to 1360.8 $\mathrm{Wm^{-2}}$ ±0.04% (Kopp and Lean, 2011).

The spectral solar irradiance (SSI) is the incident radiation observed within a particular wavelength interval measured in units of $\mathrm{Wm^{-3}}$ or $\mathrm{Wm^{-2}nm^{-1}}$. The spectral distribution of stars can be approximated as a black body. Using Planck's radiation law, the spectrum can be used to approximate the specific intensity by:

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \tag{1.1}$$

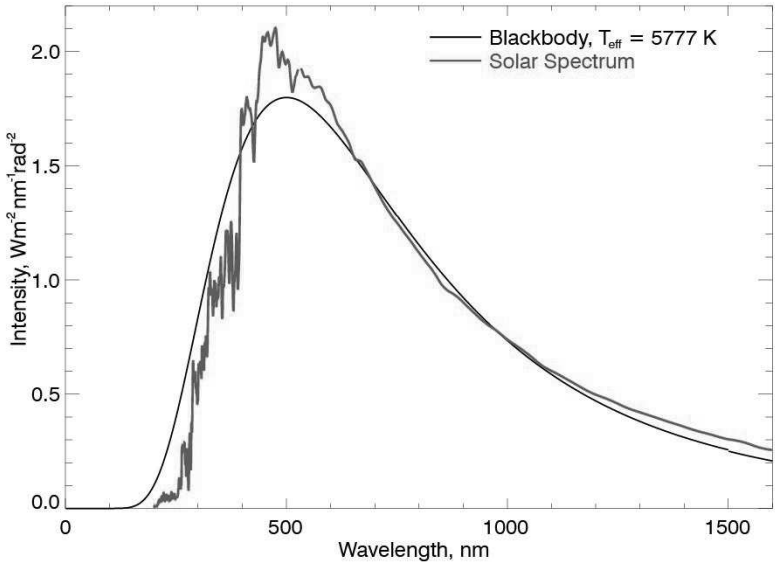


Figure 1.4: The above-atmosphere solar spectrum at a distance of 1 AU taken from the SIM instrument on SORCE (grey) on 2004 April 21. Superimposed is a Planck function black body curve with an effective temperature of 5777 K, the temperature at which the flux of the Sun equals that of the blackbody.

where λ is the wavelength of light, h is Planck's constant, c is the speed of light, k is Boltzmann's constant and T the temperature. For the Sun, TSI is best approximated for a blackbody with an effective temperature of 5777 K. Fig. 1.4 shows this Planck function (black line) along with the spectral intensity of the Sun (grey). While the broad distribution agrees, there are differences at almost all wavelengths. This is because the optical depth, τ_λ, within the solar atmosphere is highly wavelength dependent so that the solar spectrum is formed at different depths within the solar atmosphere.

The curves in Fig. 1.4 reveal information about the location of spectral emission. Fig 1.4 shows that solar radiation is not a true Planck function, with less emission in the UV and more in parts of the visible and IR. If, instead, equation 1.1 of Planck's law was rearranged to estimate the temperature at which each spectral line was produced the result would be the brightness temperature as shown in Fig. 1.5. This plot, from Harder et al. (2009), shows the solar brightness temperature between 200 and 2400 nm as observed by the SORCE/SIM instrument (Harder et al., 2005a,b). The spectral lines below 400 nm form in a region cooler than the effective temperature, indicating that they formed higher up in the solar atmosphere; generally the spectral region between 200 and 300 nm is formed in the chromosphere. At ~1600 nm, at the peak of the brightness temperature curve, the higher formation temperature indicates that they formed below the solar sur-

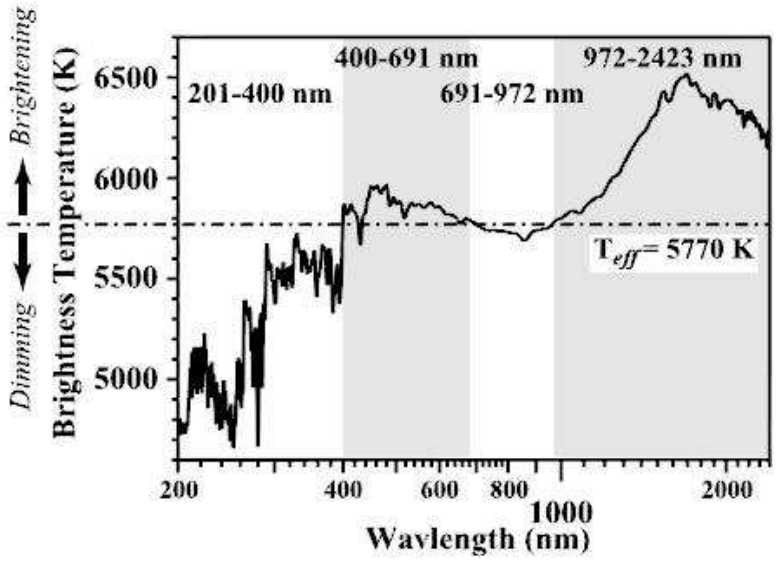


Figure 1.5: Brightness temperature of the Sun. From SIM as taken from Harder et al. (2009).

face at deeper, hotter photospheric layers. The difference is important when modelling SSI since spectral features produced in optically thin regions are not subject to the approximation of local thermodynamic equilibrium (LTE). LTE can be approximated as the domain where the mean free path of particles is small compared to that of temperature variations. If particles are not able to thermally equalise and are able escape the domain of particles with a similar kinetic energy, then they are in non-LTE and lines are formed that do not correspond to their temperature domain. The LTE approximation starts to break down in regions < 400 nm, in the higher, more diffuse layers of the photosphere and chromosphere, and it becomes much more difficult to model adequately (Fontenla et al., 1999; Unruh et al., 1999). The limitations of the LTE approximation are discussed in the context of a comparison between observed and modelled spectral irradiance (using the LTE approximation) in chapter 5.

1.4 The Solar Dynamo

Variations in TSI and SSI on time scales up to the solar cycle have been found to be largely modulated by sunspots and faculae. These features are caused by the eruption of magnetic fields through the solar surface. The production of strong magnetic fields within the Sun is an area of intensive research. A full discussion of this topic is not within the

scope of this thesis (see Charbonneau (2010) and references within), but a brief description is needed to understand the origin of surface features integral to solar variability.

Outside the nuclear burning core, the Sun is composed entirely of plasma and is divided into two distinct regions. The radiative zone, containing the fusion core, extends out to ~70% of the solar radius and exhibits solid body rotation. Beyond this, convection becomes the driving force for energy distribution and is present between the top of the radiative zone up to the photosphere, where small scale convection manifests as granulation.

The convection zone undergoes differential rotation by both latitude and depth, rotating more slowly at the poles than at the equator[2]. Helioseismology has determined that there is a significant shear profile at the boundary layer between the convective and radiative zones, known as the tachocline (Spiegel and Zahn, 1992). Here, the rapidly changing rotation rate is thought to provide a mechanism that produces a magnetic dipole. The dynamo is thought to be self-exciting, that is it replenishes itself, with a period of 22 years. This 22 year periodicity appears as two solar cycles of activity that sees two reversals in the magnetic polarity of the northern and southern hemispheres (Kundt, 1992).

It is thought that activity manifests itself at the surface due to the build up of magnetic fields deep inside the Sun, near the tachocline. At the start of a solar cycle the magnetic field is polodial, or poleward, but convective material overshooting to the base of the tachocline is stretched by the rapidly changing rotation rate (Browning et al., 2006; Charbonneau, 2010), strengthening the polodial field into a toroidal field (along latitude lines). Finally, these boosted magnetic fields, embedded within convection currents, become buoyant and rise, eventually reaching and breaking through the surface. The appearance of these strong magnetic fields at the surface are typified by sunspots.

1.5 Sunspots, pores and micropores

Sunspots are the largest and most obvious manifestations of surface magnetic activity. They consist of a darker region at the centre, the umbra, surrounded by the penumbra. Pores are essentially sunspots without the penumbral region. These features are all shown

[2]Sidereal rotation at the equator is ~25 days, though the synodic period, taking account of the Earth's orbital motion, is ~1.8 days longer. Typically a rotation period of ~27 days is quoted: a synodic period that takes account of the slower rotation at a higher latitude where sunspots more commonly form. Rotation at the poles has been measured at ~36 days (Beck, 2000).

in Figs. 1.6 and 1.7. In the left image of Fig. 1.6, from near disk centre and taken at 430.5 nm (G-band, which emanates in the photosphere), part of a sunspot umbra (darkest area) can seen in the upper left, surrounded by the lighter penumbra and granulation further out. In the bottom right is a pore grouping that lacks any surrounding penumbra. Micropores appear to be present in the middle-left of the lower part of the image. Bright dots in the granular lanes are facular regions, more clearly visible as plages in the right image taken at the Ca II K line (the emission of which stems from the chromosphere). In this region above the photosphere, the micropores are obscured and granulation is less clear.

The formation, initially, of both pores and sunspots is the same. They begin with the coalescence of tiny magnetic flux elements within granular lanes that combine into small spot-like regions. Pores form from further integration of these regions (Wang and Zirin, 1992; Sobotka, 2003) and are both smaller, typically 1-7 Mm in diameter, and brighter than umbra; umbra have white-light continuum intensities which reach as low as 0.3 (Norton and Gilman, 2004). Pores will turn into sunspots when a strongly inclined magnetic field at the pore edge becomes unstable and interacts with the photospheric medium, forming a penumbra. Beyond this, with an increasing magnetic field, sunspots can expand to sizes in excess of 40 Mm across.

Sunspots are not confined to variation in just surface area. In 1769, Alexander Wilson observed a foreshortening of the disk-centre side penumbra near the solar limb (disk edge) and concluded that sunspots are shallowing depressions in the Sun's surface. Depths ranging from 400-1000 km have been measured in the umbra (Thomas and Weiss, 2008) and 50-100 km within the penumbra (Solanki et al., 1993). While the umbra is physically lower relative to the surrounding photosphere, the effect is enhanced by a reduction in opacity resulting from the presence of a strong magnetic field. The strong magnetic field increases the magnetic pressure and so the local gas pressure decreases when in pressure equilibrium with the surrounding region. The temperature within the sunspot decreases approximately linearly with an increase in the square of the magnetic field (Martinez Pillet and Vazquez, 1990). The decrease in temperature reduces the H^- bound-free opacity, one of the most important agents of opacity in the solar atmosphere. This then produces an apparent depression in the surface at oblique angles. Temperatures are dependent on the strength and size of the sunspots. Umbral temperatures range from 3900-4800 K and penumbrae from 5400 to 5500 K (Solanki, 2003), lower than the effective temperature of the quiet Sun. Hence, this is why they appear dark and have a lower contrast relative to the quiet Sun.

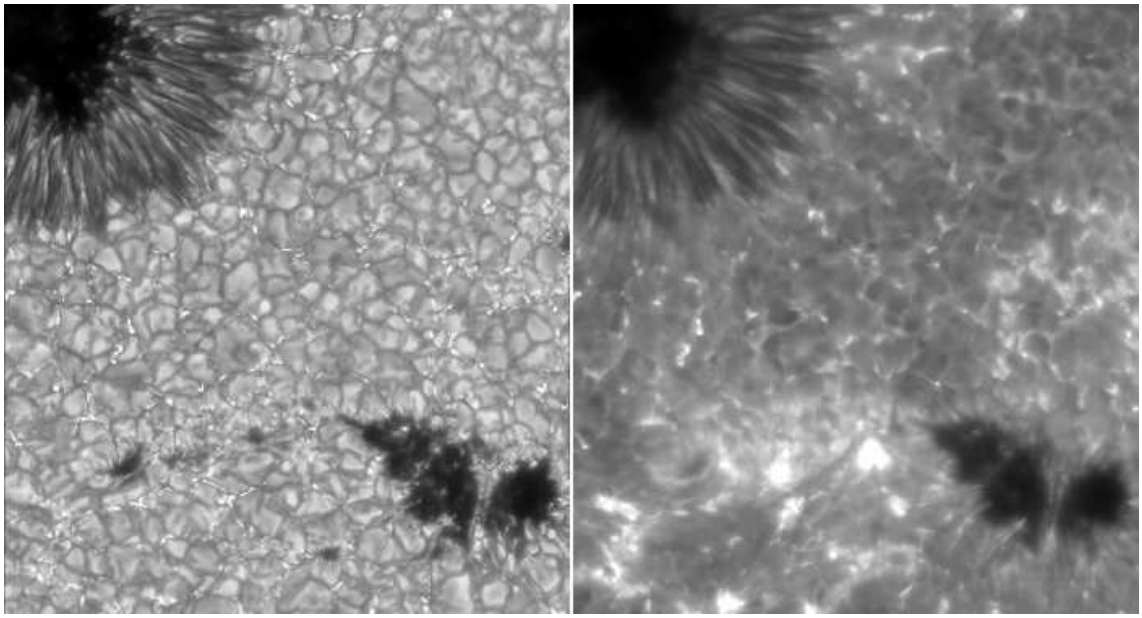


Figure 1.6: Active region filtergrams taken on 2 June 2003 with the Swedish Solar Telescope at 0.1″ resolution covering ∼30 Mm by 30 Mm (Rouppe van der Voort et al., 2005). (**left**) A sunspot (top left), pore grouping (bottom right) and micropores (bottom middle). Granulation exists everywhere outside the dark regions. The bright dots and lines between the granules are faculae and network regions. The image was taken at 430.5 nm (G-band). (**right**) The same region as the left taken in the Ca II K line at 396.9 nm. The bright regions to the left of the pore are called plages.

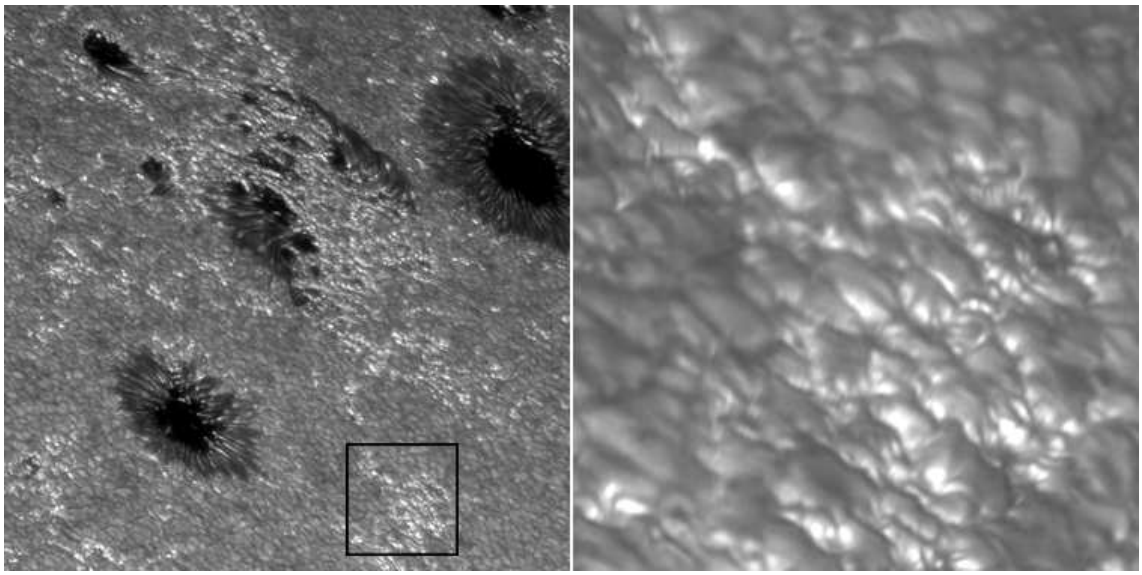


Figure 1.7: 430.5 nm (G-band) filtergram of an active region close to the limb taken by Kiselman and Löfdahl on 2003 June 29 with the Swedish 1-m Solar Telescope. (**left**) Faculae cover areas regions equal to or larger than sunspots, (**right**, boxed region in left) but the individual elements are much smaller.

The link between sunspots and magnetism was formed in 1908 when Hale discovered the Zeeman effect[3] in sunspot spectra (Hale, 1908). Magnetic field strengths within sunspot umbral regions can range from 1500-3500 G and in pores from 1600-2600 G, but these values vary both spatially within a spot and temporally over the solar cycle (Penn and Livingston, 2006). In comparison, the quiet sun exhibits a background distribution of very low magnetic fluxes with the highest frequency magnetic fluxes at around 10 G (Nordlund et al., 2009).

Comparing the magnetic field strength and the continuum intensity for regions within dark magnetic features indicates an inverse relationship. The basic idea that follows, suggested by Biermann, (1941), is that the magnetic field inhibits the convective motion of granules (and therefore the main energy transport mechanism) as a function of its strength and area covered. Thus, regions of higher field strength over a large area have a lower contrast. This is not the case for all photospheric magnetic features, such as faculae, as will be discussed in the next section.

Micropores represent the smallest dark surface feature and, while they may exhibit the initial formation traits of pores and sunspots, are not related. With diameters of less than 1000 km and magnetic field strengths of up to 1500 G, models suggest they result from the convergence of downflow-lanes surrounding a shrinking granule (Vögler et al., 2005). Linked to granulation as they are, their lifetimes are comparable (up to 20 minutes) and their temporal effect goes unnoticed in irradiance studies.

At the beginning of each solar cycle sunspots appear at higher latitudes, ~30-40°, before migrating towards the equator. This is known as Spörer's Law. Ephemeral regions, magnetic flux tube precursors to the formation of sunspot groups, appear in these higher latitudes first, at the onset of a new cycle, and often form before previous cycle active regions have disappeared (Balmaceda et al., 2007). The 'butterfly' diagram in Fig. 1.8 shows both this overlap and Spörer's law (Wilson, 1987). This figure is formed from a long period of continuous sunspot observations.

The first officially recorded dark blemishes on the solar surface come from the turn of the first millenium, but the first telescopic observations were not performed until 1610. Observations were regularly made and monthly records began in 1749. It wasn't until 1849

[3]Magnetic fields break the degeneracy of electron transition energy levels within atoms. These new transition levels show up in spectra on either side of the non-magnetised spectral line with a $\Delta\lambda$ proportional to the magnetic field strength.

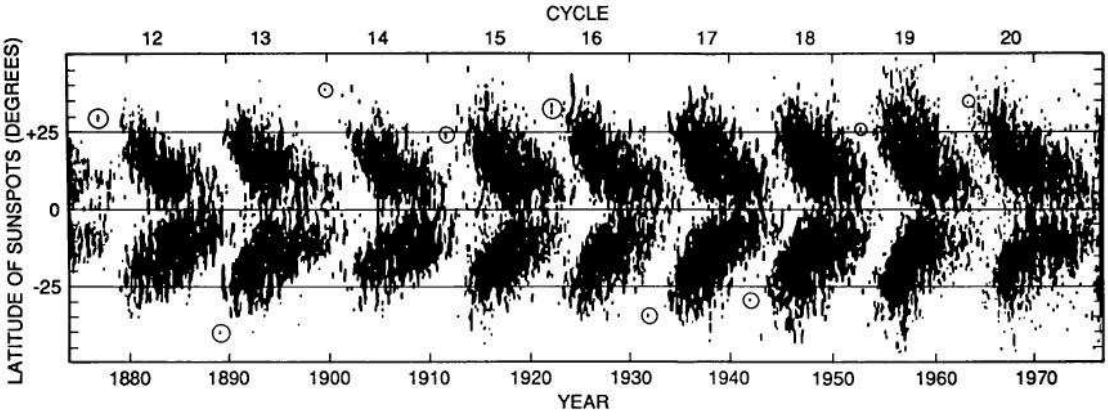


Figure 1.8: Butterfly diagram for cycles 12-20. Sunspots form at higher latitudes at the beginning of a cycle and migrate towards the equator. Circled regions indicate the first detected new cycle spots, which overlap with those in the old cycle. From Wilson (1987)

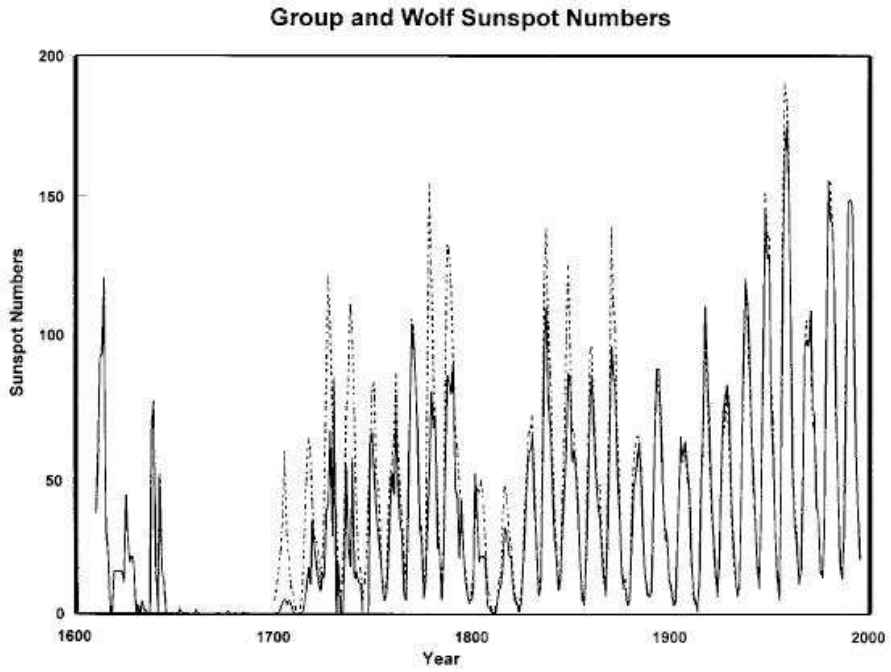


Figure 1.9: Sunspot Group (solid) and Zurich (or Wolf, dashed) number since 1610 from Hoyt and Schatten, (1998). The 11-year cycle is clear here, but superimposed on a long-term trend since the Maunder Minimum of 1645-1715.

that consistent and long term data accumulated with daily measurements being taken and the introduction of the Zurich number of sunspots. The Zurich number, or Wolf number, is now known as the International Sunspot Number and is coordinated by the Sunspot Index Data Centre in Belgium. The Zurich number and the group sunspot number, R_G, are indicators of solar activity, the latter of which fills in missing sections of the Zurich record by combining data from various sources (Hoyt and Schatten, 1998). Both records are shown in Fig. 1.9, the Zurich number with dashed lines and the group sunspot number as a solid line. The sunspot record provides the longest direct proxy record of solar activity. Two important results emerge from the record: the strong ~11-year cycle and a long-term trend of the sunspot number amplitude increasing in time. This century-scale trend will be discussed in section 1.8.

1.6 Faculae, plages and the network

Faculae represent magnetic features that have the opposite effect on solar irradiance to that of sunspots and their relatives. Faculae are regions with higher contrasts than the quiet sun, as can be seen in Figs. 1.6 and 1.7. They can be found anywhere on the solar disk, but are always found near sunspots. At solar maximum, their influence increases the SSI at almost all wavelengths leading to the 0.1% amplitude in TSI over the solar cycle.

While observations of the entire disk show large, continuous areas of faculae, high resolution images reveal they are composed of small brightenings extending up to just 0.6″ (the Sun is 32’ across) or less than 500 km across (Lites et al., 2004). Although alternative models have been proposed to describe faculae, including the hillock and cloud models (Schatten et al., 1986), Spruit (1976) was the first to describe these features and his bright wall model has become the generally accepted model and remained relatively unchanged since. It basically describes small, vertical flux tubes located on the edge of granules, evacuated and thus made optically thin from a strong magnetic field on the order of 1500 G (Rouppe van der Voort et al., 2005). This can be seen in the facular schematic of Fig. 1.10, taken from Keller et al. (2004) and adapted, where the thin layer between the granule wall and magnetic boundary, on the right, is the source of most of the brightness observed (white shading). Radiation flows in from the bright walls and heats the inner tube to temperatures above those of the surrounding quiet sun. The path through the atmosphere to the observer is optically thinner than on the left and the oblique angle allows a larger area of the facula region to be seen.

The bright wall model is consistent with observations of facular contrast across the disk. Facular regions of high magnetic field strength are darker at disk centre, at visible wavelengths, and brighten towards the limb, while lower-strength faculae remain bright at the disk centre and increase in contrast slightly towards the edge (Ortiz et al., 2002). This fits with the idea that for a stronger magnetic field a larger flux tube base forms. With an increased diameter comes a reduced heating effect from the inflowing radiation, increasing the optical thickness and decreasing the contrast. In fact, micropores may represent a contrast transition, between darker faculae and pores, beyond a diameter of 300 km (Grossmann-Doerth et al., 1994). At disk centre only the cool, low contrast (if large) base is visible, but as it rotates towards the limb, foreshortening effects reduce the base area while the hotter, brighter walls come into view and the increase in contrast differentiates faculae from the surrounding quiet sun. Therefore, the contrast of faculae are a function of diameter and disk position (Ortiz et al., 2002; Fligge et al., 2000). By comparing the left image of Fig. 1.6 at disk centre with the right image of Fig. 1.7 near to the limb, the faculae appear more prominent and easier to identify near to the limb due to their three-dimensional nature.

Plages are also regions of increased contrast, but are visible only in certain spectral bands or lines, famously the Ca II (Shine and Linsky, 1972) and Hα lines. They are found in the chromosphere co-spatially with small magnetic features, such as faculae and micropores (see Fig. 1.6 and Ortiz et al., 2002). Although their relationship is not fully understood they appear to be associated with magnetic field lines from facular regions extending above the photosphere (Martinez Pillet et al., 1997).

The last magnetic construct considered within this chapter is the network. This feature is defined in relation to the intra-network as as in Keller et al., (1994): there is a network of 0-2.5 kG field flux tubes located (Domínguez Cerdeña et al., 2006) at supergranular[4] boundaries and an intra-network magnetic flux, within the supergranules, of very low magnetic field strength close to zero (Keller et al., 1994). Network elements consist of tiny, unresolvable flux tubes or magnetic elements, similar in contrast to faculae. Unlike faculae, though, they are thought not to be related to the large-scale solar dynamo but to small-scale dynamo action within small-scale convection closer to the surface. The rate of this small-scale low-flux emergence is anti-correlated with the solar cycle, but the dis-

[4]Supergranular regions are thought to be the highest order of convection that fragment into granules closer to the surface. They are a factor of 100 larger than the small scale granulation (e.g. in Fig. 1.6) and their boundaries can be identified as darker down-flow regions.

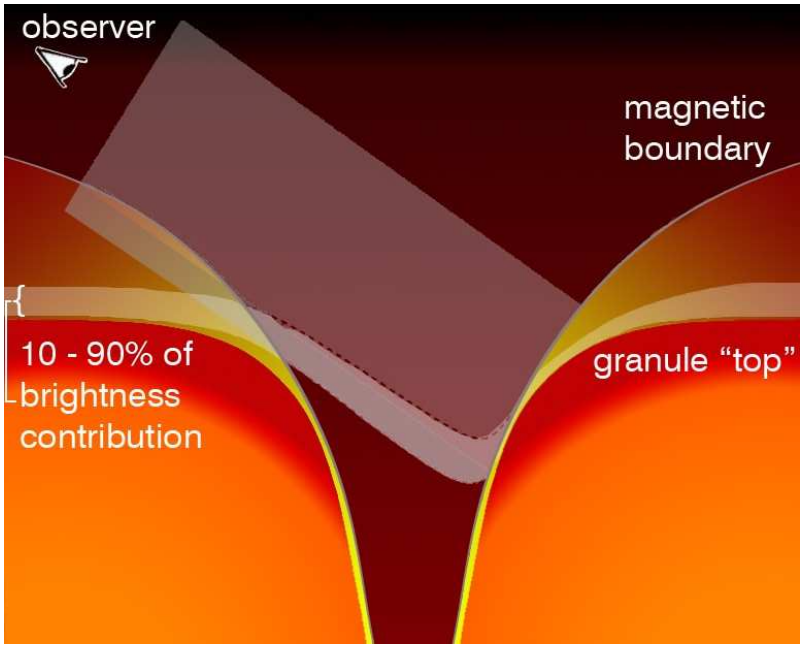


Figure 1.10: Schematic of magnetic flux tube, modified from Keller et al., (2004). Facular brightening originates at the granule wall (right) where deeper layers can be observed and where most of the brightness contribution comes from (white shading). Looking into the downward convective region at an angle increases the visible area of the bright wall making faculae more prominent when closer to the limb.

tribution of magnetic fluxes remains the same throughout, unlike for faculae (Hagenaar et al., 2003). Because the flux tubes are usually smaller than facular diameters, they are always brighter than the background quiet sun and their contrast remains relatively consistent from disk centre to limb (at least for fluxes between ~30 and 60 G, as in Ortiz et al., 2006).

The importance of the network in long term irradiance variations is not understood and so is not incorporated well within models. For the model employed in this thesis, the problem lies with a lack of resolution in the magnetograms used. While sporadic patches of concentrated kilogauss fields can be detected, within an unresolved region signed magnetic fields of opposite polarities can cancel each other out and result in lower observed magnetic fluxes (Lites, 2002; Krivova and Solanki, 2004). However, irradiance enhancement does not depend on the polarity of the magnetic field and, therefore, the effect from magnetic flux associated with the network are difficult to determine. Even at a resolution of an arcsecond the intra-network flux is not well determined and up to 99% of the quiet sun flux maybe be unresolved (Trujillo Bueno et al., 2004). It is expected that the influence of network flux tubes, along with facular regions, would change with a larger, global solar magnetic field (Foukal and Lean, 1988). The distribution of flux tube diameters is thought to be directly related to the distribution of magnetic flux in the photosphere, with the largest forming sunspots. Consequently, the distribution changes the overall contrast of the Sun and therefore solar irradiance. Since the quiet sun and the network cover the solar disk at all times, the distribution of magnetic flux in the photosphere may be an important factor in long term solar irradiance variations on both cycle and inter-cycle, secular, time scales.

1.7 Monitoring Solar Variability

If the long-term impact of the Sun on Earth's climate is to be fully understood, then a good understanding of solar variability is required. The Earth's atmosphere absorbs some of the incoming solar radiation. So, to accurately determine TSI and SSI, absolute and relative variations, above-atmosphere observations are needed. Even though sunspot numbers were known to vary over time, changes in the solar irradiance were not detected until monitoring satellites were put into orbit.

With the launch of the NIMBUS-7 satellite (Hoyt et al., 1992) in 1978, solar irradiance

science began in earnest. For the first time, using the onboard Hickey-Frieden (HF) radiometer (Hickey et al., 1988) to measure TSI, and the Solar Backscatter Ultraviolet (SBUV) for the UV, Earth's incoming solar radiation could be directly and continuously monitored from above the atmosphere. The first Active Cavity Radiometer Irradiance Monitor (ACRIM I) joined NIMBUS-7 from 1980 on board the Solar Maximum Mission (SMM; Chapman et al., 1996) and a third satellite was put into orbit alongside these two to continue the ERB mission with the Earth Radiation Budget Satellite (ERBS). This mission has continued TSI and UV observations up to the present day on National Oceanic and Atmospheric Administration (NOAA) platforms. The Upper Atmosphere Research Satellite (UARS) was launched in 1991 to replace and improve solar observations (Dessler et al., 1998). On board was the Solar Ultraviolet Spectral Irradiance Monitor (SUSIM, observing 115-411 nm) and the SOLar STellar InterComparison Experiment (SOLSTICE, observing 119-420 nm) (Brueckner et al., 1993; Rottman et al., 1993).

The SOlar and Heliospheric Observatory (SOHO; Scherrer et al., 1995) was launched in 1996 to observe and image many aspects of the Sun directly; it included the Variability of solar IRradiance and Gravity Oscillations (VIRGO; Pap et al., 1999) instrument to measure TSI. VIRGO will be discussed in more detail in chapter 3. In 2003, the SORCE satellite (Rottman, 2005) launched. Of the four instruments on board, the new SOLSTICE instrument, replacing UARS/SOLSTICE which ceased observations in 2001, and the Spectral Irradiance Monitor (SIM; Harder et al., 2005a) are of most interest in spectral irradiance studies. These two instruments cover a combined spectral range of 115-2700 nm (see sections 5 and 6 for further discusssion and comparisons). Along with the Total Irradiance Monitor (TIM; Kopp and Lawrence, 2005), SORCE has continued to monitor SSI and TSI up to present day. Other currently operating spectral instruments include the SCanning Imaging Absorption Spectrometer for Atmospheric CartograpHY (SCIAMACHY; Skupin et al., 2005), launched in 2002, which also covers a wide spectral range (240-2300 nm) and has a higher resolution than SORCE, though it is not calibrated for long-term degradation (Pagaran et al., 2009) and the data, in their present form, cannot be used for cycle length studies. The Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) mission launched in 2001 (Noël et al., 2003; Kozyra et al., 2004) and the on board Solar EUV Experiement (SEE; Woods et al., 2005) monitors spectral irradiance below 200 nm. The SOLar SPECtrum (SOLSPEC) instrument on the Solar Monitoring Observatory (SMO) on board the International Space Station has been observing the UV to far-IR (115-3000 nm) since 2008, but data are still to be released (Thuillier et al., 2009).

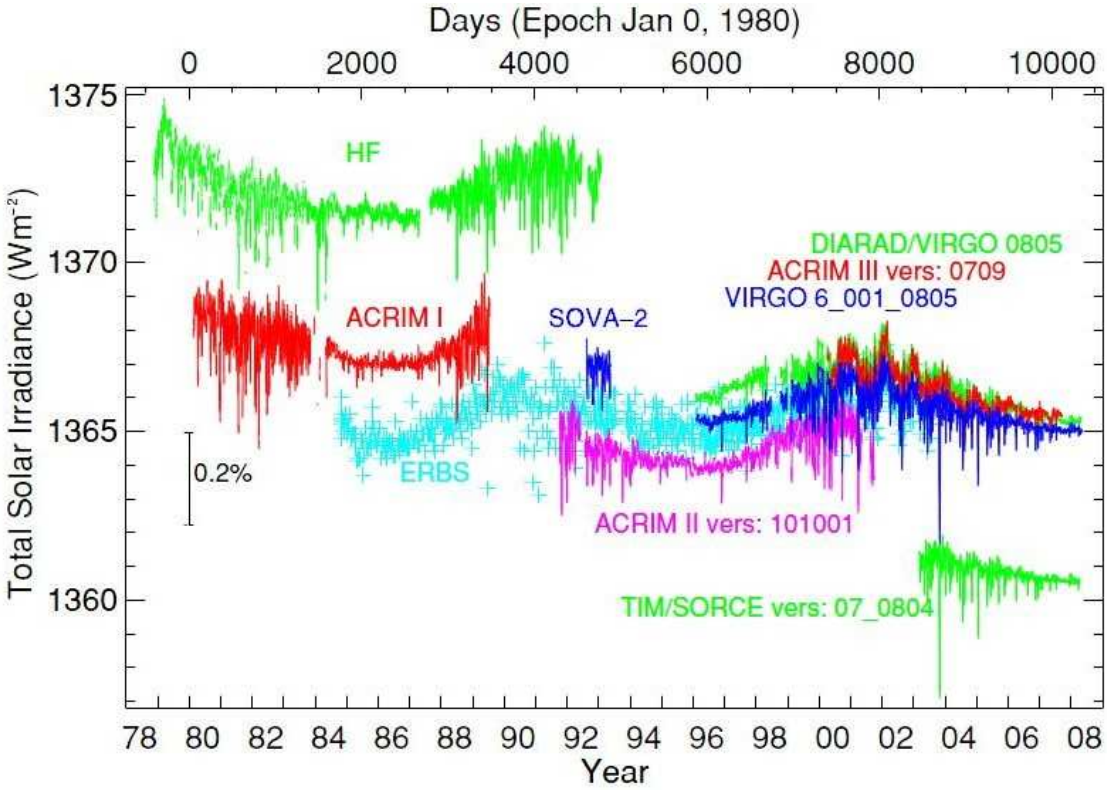


Figure 1.11: TSI data from various satellites since 1978, taken from Fröhlich (2009). Data from some of these satellites are combined to produce continuous composites in Fig. 1.12.

Although many satellites have been used to monitor spectral variability over the past 34 years, there is still much uncertainty in the absolute levels of spectral irradiance, and to what extent the Sun varies spectrally. Indeed, the solar cycle behaviour in UV wavelengths observed by older instruments do not agree with the behaviour seen by current instruments. This is discussed in chapter 5, where the SORCE/SIM data are directly compared with the SATIRE-S model (see chapter 2 for a description of the model); they both display very different cycle variability. Some of the spectral trends in SORCE/SOLSTICE disagree with SORCE/SIM. SORCE/SOLSTICE is also considered with respect to the SATIRE-S model in chapter 6. The disagreement between current observations and the models highlights the need for continued, high precision and accurate spectral irradiance observations covering extended periods of time.

1.8 The 11-year solar cycle

There are now nearly three 11-year solar cycles of irradiance data. Data began to accumulate on a continuous basis just prior to the maximum of cycle 21 (1979). Cycle 22 began at the following minimum, the time at which new cycles begin, in 1986. Cycle 22 peaked in 1989 and cycle 23 began in 1996, peaking in 2000. Cycle 24 began in January 2009 after an extended minimum.

In Fig. 1.11, from Fröhlich (2009), TSI data from various radiometers have been plotted. Absolute values of TSI, between some instruments, differ by nearly 1%, much greater than the cycle variation of 0.1%. To create a reliable and useful composite, each dataset needs to be corrected for instrument degradation and adjusted to match the same absolute level.

Approaches taken to compose the data into a single and consistent long-term dataset are detailed in Fröhlich (2006). Fig. 1.12 shows three composites that have employed different approaches to correct for these differences. The very clear and large early differences in the long-term trend, between the PMOD and ACRIM/IRMB composites, lie chiefly in the corrections by Fröhlich and Lean (1998) attributed to the data gap between ACRIM I and II, seen in Fig. 1.11, and to slips in the HF data caused by a change in the orientation of instrument (detailed in Fröhlich, 2006).

What is clear from all the composites is that during the cycle maxima, the absolute level and variability of TSI are higher, due to facular coverage and spot transits respectively.

The rotational variability is caused by the passage of active regions onto, across and off the face of the Sun and by the evolution of active regions during their transit (Foukal and Lean, 1986; Fligge et al., 2000); rotational variability has a periodicity of approximately 27 days. Very large sunspot groups cause short decreases in TSI, sometimes exceeding 0.2%, but the enhancing effect of facular and network regions over periods longer than the solar rotation, emphasised by the black line of the running mean, causes the overall enhancement of solar irradiance over the solar cycle (Lean et al., 1998; Wenzler et al., 2006; Krivova and Solanki, 2008).

An important disagreement between the composites is on the values of TSI at cycle minima. Fröhlich and Lean (1998) conclude from the PMOD composite that there has not been a change in TSI between the minima in 1986 and 1996, but that there has been a decline of 0.2 Wm^{-2} between 1996 and 2008 (Fröhlich, 2009). The ACRIM composite by Willson (1997) suggests that minima between cycle 21 and 22 had increased by at least 0.5 Wm^{-2} (Willson and Mordvinov, 2003) with a return to 1986 levels in 2008. The IRMB composite (Dewitte et al., 2004) shown in this figure has changed since the figure was published. It no longer extends back beyond 1984 and does not show any decline between 1996 and 2008 (private communication, S. Dewitte). The latest versions of all three datasets can be seen in chapter 3 where the differences are discussed in depth with respect to the SATIRE-S model.

The difference in the minima values leads to the important possibility of changes in the solar irradiance on secular time scales over hundreds of years. However, taking the composites on face value, there is a clear uncertainty in the secular behaviour of the Sun arising from the way the composites have been corrected and stitched together. In Chapter 3, the work by Wenzler et al. (2006) using the SATIRE-S model is reconsidered and extended to cover the period encompassed by the three TSI composite from 1978 to 2009. The significance of this work can been seen in the following example. As the reconstruction is now essentially independent of the composites, it provides a more robust test of which composite is likely to be correct. Previously, comparisons between the reconstruction and the composites required tailoring the reconstruction to each composite (Wenzler et al., 2009). Scafetta and Willson (2009) used a long-term version of SATIRE (Krivova et al., 2007), version -T (see section 2.1), to fill the so-called ACRIM-gap (1989-1991) and claimed that the result showed there had been an increase in TSI between the solar cycle minima of 1986 and 1996, in support of the ACRIM composite. This led Scafetta and Willson (2009) to suggest that the influence of the Sun on global warming had been significantly underestimated. However, the version of SATIRE they used was not ap-

propriate for the period replaced and SATIRE-S should have been used instead (Krivova et al., 2009), which when employed does not support the ACRIM composite. The work in this thesis provides stronger evidence that the PMOD composite is likely to be the most accurate representation of TSI, with agreement between the model and the level of solar irradiance at all three solar minimum. In addition to TSI, compared to the previously observed minima, the solar minimum of 2008 was exceptionally long and recorded unusually low levels in many solar index records, though not all showed the same relative decline as TSI (Fröhlich, 2009).

Although it appears non-trivial to create a long-term TSI composite, the creation of a long-term UV dataset is even more difficult. A composite has been produced by DeLand and Cebula (2008) covering the spectral wavelengths 120-400 nm. However, this composite only accounts for absolute differences between different instruments and does not take into account instrumental problems and trends. Absolute irradiance offsets and degradation corrections remain an issue. Corrections must be applied to every wavelength from a single instrument and each instrument does not necessarily observe exactly the same wavelengths, either because the spectral resolution or the central wavelength positions are not the same. This is true even of duplicate calibration instruments. If rapidly varying features are not sufficiently resolved then the observed wavelength interval may not vary in the same way for different satellites. This would then make it more difficult to determine the long-term variability at each particular wavelength when datasets are combined. Lockwood (2011) attempted to intercalibrate UARS/SUSIM with SORCE/SOLSTICE data to investigate UV changes in cycle 23, but this was only possible for a restricted range between 156-208 nm.

Nevertheless, observations of the UV over the past three decades have revealed complex spectral behaviour and a wavelength dependent contribution to TSI change. Fig 1.13 shows three integrated UV spectral regions from UARS/SUSIM and UARS/ SOLSTICE and the change that occurs in these regions over an 11-year period. As with TSI, variability and irradiance in the UV increases at solar maximum. While the TSI varies by 0.1% over the solar cycle, results from these two instruments have revealed that the 200-300 nm region varies on the order of a few percent, 150-200 nm by 10-20% and shorter regions by over 50% (Woods et al., 1996; Floyd et al., 2003; Krivova et al., 2006). In Fig 1.13, cycle variability is approximately 13%, 6% and 2% in the 160-165, 200-205 and 230-235 nm regions, respectively. Although the contribution to the absolute value of TSI from regions below 300 nm is small, ~1%, it provides a large contribution to TSI variability, around 20%; up to 60% of TSI variability might be accounted for below 400 nm (Lean

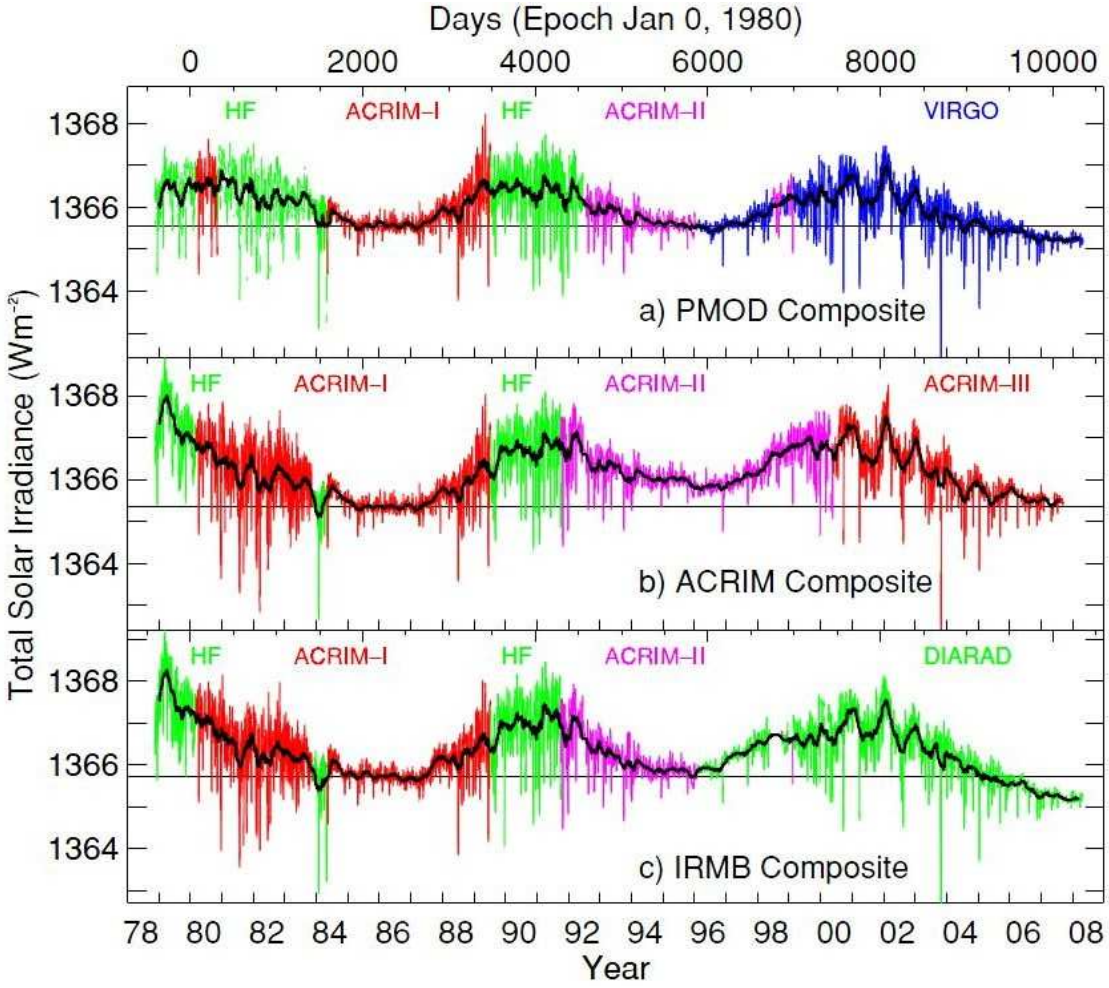


Figure 1.12: Three TSI composites as published in Fröhlich (2009): (**top**) PMOD, (**middle**) ACRIM and (**bottom**) IRMB (now outdated). The time series begins at maximum in cycle 21 and extends to the end of cycle 23. The black line in each is the running mean.

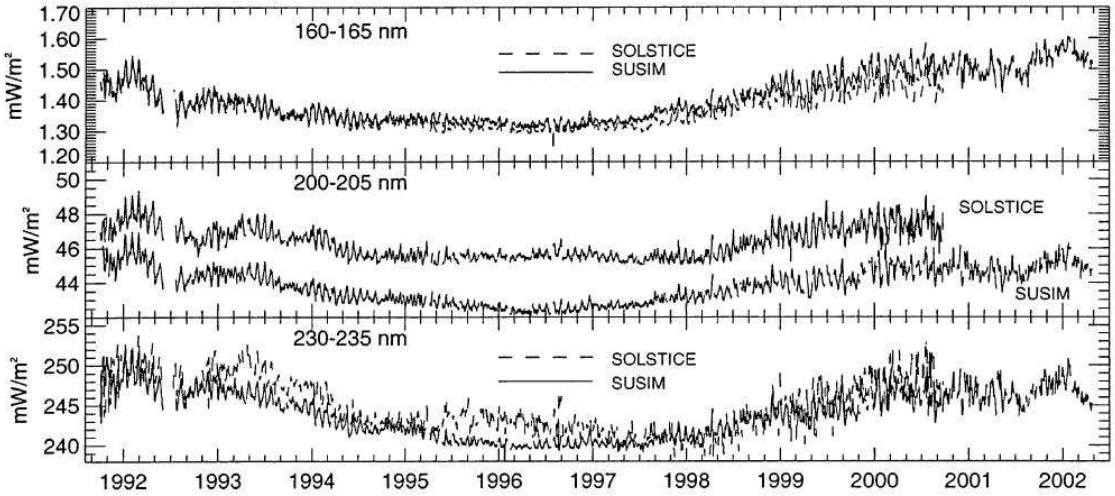


Figure 1.13: The integrated spectral regions from UARS/SUSIM and UARS/SOLSTICE at 160-165 nm, 200-205 nm and 230-235 nm, over a solar cycle length period from the end of 1991. From (Floyd et al., 2003)

et al., 1997; Krivova et al., 2006). Most radiative energy comes from the visible and near-IR and, prior to the SORCE mission, this was a region that was not well observed and thought to vary little over the solar cycle.

1.9 Changes since the Maunder Minimum and Beyond

From the possible change reflected in TSI between composite minima, as well as the increase in sunspot number amplitude over the last few hundred years as in Fig. 1.9, there are hints that the Sun's irradiance may vary over much longer periods than the solar cycle. Because of the relationship between sunspots and the magnetic fields produced by the solar dynamo, the variation in the cycle properties of sunspot number implies that there may be a more complex process than a simple 11 or 22 year cycle.

One famous example of this very long-term change in the Sun is that during the Maunder Minimum (1645-1715) almost no sunspots were observed and winter temperatures in Europe declined (Wuebbles et al., 1998; Lockwood et al., 2010). The sunspot number record is the longest record of solar activity. It is an important tool for understanding the long-term changes in the Sun. However, while linking sunspot numbers to a change in solar irradiance over the solar cycle is relatively straightforward task, relating it to a change on a longer time scale is more difficult. This is because the zero level in sunspots does not correlate with zero irradiance. To understand the behaviour of the Sun over hundreds of

years, other proxies are required.

Open solar flux on the Sun is a small subset of total photospheric surface magnetic flux and is defined as the magnetic flux passing through an arbitrarily chosen surface at 2.5 solar radii (Fisk and Schwadron, 2001; Owens et al., 2011). The closed flux is therefore confined within this boundary and is typified by coronal loops anchored to photospheric magnetic features, such as sunspots, and is the majority component of the magnetic field on the Sun . The open flux is weaker than the closed flux, but the field can survive for years, much longer than the ephermeral and active regions. Krivova et al. (2007) linked active and ephemeral regions with the open solar flux through a set of differential equations for the various categories of magnetic flux, with constraints determined from the sunspot and geomagnetic activity records. Thus, with magnetic flux presenting itself in various guises and for overlapping time periods, a possible mechanism for variation on scales longer than the 11 year cycle is realised.

By estimating the impact of ephemeral and open flux along with active regions from historical data, such as the sunspot group number, estimates for an increasing, secular change in TSI since the Maunder Minimum have been made. These estimates include increases of 1.3 Wm^{-2} (Balmaceda et al., 2007; Krivova et al., 2010), 2.8 Wm^{-2} (Lean, 2000) and even changes greater than 16 Wm^{-2} (Mendoza, 1997).

Explanations for secular trends other than solar surface magnetic activity have also been proposed. E.g., Rozelot et al. (2004) suggested that changes in the Sun's radius could account for long-term changes in the Sun. Harder et al. (2009) proposed that there may have been a change in the photospheric temperature gradient that produced a spectrally dependent change, increasing irradiance at some wavelengths and decreasing at others. A global temperature change has also been proposed (Tapping et al., 2007) with Fröhlich (2009) concluding that the recent minimum was caused by a global temperature decline of 0.2 K in the Sun's effective temperature.

The Sun is also thought to vary on even longer time scales, reaching periods of low and high magnetic activity known as grand minima and grand maxima. Indirect measurements through the ^{14}C tree ring record suggest that up to 29 grand minima-type events have occured over the past 8000 years (Usoskin et al., 2007). However, the cause of this variation is not fully understood as it has not been directly observed or successfully modelled. It has been suggested that the Sun may be entering a grand minimum state, though there is still some debate on this (Lockwood, 2010; Feynman and Ruzmaikin, 2011; Solanki and

Krivova, 2011).

1.10 Models of solar irradiance

Using satellite data to understand solar irradiance is vital, but it only allows the period after 1978 to be considered. To determine the range of solar variability and the time scales over which irradiance changes occur, models must be developed. There are various approaches to this problem, but ultimately successful models must take into consideration the causes of irradiance variation.

There are two main approaches to modelling irradiance variation. The first is proxy-based and uses regression and correlation to link spectral observations and solar proxies to create and extend a spectral dataset. Proxies such as the Ca II H&K index (Preminger et al., 2002), Mg II index (Lean et al., 1997; Fröhlich and Lean, 1998; Fligge et al., 1998; Lean, 2000) and the 10.7 cm flux (Hinteregger et al., 1981) have been utilised to represent the sunspots and faculae and their role in irradiance variations. The most widely used proxy-based model is the National Research Laboratory SSI (NRLSSI) model (Lean et al., 1997, 2005). This model is described in detail in chapter 6 where it is compared with the SATIRE-S model.

The other approach is physics-based and makes use of atmospheric models in combination with spatially resolved solar disk images decomposed into different components, usually sunspots and faculae and sub-classes of these objects. SATIRE-S is this type of model. The model atmospheres generally employed to calculate the radiative intensity are one-dimensional and semi-empirical in nature with a temperature and density structure as a function of height in the solar atmosphere. Each solar component (quiet sun, umbra, etc.) has a different model atmosphere from which emergent intensities are calculated as a function of wavelength and disk position.

One major difference in the approaches taken by different groups concerns how bright features and their contrasts are identified. Models using the Precision Solar Photometric Telescope (PSPT) full disk images infer contrasts from spectral lines such as the Ca II K line, e.g. Ermolli et al. (2009b, 2010). The SATIRE-S model, on the other hand, uses magnetograms to quantify magnetic field strength spatially and attributes a contrast according to this field strength. Thus, if the temporal variation of the magnetic field is responsible for irradiance changes then this should be a much more direct approach, though

Ca II K filtergrams cover a longer period back to early part of the 20th century (Ermolli et al., 2009b).

To understand the impact of the Sun on climate change, it is imperative that the physical mechanism behind the variation of the Sun on secular time scales is understood to a sufficient level to be able to confidently reconstruct solar variability into the past. This would then allow reconstructions that extrapolate TSI proxy data, such as sunspot number and the geomagnetic aa index, to be refined and improved (Tapping et al., 2007; Steinhilber et al., 2009; Vieira and Solanki, 2010). The potential range for a secular change in TSI over the last four hundred years, since the Maunder Minimum, could be significant in climate forcing and a better understanding of the Sun during the satellite era could lead to more precise estimates of the Sun's true effect on the Earth over that period, although there are diverging views on this. On the one hand, Shapiro et al. (2011) argue that because the Sun has been in a particularly active phase during the last century, one cannot currently extrapolate the full range of solar variability on secular scales. On the other hand, Schrijver et al. (2011) suggest that solar activity during the 2008 minimum has reached levels similar to the Maunder Minimum, a period during which sunspots were conspicuous by their absence, and cannot go much lower.

If physical models, like SATIRE-S, can accurately recreate solar irradiance, on all observed time scales, as shall be seen in chapter 3, then they can help constrain these models that try to reconstruct irradiance on significantly longer time scales for which only one or two proxies exist.

Chapter 2

The SATIRE Model: review and updates

Overview

This chapter is intended as an introduction to the SATIRE-S model. SATIRE-S will be used to reconstruct and investigate TSI and SSI throughout this thesis. Section 2.1 is a description of the model alongside a discussion of the physics it is based on. In section 2.2 I elaborate on the model's assumptions and consequent uncertainies, before discussing in detail the model's only free parameter. In section 2.3, I explain the improvements I have introduced. I finish (in section 2.4) with a discussion of different fitting procedures that can be used, and a selection of one of these, to gain the best results from SATIRE-S.

2.1 The workings of the model

The Spectral And Total Irradiance REconstruction (SATIRE) model has four components and makes one simple assumption: that all irradiance variation is due to the distribution and evolution of solar magnetic fields. There are three flavours of the model: SATIRE-S, with S denoting its use with satellite-era data, i.e. using full disk continuum intensity images and magnetograms; SATIRE-T, covering the telescopic era and using sunspot numbers; and SATIRE-M which employs cosmogenic isotope data over millenial time scales. A review of these SATIRE models can be found in Krivova et al. (2011b). This thesis focuses on the satellite era only and uses the SATIRE-S model. The schematic in Fig. 2.1 will help to explain how the model works.

SATIRE-S was originally built by Fligge et al. (2000) and further developed (Krivova et al., 2003) and employed in numerous studies (Wenzler et al., 2004; Krivova and Solanki, 2005; Krivova et al., 2006; Wenzler et al., 2009; Krivova et al., 2011a). The

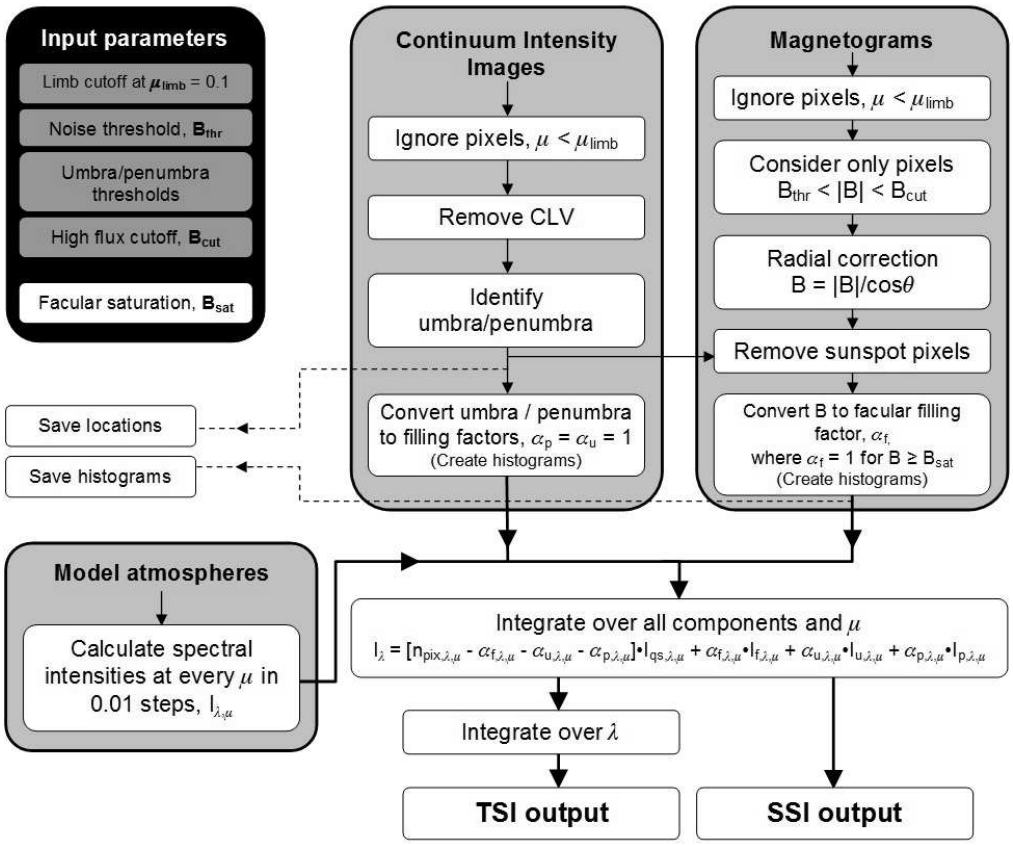


Figure 2.1: Schematic of the SATIRE-S model showing the processing order and combination of the three key ingredients needed for reconstruction: the continuum intensity images, magnetograms and model atmospheres. The facular saturation is varied as a free parameter.

assumption that magnetic fields cause irradiance changes in SATIRE-S requires a way of identifying surface magnetic features and converting their influence into an effect on irradiance. Three of the four features, or components, are identified using full-disk continuum intensity images and magnetograms: umbra, penumbra and faculae. The remaining area is attributed to the fourth component, the quiet Sun.

Sunspot umbra and penumbra are identified using the continuum intensity images, with pores generally classed as penumbral-type features. These components are identified using continuum intensity images. Two examples are given in Fig. 2.2, where masks of penumbral and umbral pixels are shown during a large sunspot transit (top left) and at solar minimum (bottom left) when no sunspots were present on the disk. Contrast thresholds are set to differentiate between the average quiet Sun and the penumbra and umbra, the latter of which is the darkest component. The value of the thresholds is dependent on the instrument used to take the continuum intensity images. To avoid the effect of limb darkening restricting the identification of sunspots, or confusing penumbra and umbral pixels, the centre-to-limb variation (CLV) is removed before the features are identified using an approximation similar to Neckel and Labs (1994). This normalises the contrast of the disk at all μ so that sunspots are easily identified everywhere on the disk. If a pixel has a contrast below either umbral or penumbral thresholds then it is assigned a filling factor of $\alpha^{u}_{i,j} = 1$ or $\alpha^{p}_{i,j} = 1$, where i and j are the x and y coordinates of the image pixel. The pixel position relative to the solar disk centre and radius allows positional information in terms of the limb angle to be retained. The limb angle, $\mu = \cos\theta$, where θ is the heliocentric angle, is the angle between the LOS and the radial vector. This is the only spatial component required as the intensities of all components depend on μ.

The smaller, bright facular components are difficult to identify in continuum images. For a large range of μ, particularly at disk centre where the bright walls cannot be seen and show no strong contrast relative to the quiet Sun (Topka et al., 1992) when observed in the visible or IR wavelengths typical of continuum intensity images. The most significant problem when trying to identify faculae in continuum images is that they are smaller than the typical area covered by a pixel[1]. As such, when averaged over the whole pixel, the large intrinsic contrast of faculae is washed out by the surrounding quiet Sun. Since faculae are pervaded by strong magentic fields, their location and concentration can instead be identified in magnetograms. Magnetograms are essentially images of the Sun where each pixel describes the LOS magnitude of the magnetic flux. For the purposes of a reconstruc-

[1]For example, a pixel at the centre of a full-disk MDI image covers approximately 1400 x 1400 km^2, compared to a typical facula which is approximately 100 km in diameter.

tion only the unsigned flux is considered as the polarity of the field does not influence the observed contrast. The procedure to isolate and identify facular pixels is as follows.

First, pixels at or below the level of the magnetogram noise must be removed. The noise level, σ_{mag}, is determined in the magnetograms by using a period of very low activity to identify the lowest level of unsigned magnetic flux. The noise level is calculated at low periods of activity to avoid active-regions, and generally higher levels of magnetic flux, that would lead to an incorrect, overall higher calculated level of noise. The spatial distribution of the noise is calculated, in a similar manner to that used by Ortiz et al. (2002), by using a small boxed region centred on each pixel (e.g. 100×100 pixels was used to calculate σ_{mag}(i,j) for MDI magnetograms in chapter 3). If the value of the pixel exceeds three standard deviations of all pixels in the box, the central pixel is replaced with the

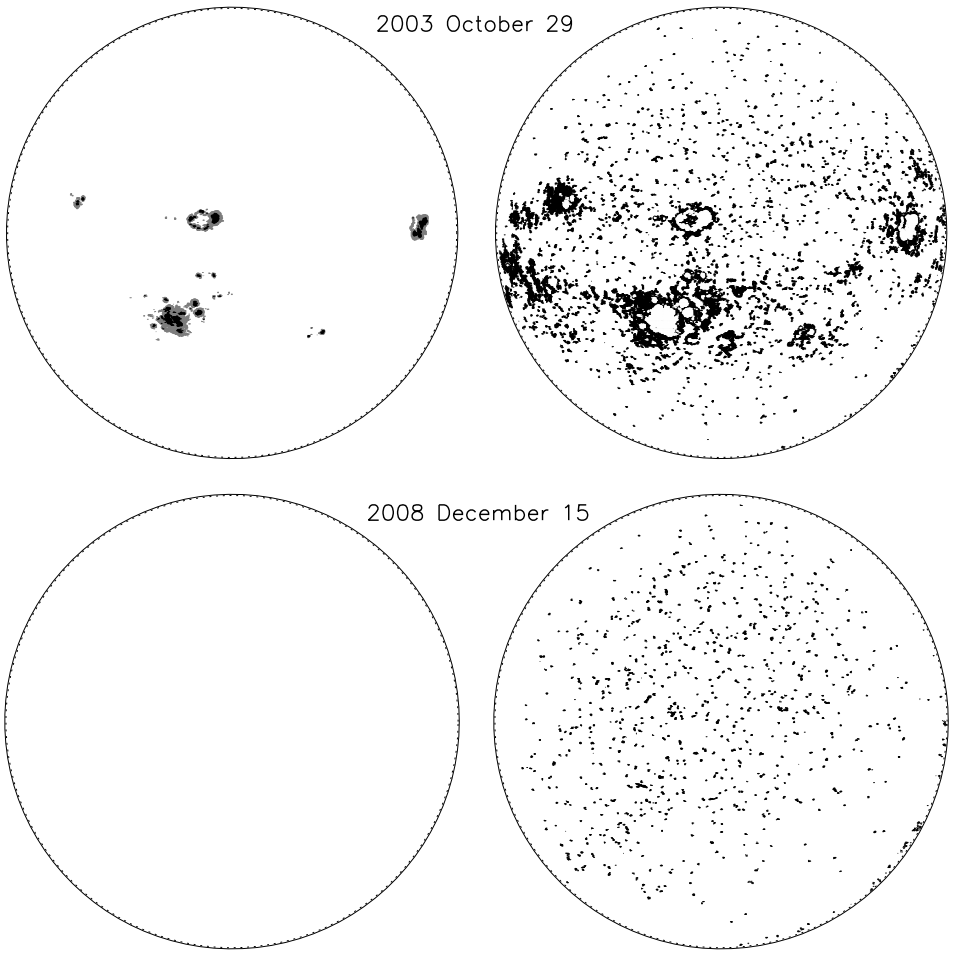


Figure 2.2: Sunspot masks (umbra in black, penumbra in grey) from MDI continuum intensity images (left) and facular masks from MDI magnetograms (right), as derived from SATIRE-S during high activity (top) and low activity during a solar minimum (bottom).

mean within the box. Repeating this two further times on the same magnetogram and then taking the median of several low-activity magnetograms produces a map of the lowest level magnetic flux detected by the instrument. This is taken as the noise level in the instrument, σ_{mag}. A threshold level, $B_{thr} = 3\sigma_{mag}$, is used that should remove approximately 99% of noise pixels. This procedure is discussed in more detail when applied to MDI magnetograms in Chapter 3.

In Kitt Peak (KP) magnetograms (see Chapter 3), the spatial distribution of B_{thr} is relatively uniform so a constant is used as a threshold (Wenzler et al., 2006). In SoHO/MDI this is not the case (Ortiz et al., 2002), so a position-dependent threshold is required (see Chapter 4 for a detailed discussion of MDI magnetograms). If any non-sunspot pixel on the disk has a value of $B > B_{thr}$ it is considered to contain faculae, unless it is an individual pixel that does not neighbour any other facular pixels, which results in it being rejected as noise. Two examples of facular masks derived from SATIRE-S are shown on the right of Fig. 2.2, with high activity in the top right and solar minimum activity in the lower right. In the top right, the sunspot pixels from the top left sunspot mask have been removed, as prescribed in Fig. 2.1. A modification to the model removes pixels with flux above an upper threshold, B_{cut}, as they were found to be misidentified pixels. This is discussed and justified in section 2.3.

The magnetograms must also be corrected for the line of sight (LOS) since the pixel-averaged magnetic flux will be measured as $B_{i,j}=<|\mathbf{B}_{i,j}|\cos\alpha>$, where α is the angle between the LOS and the radial magnetic field vector, **B**. It is used to make a first order approximation, $B_{i,j}/\mu$, by assuming that $\alpha \approx \theta$, i.e. that the field is predominately radial. Due to foreshortening at the disk edge, large errors can be incurred when dividing by μ as this will enhance the noise. This is why regions close to the solar limb, i.e. $\mu < 0.1$, are excluded from the beginning, as detailed in the schematic of Fig. 2.1 and marked as dotted lines at the solar limb in Fig. 2.2.

Faculae do not always physically fill the entire area in a pixel, so the relative area of the pixel filled by faculae must be attributed a filling factor, $\alpha^{f}_{i,j}$. In SATIRE-S, B_{thr} acts as the lower limit and above this it is assumed that this filling factor is dependent solely on a linear relationship with $B_{i,j}$, up to a saturation value, B_{sat}. This is illustrated in Fig. 2.3 which has been taken from Fligge et al. (2000) and adapted for this thesis. For $B \geq B_{sat}$, pixels are considered to be completely filled by faculae and all flux tubes are assumed to be equally bright.

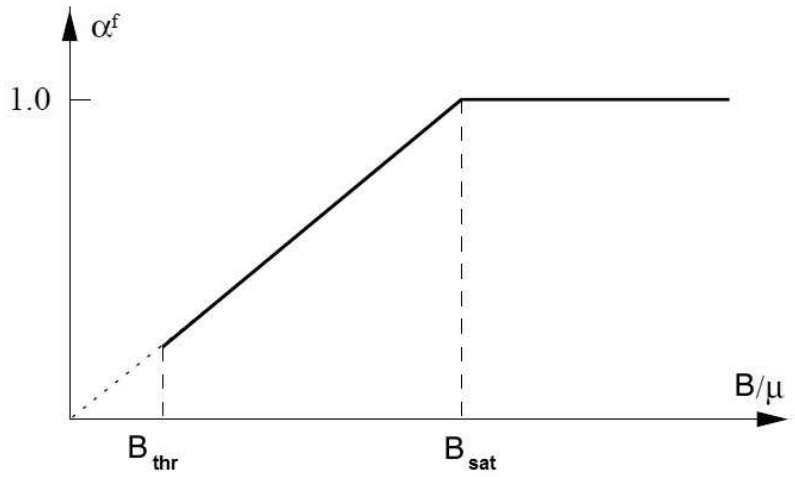


Figure 2.3: Determining the filling factor, $\alpha^{f}_{i,j}$ for facular pixels using the magnetic flux, after correcting for the limb angle. Taken and modified from Fligge et al., (2000).

The value of B_{sat} is treated as a free parameter. B_{sat} is varied to attain the best agreement with observational TSI data. In terms of the reconstruction, B_{sat} changes both rotational (~27 day) and cyclical variability (~11 year). The effect of B_{sat} on the reconstruction is discussed in detail in section 2.2.

Where $\alpha^{f}_{i,j} < 1.0$ the remaining filling factor is attributed to the quiet sun, $\alpha^{qs}_{i,j} = 1 - \alpha^{f}_{i,j}$. All remaining pixels on the solar disk that are not facular, penumbral or umbral are given a quiet sun filling factor of $\alpha^{qs}_{i,j} = 1$.

The reconstruction becomes a case of calculating the filling factor of each component for each limb angle and associating emergent intensities to each of these. The emergent intensities are calculated from model atmospheres at 11 limb angles between $\mu = 1.0$ and $\mu = 0.05$ and interpolated onto intervals of $\Delta\mu = 0.01$. In the top plot of Fig. 2.4, the disk-centre ($\mu = 1$) intensities of each component can be seen between 200 and 1600 nm; the bottom plot shows the contrast relative to the quiet sun at disk centre (solid line), $\mu = 0.7$ (dashed line) and close to the limb ($\mu = 0.05$, dotted line). Note the strong limb dependence of faculae compared to the weak dependence of umbra and penumbra. The quiet sun (black), penumbral (blue) and umbral (red) intensities in Fig. 2.4 are derived from ATLAS9 (Kurucz, 1993) with spectral intensities calculated at temperatures of 5777, 5450 and 4500 K respectively. The facular and network intensities (green) use a modified FAL P model atmosphere (Fontenla et al., 1993) that was adapted by Unruh et al. (1999). The main modification was to remove the temperature reversal in the chromosphere, at the temperature minimum, so that the model would work with the ATLAS9

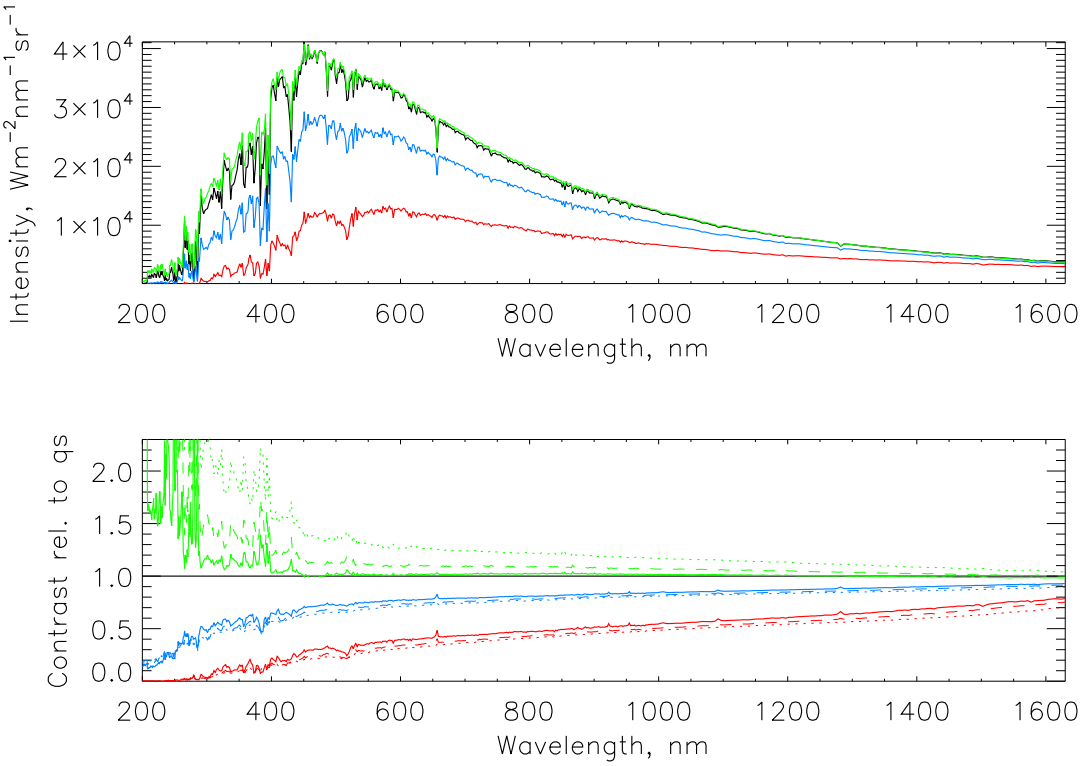


Figure 2.4: The disk-centre spectral intensity (**top**) and contrast relative to the quiet Sun (**bottom**) for each model component: quiet Sun (black), penumbra (blue), umbra (red) and faculae and network (green). In the bottom plot, solid lines are for disk-centre components, dashed lines at $\mu = 0.7$ and dotted lines at $\mu = 0.05$.

code (see Unruh et al., 1999). The model assumes LTE at all atmospheric heights, which is not adequate for regions below ~350 nm and within the cores of stronger lines. To remedy this, Krivova et al. (2006) used correlated UARS/SUSIM observations to empirically correct the model at these shorter wavelengths. This modification is not introduced for the spectral comparisons of chapter 5 but becomes relevant in chapter 6 where it is discussed in more detail.

Finally, the intensity contribution of each pixel to the total and spectral irradiance is calculated. This is first done by summing up the filling factors of each component within each annulus between μ and $\mu - \Delta\mu$ in $\Delta\mu = 0.01$ steps. For example, $\mathrm{A}_\mu^f = \sum \alpha_{\mathrm{i,j}}^f$ within each annulus. By doing this for all components at all limb angles,

$$I^{tot}(\lambda, t) = \sum_\mu [A_\mu^u(t) I_\mu^u(\mu, \lambda) + A_\mu^p(t) I_\mu^p(\mu, \lambda) + A_\mu^f(t) I_\mu^f(\mu, \lambda) + (N_\mu - A_\mu^u(t) - A_\mu^p(t) - A_\mu^f(t)) I_\mu^{qs}(\mu, \lambda)], \quad (2.1)$$

where $I_\mu^{u,p,f,q}(\lambda)$ are the intensity contribution by wavelength of the umbra (u), penumbra (p), facular (f) and quiet sun (qs) pixels and N_μ is the number of pixels in each annulus. This equation is given in terms of μ instead of i and j because the intensity as a function of limb angle is the important criterion. When equation 2.1 is applied to a series of temporally matched magnetogram and continuum intensity images, spectral irradiance variations can be reproduced. The irradiance change by wavelength, relative to an inactive, quiet Sun, is calculated by

$$\frac{\Delta S_\lambda}{S_\lambda} = \frac{\sum_\mu I^{tot}(\mu, \lambda) - \sum_\mu I^{qs}(\mu, \lambda)}{\sum_\mu I^{qs}(\mu, \lambda)}. \quad (2.2)$$

$\Delta S_\lambda / S_\lambda$ is calculated for wavelengths between 9-160 000 nm, though in practise the derived intensity contrasts are only useful for wavelengths above ~200 nm. This is discussed further in chapters 5 and 6.

When integrated over the full spectrum, SATIRE-S produces total solar irradiance. An example of this is shown in Fig. 2.5. In this figure, TSI from SORCE/TIM (green), between

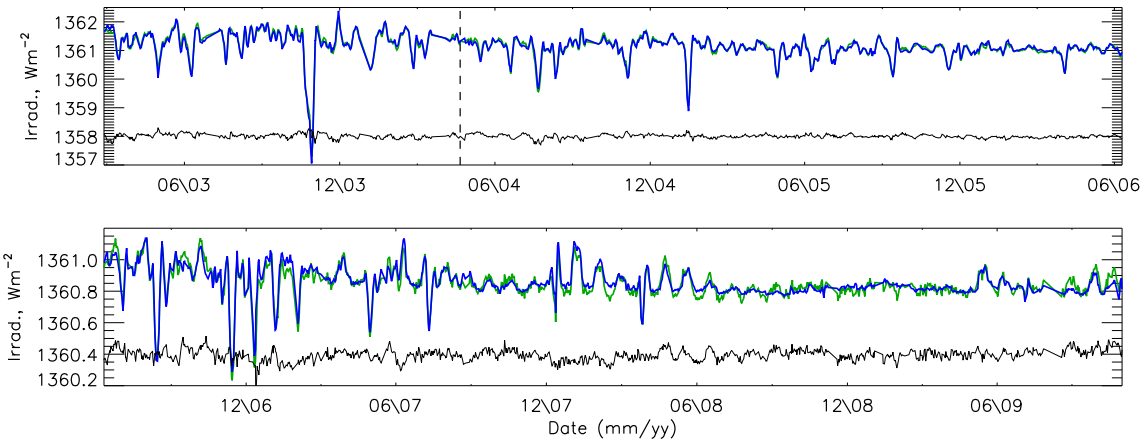


Figure 2.5: Irradiance lightcurves of TIM (green) and SATIRE-S integrated over 200-160,000 nm (blue) between 2003 February 25 and 2009 October 2009 31. The upper plot shows the first half of the period and the irradiance scale covers 5.5 Wm^{-2} while the lower half is on a reduced irradiance scale of 1.0 Wm^{-2} to highlight the differences in the quieter minimum period of 2008 and 2009. The black line shows the difference between TIM and SATIRE-S TSI, shifted up by 1358 Wm^{-2} in the upper frame and by 1360.4 Wm^{-2} in the lower. The vertical dashed line marks the date from which SORCE/SIM observations are available.

2003 and 2009, is overplotted with SATIRE-S (blue) normalised to TIM at the solar minimum of December 2008. SATIRE-S is constructed here using MDI data only and B_{sat} set to 443 G. The correlation coefficent is 0.987 and therefore the model accounts for over 97% of the variability during this time period. Using SATIRE-S, Krivova et al. (2003) was able to account for 92% of the variability in the ascending phase of cycle 23 and Wenzler et al. (2006) used SATIRE-S to reproduce 83% of TSI variability between 1978 and 2003. These results show that SATIRE-S recreates TSI well during active periods, as shown in the upper frame, but during the minimum (lower frame, with a smaller scale), any low-level decaying weak magnetic flux of active regions remaining from the end of the current solar cycle are not picked up in the five-minute MDI magnetograms that were used, so that the small rotational variation is not well reproduced, from e.g. May 2008 to April 2009.

For the rest of this chapter, SATIRE-S is discussed in terms of TSI only. There is a detailed discussion of the TSI reconstruction in chapter 3 together with comparisons with TSI observations where very good agreement is found between the model and observations on all time scales.

2.2 Assumptions and uncertainties of the model

2.2.1 Treatment of faculae

The treatment of faculae in SATIRE-S involves several assumptions to approximate their true physical nature. In reality the relationship between magnetic flux and the brightness of faculae is far more complicated than the simple approximations employed in SATIRE-S. This section acts to explain some of these approximations. Despite the simplifications, very good agreement is found with TSI and SSI observations on all time periods considered, except for cycle variability in comparison to the SORCE/SIM instrument (see Chapter 5).

The main assumption for faculae is the linear relationship between magnetic flux and the facular filling factor, as in Fig. 2.3. Through the free parameter, B_{sat}, it has a significant impact on determining a good reconstruction. B_{sat} sets the rate at which faculae fill a pixel with increasing magnetic flux. If the value of B_{sat} in Fig. 2.3 was reduced, i.e. moved to the left, (for the same magnetogram) it would mean that more pixels have a facular filling factor of 1, i.e. that they are completely filled ($B \geq B_{sat}$). In addition, all pixels with magnetic fluxes between B_{thr} and B_{sat} will have larger filling factors than for the original, higher value of B_{sat}. The opposite is true if B_{sat} is increased, so that the sum of all facular filling factors would reduce. Changing B_{sat} has two main consequences for the reconstruction.

Figure 2.6 illustrates the effect that varying B_{sat} has on rotational and cyclical variability, relative to some arbitrary value of B_{sat} = 450 G. The top of Fig 2.6 shows a smoothed timeseries while the bottom plot is the detrended timeseries (detrended is the original[2] minus smoothed time series). Three timeseries are shown with different values of B_{sat}, with B_{sat} = 400 G in red, 450 G in black and 500 G in blue. All time series are normalised to December 2008. In the top plot, the difference between the reconstructions with B_{sat} = 400 G and with B_{sat} = 450 G is shown as a thin red line. Similarly, the difference between reconstructions as with B_{sat} = 450 G and B_{sat} = 500 G is shown in blue. The differences are shifted up to 1360.7 and a dotted line shows the level relative to the difference at December 2008. The difference in rotational variability is very small and has been omitted in the lower plot. Slight differences can be seen at the peaks and troughs of the detrended lightcurve in the lower plot resulting from the change in B_{sat}.

[2] 'Original' means the data 'as is'. It is used, when discussing timeseries to differentiate between detrended and smoothed timeseries.

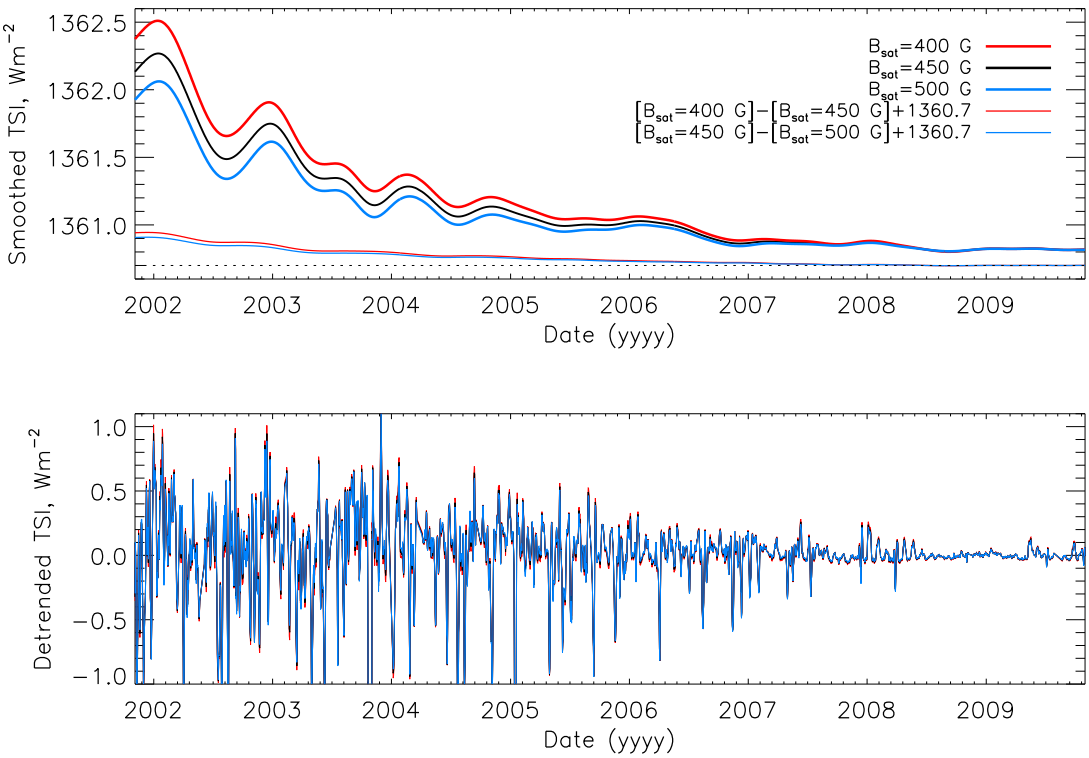


Figure 2.6: SATIRE-S reconstructions between 2001 and 2009, based on MDI data. Reconstructions have B_{sat} set to 400 (red), 450 (black) and 500 G (blue) and are normalised to December 2008. The **top** plot shows the smoothed time series, while the lower shows the detrended (original minus smoothed timeseries). The difference relative to the reconstruction using B_{sat} = 450 G are shown by the thin lines below the smoothed time series in the top plot. The dotted line aids the eye when considering the differences relative to December 2008.

Figure 2.6 shows that both cycle and rotational variability are affected by a change in B_{sat}. The changes to the cycle amplitude that result from a change in B_{sat} are more pronounced than those to the rotational variability. This is because rotational variability is caused by active-region passages which contain a distribution of magnetic fluxes with more pixels with values of magnetic flux above B_{sat}. The steady decrease toward solar minimum in 2008, from solar maximum in 2002, is caused by a roughly uniform global decrease in surface magnetic flux which consists of a magnetic flux distribution with magnetic flux values generally lower than in active-regions. The filling factor of pixels with a magnetic flux above B_{sat} do not change, where as those below will. The distributions overlap, so changing B_{sat} affects both rotational and cycle variability in the same way, e.g. by increasing variability by lowering B_{sat}, but with the consequence of acting differently on different

time scales. The difference plots shown in Fig. 2.6 illustrate this nicely.

In the top plot of Fig. 2.6, it can be seen that decreasing B_{sat} from 450 G to 400 G increases the cycle amplitude by 0.25 Wm^{-2}, or about 16%, while increasing B_{sat} to 500 G reduces the cycle amplitude by 0.21 Wm^{-2} (~14%). The effect on rotational amplitude (lower plot) depends on the phase of the cycle, but even near cycle maximum, the rotational amplitude increases by no more than ~0.15 Wm^{-2} in absolute terms. It makes sense that rotational variability sees a smaller increase in variability because a smaller portion of the magnetic flux responsible for rotational variability is below B_{sat} than is the case for magnetic flux responsible for cycle variability.

Using the MDI instrument, with a spatial resolution of 2″ at the wavelength of 676.8 nm, Ortiz et al. (2006) found that contrast, i.e. the intensity relative to the surrounding quiet Sun intensity, as a function of magnetic flux increases up to a maximum before decreasing at higher levels of magnetic flux. Their results showed that the rate at which the contrast increases, the value of magnetic flux at maximum contrast and the rate of decrease after maximum contrast also depends on μ. The relationship between magnetic flux and contrast has also been noted in many other publications, e.g. Topka et al. (1997); Vögler (2004); Kobel et al. (2011). SATIRE-S does not consider contrast as a function of magnetic flux; the filling factor acts as a substitute for this. Therefore, the decrease in contrast at higher magnetic fluxes is not taken into account and, in addition, will also be dependent on the wavelength[3] and limb angle.

A limb-dependent contrast relationship was reported by Ortiz et al. (2002). They showed a centre-to-limb variation of facular contrast that increases to a maximum contrast at a limb angle that is dependent on the level of magnetic flux, before decreasing again. The model atmospheres in SATIRE-S are one-dimensional and so do not take into account the three-dimensional geometry of faculae. As such, the contrast of faculae is assumed to increase rapidly and monotonically, though not linearly, towards the limb. This is demonstrated by the facular contrasts of the lower plot of Fig. 2.4 with the solid, dashed and dotted green facular contrasts at limb angles of 1.0, 0.7 and 0.05 respectively. The rapid increase in contrast towards the limb is almost certainly an unrealistic approximation, so that some spectral regions overestimate spectral irradiance as facular regions move onto and off the solar disk (see Unruh et al. (2008) and Chapter 5). Initial tests done within the framework of this thesis, and preceeded by the work of Daniel Siegel (undergraduate

[3]There have only been a limited number of studies, in very few wavelengths, most of which disagree on how exactly contrast varies with limb angle (see Unruh et al. (2000) and references therein).

project), show that either keeping the contrast constant, or declining, for $\mu \leq 0.4$ improves rotational agreement with observations, though only slightly.

Even with the approximations used in the model, it already produces surprisingly good agreement with observations. Adding in an additional free parameter, for where contrast ceases to increase at some arbitrary limb angle, would do little to improve it. Therefore, it is left as is until a more satisfactory solution can be found (see Chapter 8).

2.2.2 Penumbral and umbral parameters

The exact choice of umbral and penumbral temperatures is relatively unimportant in comparison to the effect that the free parameter, B_{sat}, has on the reconstruction. To show this, in Fig. 2.7, the regression slope between the model and SORCE/TIM TSI observations has been calculated for a range of umbral and penumbral temperatures over the time period considered for Fig. 2.6. In Fig. 2.7 the white curves represent the mean temperature of the spot as a whole, $T_{sunspot} = \sqrt[4]{(T^4_{umbra} + 4 \times T^4_{penumbra})/5}$. The ratio of umbral to penumbral area of 1:4 is used here, e.g. Brandt et al. (1990). Light blue shading represent regions close to a unity regression slope, i.e. a slope equal to 1.00. Lower values deviating away from unity are shaded in yellow and white. The thin black lines are iso-contours. For all reconstructions here, B_{sat} = 443 G, a value chosen for a best fit to SORCE/TIM using an umbral and penumbral temperature of 4500 and 5450 K, respectively, and indicated by the white square in the plot. This is why the iso-contour with value 1.00 crosses the square. These temperatures also agree with the range of observations, as in section 1.5.

Fig. 2.7 qualitatively shows that as long as $T_{sunspot}$ remains relatively constant, i.e. for the chosen values of around 5300 K, then good agreement with observations can be maintained. This is shown by the relatively good agreement between the gradients of the iso-contours and white curves of constant sunspot temperature. To provide a way of comparing with B_{sat}, a change of 0.01 in the regression slope can be simulated by a change of 7 G in B_{sat}. The regression slope is more sensitive to a change in $T_{penumbra}$ than a change in T_{umbra}, the latter of which can vary by 600 K (from 4200 to 4800 K), and the former by approximately 100 K, while still remaining within 1.00 ± 0.01 iso-contours. This is hardly an unexpected result given the assumption that the penumbral area is four times the area of the umbra. Given the relative insensitivity of the sunspot component temperatures in themselves and with B_{sat}, they can be fixed with confidence, so that varying B_{sat} to find a best fit with observational data (see section 2.4) will not be significantly affected

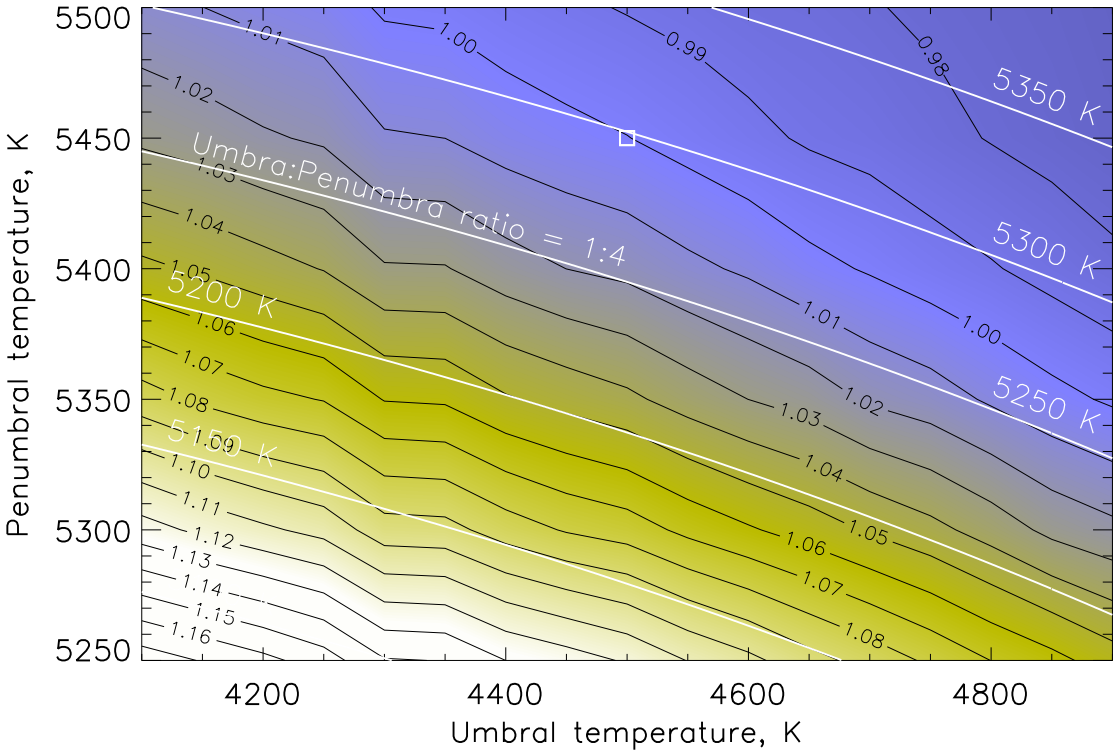


Figure 2.7: Contour plot of umbral temperature against penumbral temperature for a reconstruction using only MDI data and B_{sat} fixed at 443 G. Sunspots have an umbral to penumbral area ratio of 1:4. Lighter blue indicates value of the regression slope close to 1.00 while yellow and white are larger deviations away from 1.00. The white curves represent the averaged sunspot temperature and the thin black lines represent iso-contours. The white square represents the location of sunspot temperatures used for reconstructions in this thesis.

by fixing the umbral and penumbral temperatures.

It should also be noted that there is some degeneracy of the temperature with the area covered by each component. However, Wenzler (2005) determined the area of sunspot components in Kitt Peak data to agree with the established ratio of umbral to penumbral areas of 1:4 (e.g. Brandt et al., 1990) and with direct observations of sunspot areas. In this thesis, SoHO/MDI umbral and penumbral thresholds are set so that areas of both umbra and penumbra agree with Kitt Peak, as explained in Chapter 3.

2.3 Improvements and modifications

The additions to the SATIRE-S code discussed in the following sub-sections are not intrinsically a part of the underlying physics of the model. They are implemented to ensure consistency when spectral comparisons are made later in the thesis, and to account for artefacts that appear from the processing of the continuum intensity images and magnetograms in preparation for, and within, the model code. Since the changes are embedded within the SATIRE-S code, the adjustments are discussed within this chapter.

2.3.1 Integration method

SATIRE-S originally employed the *int_tabulated* IDL routine which uses a five-point Newton-Cotes integration formula. When spectral comparisons between model data and real observations (see Chapter 5) were initially performed, the sum of spectral sub-regions did not add up to a direct integration over the whole spectral interval. To function effectively, the Newton-Cotes method requires the interval being integrated over to be divided into equal parts, but the wavelength scale used for the spectra in both model and observations is non-linear.

In contrast, the *trapezium rule* integration method is simple in its approach, can be well understood and the summation over sub-regions *does* equal the integration of the entire spectral region. The use of polynominals to approximate the tabulated data in the Newton-Cotes method introduces errors, the source of which is not explicity clear, and thus, when comparing the outputs of both methods, a difference emerges.

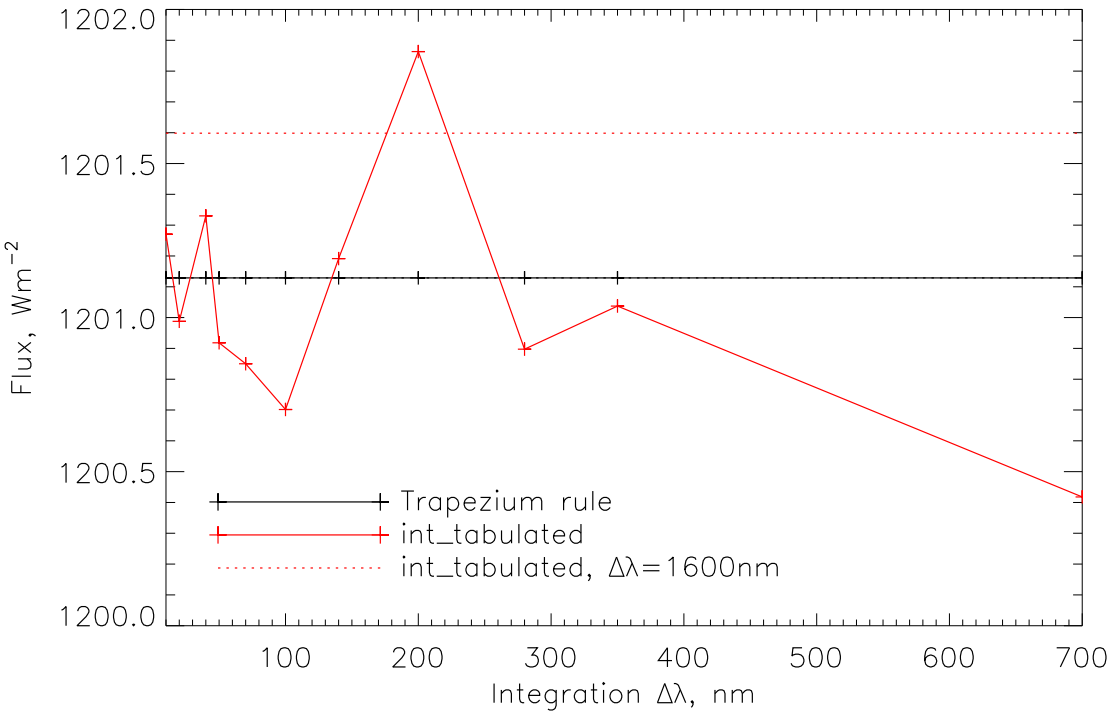


Figure 2.8: A demonstration of the different results gained by using the *trapezium rule* compared to *int_tabulated* by integrating over different spectral widths. The summation of integrals over different sized sub-interval of the full 200-1600 nm region are plotted. The red, crossed line is for *int_tabulated* and the black, crossed line for the *trapezium* method. Integrating over 200-1600 nm directly produces the red dotted line when using *int_tabulated*, whereas using the *trapezium* method gives the same result as the black, crossed line.

By integrating over sub-intervals within an arbitrarily large spectral region the difference in output from the two integration methods can be shown. For example, the integral of the spectral region 200-1600 nm in SATIRE-S (using data on 1996 May 20) equals ~1201.1 Wm^{-2} when using the *trapezium rule* method. The summation of 140 integrals over 10 nm sub-intervals, for example, should (and does) give the same result. In Fig. 2.8, the summation of equal-width sub-intervals of 11 different sizes is performed over 200 to 1600 nm using the *trapezium* method (black) of integration and using *int_tabulated* (red). The dotted red line, at 1201.6 Wm^{-2}, is the value of a direct integration over 200-1600 nm using the *int_tabulated* method. This is 0.5 Wm^{-2} higher than when using the *trapezium rule* method. From this figure the summation of sub-intervals when using the *trapezium rule* is the same whatever interval size is used to integrate over the full 200-1600 nm spectral region. This is not the case using *int_tabulated*. There does not appear to be any correlation between interval size and the difference relative to integration using the *trapezium rule*, though generally larger bins give less reliable results. The 200-1600 nm region considered here, and the chosen sub-intervals, are just examples. In this case the deviation of *int_tabulated* from the *trapezium rule* can exceed 0.7 Wm^{-2} and the range of possible integral values varies up to twice that (i.e. the difference between 200 nm intervals and 700 nm intervals). Although this is only about 0.1%, a difference of 1.7%, 23 Wm^{-2}, has been found when integrating over the full range from 9-160 000 nm.

While these differences are small in most cases, it means that additional flux is being introduced in an unclear way that can change the relative variation of the different spectral regions being considered. It is also unsatisfactory to have the sum of the parts not equal the whole. These problems disappear when using the *trapezium rule*. Therefore, the *trapezium rule* replaces *int_tabulated* in SATIRE-S and is used for all integrations performed in this thesis.

The change of integration method has an effect on the reconstruction, most notable in the long-term gradient. Because B_{sat} is fixed by varying it to gain a fit to TSI data, the change in integration method requires that the free parameter is reconsidered. In the top panel of Fig. 2.9, model reconstructions between February 2003 and October 2009 are plotted, normalised to the SORCE/TIM TSI data (not shown) during the 2008 minimum. Both integration methods have B_{sat} = 443 G, chosen so that the *trapezium* method (blue) will agree with SORCE/TIM with a unity regression (see Chapter 3). From this plot, it can be seen that a reconstruction using *int_tabulated* (red) produces a lower cycle gradient for a given B_{sat} than using the *trapezium rule*.

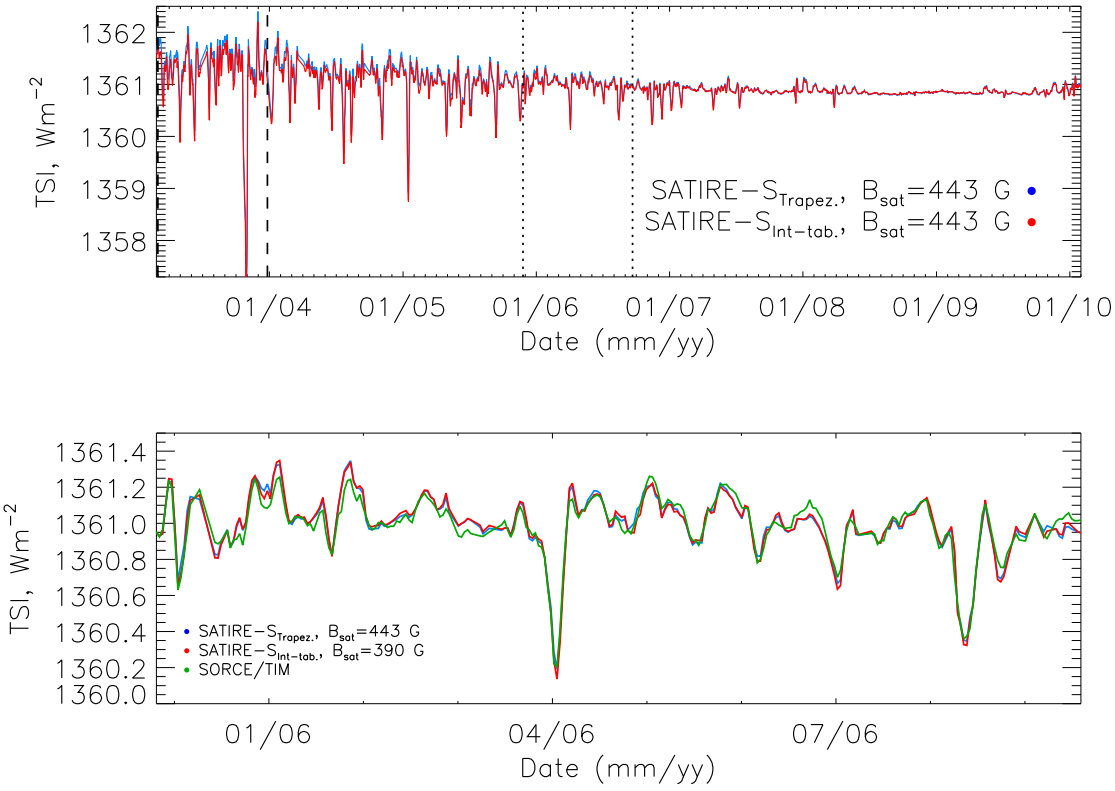


Figure 2.9: The effect of changing integration method on TSI reconstructions. (**top**) The long term trend of each method using the same model settings and normalised to SORCE/TIM at December 2008. (**bottom**) To show the effect on the short-term variability, the free parameter is set to gain a unity regression with SORCE/TIM and normalised to the shorter period shown. The shorter period is between the dotted lines of the top plot.

By decreasing B_{sat}, a unity regression with a reconstruction using *int_tabulated* is found at 390 G. The correlation coefficient between the two is, as expected, very good (R_c = 0.997), but there are minor changes to the short-term variation. The period between the vertical dotted lines in the upper plot is replotted in the lower plot of Fig. 2.9, along with SORCE/TIM (green). In this plot, both reconstructions are normalised to SORCE/TIM over the shorter period plotted. There are clearly differences, but these are all very small[4].

Therefore, ultimately, once the free parameter is set to give the best agreement with TSI observations the difference between the two approaches is minor. The benefit of changing integration method is mainly found when doing spectral comparisons. TSI integration within SATIRE-S should be consistent with the integral of SSI and for this reason the change is implemented. The result is that the chosen value of the free parameter is different to, and cannot be compared with, previously published reconstructions using SATIRE-S. Due to the nature of fitting to TSI, the actual irradiance trends are very similar.

2.3.2 Removal of umbral limb-artefacts

To prepare the continuum intensity images for use in SATIRE-S, all pixels in the continuum intensity images must be aligned with corresponding pixels in the magnetograms so that magnetic flux belonging to sunspot pixels is not assigned as facular pixels. Even just a few minutes difference between the continuum intensity image and magnetogram can result in an offset that introduces errors. This is not a problem when using Kitt Peak 512-channel Diode Array Magnetograph (KP/512) and Spectromagnetograph (KP/SPM) continuum intensity images and magnetograms (see Chapter 3) which are observed simultaneously. For SoHO/MDI, continuum intensity images and magnetograms are rarely available within a few minutes of each other. Therefore, MDI continuum intensity images have to be rotated to align pixels in both images. A limit of 12 hours time difference between MDI continuum intensity images and magnetograms is set to avoid the introduction of large differences resulting from surface evolution or from rotation of active-regions off the visible disk.

The *solarsoft* suite of routines is used to rotate MDI continuum images. For periods more than a couple of hours apart, this can lead to the introduction of limb artefacts, an example of which is shown in Fig. 2.10. In this image the dashed circle marks the boundary of the

[4]The correlation coefficient between the model, using the *trapezium rule*, and SORCE/TIM is marginally lower than using the *int_tabulated* approach, 0.980 and 0.982.

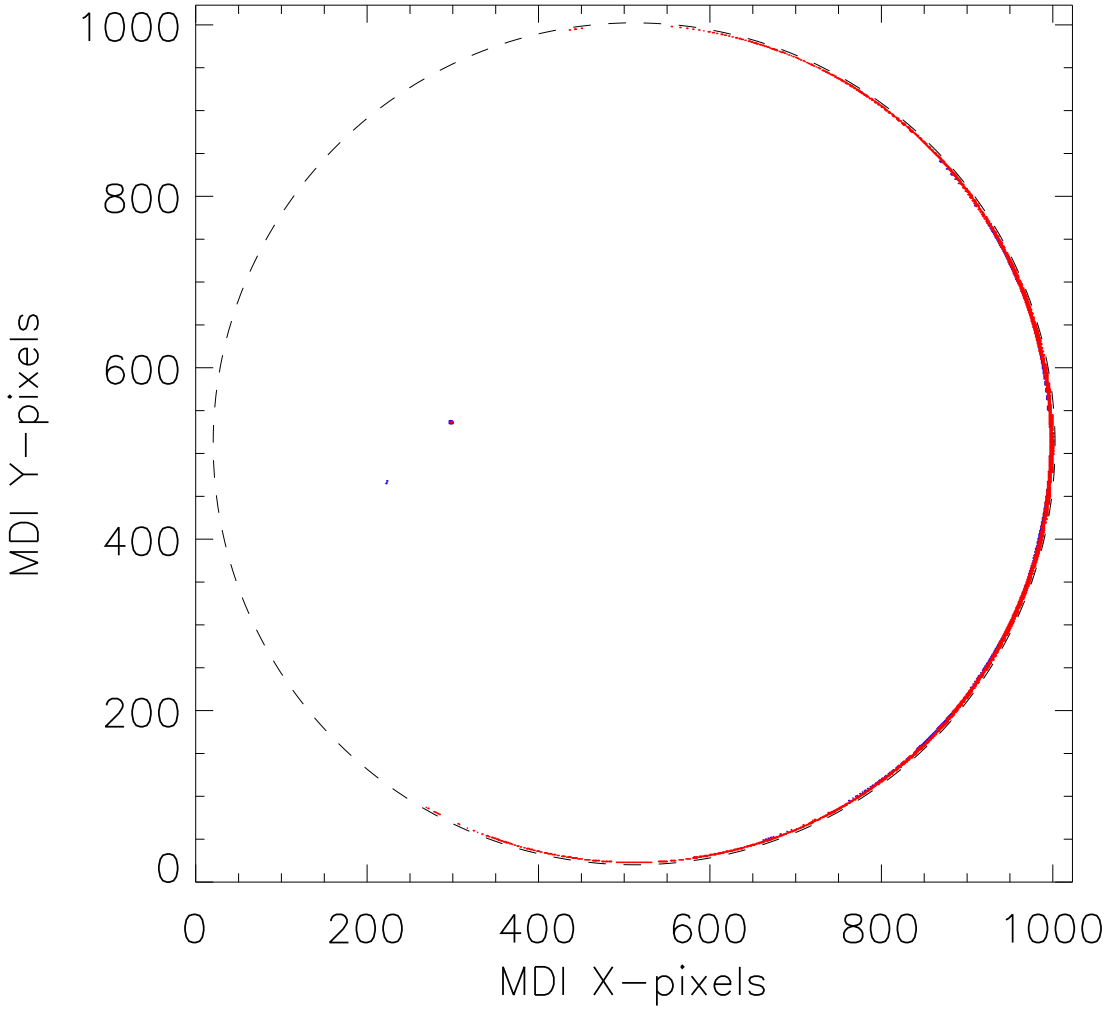


Figure 2.10: Sunspot mask (2006 March 16) extracted from SATIRE-S. The continuum intensity image used to produce this mask was rotated by 4.8 hours to bring pixels into line with those in the magnetograms. The act of rotating the images has produced an artefact at the western limb that must be removed to prevent being falsely attributed as sunspot pixels (penumbral pixels are blue, umbral are red). Note the two small sunspots just east of disk centre.

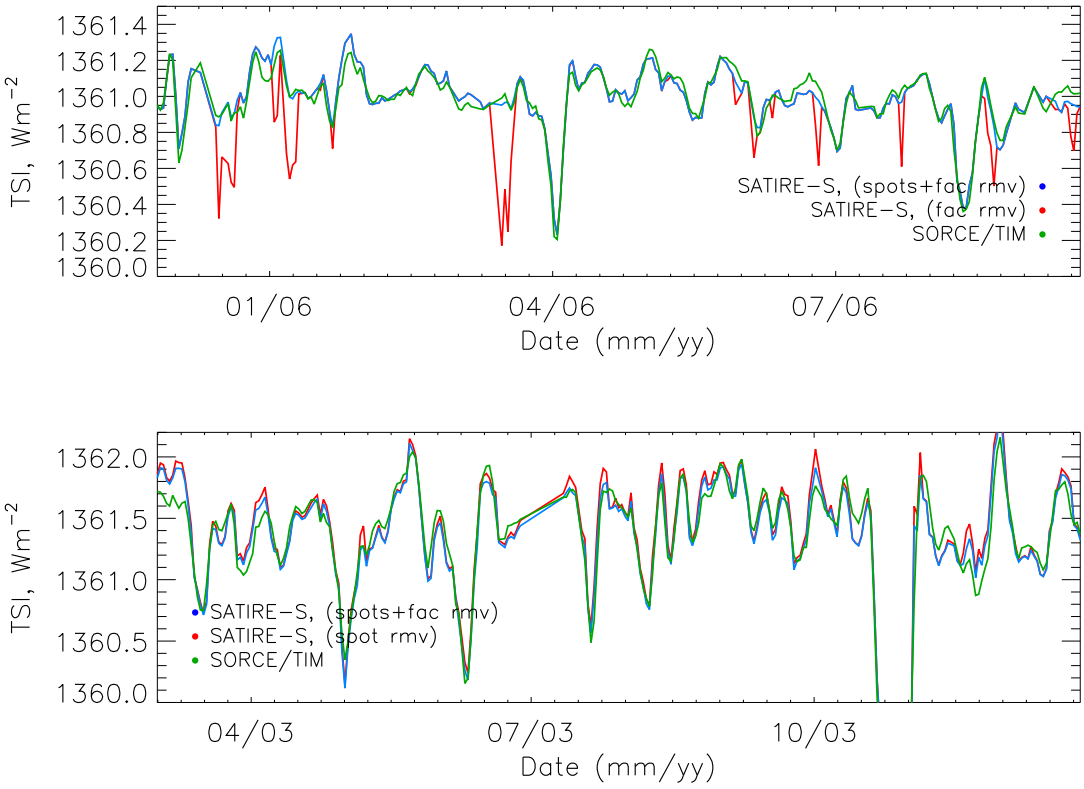


Figure 2.11: The effect of removing problem pixels from MDI data in SATIRE-S: (**top**) reconstructions having removed only erroneous faculae (red); (**bottom**) reconstruction after removing false sunspot pixels only (red). In addition, both plots contain SORCE/TIM (green) as a reference and also the reconstruction after removing both erroneous faculae and false sunspot pixels (blue). The periods considered in these plots are delimited by vertical dotted (top) and dashed (bottom) lines in the upper plot of Fig. 2.9.

solar disk with umbral pixels shown in red and penumbral in blue. There are two small sunspots in the eastern hemisphere, but in the west almost the entire limb from pole to pole has an apparent thin lining of spurious umbral pixels along with a few penumbral pixels. Off-limb, the pre-rotated image has a relative contrast set at zero. Therefore, the likely cause of these limb artefacts is from a coding error that rotates zero-contrast off-limb pixels onto the solar disk. These are then identified in SATIRE-S as umbral pixels. This type of artefact occurs in up to 20% of rotated images.

Reconstructing TSI without removing these erroneous umbral pixels results in spurious decreases that mimic sunspot crossings. This is illustrated by the red line of the top plot in Fig. 2.11. This is the same period as in the lower plot of Fig. 2.9, along with the sunspot and facular (see section 2.3.3) corrected SATIRE-S (blue) and SORCE/TIM (green) ob-

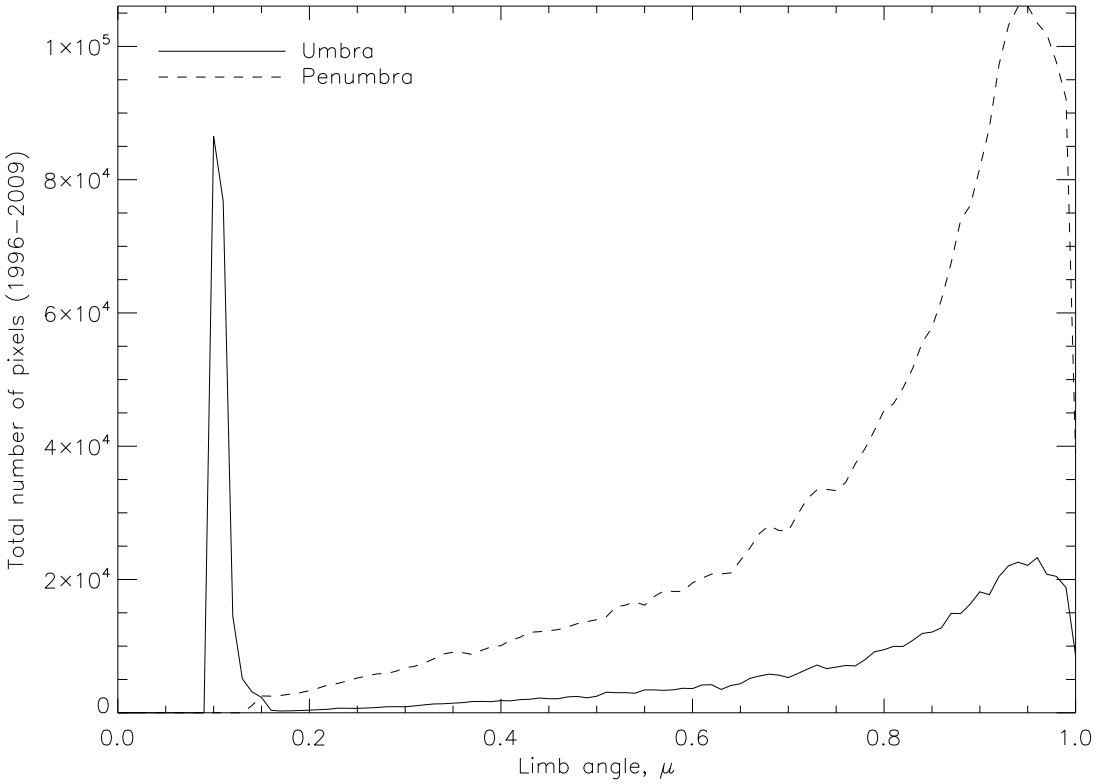


Figure 2.12: The total number of umbral (solid line) and penumbral (dashed) pixels as a function of limb angle. Artefacts at the limb, introduced by rotating the Sun in continuum images, appear as a peak at the limb below $\mu = 0.2$.

servations. From this short stretch it is clear that the limb artefacts occur quite frequently and further analysis has shown that the number of false umbral pixels is related to the time difference between continuum intensity image and magnetogram observations. Obviously, these false umbral pixels need to be removed.

In Fig. 2.12, the total number of umbral and penumbral pixels is shown as a function of limb angle, taken from 4208 images between May 1996 and October 2009. Because pixels are grouped as a function of limb angle in bins of $\Delta\mu = 0.01$, the number of pixels peak around disk centre, with the number of pixels declining to near zero at $\mu = 0.2$ for both penumbral and umbral pixels. Penumbral pixels then decline to zero before the cutoff at $\mu = 0.1$. On the other hand, umbral pixels rapidly increase in frequency with decreasing limb angle from $\mu = 0.16$. This suggests that the number of penumbral pixels introduced to the disk through the rotation of continuum images is marginal.

Fig. 2.12 shows that penumbral pixels at the limb do not need to be removed, but umbral

ones do. The choice is made to remove only umbral pixels at and below the minimum point in the distribution. In Fig. 2.12 this lies at $\mu = 0.18$, but the original calculation, which included more dates (which have been excluded from the reconstructions due to further problems not considered in this thesis) found that this lay at $\mu = 0.19$. This higher value is used because there are some spurious umbral pixels at this limb angle and the dates not used may be included at a later stage. Choosing this slightly higher value does not cause any significant difference. The spurious umbral pixels are replaced with quiet Sun pixels. The resulting reconstruction is a significant improvement as can clearly be seen in the blue line in the upper plot of Fig. 2.11. Prior to correcting for faculae (see below), for the entire overlap period with SORCE/TIM, the correlation coefficient improves from 0.947 to 0.974 and as such the model can be said to account for an additional 5% of variability in TSI.

2.3.3 Removal of high magnetic flux pixels

The investigation into suprious umbral pixels naturally led to a similar investigation of facular pixels. The result of this investigation was to set all pixels identified with a magnetic flux above 800 G to quiet Sun pixels.

False facular pixels first became conspicuous when inspecting detailed histograms of magnetic flux. In these histograms, the number of faculae at each limb angle are grouped in 5 G bins from 0 to 1200 G. Although erroneous pixels unassociated with sunspots were not numerous, because all magnetic flux above 1197.5 G is binned together, the excess became noticable in these histograms. Dates on which the largest concentrations appeared were associated with overenhancements in the TSI reconstruction relative to TSI observations. Some of these overestimates can be seen when comparing the model, without these faculae removed, with the SORCE/TIM data, red and green respectively, in the lower plot of Fig. 2.11. They usually occur just prior to and after a sunspot transit, i.e. when the active-regions are near either limb. A good example of this can be seen either side of the small sunspot transit in August 2003 in the lower plot of Fig. 2.11. Here, the red lines show a higher increase, compared to SORCE/TIM, just prior to and just after the decrease in irradiance caused by a spot.

Further investigation, comparing the contrast of pixels in continuum intensity images with those in magnetograms, led to the result that higher magnetic fluxes, especially above 800 G are usually darker than the mean quiet Sun contrast even close to the limb. A conserva-

tive and arbritary decision was made to set all magnetic pixels above 800 G to quiet Sun pixels, calling this cutoff, B_{cut}. It was initially thought that an over-enhancement of magnetic flux due to the $1/\mu$ line-of-sight correction was responsible for allowing relatively low contrast pixels to be detected as bright faculae. However, this was not the reason for these false facular pixels appearing.

Kok Leng Yeo, at the Max Planck Institute for Solar System Research in Germany, using high-resolution SDO/HMI continuum images and magnetograms, has recently provided an explanation (private communication). He found that the magnetic flux in canopies above pores and sunspots becomes preferentially aligned with the LOS near the limb and is detected as a region of high magnetic flux and falsely attributed to faculae. He also found that falsely attributed facular pixels occured at magnetic flux levels in HMI as low as 400 G. In MDI, the magnetic fluxes are reported ~30% higher. By reassigning all facular pixels to quiet sun pixels above some level, B_{cut}, the best correlation between SATIRE-S and SORCE/TIM is indeed found when $B_{cut,MDI}$ = 550 G. However, at ~550 G real magnetic features closer to disk centre are also removed. Therefore, the cutoff level of 800 G was retained as a compromise and is used throughout this thesis. By removing these higher magnetic flux pixels the improvement, although smaller than accounting for the false sunspot pixels, is that TSI variability is accounted for by a further 1% with the correlation coefficient improving from 0.974 to 0.980. Returning to the example discussed earlier in this subsection, in the lower plot of Fig. 2.11 the over-enhancements just prior to and just after the sunspot transit of August 2003 is in much better agreement with SORCE/TIM after this correction. Other examples like this can also be seen in this figure.

2.4 Optimising SATIRE-S using the free parameter

As has been made clear in this chapter, the long and short-term TSI variability of a SATIRE-S reconstruction is highly dependent on the value of B_{sat}, the only parameter allowed to vary freely. The only way to set the optimal value at present is to use an observational record of TSI. The chosen fitting technique turns out to be important. This section focuses on why the approach of using a unity regression fit to observational data, i.e. attaining a regression slope of 1.00 with a TSI record, is used instead of a least squares approach. In addition, the *fitexy* regression routine (Press et al., 1992) is used instead of the standard *regress* IDL routine.

The TSI output from SATIRE-S is the relative variation in ppm, with respect to a completely inactive, quiet Sun. The data are then normalised to the observed TSI and compared. Using the chosen optimisation technique, whether that is to find a unity regression or minimisation of the residuals of least squares, B_{sat} is varied until the best-fit is reached. The variation of just this one free parameter means that only the facular contribution changes.

The differences between each approach are easier to explain by example. Therefore, fitting will be discussed in the context of a fit to SORCE/TIM TSI data using a reconstruction of SATIRE-S based on MDI data only (see Chapter 3 for more information on the MDI and SORCE/TIM datasets). This section focuses entirely on the difference between simple fitting techniques. Three approaches are considered here: using the method of *least squares* by minimising the sum of the residuals; and gaining a unity regression between model and observation using either the IDL standard regression routine, *regress*, or the *fitexy* regression routine that allows x and y axis errors to be prescribed.

A summary of the results from each method is shown in Tables 2.1 and 2.2. In Table 2.1, best-fits using the original time series are shown for the given optimal value of B_{sat}, alongside the cycle amplitude, and the correlation coefficient, r_c, and its square[5]. In Table 2.2 the optimal value of B_{sat} for smoothed and detrended datasets are given, along with the cycle amplitude that would result if that value of B_{sat} were used to reconstruct without smoothing or detrending (detrended results are given to the nearest 20 G). The cycle amplitude is calculated by taking the difference of the six-month means around February 2002 and December 2008. In Table 2.2, using only original data, the suggested optimal fit ranges over nearly 20 G in B_{sat}, or about 7% in cycle amplitude. As a point of interest, the correlation coeffcient reaches a maximum at B_{sat} = 443 G (although the differences are marginal), i.e. at the optimal value of B_{sat} using *fitexy*, though this is not considered as a fitting criterion. The different values of B_{sat} pose a problem because in Chapter 3, the fit between SORCE/TIM and SATIRE-S using MDI over 2003-2009 will determine the properties of the entire reconstruction of TSI and SSI spanning three cycles between 1974 and 2009.

An understanding of why the fitting methods give different values of B_{sat} for the best-fit results, is needed. The *regress* routine assumes that the errors exist only in the y-axis variable, whereas *fitexy* allows the user to define the errors in both axes. Tests show that results from *fitexy* tend to those of *regress* if the errors in one axis are significantly larger

[5]The value of r^2 estimates how well one dataset captures variability in the other.

Table 2.1: Best-fit model parameter for different fitting methods when comparing SATIRE-S with SORCE/TIM TSI data between 2003 and 2009. Column 1 is the fitting technique, column 2 is the value of B_{sat} when the best fit is found, column 3 is the cycle amplitude between 2002 and 2008 and column 4 are the correlation coefficients.

Daily dataset comparison

Method	B_{sat} [B]	Cycle amp. [Wm^{-2}]	$r[r^2]$
(1)	(2)	(3)	(4)
regress (TIM, x-axis)	429	1.45	0.979 [0.959]
fitexy	443	1.38	0.979 [0.959]
Least squares	448	1.36	0.979 [0.959]

Table 2.2: Best-fit model parameter for different fitting methods when comparing SATIRE-S with SORCE/TIM TSI data between 2003 and 2009. Column 1 is the method to attain the best fit with observation, column 2 is the best-fit value of B_{sat} for the smoothed dataset, column 3 is the amplitude of the original data using the smoothed B_{sat} and columns 4 and 5 are the same as 2 and 3 but using detrended data and to the nearest 20 G of B_{sat}.

Smoothed and Detrended dataset comparison

Method	$B_{sat,sm}$ [B]	A_{sm} [Wm^{-2}]	$B_{sat,dt}$ [B]	A_{dt} [Wm^{-2}]
(1)	(2)	(3)	(4)	(5)
regress (TIM, x-axis)	431	1.44	620	0.78
fitexy	432	1.43	600	0.83
Least squares	432	1.43	580	0.89

than the other (in this case more than approximately five times that of the other). While the systematic error in the SORCE/TIM data is published, it is very difficult to determine an error from the model because it is not clear how the errors on the continuum intensity images and magnetograms propagate through the model. In addition, there are no uncertainities on the model atmospheres which may also have an important effect; how closely the use of B_{sat} and a linear interpolation between magnetic flux and the facular filling factor reflect reality cannot be determined at present. There may be systematic problems related to the model atmospheres, but the lack of noise in the time series (e.g. Figs 2.5, 2.9 and 2.11) suggests that there is little noise introduced from the continuum intensity images and magnetograms. Therefore, in this section and for the thesis as a whole, the assumption is made that the error is the same in observations and model.

The detrended and smoothed results in table 2.2 also reveal further insight. The results

show that the optimal value of B_{sat} in table 2.1 is a trade off between cyclical and rotational variability. All three methods give quite similar results. There is very good agreement between all three approaches using the smoothed time series. with all results around 432 G; *regress* is 1 G lower. The B_{sat} values from the smoothed data are close to the values of the original results. This means that the long-term trends are well captured when fitting the original datasets through these methods. For the detrended results, all to the nearest 20 G, B_{sat} is between 150 and 190 G higher. Least squares gives the lowest value, at 580 G, and *regress* the highest, at 620 G. This large increase occurs to accommodate an overestimation of rotational variability.

The agreement between all three approaches using the 'original' data is reasonably good (see Tab. 2.1). It should be noted that by using the residuals between model and observation as an estimate of the uncertainty as an input in *fitexy*, the formal error on the regression using *fitexy* leads to a 3σ error of $\sim$16 G. This uncertainty covers the optimal fits when using original data in both of the other methods (i.e. in table 2.1).

In summary, there is little difference between all three approaches and all methods fall within the 3σ error of *fitexy*. Also, *fitexy* returns intermediate results compared to *regress* and least squares when timeseries are detrended. Therefore, *fitexy* is a reasonable choice to do further fitting and comparisons and this approach is used as a standard throughout this thesis.

2.5 Conclusions

In this chapter, the functionality of the SATIRE-S model has been outlined in detail. In addition, the assumptions underlying the free parameter have been discussed. While the free parameter does not explain how magnetic flux and the contrast of faculae relate to each other, it serves as part of a useful approximation that recreates solar irradiance variations well. The choice of sunspot characteristics was also discussed and justified.

Small improvements to the model were identified and addressed, including the removal of umbral and facular artefacts arising from the preprocessing of continuum intensity images and line-of-sight misidenification of sunspot canopy magnetic flux, respectively. The integration method previously used was compared to that of the trapezium rule method. The trapezium rule method was found to be more consistent than int_tabulated and now

replaces it in the model routines.

Finally, the method by which the free parameter is fixed, namely through fitting to TSI observational data, was discussed and the decision was made to use the *fitexy* routine. Finding a unity regression with *fitexy* results in a value of the free parameter that sits in-between those of the other methods considered.

This chapter, therefore, sets the scene for the comparisons of TSI and SSI observations with models that will be performed in the remainder of this thesis.

Chapter 3

Reconstruction of total solar irradiance, 1974-2009

Overview

In this chapter[1], the SATIRE-S model is used to construct TSI over three solar cycles, between 1974 and 2009. This chapter extends the work of Wenzler et al. (2006) from 2003 to 2009 and, for the first time, makes the reconstruction independent of all three TSI composites. This allows a more robust investigation of the different secular trends exhibited by each composite, especially in light of the recent unusual solar minimum. The results indicate that if the model reconstructs solar irradiance accurately, then the PMOD composite is likely to be the best composite of TSI on cyclical and secular timescales.

Full disk continuum intensity images and magnetograms from three instruments are used to create the reconstruction. These data are described first, followed by a detailed consideration of their long-term behaviour, which is then accounted for. The approach taken to combine these datasets is described first and then the reconstruction is compared with three TSI composites to investigate how well the model recreates TSI and to see which composite best agrees with the model. The model dataset constructed in this chapter is independent of TSI composites after 1989. As the model produces SSI by design (see section 2.1), the work done in this chapter forms the basis of the spectral dataset used in chapters 5, 6 and 7.

[1]This chapter has been accepted for publication in the Astronomy & Astrophysics journal. The following people also contributed to the work and were named as co-authors: Yvonne C. Unruh, Natalie A. Krivova, Sami Solanki, Thomas Wenzler, Daniel J. Mortlock, Andrew H. Jaffe.

3.1 Data

3.1.1 Data used in SATIRE-S

In this chapter we reconstruct solar irradiance using full-disk continuum intensity images and magnetograms employed in SATIRE-S, taken from two locations: the National Solar Observatory Kitt Peak Vacuum Tower (KP) and the Solar and Heliospheric Observatory (SoHO). KP data were taken using the 512-channel diode array (KP/512) that began operation in 1974. KP/512 was replaced by the spectromagnetograph (KP/SPM) which operated between 1993 and 2003. The Michelson Doppler Imager (MDI) operated onboard the SoHO satellite between 1996 and 2011.

3.1.1.1 SoHO/MDI

MDI (Scherrer et al., 1995) records 1024 × 1024 pixel full-disk filtergrams and magnetograms in the Ni I absorption line at 676.78 nm with 4″ resolution. As discussed in section 2.1, simultaneous continuum intensity image and magnetogram data are usually not available to download, so continuum intensity images are rotated to align features in both images. Continuum intensity images have a measurement uncertainty of 0.3% and only a single image is needed to identify the sunspot components. For the magnetograms, we attain a reduced noise level by using 5-minute averaged magnetograms. The calibrated 1-minute magnetograms have a reported uncertainty of ∼30 G; averaging over five-minutes reduces the uncertainty to ∼15 G (these are recalibrated values following Tran et al. (2005); see below).

In this study, images are used from almost the entire working period of the MDI instrument from 1996 May 20 to 2009 October 31, using 4140 image pairs[2].

Previous studies using SATIRE-S have employed level 1.5 magnetograms (Fligge et al., 2000; Krivova et al., 2003; Wenzler et al., 2004). Recalibrations of the data have now produced the level 1.8.2 magnetogram data, as used in, e.g. Krivova et al. (2011a). The most significant change came after Tran et al. (2005) compared the sensitivity of MDI magnetograms with those of the Mount Wilson Observatory. This work led to a position-dependent correction[3] that has increased the estimated observed magnetic flux by a factor

[2] MDI data after 2009 October 31 have artefacts that need to be considered before the data can be used. This is outside the scope of this work.

[3] All modifications to produce level 1.8.2 magnetograms are described in more detail at

of ∼1.6, necessitating a change in the SATIRE-S parameter B_{sat}.

SoHO was subject to a number of operational problems that MDI was not exempt from. It appears that these, along with some operational changes to the instrument, have resulted in changes to the magnetograms over time. The causes and effects of these changes are briefly discussed in section 3.2.5, with an in-depth analysis forming the basis of chapter 4.

3.1.1.2 KP/SPM

KP/SPM (Jones et al., 1992) operated between 1992 and 2003 taking 1788 × 1788 pixel images in the Fe I line at 868.8 nm with a 1.14″ pixel size. Continuum intensity images and magnetograms were extracted from slit spectra that scanned across the solar disk in one hour. Jones et al. (1992) and Wenzler et al. (2004) both determined a background noise level[4] of ∼5 G = $\sigma_{mag,SPM}$.

Wenzler et al. (2005) identified 2055 sets of images between 1992 November 21 and 2003 September 21 that were relatively free of atmospheric effects. These images are used in this study.

3.1.1.3 KP/512

Observations by the KP/512 magnetograph (Livingston et al., 1976) were taken between 1974 and 1993. Full-disk images are 2048 × 2048 pixels with a 1″ pixel size, but with a generally siginificantly lower spatial resolution due to atmospheric seeing. Continuum intensity images and magnetograms were built simultaneously from scans lasting approximately 40 minutes. Continuum intensity images are limited to a 4-bit intensity scale providing only 15 discrete levels that make the separation of penumbra from umbra, or sunspots from limb-darkening, difficult. A brief description of the procedure employed by Wenzler et al. (2006) to resolve this is given in section 3.2.1. Jones et al. (1992) and Wenzler et al. (2006) both found that the magnetograms had a background noise level of ∼7-9 G; Wenzler et al. (2006) found a mean of 8.1 G = $\sigma_{mag,512}$, which we adopt here.

http://soi.stanford.edu/magnetic/Lev1.8/

[4]Note that terms with the subscripts MDI, SPM and 512 mean they are set for reconstructions using only that specific dataset.

As the observatory is ground-based, seeing and weather conditions in general have a significant effect on the quality of the images. Work done by Wenzler et al. (2006) identified 1734 dates with usable data over the period 1974 February 1 to 1992 April 4. In 1992 there was a change of instrument and KP/512 observations are only available as an overlap with the superceding instrument, KP/SPM, on 45 days between 1992 November 28 and 1993 April 10.

Wenzler et al. (2006) found that a conversion factor, f_{512}, was necessary in order to multiply magnetic flux in KP/512 magnetograms to bring them into agreement with KP/SPM. The official value of f_{512} quoted in the data release was 1.46, but Wenzler et al. (2006) found that to get best agreement with PMOD a value of 1.60 was needed. The conversion factor allows other parameters, such as the magnetic flux level B_{thr} and B_{sat}, to be the same in both KP/SPM and KP/512 reconstructions. We show, in section 3.2.2, that f_{512} is dependent on B_{sat} and has no freedom to be chosen or vary independently.

3.1.2 Observational TSI composite

TSI observations are needed to fix the free parameter, B_{sat} (section 2.2.1), in SATIRE-S. They also provide a basis with which to test the model reconstruction on various timescales. We describe in detail here the TSI composites used for the model comparison later in the chapter.

The current version of the PMOD composite used here, d41_62_1003 (Fröhlich, 2000), covers the period 1978 to present. It is composed of data from four radiometers: NIMBUS7/HF, SMM/ACRIM I, UARS/ACRIM II and the PMO6V radiometer on the SoHO/VIRGO instrument (see references in Fröhlich and Lean, 1998). The ACRIM composite (Willson and Mordvinov, 2003) also uses NIMBUS7/HF and SMM/ACRIM I, but where PMOD uses SoHO/VIRGO from 1996, the ACRIM composite continues to use UARS/ACRIM II until 2000 before switching to ACRIMSAT/ACRIM III. The latest version of the IRMB composite (Dewitte et al., 2004) begins in 1984, employing NIMBUS7/HF, SMM/ACRIM I, UARS/ACRIM II and then the DIARAD radiometer, on the SoHO/VIRGO instrument, from 1996.

Due to the more accurate absolute irradiance level of SORCE/TIM (see section 1.3), all composites and SATIRE-S are normalised to the absolute level of TIM at solar minimum in 2008.

3.1.3 SORCE/TIM

The Total Irradiance Monitor (TIM), a radiometer on the SORCE satellite, has been observing TSI since February 2003 with an absolute accuracy of 350 ppm, an instrumental noise of ~2 ppm and a stability of better than 10 ppm/yr (Kopp and Lawrence, 2005; Kopp et al., 2005; Kopp and Lean, 2011). Apart from the lower absolute value of TSI, TIM also displays a shallower gradient than PMOD during the declining phase of cycle 23. In this thesis we use version 11, level 3, six-hourly data[5] averaged onto a daily cadence.

By the time solar minimum had been reached, in December 2008, TIM had been operating for about 5.5 years. This equates to a long-term 1σ uncertainty in the change in TSI of ~55 ppm or ~0.07 Wm^{-2}. This is about the same as the difference in TSI variations seen between VIRGO and TIM over the same period and means that the gradients in TSI agree with each other within this uncertainty, not even taking into account the error on VIRGO (private communication, Greg Kopp).

3.2 Homogenising KP/512, KP/SPM and MDI data

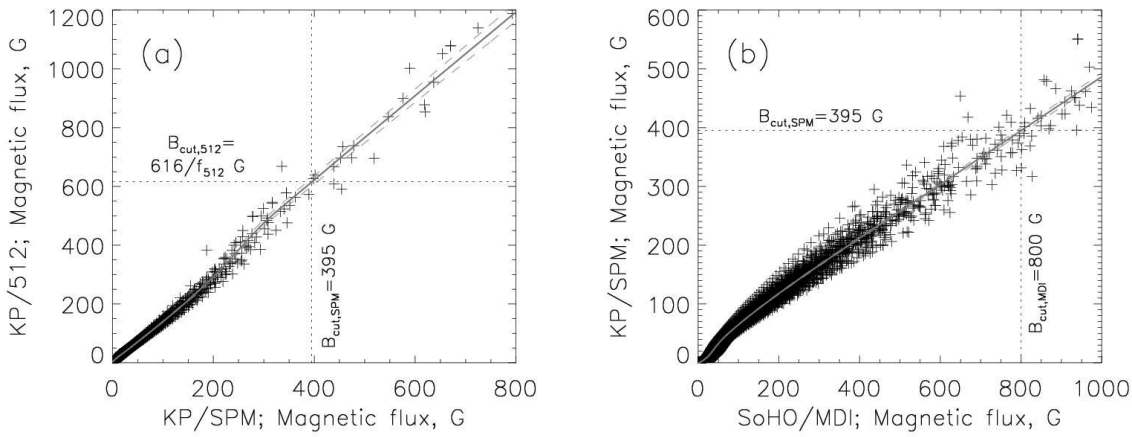


Figure 3.1: Histogram equating plots. (**a**) 18 contemporaneous KP/SPM and KP/512 magnetograms and (**b**) 84 contemporaneous SoHO/MDI and KP/SPM magnetograms. The grey line is the interpolated mean of the black crosses with the dashed grey lines the standard error. Dotted lines highlight the relationship of B_{cut} between the datasets.

[5]SORCE/TIM data downloaded 2011 July 11

In this chapter, we combine KP/512 and KP/SPM with SoHO/MDI data. The studies of Krivova et al. (2003) and Wenzler et al. (2004, 2005, 2006) also used these data, but as a result of the recalibration of MDI data (see section 3.1.1.1) the settings in SATIRE-S used by these authors are no longer appropriate and need to be reconsidered.

In this section, we discuss how the parameters in SATIRE-S are determined so that consistency is achieved across all three datasets.

3.2.1 Sunspot areas and temperatures

The study of Wenzler et al. (2006) combined KP/SPM and KP/512 data. They set sunspot thresholds so that agreement was found, in both datasets, with an independent ground-based composite record of sunspot area (Balmaceda et al., 2005). The KP/SPM penumbral threshold was set in Wenzler et al. (2004) at 0.92 so as to obtain good agreement with total sunspot area. Wenzler (2005) confirmed that the ratio of umbral area to total sunspot area is 0.2 and used this to determine a KP/SPM umbral threshold of 0.70.

In this study, we use the KP/512 sunspot data computed by Wenzler et al. (2006) for the KP/512 part of the reconstruction. The way these were obtained by Wenzler et al. (2006) is as follows. Due to the 4-bit digitization of the KP/512 continuum intensity images, only the combined total umbral and penumbral sunspot areas could be identified. Therefore, Wenzler et al. (2006) used the established ratio of umbra to sunspot area of 0.2 to estimate umbral areas. Pores could not be identified, so that a simple multiplication of the total sunspot area brought the KP/512 sunspot areas into agreement with an independent sunspot area composite.

Following this, we vary MDI contrast thresholds to bring them into agreement with the spot ratio of 0.2 and KP/SPM. Using 847 continuum images from KP/SPM and MDI between 1999 and 2003, we compare areas and find that the penumbral and umbral thresholds should be 0.89 and 0.64 respectively. The difference between KP and MDI thresholds is partly due to the different wavelengths at which the instruments operated, as well as differences in stray light.

As mentioned in section 2.2.2 and reiterated for completeness, we use temperatures of 5450 K for penumbra and 4500 K for umbra here, as in Wenzler et al. (2006).

3.2.2 The KP/512 magnetogram conversion factor, f_{512}

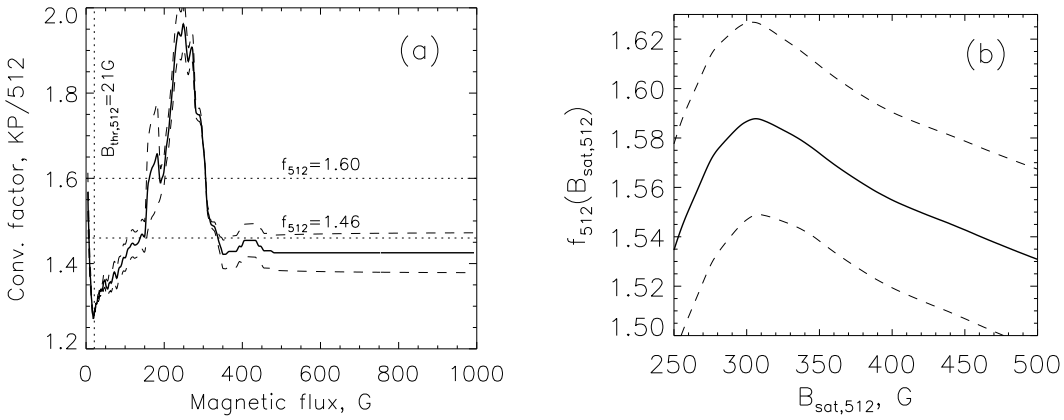


Figure 3.2: (**a**) The factor by which magnetic flux in KP/512 magnetograms should be multiplied in order to obtain the same magnetic flux as in KP/SPM magnetograms, as a function of magnetic flux (see text). (**b**) The best global conversion factor, f_{512}, by which complete KP/512 magnetograms should be multiplied as a function of $B_{sat,512}$, the SATIRE-S free parameter. Dashed lines in both panels represent the standard error as shown in Fig. 3.1a.

The introduction of f_{512} (see section 3.1.1.3) by Wenzler et al. (2006) to easily combine KP/512 and KP/SPM created an additional free parameter in the reconstruction. We are able to remove the freedom in this parameter by considering how f_{512} relates to B_{sat}. This is done as follows.

Magnetic flux is registered differently in KP/512 and KP/SPM magnetograms, so we use the method of histogram equating (Jones and Ceja, 2001; Wenzler et al., 2004) to compare the signal observed in the magnetograms. In Fig. 3.1a, we compare unsigned magnetic flux in 18 KP/SPM and KP/512 magnetograms taken on the same days in 1992/3. Crosses are the comparative magnetic flux calculated by sorting pixel values into ascending order and taking the mean in each of 1000 equally divided bins. By interpolating the KP/512 values onto a linear KP/SPM scale and averaging, the solid grey line results. This gives the corresponding mean level of magnetic flux in KP/SPM as observed by KP/512; the grey dashed lines show plus and minus the standard error on the mean. Then, the mean level of magnetic flux over 10 G intervals centred on each level of magnetic flux along the grey line is calculated. The ratio of these mean magnetic flux values in KP/512 to those of KP/SPM result in a conversion factor at every level of magnetic flux, as shown in Fig. 3.2a. In this plot, the solid line is the factor required to convert magnetic flux in KP/512 to bring it in to agreement with KP/SPM, as a function of magnetic flux. The

dashed lines show the standard errors, also from Fig. 3.1a.

The official conversion factor value of $f_{512} = 1.46$ and the factor value of $f_{512} = 1.60$ found by Wenzler et al. (2006) are indicated by horizontal dashed lines in Fig. 3.2a. It now becomes clear where the official conversion factor and that of Wenzler et al. (2006) come from: at high flux levels above 350 G the conversion factor is close to the official factor; below this it varies between 1.25 and 1.90, a range that very nearly averages to the value of Wenzler et al. (2006).

In SATIRE-S, magnetic flux below B_{thr} is ignored and all magnetic flux above B_{sat} is considered to have the same effect on irradiance, so only fluxes between these two values are important in determining f_{512}. An approximate value for f_{512} can now be calculated by simply taking the mean of the conversion factor between $B_{thr,512}$ and $B_{sat,512}$. $B_{thr,512}$ is fixed, therefore f_{512} depends only on the chosen value of $B_{sat,512}$. The result is that f_{512} is completely determined by $B_{sat,512}$, has ceased to be an independent free parameter, and allows us to combine KP/SPM and KP/512 images while maintaining the same value of the single free parameter, B_{sat}.

Although $B_{thr,512} = 21$ G was determined using $f_{512} = 1.46$, it changes very little with f_{512}, so we leave it at this value. In Fig. 3.2b, the resulting conversion factor as a function of $B_{sat,512}$ is shown for the range of $B_{sat,512}$ values of 250 to 500 G (solid line) with the propagated standard errors (dashed lines) from the left plot at approximately $\pm$ 0.04. In section 3.3.1 we find that the required value of $B_{sat,512}$ is 330 G. The resulting value of f_{512}, 1.580, is close to the best value determined by Wenzler et al. (2006) when the model was fitted to the PMOD composite. Now, however, f_{512} is determined independently of the PMOD composite.

3.2.3 Removing dark high-flux magnetogram pixels

In chapter 2, it was noted that MDI magnetogram pixels above 800 G were incorrectly identified as faculae. For consistency, this result needs to be propagated through to KP/SPM and KP/512 magnetograms.

Figure 3.1b is constructed in the same way as Fig. 3.1a, as described in the previous subsection. To calculate the cutoff, $B_{cut,SPM}$, we compare 84 magnetograms observed within two hours of corresponding MDI magnetograms between 1999 and 2003. Using such

close temporal proximity reduces the effect of evolving active regions. This indicates that when $B_{cut,MDI}$ = 800 G then $B_{cut,SPM}$ = 395 G. We repeat the same process for KP/512 using Fig. 3.1a. It follows that $B_{cut,512} = 616/f_{512}$ G, where f_{512} is the conversion factor (see section 3.1.1.3). This establishes that B_{cut} is dependent on the chosen value in MDI. While this was an arbitrary choice, it is not allowed to freely vary and is fixed.

3.2.4 Correcting pre-1990 KP/512 magnetograms

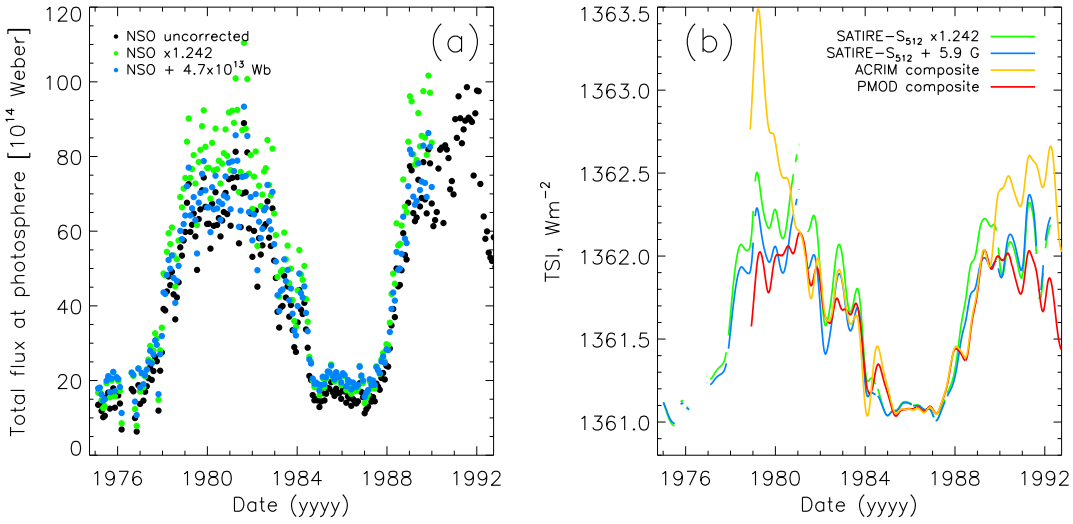


Figure 3.3: (**a**) Total photospheric flux, obtained from KP/512 Carrington maps, before applying a correction (black dots), and after either a scaling (green) or a shift is applied (blue). (**b**) SATIRE-S reconstruction of TSI using KP/512 continuum intensity images and magnetograms only, with either a scaling or a shift compared to the ACRIM and PMOD composites. All datasets are smoothed and normalised to the minimum of 1986. Gaps in the curves are present when data gaps are larger than 27 days. See main text for scaling and shifts applied to KP/512 Carrington maps and magnetograms.

Arge et al. (2002) reported that, prior to 1990, the total photospheric magnetic flux from KP/512 Carrington maps (NSO) was lower compared to that from Wilcox (WSO) and Mount Wilson (MWO) Solar Observatories Carrington maps. They used their analysis to develop a correction procedure, but this was only applied to a limited number of magnetograms, most of which were recorded during inactive times of the solar cycle. Wenzler et al. (2006) needed to use all available KP/512 magnetograms. They found a straight-line fit between the corrected and uncorrected total Carrington-rotation integrated photospheric flux, (see Fig. 3 of Wenzler et al., 2006). The linear fit was excellent (correlation coefficient, r_c=0.98), so all pre-1990 magnetograms were scaled by the regression slope

value of 1.242 (the intercept of the fit was 0.306).

In Fig. 3.3a we show the effect of scaling by a factor of 1.242 (green dots) compared to uncorrected Carrington maps (black dots), shown as the total photospheric flux of each Carrington rotation between 1974 and 1992. The uncorrected Carrington maps can be seen compared to the WSO and MWO in Fig. 1 of Arge et al. (2002). Figure 5 of Wenzler et al. (2006) shows the effect after scaling KP magnetograms by a factor of 1.242 and while the scaling seems to have improved the NSO results during the 1974-1976 minimum relative to WSO and MWO, it appears that the effect is too large between 1987 and 1989. The effect on the maximum of cycle 21 seems to be reasonable but in general the scaled level sits above MWO and WSO level of total flux at all other times. This may be due to the fact that the scaling factor is mainly derived from Carrington maps recorded during solar activity minima. Extrapolating the scaling factor to more active periods may cause magnetic flux at cycle maximum to be overestimated. Therefore, we also consider an alternate correction to apply.

While a general correction to pre-1990 data would likely take the form of a time-dependent shift and multiplication, this would require a detailed analysis of all magnetograms. Here, we consider the effect of applying both a constant shift or a scaling. To calculate the shift, we take the difference between uncorrected and corrected total NSO flux below 3×10^{15} Weber during low activity times. The value of 3×10^{15} Weber is arbitrary but encompasses the majority of low activity Carrington maps. The mean difference is 4.7×10^{14} Weber. The effect of applying this shift to the Carrington maps is also shown in Fig. 3.3a as blue dots. At solar minima, the shift results in similar values to the scaling, but the cycle amplitude is lower and agrees better with MWO and WSO during the early part of cycle 22 maximum. Unfortunately, the shift derived for the synoptic charts cannot be converted to a shift in the KP/512 magnetograms: total magnetic flux reported in KP/512 magnetograms can exceed NSO Carrington maps by a factor of well over 100 due to massive flux cancellation in the lower-resolution Carrington maps.

Instead, we take a different approach to calculate a constant shift for the pre-1990 KP/512 magnetograms. First, we assume that the scaling factor of 1.242 found by Wenzler et al. (2006) is appropriate for low-activity magnetograms. We can use the difference between low-level magnetic flux in uncorrected and scaled magnetograms as a constant shift. For this we use the threshold level determined by Wenzler et al. (2006) for the KP/512 magnetograms, i.e. $B_{thr,512} = 3\sigma_{mag,512} = 3 \times 8.1$ G = 24.3 G. This value was found by using only magnetograms during the solar minimum of 1986 and, therefore, is appropriate for

use in calculating a constant shift. The difference between scaled and unscaled $B_{thr,512}$ is 5.9 G. In Fig. 3.3b, reconstructions using KP/512 magnetograms are shown; those shown in green are for scaled magnetograms and those in blue for shifted magnetograms. All datasets in this plot are shown as smoothed time series and normalised to the solar minimum of 1986. All smoothings are performed with a Gaussian window with a full-width half-maximum of 127 days. The result of adding a constant instead of scaling is to reduce the cycle amplitude prior to 1990.

It is not immediately clear which approach, scaling or shifting, should be adopted. Therefore, we use the TSI composites as a guide, assuming that where they agree they can be used as a reasonable estimate of the true value of TSI at that time. In Fig. 3.3b, the ACRIM (yellow) and PMOD (red) composites are also shown. The IRMB composite is no longer extended back to 1978 and so is not considered here. For the period of time covered in this plot ACRIM and PMOD use the same TSI data sources post-1983, though the way they have been combined is different and after 1989 the trends clearly differ. Therefore, the agreement over 1984 - 1989 is expected. Even though ACRIM and PMOD use different datasets prior to 1984, the two composites agree well between 1981 and 1983. Before 1981, instrument degradation corrections are applied to PMOD, but not to ACRIM. Neither scaling nor shifting fully corrects the KP/512 magnetograms (private communication, Jack Harvey), though together they should give a good indication of the expected range of plausible reconstructions, with shifting minimising the cycle amplitude and scaling maximising the cycle amplitude. As the reconstruction with a shift applied clearly agrees better with ACRIM and PMOD over the 1981 to 1989 period than using the scaling, we choose to adopt the shift for the final reconstruction (see section 3.3). This then assumes that shifting is a more appropriate approach to reconstructing TSI than using a scaling. When considering the range of uncertainty on the reconstruction, we include both the shift and scaling as discussed in section 3.3.2.

3.2.5 SoHO/MDI magnetograms

The two KP datasets, KP/512 and KP/SPM, have been set up to use the same parameters. But to combine MDI with these datasets, an overlap period is needed to find the appropriate B_{sat} through linear regression. It is, therefore, important that the images used in SATIRE-S are stable over time. For MDI this does not appear to be the case.

The lowest level of magnetic flux in the MDI magnetograms, $B_{thr,MDI} = 3\sigma_{mag,MDI}$ is de-

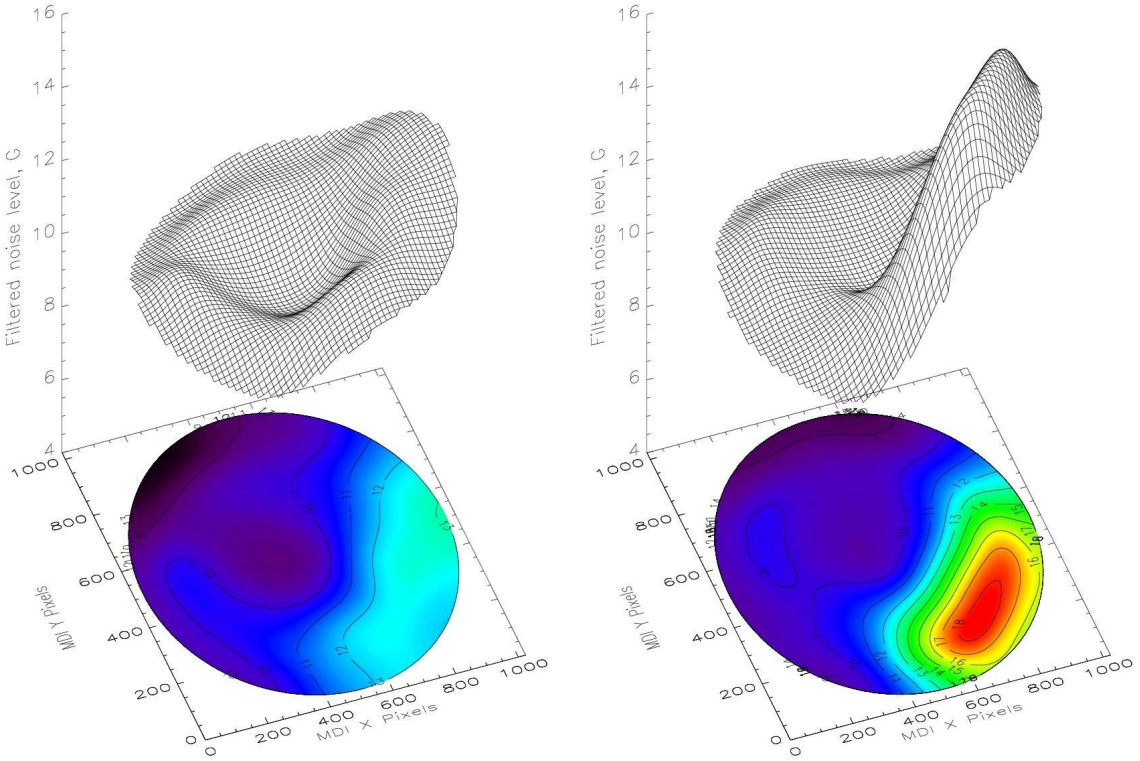


Figure 3.4: Mean background magnetic flux distribution of magnetograms from 1996-7 (left) and 2008 (right). Both plots have the same colour scale and contours are in 1 G levels. X and Y axes are the instrument CCD x and y coordinates with a fixed solar rotation axis parallel to the x-axis.

pendent on disk position. The changes to the MDI magnetograms from level 1.5 to 1.8.2 required a re-examination of the mean background flux to recalculate $\sigma_{\mathrm{mag,MDI}}$ as a function of disk-position.

We used the approach described in section 2.1 as applied to two sets of 29 magnetograms from both 1996/7 and 2008, chosen for the minimal activity present on the disk. The contour and surface plots of the level 1.8.2 magnetograms in Fig. 3.4 show the mean background flux distribution, $\sigma_{\mathrm{mag,MDI}}$, in 1996-7 (left) and 2008 (right).

Two things are clear from Fig. 3.4: the distribution is not uniform across the disk and it changed between 1996 and 2008. The non-uniformity across the disk, seen as an increase from the upper-left to lower-right in both plots, can possibly be explained by the inhomogenous nature of the Michelson interferometers in the field of view: slightly different wavelengths are observed across the disk. However, the increase over time of the lower-right quadrant is of most concern here. Here we consider possible causes for such a change and then reject data considered to be problematic. In the next chapter an investiga-

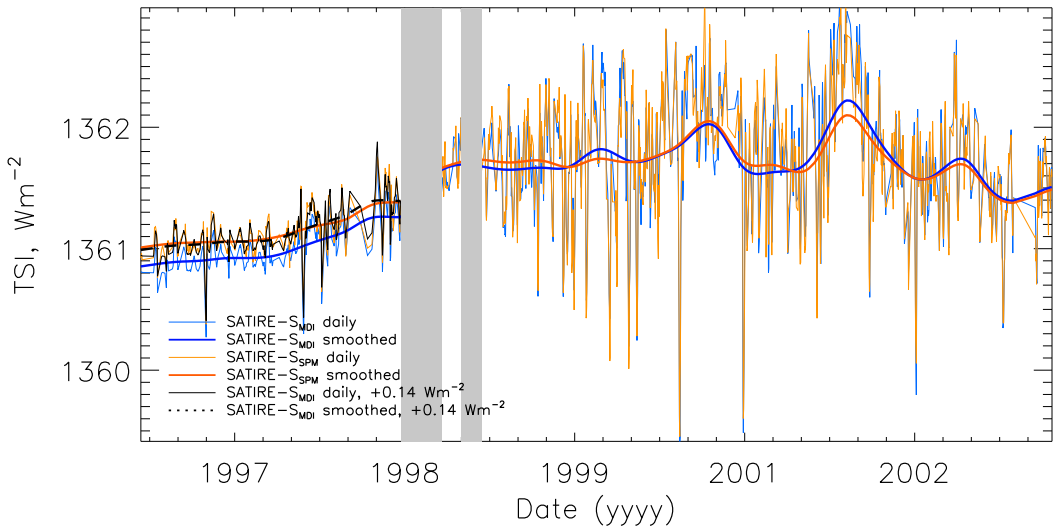


Figure 3.5: Reconstructions of KP/SPM (orange) and SoHO/MDI (blue) for the overlap period May 1996 to September 2003; KP/SPM is normalised to MDI after February 1999. Daily data are shown as thin lines, smoothed as thick, with the grey boxes covering periods when SoHO was out of contact. Prior to the loss of contact, MDI is also plotted shifted by a constant value (daily data, black thin solid line; smoothed, dashed thick line).

tion is made of findings below, and the subsequent effects they have on the reconstruction, along with an attempt to correct for them.

Aside from several shorter periods without observation, there were two extended periods, between June and October 1998 and from December 1998 to February 1999, when SoHO lost contact and was in an uncontrolled state. During the early years of operation, prior to this loss of contact, MDI was deliberately defocused to ensure the quality of high resolution observations and prevent spatial aliasing in full-disk observations. When contact was re-established, following an analysis that showed this defocusing was not required, the instrument was fully refocused, which significantly changes the point spread function (PSF). Subsequent periodic refocusing was performed to correct a drift and maintain the focus, which should basically maintain the PSF at roughly the level achieved after the loss of contact. This, among other effects such as e.g. the front window darkening, thermal effects, a small change in the observed wavelength and any unknown effects during the period without contact, may have contributed to a change in the registered magnetic flux as a function of disk-position (private communication, Rock Bush). We found that dates when focus changes were performed correlate with a change in distribution of flux in the lower right quadrant. Indeed, by bringing the instrument into focus and narrowing the PSF, the expected outcome would be an increase in the observed magnetic flux, and this

is exactly what can be seen in Fig. 3.4. We are mainly interested in global sensitivity to the magnetic flux through the effect they have on a reconstructed irradiance, which can be judged by comparing reconstructions using MDI and KP/SPM data prior to and after the loss of telemetry.

In Fig. 3.5, daily and smoothed reconstructions for the overlap period of MDI (blue) and KP/SPM (orange) are shown. The two reconstructions have a unity regression (see section 3.3.1) post-SoHO gaps (grey shading), but show a shift prior to the gaps. It can be rectified by a simple shift in the MDI-based reconstruction of 0.14 Wm^{-2}. The shifted reconstruction is shown as daily (thin, black) and smoothed (dashed) lines. The agreement of the trends in the shifted MDI and KP/SPM reconstructions is excellent.

Note that a defocus also reduces contrast in the continuum intensity images (Danilovic et al., 2008), which would also introduce a bias between sunspot and umbral areas if determined from the period prior to loss of contact relative to determining it afterward. The chosen thresholds are based on continuum intensity images from after the loss of telemetry (see section 3.2.1), so this is not a concern here.

With both of these results, along with the knowledge that the source of the difference almost certainly lies with MDI given the known changes over the SoHO data gaps, we decide to only use MDI data from after the data gaps and use KP/SPM to reconstruct prior to this. In chapter 4, a first-order correction for the focus change is attempted.

3.3 Reconstruction of TSI for cycles 21-23

In this section, we explain how we produce the three-cycle reconstruction, independently of any TSI composite after 1989, and how the errors have been formulated. The ultimate aim here is to produce a composite-independent reconstruction of TSI. This eliminates the need to tailor the reconstruction to each composite to test their validity (see section 1.8). The SORCE/TIM data, unused in the composites at the time of writing, is therefore an ideal candidate for use as an independent basis to fix the model's free parameter.

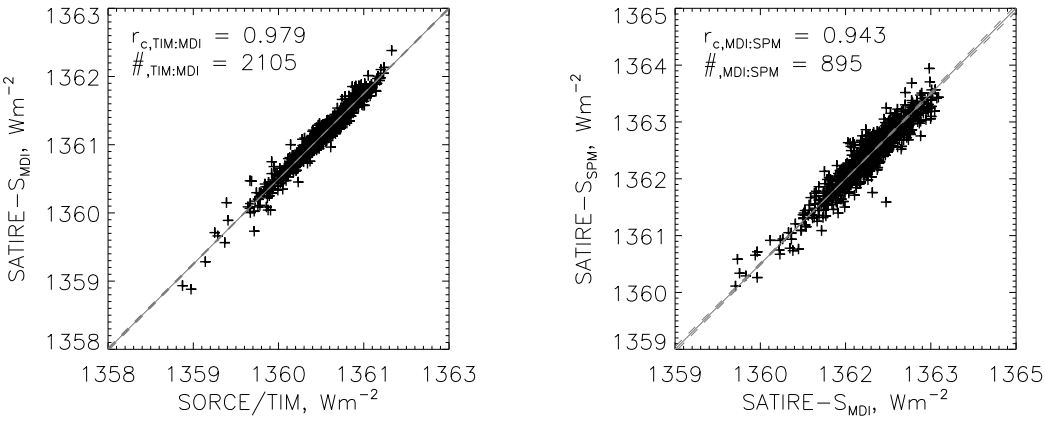


Figure 3.6: Regression of (left) TIM and MDI and (right) MDI and KP/SPM. The regression line is shown by the solid grey line and 1σ uncertainty is represented by the dashed grey lines.

3.3.1 Combining MDI and KP/SPM data

SATIRE-S_{MDI} is normalised to TIM for 2105 days between 2003 February 25 and 2009 October 31. By varying $B_{sat,MDI}$ the resulting best fit unity regression is found at 443 G[6]. This is shown as the solid line in the left regression plot of Fig. 3.6. The correlation coefficient, $r_{c,TIM:MDI}$, is 0.979.

Note that a 30-day period centred on the huge sunspot group passage that reached disk centre around 2003 October 29 has been omitted from this fit (i.e. the first major deviation in Fig. 2.5). We find that the fit is relatively stable with $B_{sat,MDI}$ = 443 G for the entire overlap period with TIM except during this passage. SATIRE-S generally reproduces the influence of sunspots well, but this sunspot group seems to have been so exceptionally large that the assumed sunspot properties appear to have changed. SATIRE-S significantly overestimates the darkening produced by this sunspot group and this leads to a change in $B_{sat,MDI}$ that causes the rest of the time series to be less well reproduced. The spot is only removed when determining $B_{sat,MDI}$, but considered in the comparisons with the TSI composites (section 3.4).

Next, we combine SATIRE-S_{SPM} with SATIRE-S_{MDI} using the overlap period of 895 days between 1999 February 18 and 2003 September 21. SATIRE-S_{SPM} is normalised to SATIRE-S_{MDI}. Using SATIRE-S_{MDI} with $B_{sat,MDI}$ = 443 G, a unity regression fit for

[6]This is done using the *fitexy* linear regression routine; see section 2.4 for a discussion on using this approach.

SATIRE-S_{SPM} is achieved when $B_{sat,SPM}$ = 330 G. The regression plot is shown on the right-hand side of Fig. 3.6. The correlation coefficient of this fit, $r_{c,MDI:SPM}$, is 0.943.

With SATIRE-S_{SPM} established, the link to SATIRE-S_{512} is automatically prescribed. The same normalisation factor found to bring SATIRE-S_{SPM} in to line with SATIRE-S_{MDI} is used for SATIRE-S_{512}. The conversion factor, f_{512}, for $B_{sat,512}$ = 330 G (see Fig. 3.2b) is 1.580, to the nearest 0.005. The entire time series is normalised to TIM at the December 2008 minimum as a reference period.

Note that the method to combine MDI with KP/SPM differs to that of combining KP/SPM with KP/512. In the latter case, magnetograms from both datasets are compared and adjusted so that the same value of the free parameter can be used. This method is not possible with MDI because the spatial response is not uniform, as mentioned in section 3.2.5 and discussed in detail in the following chapter.

We use the TSI composites to justify the addition of a constant to pre-1990 KP/512 magnetograms (see section 3.2.4) and, therefore, the reconstruction is independent of the composites after 1989. We denote this full reconstruction as SATIRE-S_{Ind}. In Fig. 3.7 the complete SATIRE-S_{Ind} reconstruction is shown as daily (thin lines) and smoothed (grey lines) time series. All that remains is to estimate an uncertainty on this result.

3.3.2 Uncertainty estimation on the reconstruction

As the model uncertainties are difficult to quantify (see section 2.2), we try and estimate the uncertainty in the SATIRE-S reconstructions by exploring the change in the modelled TSI that results from variations in the free parameter, B_{sat}, that fall within the allowed fit parameters, such as e.g., quantified by a 1σ change in the regressions. We use the difference between the datasets being regressed as an estimate of the standard deviation for both datasets in the *fitexy* routine.

Starting with SATIRE-S_{MDI} we found that the formal 1σ error on the unity fit gradient with SORCE/TIM, $\pm$0.006, corresponds to $\pm$5 G in $B_{sat,MDI}$. This uncertainty also encompases the best value for $B_{sat,MDI}$ when a least-squares fit, rather than a regression, is used to find agreement between TIM and SATIRE-S (see section 2.4).

The uncertainty over the KP/SPM part of SATIRE-S_{Ind} is determined from the uncer-

tainty range on the unity regression fit between SATIRE-S_{MDI} and TIM. In other words, SATIRE-S_{SPM} is fitted to SATIRE-S_{MDI} using $B_{sat,MDI}$ = 438 and 448 G for the upper and lower boundaries, respectively. Then, the 1σ errors on these unity regression fits are used to find an expanded uncertainty range prior to 1999. The 1σ error on the unity regression gradient is 0.017 and the resulting $B_{sat,SPM}$ is 307 G and 365 G for the upper and lower range, respectively.

The uncertainty range of $B_{sat,512}$ in SATIRE-S_{512} includes the standard error of approximately ±0.04 on the associated conversion factor, f_{512} (see section 3.2.2). We take values of f_{512} that exacerbate the uncertainty range. For $B_{sat,512}$ = 365 G, this equates to f_{512} = 1.600 and for $B_{sat,512}$ = 307 G, f_{512} = 1.550.

Finally, we must consider the effect of using a scaling or constant shift correction to KP/512 magnetograms prior to 1990 (i.e. recall the comparison in section 3.2.4 and shown in Fig. 3.3b). We consider both approaches so that on each date the pre-1990 correction that produces the largest deviation from SATIRE-S_{Ind}, either highest or lowest in value, is adopted as the upper and lower uncertainty, respectively, to maximise the range.

The solar minimum of December 2008 is used as the reference date, assuming no long-term uncertainty at that time. Therefore, the two extreme reconstructions that constitute the uncertainty range are normalised to that time, though this does not mean that there is no uncertainty on the absolute value of TSI then. The resulting reconstruction is shown in Fig. 3.7 along with the smoothed uncertainty range shown as a contour around the smoothed dataset. While SATIRE-S_{Ind} pre-1990 is not independent of the TSI composites, the uncertainty range is, as the uncertainty includes both the maximum range produced by shifting or scaling KP/512 magnetograms prior to 1990.

3.4 Comparison of SATIRE-S_{Ind} with TSI Composites

In this section we compare SATIRE-S_{Ind} to the three TSI composites in detail. This returns us to a debate over which composite (ACRIM, IRMB or PMOD) provides the best record of TSI (see section 1.8). Wenzler et al. (2009) used SATIRE-S to do this by optimising the model for each composite and found that SATIRE-S fitted to PMOD gave the best correlation coefficients, while ACRIM and IRMB displayed trends that PMOD and SATIRE-S did not show.

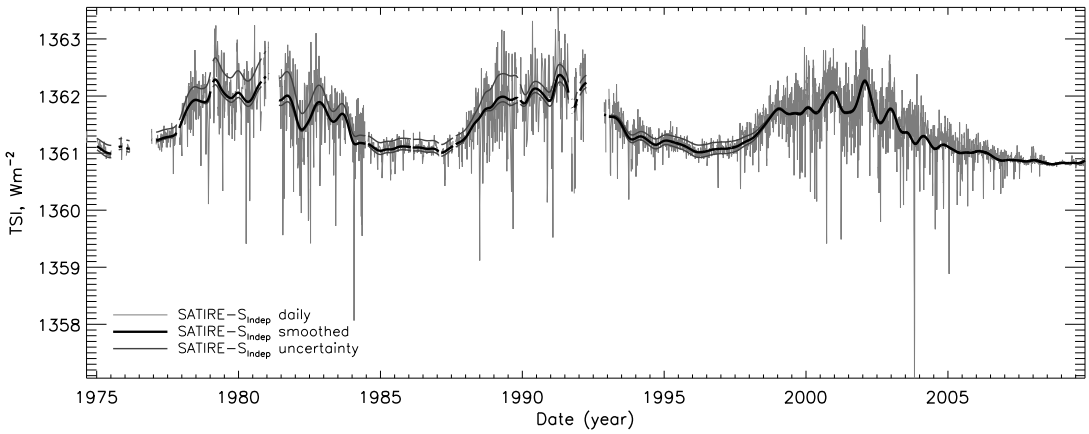


Figure 3.7: The entire SATIRE-S$_{Ind}$ dataset. Daily data are shown as solid grey lines when dates are contiguous. The thick black line is smoothed data and the uncertainty range is shown, smoothed only, as thin black lines. Gaps in the curves are present when data gaps are larger than 27 days.

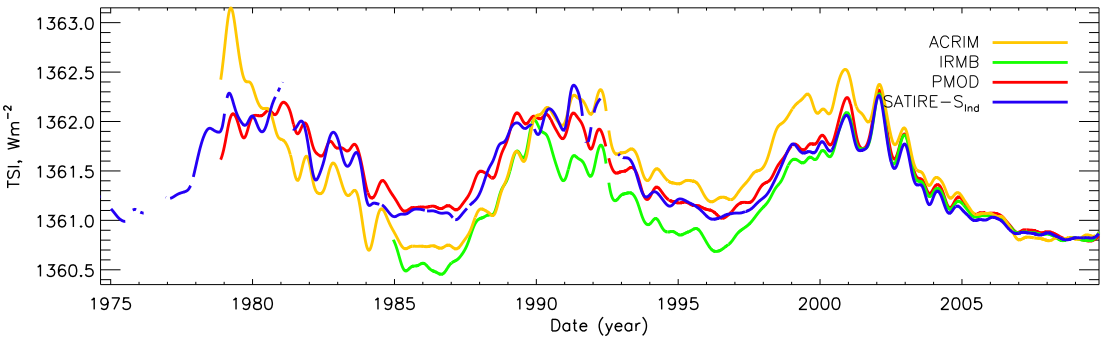


Figure 3.8: Comparison of smoothed TSI composites and SATIRE-S$_{Ind}$, all normalised to SORCE/TIM at the minimum of December 2008. ACRIM is in yellow, IRMB in green, PMOD in red and SATIRE-S$_{Ind}$ in blue. Gaps in the curves are present when data gaps are larger than 27 days.

SATIRE-S$_{Ind}$ provides a much stronger test as it covers a longer period and is independently optimised. In Fig. 3.8, all data are shown together, normalised to the mean value of SORCE/TIM (not shown) over the six month period centred on the December 2008 solar minimum. The data have been smoothed to aid clarity, with PMOD in red, IRMB in green, ACRIM in yellow and SATIRE-S$_{Ind}$ in blue. In Fig. 3.9, the difference between the composites and SATIRE-S$_{Ind}$ are shown for daily data for the period 1978-2009. Before 1984, data are shown on each date that data are available in PMOD, ACRIM and SATIRE-S$_{Ind}$; after 1984, data are shown when available in all four datasets. Table 3.1

Table 3.1: Comparison of SATIRE-S_{Ind} with TSI composites using the original time series (upper half of the table) and detrended (lower half). Comparisons are made only for dates when data in all datasets are available: since the IRMB composite covers a shorter period, the table is further split into two to make fair comparisons over the longer and shorter periods.

Daily dataset comparison (1978 - 2009)				
Series 1	Series 2	$r_c[r_c^2]$	Slope	No. points
ACRIM	SATIRE-S_{Ind}	0.879 [0.772]	0.832	5663
PMOD	SATIRE-S_{Ind}	0.936 [0.875]	1.000	5663
ACRIM	PMOD	0.866 [0.750]	0.830	5663
Daily dataset comparison (1984 - 2009)				
ACRIM	SATIRE-S_{Ind}	0.903 [0.815]	0.835	5150
IRMB	SATIRE-S_{Ind}	0.924 [0.854]	0.978	5150
PMOD	SATIRE-S_{Ind}	0.942 [0.888]	0.992	5150
ACRIM	IRMB	0.924 [0.853]	0.856	5150
PMOD	IRMB	0.922 [0.850]	1.013	5150
ACRIM	PMOD	0.894 [0.799]	0.841	5150
Daily dataset comparison, detrended (1978 - 2009)				
ACRIM	SATIRE-S_{Ind}	0.866 [0.750]	0.986	5663
PMOD	SATIRE-S_{Ind}	0.873 [0.762]	1.039	5663
ACRIM	PMOD	0.830 [0.689]	0.947	5663
Daily dataset comparison, detrended (1984 - 2009)				
ACRIM	SATIRE-S_{Ind}	0.889 [0.791]	0.992	5150
IRMB	SATIRE-S_{Ind}	0.940 [0.883]	1.036	5150
PMOD	SATIRE-S_{Ind}	0.877 [0.769]	1.032	5150
ACRIM	IRMB	0.945 [0.892]	0.957	5150
PMOD	IRMB	0.911 [0.831]	0.994	5150
ACRIM	PMOD	0.847 [0.718]	0.960	5150

summarises comparisons between the TSI composites and SATIRE-S_{Ind}. It is divided in two halves deliniated by the double line: the top half shows the results of using the daily dataset while the bottom half is for detrended data only so that the rotational variability alone can be considered. These two halves are further divided into comparisons spanning 1978-2009 and 1984-2009, the latter period being the one to which the IRMB composite is restricted. All comparisons use only dates when data are available in all datasets. This means that the correlation coefficients, r_c, and slopes listed in Table 3.2 are strongly weighted by cycle 23, as most data points in SATIRE-S_{Ind} are during that period.

It is clear in Fig. 3.8 that PMOD has a trend at the lower extreme of the composite range at nearly all times and SATIRE-S_{Ind} is in good agreement with PMOD at all times, except

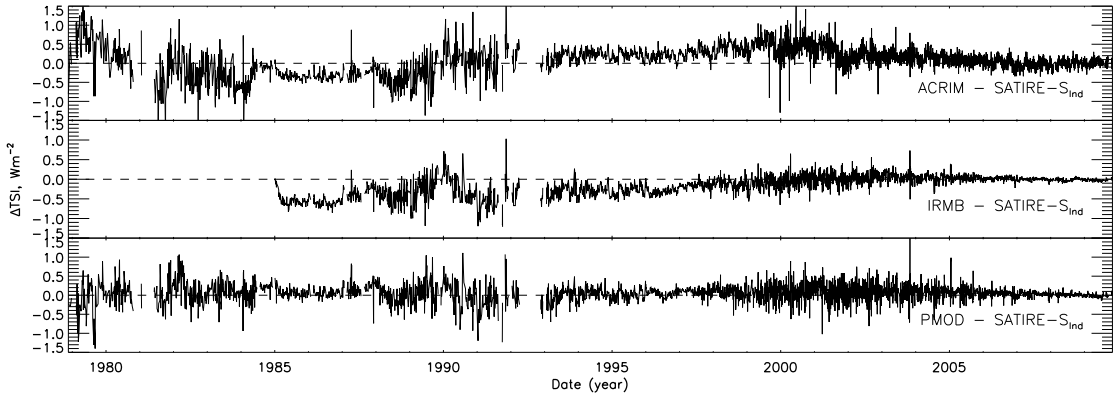


Figure 3.9: Difference plots of (**top**) ACRIM - SATIRE-S_{Ind}, (**middle**) IRMB - SATIRE-S_{Ind}, (**bottom**) PMOD - SATIRE-S_{Ind}. The plots cover the period 1984-2009 for dates when data are available in all four datasets and prior to this when data are available in all datasets excluding IRMB.

around the peak of cycle 22. The better agreement with PMOD is shown by value of the regression slope between 1984 and 2009 for the daily non-detrended results, 0.992, and the highest correlation coefficient, 0.942. SATIRE-S_{Ind} only displays a secular change between minima over cycle 23, like PMOD, and does not show the more marked secular variations in the IRMB and ACRIM composites. Interestingly, IRMB and SATIRE-S_{Ind} appear to have the best agreement in the declining phase of cycle 23, with a gradient in better agreement with SORCE/TIM than either PMOD or ACRIM (for SATIRE-S this is by construction). The lower regression slope in ACRIM is caused by a difference in long-term trends, relative to SATIRE-S_{Ind}. The correlation coefficient between SATIRE-S_{Ind} and PMOD is higher than between SATIRE-S_{Ind} and other TSI composites and between the composites themselves.

The agreement between the detrended TSI records, which are measures of rotational modulation of irradiance and the effects of active region evolution, is generally good. IRMB has the best agreement with SATIRE-S_{Ind}, with a correlation coefficient of 0.940, then ACRIM and finally PMOD. The agreement between SATIRE-S_{Ind} and IRMB is almost as good as the best agreement between TSI composites, i.e. r_c = 0.945 between IRMB and ACRIM. It is also clear that the scatter in the detrended datasets in Fig. 3.9 between IRMB and SATIRE-S_{Ind} during cycle 23 is much lower than with the other TSI composites, though the longer-term trend during the rising phase is not the same, especially before 1997; the entire period of cycle 23 utilises the DIARAD instrument on SoHO. This goes some way to explaining why IRMB has a relatively good regression gradient with SATIRE-S_{Ind} when considering the data that have not been detrended, despite the large disagreement in the smoothed long-term trend shown in Fig. 3.8. These detrended

comparisons show that SATIRE-S_{Ind} is able to recreate rotational variability to a level of agreement that is seen in the comparisons between the TSI composites.

The regression slopes also indicate that the rotational variation in SATIRE-S_{Ind} falls between the TSI composites, with ACRIM having the largest regression slope, and therefore amplitude, and PMOD and IRMB the lowest. It should be borne in mind that the detrended results may be affected strongly by active regions. It has been documented, e.g. Unruh et al. (2008), that TSI can be overestimated in SATIRE-S when large facular regions transit the edge of the solar disk (this is also discussed in chapter 5). Some of this overestimation has been improved upon through the modification described in section 2.3.3. It may be the case, though, that SATIRE-S_{Ind} still overestimates rotational varability and that is why the regression slopes suggest that SATIRE-S_{Ind} has a higher amplitude than PMOD and IRMB.

In general, the composites and reconstruction agree well with each other on a rotational basis. The major cause of differences between composite datasets are likely to be reflected by the use of different data sources in cycle 23. Note that the regressions and correlation coefficients also improve for the shorter time period, consistent with the fact that the older data are less certain. Assuming, for the sake of argument, that SATIRE-S_{Ind} is a good representation of TSI behaviour we find that PMOD is the most reliable composite of TSI, taking into account rotational, cyclical and secular variations, although on rotational time scales IRMB is superior.

Given that PMOD is the composite by far the closest to SATIRE-S_{Ind} overall, we concentrate in the following on an in-depth comparison of the two. Figure 3.10 shows in the top panel PMOD (light red, daily; red, smoothed) and SATIRE-S_{Ind} (light blue, daily; blue, smoothed) irradiance. The bottom plot shows the difference between the two, PMOD minus SATIRE-S_{Ind} (black, smoothed). Also indicated are the smoothed uncertainty range for SATIRE-S_{Ind}, as thin, light blue and black lines in top and bottom plots respectively, and PMOD black and red error bars from Fröhlich (2009). The values of these errors are given in the middle table of Table 3.2. These error bars were derived by Fröhlich (2009) as an estimate of the long-term stability in the instruments and the calculation is described within that publication. The error bar at the 1986 minimum is very small because this minimum was used as the reference date for comparison with other minima. In the plot, PMOD is normalised to the mean of SATIRE-S_{Ind} at the 2008 minimum, 2008 December 15 $\pm$90 days, to help discuss the change between cycle minima and bring irradiance levels in line with SORCE/TIM. Dashed lines in blue and red guide the eye to minima levels in

1986, 1996 and 2008.

The immediate impression is that across the 5899 days, where data are available in both datasets, there is very good agreement on all timescales for the majority of the time period. The regression slope between SATIRE-S_{Ind} and PMOD is 0.995 with a correlation coefficient of 0.958 meaning that over the 31 year period of overlap, SATIRE-S_{Ind} can explain almost 92% of variability in the TSI composite. Note that the correlation coefficients and regression slopes between PMOD and SATIRE-S_{Ind} differ in Tables 3.1 and 3.2 due to a change in the intervals considered.

In the lower part of Table 3.2, comparisons have been made for the full cycle and each individual cycle. The agreement is highest during cycle 23 during which SATIRE-S_{Ind} is able to explain more than 96% of the variation. We note that the difference plot shows that the gradient in TSI during the rising and falling phases of cycle 23 is slightly lower in SATIRE-S_{Ind}, a consequence of using TIM to fit the reconstruction, which shows a gentler decline of TSI with time than PMOD.

A significant result is that SATIRE-S_{Ind} is able to reproduce the secular, inter-cycle, change in PMOD TSI, mimicking the ~0.2 Wm^{-2} decline between 1996 and 2008. There is a slight difference between minima values in 1996, relative to 2008: within the uncertainty range, SATIRE-S_{Ind} suggests a decline between cycle 23 minima of between 0.11 and 0.32 Wm^{-2} encompassing the result from PMOD. The best fit from SATIRE-S_{Ind} also falls within the error bars of PMOD in 1996. SATIRE-S_{Ind} also agrees with PMOD on the secular change over cycle 22 and possibly shows a slight secular increase over cycle 21 as can be seen in Table 3.2, though data are sparse prior to 1977 and there is no change within the uncertainty range. These results, summarised in Table 3.2, calculate the 1976 minimum using the mean of data between June 1975 and July 1976, while all other maxima and minima are for six month periods centred on the maximum and minimum months of the sunspot cycle from the Solar Influences Data Centre (SIDC) smoothed monthly sunspot dataset (SIDC-team, 2011).

Table 3.2: Best-fit comparison between PMOD and SATIRE-S for dates when both datasets overlap (except the minimum of 1976). The uppermost table shows the absolute values at activity cycle maximum and minimum. The middle table gives irradiance values relative to the December 2008 minimum with uncertainty estimates (SATIRE-S_{Ind}) and errors (PMOD). The bottom set gives the results of comparing different cycles of the reconstruction, including the regression slope and correlation coefficients.

Cycle Maximum[1] and Minimum TSI, Wm^{-2}							
Series	Min C20/21[2]	Max C21	Min C21/22	Max C22	Min C22/23	Max C23	Min C23/24
SATIRE-S_{Ind}	1361.06	1362.01	1361.08	1361.96	1361.01	1361.74	1360.82
PMOD	-	1362.02	1361.14	1362.10	1361.07	1361.82	1360.82
Cycle Maximum[1] and Minimum TSI Relative to 2008 Minimum, Wm^{-2}							
Series	Min C20/21	Max C21	Min C21/22	Max C22	Min C22/23	Max C23	Min C23/24
SATIRE-S_{Ind}	$0.24^{+0.14}_{-0.06}$	$1.19^{+0.40}_{-0.09}$	$0.26^{+0.14}_{-0.06}$	$1.14^{+0.42}_{-0.09}$	$0.20^{+0.12}_{-0.09}$	$0.92^{+0.02}_{-0.02}$	$0.00^{+0.00}_{-0.00}$
PMOD	-	1.21	0.32±0.01	1.28	0.25±0.10	1.00	0.00±0.16

Statistics of Full and Cycle Periods					
Period	Start date	End date	$r_c[r_c^2]$	Slope	No. points
Full	1978/12/10	2009/10/31	0.958 [0.919]	0.995	5899
Cycle 23	1996/05/15	2008/12/15	0.981 [0.963]	0.978	3760
Cycle 22	1986/09/15	1996/05/14	0.917 [0.840]	1.040	1148
Cycle 21 (part)	1978/12/09	1986/09/14	0.898 [0.806]	1.050	740

[1] TSI values are calculated from ±90 days centred on the dates of maximum and minimum smoothed monthly sunspot number from SIDC data. [2] Minimum values are estimated between June 1975 and July 1976 due to the lack of data around the sunspot minimum of June 1976.

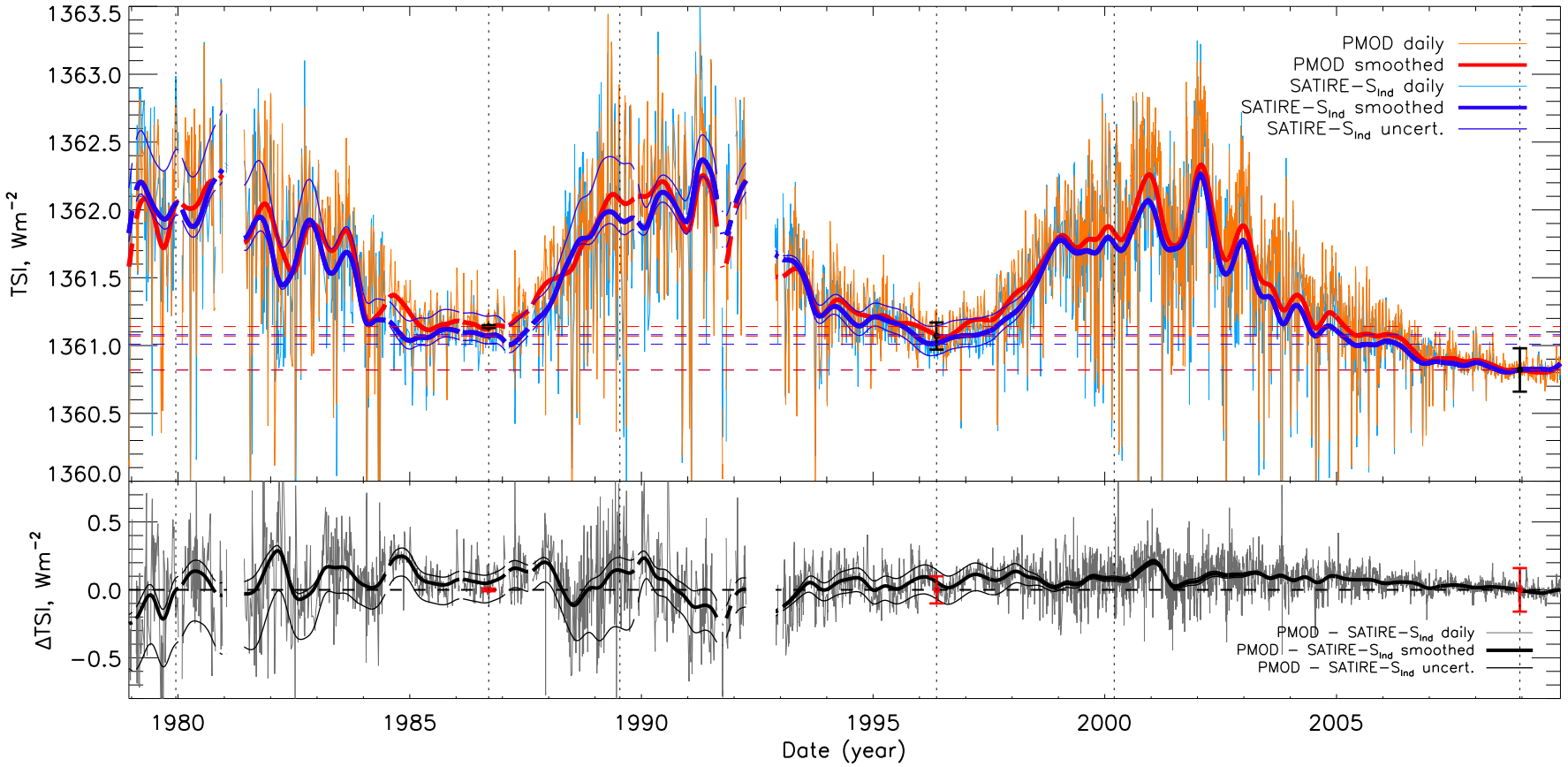


Figure 3.10: In the upper plot, PMOD (daily data, light red; smoothed, thick red line) and SATIRE-S_{Ind} (daily data, light blue; smoothed, thick blue) between 1978 and 2009 normalised to SORCE/TIM at December 2008. The thin blue lines mark the uncertainty range of SATIRE-S_{Ind} (only smoothed values plotted). In the lower plot, the difference between PMOD and SATIRE-S_{Ind} is shown (daily, grey; smoothed, black) along with the difference of the uncertainty with respect to PMOD. The black and red error bars are the errors from Fröhlich (2009) in the upper and lower plots, respectively. Dotted horizontal lines indicate cycle maxima and minima. Gaps in the curves are present when data gaps are larger than 27 days.

Although there is good agreement in the long-term trends, there are periods of approximately one to two years, during the maxima of cycles 21 and 22, where disagreement reaches ~0.3 Wm^{-2}. This is likely to be the result of the poorer data quality of KP/512 and the larger uncertainty in PMOD during the first two cycles. The increased difference in 1990/1 occurs in KP/512 data in the period where no additive correction is applied to the magnetograms (see section 3.2.4). There is an implication that a scaling correction per se should not only be applied to pre-1990 magnetograms and that a more sophisticated image correction procedure should be used, rather than a constant shift, scaling or combination of the two. A fuller investigation should be performed to understand the time-dependent problems in the KP/512 magnetograms, expanding on the work done by Arge et al. (2002).

There is also clearly a different six-monthly trend apparent in the period from late-1996 to 1998 where the gradient in SATIRE-S_{Ind} is lower than PMOD. Although MDI data over this period have not been employed for the final reconstruction, for this small stretch, both SATIRE-S_{MDI} and SATIRE-S_{SPM} show a very similar trend (see Fig. 3.5). Krivova et al. (2011a) discuss this difference in some detail and also compare other radiometric observations that show significant deviations from PMOD. It is therefore very likely that the degradation corrections applied during this period to VIRGO data may profit from being revisited.

3.5 Discussion and conclusions

In this chapter we present the first reconstruction using the SATIRE-S model to fully cover solar cycles 21, 22 and 23. The reconstruction is compared with three TSI composites constructed by Willson and Mordvinov (2003), Dewitte et al. (2004) and Fröhlich and Lean (1998) and referred to as ACRIM, IRMB and PMOD respectively, the last of which is considered in more detail. The modelled time series was constructed without reference to any of the composites, except prior to 1990. This means that the entirety of cycle 23 is independent of all the TSI composites. This provides an opportunity to test the model's assumption: that changes in irradiance are only caused by the evolution of surface magnetic flux.

The SATIRE-S reconstruction is found to agree best with the PMOD composite on time scales longer than rotational. On shorter time scales the IRMB composite displays a bet-

ter agreement, though all three composites are similar on this shorter time scale. The model is not just extremely successful at recreating rotational and cyclical variability in PMOD, but also in reproducing the secular trend, particularly apparent in the decline in TSI of ~0.2 Wm^{-2} between the minima of cycle 23. The difference between the model and PMOD minima of cycle 21/22 and cycle 22/23, relative to the 2008 minima, is less than or equal to 0.06 Wm^{-2} and within the uncertainty of the model and of the observations. The fact that the model and the PMOD composite agree very well provides support for both of them.

The model is able to account for 92% of the variability in PMOD since 1978 and 96% of the variability over cycle 23. Due to the general independence of the model, these results support the PMOD composite as being the most realistic record of TSI. It also suggests that rotational, cyclical and secular variations in TSI are mainly the result of the evolution of photospheric magnetic fields. This finding is in opposition to Fröhlich (2009, 2011) and Steinhilber (2010), who proposed that secular changes in TSI are caused by a global temperature change (see also Tapping et al. (2007) and references therein).

It is clear from Fig. 3.10 that there is disagreement between the model and PMOD on periods of one to two years, most significantly over the maximum years of cycle 21 and 22 during which KP/512 are employed in the model. KP/512 data are subject to many more problems than KP/SPM and SoHO/MDI. These have been discussed in sections 3.1.1.3, 3.2.2 and 3.2.4.

Some of the differences are also a result of inadequate degradation corrections in PMOD. During the period prior to 1982 the NIMBUS-7/HF radiometer was the only instrument in operation. The degradation correction applied to data from this instrument was based on the work done on the VIRGO/PMO6V radiometer in conjunction with ACRIM I data (Fröhlich, 2003). While we accept that this is a reasonable correction, we believe that some uncertainty in PMOD remains during this early period.

Krivova et al. (2011a) found that the degradation correction of VIRGO during the initial years of observation over 1996-1998 may not be entirely adequate. Our study adds to that comparison, which also used UARS/ACRIM II and DIARAD and PMO6V radiometers within the SoHO/VIRGO instrument. We find that SATIRE-S reconstructions using either SoHO/MDI or KP/SPM images disagree in an almost identical manner to the trend in PMOD during the first 18 months of the rising phase of cycle 23. We did find, however, that the magnetic flux in SoHO/MDI magnetograms is different before and after the pe-

riod when SoHO was not in operation in 1998. This can be attributed to instrument focus changes. While they can be used in shorter reconstruction stretches which agree with the reconstruction using KP/SPM (see Fig. 3.5), MDI magnetograms should not be used to test secular variation, without adequately correcting for this effect.

The SATIRE-S_{Ind} full reconstruction covers 6379 days between 1974 and 2009. The daily and smoothed data, along with the uncertainty range will be made available at http://www.mps.mpg.de/projects/sun-climate/data.html.

While TSI is an important observable of changes in solar irradiance, many recent climate modelling attempts attribute solar spectral irradiance as physically important when trying to understand the impact the Sun has on the Earth's climate. By design, SATIRE-S automatically produces spectra over climate-relevant wavelengths. In chapter 5 the spectral irradiance resulting from SATIRE-S_{Ind} will be compared to the results of SORCE/SIM between 200 and 1630 nm. SORCE/SIM was used in conjunction with SORCE/SOLSTICE in Haigh et al. (2010) to investigate the unexpected spectral forcings on stratospheric ozone concentration and temperatures. Therefore in chapter 6, SATIRE-S_{Ind} SSI will be employed in the same 2D atmospheric model used in that study to understand how the variability in SATIRE-S_{Ind} SSI differs to that observed by the SORCE instruments in the context of the effect different spectral irradiance changes have on the Earth's atmosphere.

Chapter 4

Analysis of temporal changes in MDI magnetograms

Overview

In this chapter, I expand on section 3.2.5, by analysing MDI magnetograms in greater detail. The aim is to provide evidence that changes in the focus of the MDI instrument have contributed to a change in the reported magnetic flux in MDI magnetograms. As a consequence, this leads to an incorrect reconstruction of irradiance on cyclical time scales when using only MDI data. In addition, it is apparent that early degradation of the front window also resulted in a significant change to MDI magnetograms prior to the loss of contact in 1998. A simplistic approach to correct the MDI magnetograms for focus changes is then applied before producing a full cycle reconstruction using only corrected MDI magnetograms. The correction has some limited success.

Comparisons of reconstructions, using only MDI or KP data, are also made with the PMOD composite to investigate the validity of the regression fitting approach taken to combine them in SATIRE-S_{Ind}. This is important because the way the datasets are combined influences the reconstructed change between the cycle minima of 1996 and 2008. The results from both analyses further support the claim of the previous chapter, that solar irradiance has declined between the cycle minima of 1996 and 2008.

4.1 Changes in magnetic flux recorded in MDI magnetograms

In section 3.2.5, only the change in the mean background magnetic flux was considered. Figure 3.4 showed that the spatial distribution of the mean background magnetic flux had changed between 1996 and 2008, with a focus change of the MDI instrument a suggested cause. In Fig. 4.1, a contour plot is shown of the difference between the mean background magnetic flux in 2008 and 1996, i.e. the difference between right and left plots of Fig. 3.4. The difference plot shows, once again, that the change is not uniform across the disk and that there is a significant increase in the lower right quadrant. This plot is also used to identify regions used in the analysis of this chapter, indicated by curved regions enclosed by dashed lines in each quadrant. These regions have been arbitrarily chosen so that several equally sized and evenly spaced regions on the solar disk can be considered and compared to the lower right region which shows a peak in the change between 1996 and 2008. As will be seen, the cause of the change in behaviour of each quadrant in response to focus changes is unknown but may be related to the variable wavelength as a function of position on the detector or from spatial degradation of the instrument, as discussed in section 3.2.5. Note that the mean background magnetic flux shows only that the lowest reported magnetic fluxes in MDI magnetograms have changed, but does not show how higher magnetic fluxes may have been affected.

In Fig. 4.2, histogram equating between MDI and KP/SPM magnetograms has been used to determine the multiplication factor (referred to as *factor* from here on) that would be required, as a function of magnetic flux, to bring magnetic flux in KP/SPM magnetograms into agreement with MDI magnetograms (see section 3.2.2 for a description of the method used). The factors have been calculated over a time period spanning 1996 May 20 and 2003 September 21 when MDI and KP/SPM overlap. The upper four plots show the factors resulting from a consideration of the separate regions contained within the curved, dashed boundaries in the upper left (UL), upper right (UR), lower left (LL) and lower right (LR) of Fig. 4.1. Solid lines are the average of the 282 near-contemporaneous dates for which KP/SPM and MDI magnetograms are available prior to 1999 February 18 and the dashed lines for the 895 dates after this. The earlier period includes all overlapping dates and so are not comparable to the the dates used to produce the mean background magnetic flux in Figs. 3.4 and 4.1. By using all available dates to compare MDI and KP/SPM magnetograms, higher levels of magnetic flux can be compared that are not present at solar minimum times. Relative to KP/SPM, the registered magnetic flux in MDI increases with

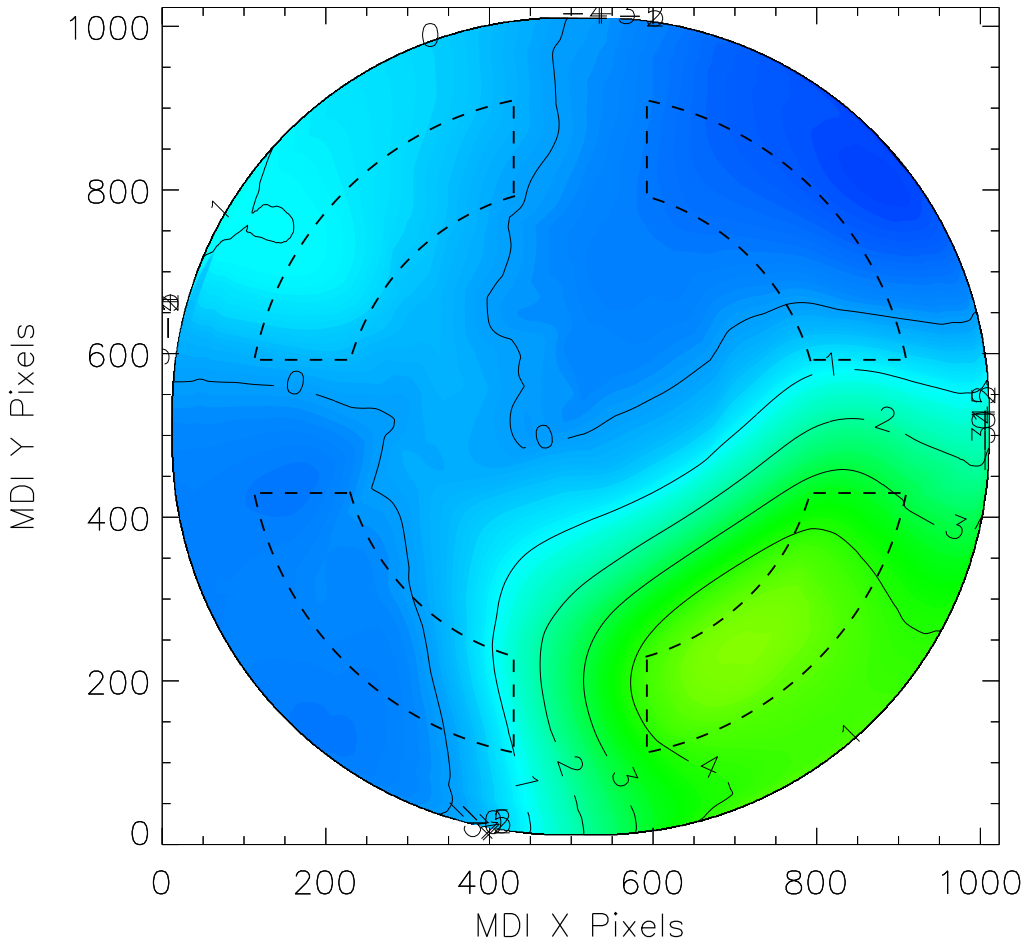


Figure 4.1: The difference between the distribution of the mean background magnetic flux between the minima of 1996 and 2008, i.e. the difference of left and right plots in Fig. 3.4. Contours are 1 G levels. The curved, dashed lines denote the region considered in the analysis of this chapter (see main text).

increasing magnetic flux from a minimum point at around 20 G up to approximately 200 G, after which the ratio of magnetic fluxes remains relatively constant. In these four plots there has been an increase in the reported magnetic fluxes of MDI, relative to KP/SPM, between those prior to 1999 February 18 and those after, in all four regions and at almost every level of magnetic flux, except in the LL at around 50 G.

As Fig. 4.2 compares data from two instruments, the resulting factors could be the result of changes in one or both of the instruments. While we can be confident that MDI has changed (see Figs. 4.1 and 3.4), it is not possible to confirm if there has been a change in the KP/SPM magnetograms in the same manner as for MDI. For MDI, two minima periods were used when only a few small active regions were present, thus avoiding large active regions biasing the resulting spatial distributions. Since KP/SPM ran for 11 years between periods of maximum activity in cycles 22 and 23 and includes only one minimum, that approach cannot be used with the KP/SPM data. However, the distribution of magnetic flux reported by Wenzler et al. (2004) was relatively uniform across the disk, there were no reported changes in the focus of the instrument and, as a ground-based instrument, it is not subjected to degradation from the harsh environment of above-atmosphere flight. The assumption is therefore made that the reported magnetic fluxes in KP/SPM magnetograms remained sufficiently stable during the overlap period with MDI.

The result of assuming no change in KP/SPM magnetograms is that the difference between the pre- and post-February 1999 factors can be wholly attributed to changes in the MDI instrument alone. From Fig. 4.2, reported magnetic fluxes below 50 G appear to have changed little in the LL and LR, while in the UL and UR they appear to have increased by up to ~20 - 30%. Above 50 G, factors increase along with increasing magnetic flux in all quadrants up to ~200 G. Above 200 G, the factor remains relatively constant, with reported magnetic fluxes increased by between approximately 10 and 35%.

In chapter 3, a position-dependent threshold, $B_{thr,MDI}(i, j)$, is used to remove noise pixels from MDI magnetograms when reconstructing TSI. $B_{thr,MDI}$ is three times the level of the 2008 mean background magnetic flux (i.e. right plot of Fig. 3.4). If the reported magnetic flux in MDI magnetograms had only been affected at the level of the mean background magnetic flux, then the application of this position-dependent threshold should mitigate any changes that occur at these low levels of magnetic flux from affecting the reconstruction, since the mean background magnetic flux is generally higher in 2008 than in 1996. However, as indicated in the results shown in Fig. 4.2, the change in the level of reported magnetic flux in MDI magnetograms after February 1999 affects all levels of magnetic

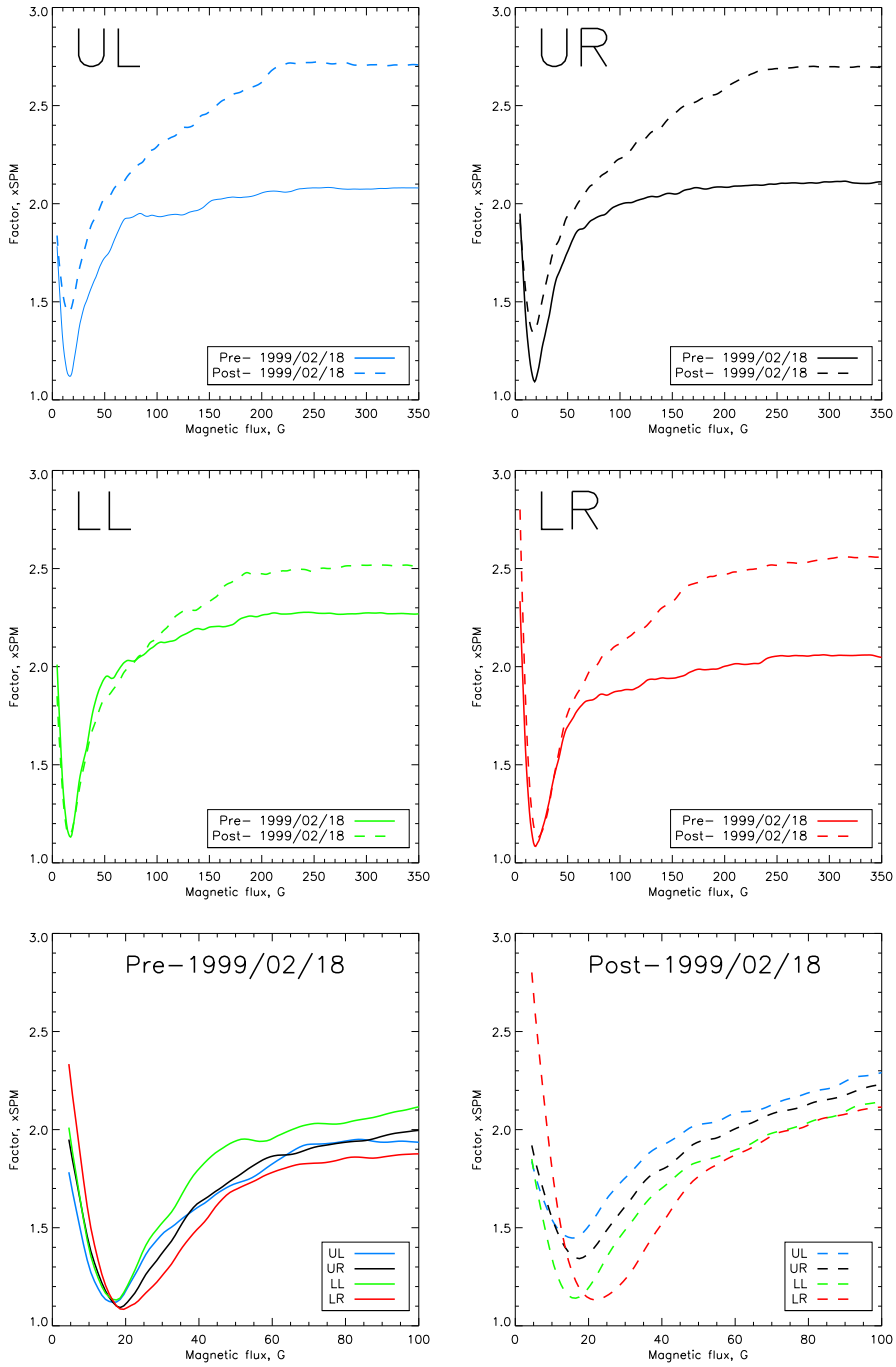


Figure 4.2: Multiplication factors (y-axis), derived from histogram equating, to bring KP/SPM magnetogram fluxes (x-axis) into agreement with MDI (or vice versa as division factors). The upper four plots, clockwise from top left, are from the regions bounded by the dashed lines in Fig. 4.1 in the upper-left (UL), upper-right (UR), lower-right (LR) and lower-left (LL) quadrants. In these four plots, the solid lines are derived from dates on which MDI and KP/SPM magnetograms exist prior to February 1999 and the dashed lines are from dates after this time. The bottom two plots show the pre- and post-1999 factors i.e. the solid lines (left) and dashed lines (right) from the upper four plots; the colours are indicated in the legend and reflect the colours used in the upper four plots; the range of the x-axis has been reduced.

flux in a non-linear, position-dependent way and is therefore the reason why the change between minima cannot be reconstructed using uncorrected MDI magnetograms.

The enhancement of magnetic flux at all positions raises the question of why the change in the mean background magnetic flux is only apparent in the LR quadrant. To understand this better, the two bottom plots in Fig. 4.2 show a comparison between the factors before the focus change in February 1999 (left plot) and after the focus change (right plot). These are the same distributions as plotted in the top four plots. The x-axis shows only magnetic flux up to a level of 100 G. The plots show that there is a minimum value in the distribution of factor values: this is an indicator of the noise level in the MDI magnetograms. Histogram equating assumes that sorted magnetic fluxes in KP/SPM should relate in the same order to sorted magnetic fluxes in MDI. Below the minimum point, the factors increase rapidly because KP/SPM can report lower magnetic fluxes than MDI; the noise level in MDI has been reached and the corresponding low magnetic flux in KP/SPM is reported at the level of the noise in MDI.

The reduced range in the lower two plots of Fig. 4.2 reveals that the minimum value is reached at different levels of magnetic flux in each quadrant and, significantly, the minimum value of the factors in the LR quadrant shows an increase after February 1999 relative to the other three quadrants and relative to the period before February 1999. This is in approximate agreement with the change in the levels of magnetic flux in the four curved regions of Fig. 4.1 and shows that the magnetic flux level at the minimum values of the factors in Fig. 4.2 are related to the level of the mean background magnetic flux in Fig. 3.4.

These plots neither explain why the changes in the minimum level of the factors have occured, nor why change has caused an increase to reported magnetic fluxes at almost all levels of magnetic flux. However, the explanation given in section 3.2.5 remains valid: that a change to a better focus could reduce the point spread function (PSF) and therefore result in higher magnetic fluxes being reported at all levels. This, combined with a slight shift in the observed wavelength as a function of disk position, could help explain why the minimum level of the factors has changed in the LR quadrant in a different way to the rest of the disk.

To determine if the focus change is indeed causing a change in the LR quadrant different to the other quadrants, the following approach is taken. All unsigned magnetic flux pixels confined within the regions bounded by curved, dashed lines of Fig. 4.1 are binned

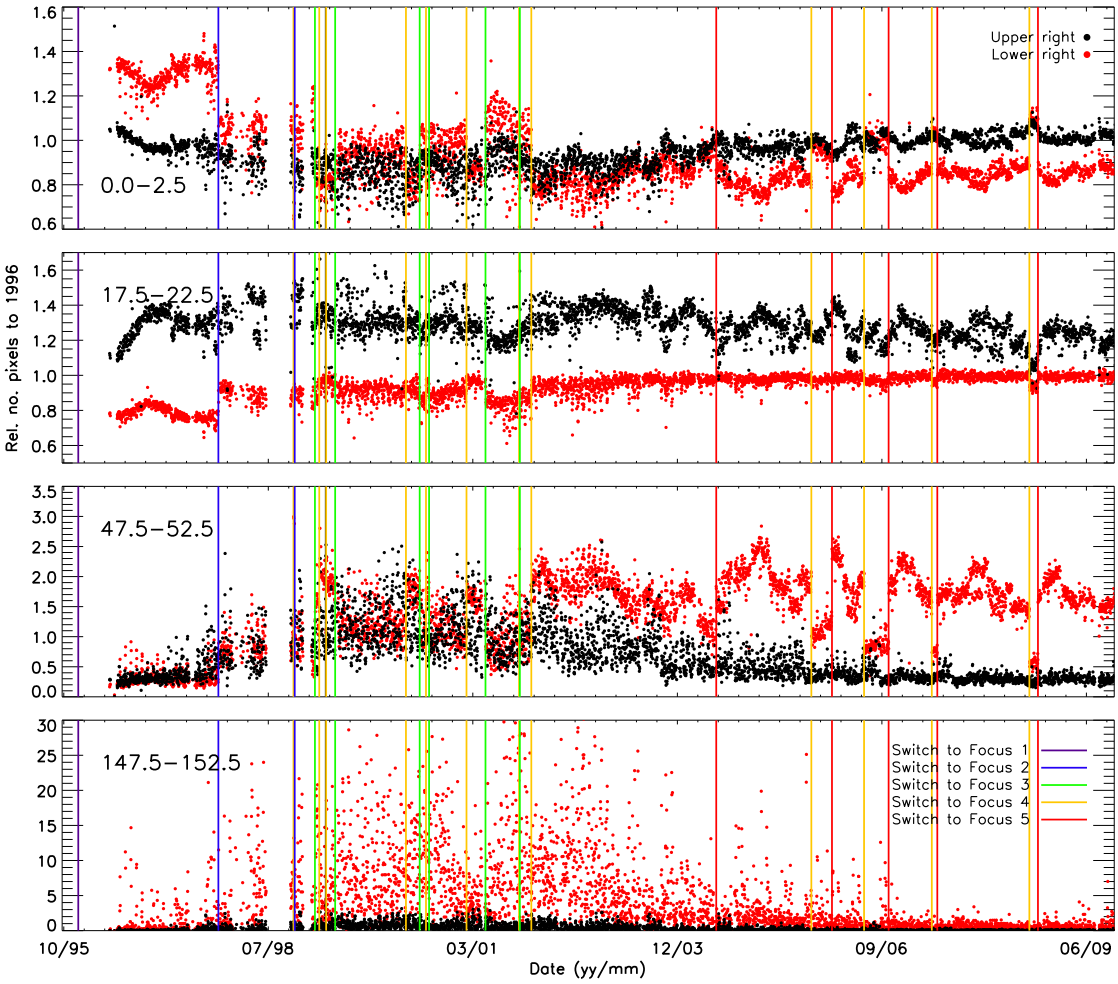


Figure 4.3: The number of pixels (within the lower-right (red) and upper-right (black) curved, dashed regions of Fig. 4.1) relative to the mean number of pixels in specified bins of magnetic flux between May 1996 and May 1997. These are shown for bins of 0.0-2.5 G (top), 17.5-22.5 G (upper middle), 47.5-52.5 G (lower middle) and 147.5-152.5 G (bottom). Dates on which a focus change on the MDI instrument were performed are highlighted by the vertical coloured lines, which are described in the legend of the lowest plot.

into histograms with bins in 5 G intervals, except for the first bin which encompasses the range 0.0-2.5 G. The mean number of pixels in the bounded regions between May 1996 and May 1997 is used to normalise every bin and provide a relative variation of pixels at that level of magnetic flux, although due to a few dates with very large deviations from the bulk of points they do not appear normalised to a value of 1.0. If magnetic flux was recorded at the same level at all positions on the disk, the long-term trend, in each region and at each interval level of magnetic flux, would be similar. Shown in Fig. 4.3 are the relative change in the number of pixels at four 5 G intervals of magnetic flux. Only the results from the UR and LR are shown, for the sake of clarity, as the long and short-term trends are similar in the UL, UR and LL at each level of magnetic flux, except prior to 1999. Overlaid on the plots are vertical coloured lines representing the dates on which the MDI instrument changed to a different focus setting, denoted 1 to 5 and coloured as in the key at the upper right corner of the bottom plot. Focus changes were performed on an as-required basis after 1999 to rectify a focus drift and thereby maintain a relatively constant level of focus.

In the upper plot, the relative number of unsigned magnetic flux pixels between 0.0 and 2.5 G is shown. This low level of magnetic flux is far below the noise level of ~13 - 18 G and so can be said to indicative of the change in the number of noise pixels. The trend at this level of magnetic flux is a decline in both regions between 1996 and 2001 as solar activity increases, after which an increase is seen, though the relative number of pixels never returns to the 1996-7 level in the LR region. The decline in the 0.0 to 2.5 G interval can be explained from the fact that there are only a finite number of pixels on the disk, so as solar activity increases, ever higher proportions of the solar disk are covered by strong magnetic fields and, therefore, fewer pixels will report low magnetic flux. This explanation is supported by the almost mirrored short-term trend in the high-flux interval of the lower plots.

In the lower three plots, intervals of higher level magnetic flux between 17.5 and 22.5 G, 47.5 and 52.5 G, and 147.5 and 152.5 G are shown. These levels of magnetic flux are, respectively, the level of the noise, just above three times the noise level (i.e., at the level of the position-dependent threshold, $B_{thr,MDI}$) and an arbitrary higher level of magnetic flux. All three display the opposite trend in time with respect to the top plot. This opposite trend is indeed seen in all 5 G bins above 17.5 G, but with the discontinuities clearest at around the 50 G level.

In each of the three lower plots, the trends exhibited from the UR and LR regions are

different. The LR appears to display a jump around February 1999 and the level in 2008 never recovers to the same level as in 1996; the UR region, by contrast, does not appear to display such a large jump around February 1999. There is more variability in the UR in the 17.5-22.5 G interval after 2002, unassociated with focus changes, which is not present in the LR, but the opposite is true of the 47.5-52.5 G level. The highest level plot, 147.5-152.5 G shows that the magnetic flux in the LR is registered very differently over the cycle with respect to the UR. Together, these plots suggest that the effects of early degradation, prior to 1998 (see below), and focus changes affect different regions of the disk at different levels of magnetic flux. The consequence on the reconstruction is therefore complicated.

Jumps in the relative number of pixels for the LR region occur on dates when focus changes occur. This is most clear at the switch to focus 2 (1997 November 3, first blue line), and the switches between focus 4 and 5 after 2004 July 3, though jumps at almost every focus change can be seen for the LR region. The first switch to focus 3[1] is correlated with a distinct and opposing jump in the number of pixels in the 0-2.5 and 47.5-52.5 G intervals. This focus change was performed to bring the instrument into focus from a defocused setting and results in the expected enhancement of reported magnetic flux with a reduced PSF.

In general, the UR displays fewer artifacts than the LR, except for some wave-like behaviour after 2001 and an initially rapid change in the early period of operation prior to 1998. In contrast, the LR displays relatively large jumps at focus changes and quite alot of wave-like variability over most of the 1996 to 2008 period. The jumps are not as strong in the UR region and are more difficult to determine as there appear to be jumps not associated with focus changes.

In addition to the deliberate focus changes initiated by the MDI instrument team, the degradation of the front window, through darkening and thermal effects, caused a more gradual change in the focus (private communication, Rock Bush). The effect of this degradation manifests itself in the first 18 months of Fig. 4.3 as an early wiggle before the switch to focus 2. This wiggle is particularly distinct in the LR region at 0-2.5 G and, with the opposite trend, in the UR at 17.5-22.5 G, though it can be seen in both quadrants at both these levels. The degradation trend also appears in 47.5-52.5 G plot, though the change is small relative to the rest of the time series.

[1]The first switch to focus 3 occured on 1999 February 18, the date from which MDI magnetograms are used in SATIRE-S_{Ind}.

Degradation is generally only significant for space-based instruments in the early period of operation. It certainly appears that in general the quickest changes did occur during the first 18 months of operation. On a long-term basis, degradation may have continued at a reduced rate, but it is not really possible to identify degradation-only changes after 1999 in Fig. 4.3. There are clear changes in the LR and UR regions that occur between focus changes (e.g. in 2003 and 2004) and may be the result of solar activity or instrumental effects. These non-focus changes appear to happen in both quadrants at the same time, and at different bin levels, so are likely to be related, but manifest at different levels of magnetic flux in the different regions because the same magnetic flux is registered differently in these regions.

Figure 4.3 provides some evidence that degradation has had an effect on the pre-1999 MDI magnetograms, in addition to the deliberate focus changes from the MDI team. Overall, focus changes and degradation prior to 1999 February 18 appear to have had a significant effect on reported magnetic fluxes in all regions of the disk, relative to after this date. This latter period then provides a relatively stable set of data with which to reconstruct long-term changes in solar irradiance.

4.2 A simple approach to correcting MDI magnetograms

It is difficult to account for the instrument degradation without modelling the internal systems and external influences fully. Degradation and focus changes appear to be important in different ways and separating the effects of each is not within the scope of this work. However, in principle some of the effects of focus changes can be accounted for by making a couple of assumptions. This will allow a rough check to see if correcting for a focus change will have the expected effect of raising the 1996 minimum when reconstructing using on MDI data.

If, in addition to the assumption that KP/SPM has been stable throughout the overlap period with MDI, the assumption is made that the focus change is solely responsible for changes in MDI, then a partial correction to the MDI magnetograms can be made. The assumption that the focus change is the only effect is incorrect, but it is adopted for the purposes of this correction.

Here, the same approach of histogram equating is taken as used to produce the factors

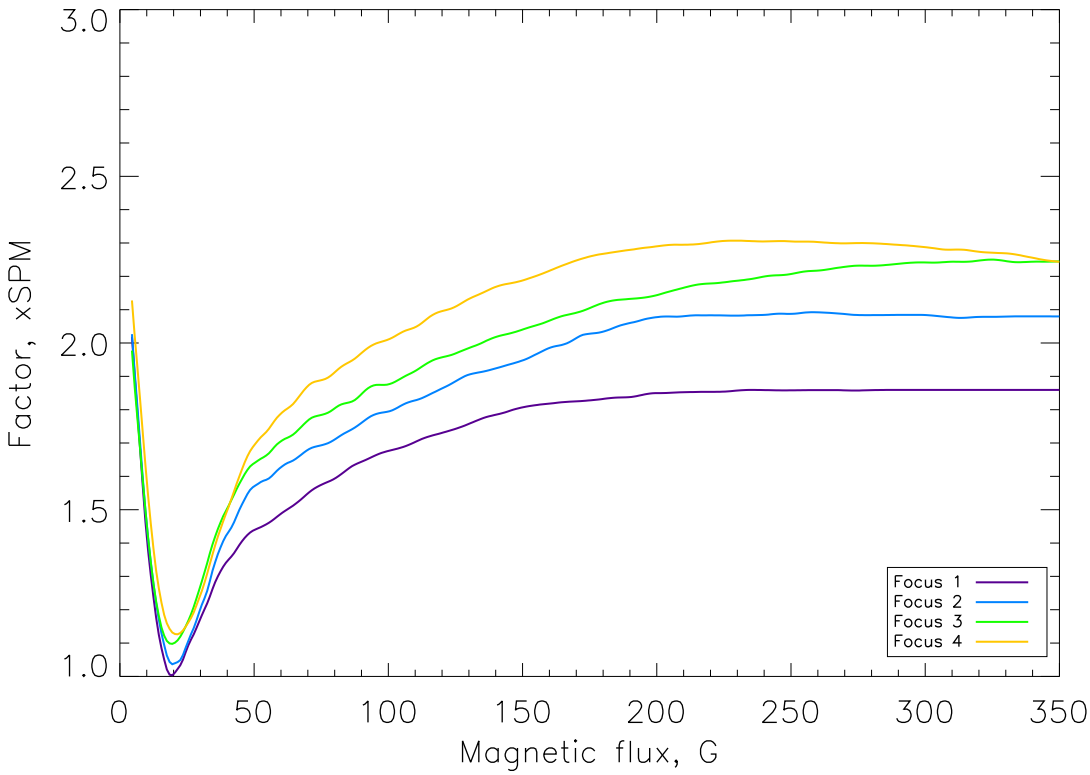


Figure 4.4: Histogram equating between MDI and KP/SPM for all pixels with a limb angle ≥ 0.1 for all dates that MDI and KP/SPM magnetograms exist when MDI used focus setting 1 (purple), 2 (blue), 3 (green) and 4 (yellow). The y-axis indicates the factor required to multiply magnetic flux in KP/SPM magnetograms to bring it into agreement with MDI magnetograms.

shown in Fig. 4.2. In Fig. 4.4, factors are plotted for each focus setting, relative to KP/SPM, by considering only dates in MDI when that focus was set. These are now full disk factors, considering all pixels within a limb angle of 0.1. There are a variable number of dates available to calculate these factors: at focus setting 1 (purple) there were 189 dates available; focus setting 2 (blue) had 91; focus setting 3, 447; and focus setting 4, 450. There was no overlap with MDI during focus setting 5. The lack of an overlap between KP/SPM and MDI magnetograms observed with focus setting 5 will mean that a full correction to the MDI magnetogram time series will not be possible. Figure 4.4 shows that for higher focus settings, selected later in time, the reported magnetic flux in MDI magnetograms increases.

If KP/SPM has not changed over time then these factors allow a way to make MDI magnetograms with different focus settings agree and produce a corrected reconstruction. For each 1 G level, all magnetic flux in magnetograms with focus settings 1, 2 and 3 are converted into the fourth setting using the ratio of the factors shown in Fig. 4.4. Then, these corrected magnetograms are used to produce a reconstruction using the SATIRE-S model. Given that each setting, in numerical order, has higher factors than the last, it would be expected that a correction to MDI magnetograms of lower settings should increase the reconstructed irradiance relative to the later settings and would therefore increase the value of TSI during the first half of the solar cycle.

In Fig. 4.5, the smoothed corrected reconstruction using MDI magnetograms only (orange) is shown in comparison to the smoothed uncorrected reconstruction (black) and smoothed SATIRE-S_{Ind}, which uses a combination of KP/SPM and MDI data (blue). Smoothings are performed to retain clarity. Data were only considered where they existed in all datasets and gaps exist in the smoothed curves when data gaps exceed 27 days. All three reconstructions are normalised to the mean value of SORCE/TIM TSI over the six month period centred on December 2008. After approximately 2002, all reconstructions agree with each other because focus settings 1, 2 and 3 are corrected to agree with setting 4 only and this period uses focus settings 4 and 5 only.

As expected, the TSI calculated using corrected magnetograms is higher at all times relative to the uncorrected reconstruction. The mean value over the first six months of the corrected reconstruction is 0.07 Wm^{-2} higher than that of the uncorrected reconstruction, which accounts for approximately half of the difference between the uncorrected and the SATIRE-S_{Ind} reconstructions. By the time contact is lost in 1998, the correction has brought the reconstruction to almost the same level as SATIRE-S_{Ind}. After contact is

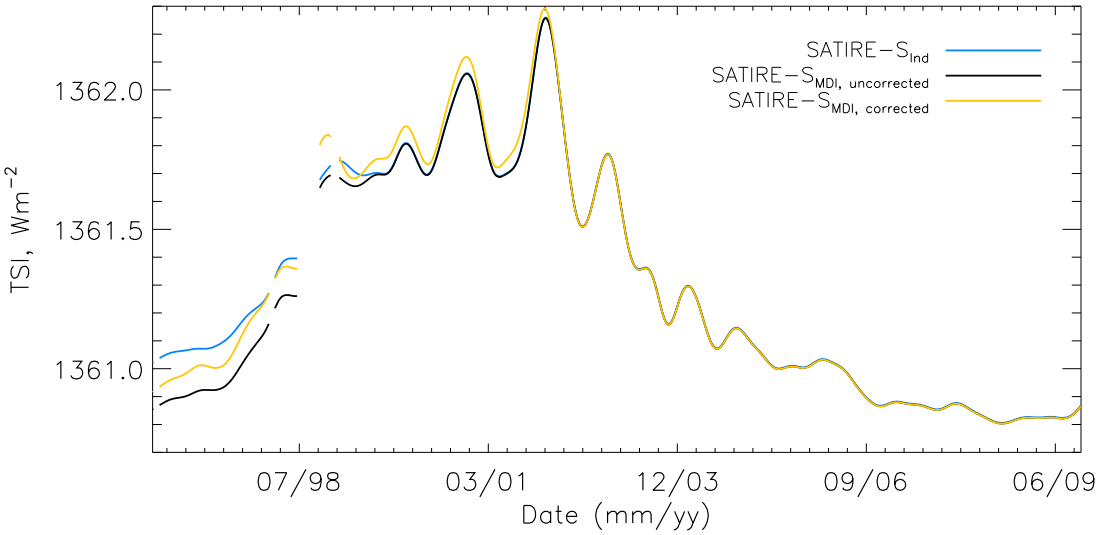


Figure 4.5: Smoothed reconstructions of TSI between 1996 and 2008. Shown are reconstructions using original, uncorrected MDI magnetograms only (black), the SATIRE-S_{Ind} reconstruction from the previous chapter (blue) and a reconstruction using corrected magnetograms (orange), as described in the main text.

regained the corrected reconstruction is slightly higher than SATIRE-S_{Ind}, but the difference between SATIRE-S_{Ind} and the corrected reconstruction after February 1999 is less than between them prior to loss of contact. It is unclear why there is such a large increase in reconstructed irradiance during the brief period of contact with SoHO around the end of 1998.

An important caveat is the inability to correct magnetograms relative to focus setting 5 so that all MDI magnetograms have consistent levels of magnetic flux. If this could be done, two significant outcomes would result. First, Fig. 4.3 suggests that the switch from focus setting 4 to 5 may increase the reported magnetic flux, at least in the LR quadrant. If all magnetograms were corrected relative to setting 5, the entire reconstructed period prior to July 2004 (when the first switch to setting 5 was made) would be further elevated relative to the minimum of 2008, further reducing the difference between the corrected reconstruction and SATIRE-S_{Ind} in 1996. A correction relative to focus setting 5 should result in a higher B_{sat} being required to reduce the cycle amplitude, which would compensate the increased irradiance at cycle maximum in the corrected Fig. 4.5. Crucially, it would not significantly affect the minimum value in 1996 since the effect on lower fluxes from a change in B_{sat} is less than on higher fluxes. Unfortunately, the influence of focus setting 5 cannot be estimated.

It is worth pointing out that if we were to use all the corrected MDI magnetograms (including those from before 1999) to produce a new version of SATIRE-S_{Ind} (see chapter 3), B_{sat} for KP/SPM would need to be redetermined. If KP/SPM is fitted to the corrected MDI reconstruction, following the same method of section 3.3.1, the value of $B_{sat,SPM}$ is 308 G, compared to the value of 330 G used SATIRE-S_{Ind}. This lower setting of $B_{sat,SPM}$ produces a higher cycle amplitude in previous cycles and a lower cycle minimum in 1996. However, this is just within the range of uncertainty prescribed for SATIRE-S_{Ind}.

The correction before the the loss of contact with SoHO was only partially successful. This is most likely due to the early period of degradation, prior to the change to focus setting 2. Before 1999, the smoothed uncorrected and the SATIRE-S_{Ind} reconstructions in 1996 and 1997 are very similar (Fig. 4.5). After the correction, the trend displays stronger gradients over shorter periods and is reminiscent of the degradation effects seen in Fig. 4.3 during the same period. This, then, provides evidence that changes resulting from instrument degradation have also, in addition to focus changes, had an effect on MDI magnetograms prior February 1999.

Less important problems with this correction approach relate to not correcting for the spatial variation in reported magnetic flux with time. It is clear from Figs 4.1 and 4.2 that for a proper correction the spatial variation should be taken into account as well. It is likely that correcting for spatial variation would have more effect on rotational variability than the long-term trends, as active regions cross the disk and pass through regions that report magnetic flux differently. The factors used in the correction were derived from positions relative to the centre of the solar disk, but the variation in reported magnetic flux will be determined from the position relative to the centre of the detector. If the variation across the detector is relatively smooth then this will not have a significant effect, as the Sun usually sits in the middle of the detector view, but it will mean that the factors relating KP/SPM and MDI will be slightly less accurate.

4.3 Comparison of KP/SPM and MDI reconstructions with the PMOD Composite

The previous subsection assumed that KP/SPM had no long-term variations and a basic attempt at a correction to MDI suggested a reduced change between the cycle minima of

Table 4.1: Best-fit comparison of PMOD with SATIRE-S using either MDI or KP/SPM continuum intensity image and magnetogram data.

Daily dataset comparison						
Series	B_{sat}, G	r_c	Slope	Start date	End date	No. points
SATIRE-S_{MDI}	463	0.969	1.000	1999/02/18	2003/09/21	1551
SATIRE-S_{SPM}	354	0.944	1.000	1999/02/18	2003/09/21	965
SATIRE-$S_{SPM,w/96-99}$	326	0.951	1.001	1992/11/21	2003/09/21	1996
SATIRE-$S_{SPM,w/o96-99}$	323	0.948	1.001	1992/11/21	2003/09/21	1532
SATIRE-S_{MDI}	407	0.975	1.002	2003/02/25	2009/10/31	2123
SATIRE-S_{MDI}	423	0.983	1.001	1999/02/18	2009/10/31	3482

1996 and 2008 relative to SATIRE-S_{Ind}. Although the correction was acknowledged to be inadequate, the result throws some doubt on the approach of attaining a unity regression between MDI and KP/SPM in order to construct SATIRE-S_{Ind}. This is because changes in MDI magnetograms during the overlap period may give an inaccurate gradient with a unity fit. It is also possible that KP/SPM had time-dependent changes. It is therefore worth considering both KP/SPM and MDI separately with respect to a third dataset. For this the PMOD composite is used, which was shown to be the best overall TSI composite available, though one must bear in mind that this conclusion was reached using SATIRE-S_{Ind}. Separate comparisons of KP/SPM and MDI reconstructions with PMOD will allow an assessment of whether using regression fitting between KP/SPM and MDI is reasonable.

The approach taken to fit datasets here is the same as in section 3.3.1, which uses a process of finding a unity regression between the datasets being compared to determine the best fit. The results of the fits are summarised in Table 4.1.

The free parameter for the MDI reconstruction in SATIRE-S_{Ind} was determined using SORCE/TIM. Then a reconstruction using KP/SPM data was fitted to the, now fixed, MDI reconstruction for an overlap period between 1999 and 2003. The free parameter for the MDI fit to SORCE/TIM and for KP/SPM to MDI were $B_{sat,MDI}$ = 443 G and $B_{sat,SPM}$ = 330 G, respectively. We use these values as a reference for the discussion in this section.

In Tab. 4.1, rows one and two show the fits between PMOD and both SATIRE-S_{MDI} and SATIRE-S_{SPM} for the same overlap period as for MDI and KP/SPM datasets in SATIRE-S_{Ind} . When fitting to PMOD $B_{sat,SPM}$ = 354G, 24 G higher than for the equivalent fit with SORCE/TIM as a basis. This is consistent with the fit of SATIRE-S_{MDI} to PMOD, which

increases by the similar value of 20 G. This check for consistency is important because the period of overlap only covers a small part of the full range of variability in the solar cycle and would be expected to be more affected by the rotational variability than a fit over a period with a large range of cycle variability such as a full ascending or descending period of the cycle.

Rows three and four are a comparison over the full reconstruction period of SATIRE-S_{SPM} with PMOD, almost a full cycle between 1992 and 2003. This now includes the majority of the cycle variability. The first fit between SATIRE-S_{SPM} and PMOD over this period includes all dates when data from both are available. This reconstruction is shown in Fig. 4.6 as a daily (purple) and smoothed (dark green) time series along with PMOD (orange, daily and red, smoothed) and SATIRE-S_{MDI} (light blue, daily and blue, smoothed) with $B_{sat,MDI}$ = 423 G (see below). The second fit of SATIRE-S_{SPM} to PMOD (line four of the table) ignores dates between 1996 May 20 and 1999 February 18 during which the PMOD composite may not have fully accounted for instrument degradation (see section 3.5 for further discussion). The values of $B_{sat,SPM}$, 323 and 326 G, for the overlap with and without the 1996-1999 period respectively, agree well with the value of $B_{sat,SPM}$ needed to achieve unity regression with SATIRE-S_{MDI} in SATIRE-S_{Ind}, i.e. $B_{sat,SPM}$ = 330 G. Although this result was gained by fitting PMOD, it supports the value determined for the SATIRE-S_{Ind} reconstruction. The result is also more robust as it covers a larger range of cyclical variability.

The large difference in the value of $B_{sat,SPM}$ when fitted to PMOD over 1992-2003 ($B_{sat,SPM}$ = 323 G) and 1999-2003 ($B_{sat,SPM}$ = 354 G) suggests that the 1999-2003 period may not allow a robust connection between MDI and KP/SPM in SATIRE-S_{Ind} and might lead to a value of B_{sat} that is too low. However, as was discussed in chapter 2, SATIRE-S reconstructs rotational variability that is slightly larger than that of the observations, so by determining $B_{sat,SPM}$ over the longer period, the cycle variability is better accounted for. This further supports the value of $B_{sat,SPM}$ in SATIRE-S_{Ind}.

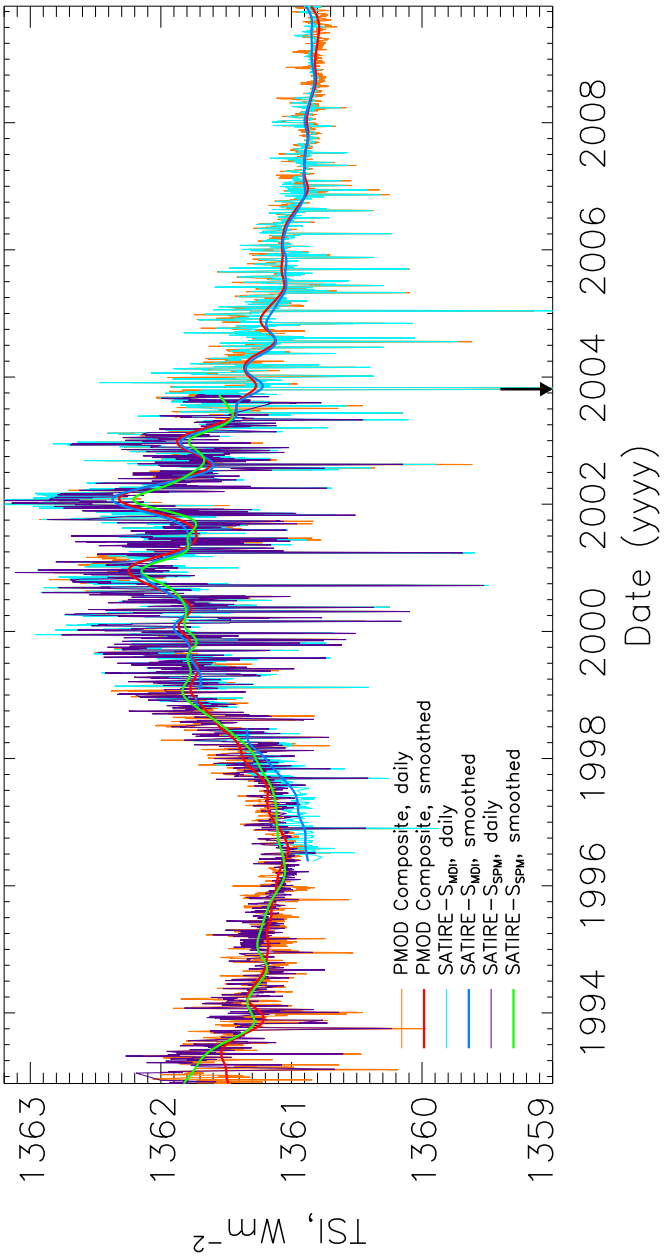


Figure 4.6: Daily (thin lines) and smoothed (thick) time series between 1992 and 2009. The PMOD composite is shown in orange and red, for daily and smoothed time series respectively, SATIRE-S_{MDI} is shown in light and dark blue and SATIRE-S_{SPM} in purple and green. Gaps in SATIRE-S_{MDI} occur when data are missing for at least 27 consecutive days. The arrow indicates a spot transit that caused TSI to drop to ~1358 Wm^{-2}. PMOD is normalised to SORCE/TIM at December 2008. SATIRE-S_{MDI} and SATIRE-S_{SPM} are normalised to PMOD over their respective overlap periods.

For completeness, the last two rows of the table give the best unity fits of SATIRE-S_{MDI} to the PMOD composite. Two periods are compared: 2003-2009, the period used to determine $B_{sat,MDI}$ when fitting SATIRE-S_{MDI} to SORCE/TIM for the construction of SATIRE-S_{Ind}; and 1999-2009, the longest continuous period of SATIRE-S_{MDI} of relative stability following the SoHO telemetry gaps of 1998/9. Both values of $B_{sat,MDI}$, 407 and 423 G, are lower than the fit to SORCE/TIM, the former of which is directly comparable to SORCE/TIM. The significantly lower result of 407 G is not surprising as SORCE/TIM has a shallower gradient than PMOD during the declining phase of cycle 23. The higher value of $B_{sat,MDI}$ = 423 G for the longer period reflects the inclusion of the cycle maximum period, 1999-2003, which had a much larger value of $B_{sat,MDI}$ = 463 G for SATIRE-S_{MDI} with respect to PMOD (first row of the table). The reconstruction with $B_{sat,MDI}$ = 423 G is shown in Fig. 4.6. If SATIRE-S_{SPM} had been attached to SATIRE-S_{MDI} with either $B_{sat,MDI}$ = 407 or 423 G (as implied by fits to PMOD rather than TIM), the reconstructed cycle amplitude would have increased by $\sim$0.1 Wm^{-2}. The difference in TSI between 1996 and 2008, however, changes only marginally with respect to SATIRE-S_{Ind}, decreasing by approximately 0.01 Wm^{-2}. This means the change between cycle 23 minima is relatively stable and robust using the method employed here.

4.4 Conclusions

In this chapter, the temporal variation of MDI magnetograms and the resulting consequences in the model reconstruction have been investigated. After discussing the spatial and temporal variation in different quadrants of the solar disk, a reconstruction attempting to correct for focus-induced changes was performed. This was followed by checking the stability of the unity regression fit between SATIRE-S_{MDI} and SATIRE-S_{SPM}, over the period 1999-2003, by comparing these reconstructions with the PMOD composite over longer time periods.

There were three main results from the comparison of SATIRE-S_{MDI} and SATIRE-S_{SPM} with the PMOD composite using different time periods. Firstly, varying $B_{sat,MDI}$ alters $B_{sat,SPM}$ consistently in the same direction and with roughly the same magnitude of change. Second, it was found that the change between minima is relatively invariant to changes in the value of $B_{sat,MDI}$ when SATIRE-S_{MDI} and SATIRE-S_{SPM} are combined in this way. Finally, the best fit value of the free parameter for SATIRE-S_{SPM} when fitted to the PMOD composite agrees with the value previously determined in section 3.3.1. These results

support the approach taken in chapter 3 to construct a long-term reconstruction by fitting datasets together using a unity regression fit for the relatively short period of 1999-2003.

From the analysis of the original and corrected MDI magnetograms, there is some evidence to suggest that deliberate or degradation-induced focus changes are responsible for the majority of the change in the response of the MDI instrument. The consequence is that this has led to a change in the reported magnetic fluxes in MDI magnetograms. In particular, we found that magnetic fluxes reported after the loss of contact with SoHO were generally higher. Using MDI magnetograms without accounting for changes in the instrument underestimates the change between the minima during cycle 23.

While it is difficult to produce an adequate correction to the MDI magnetograms, the settings for a KP/SPM reconstruction, when fitted to PMOD or the modified MDI magnetograms, remain within the uncertainty range for the free parameter as determined in the previous chapter.

Chapter 5

Comparison between SORCE/SIM and the SATIRE model

Overview

In this chapter[1], the spectral irradiance SATIRE-S reconstruction, between 200 and 1630 nm, as prepared in chapter 3, are compared with data from the Spectral Irradiance Monitor (SIM) instrument, on SORCE, over the period 2004 April 21 to 2009 October 31. I discuss the spectral changes in flux, as well as the integrated change, and the rotational and long-term trends during this period of decline from moderate activity to the recent solar minimum of 2008.

The rotational variability of SORCE/SIM is well reproduced by SATIRE-S, especially between 400 and 1200 nm. The visible and IR regions have previously not been observed on a continuous basis to monitor cyclical trends. The magnitude of change in the long-term trends is many times larger in SIM than SATIRE-S at almost all wavelengths, especially in the UV where previous observations exist which do not show the same cycle variability. Cycle trends in SIM oppose SATIRE in the visible between 500 and 700 nm and again between 1000 and 1200 nm. Other remaining issues with both SIM data and the identified limits of the model are also discussed in this chapter.

[1]This chapter formed part of a paper published in Astronomy & Astrophysics journal entitled 'Solar irradiance variability: A six-year comparison between SORCE observations and the SATIRE model' (Ball et al., 2011). The paper has been updated and uses a corrected noise threshold and an improved fitting procedure. The following people contributed to the work and were named as co-authors: Yvonne C. Unruh, Natalie A. Krivova, Sami Solanki and Jerald W. Harder.

5.1 SORCE/SIM

The Spectral Irradiance Monitor (SIM) was launched in January 2003 on board the SOlar Radiation and Climate Experiment (SORCE) satellite (Rottman, 2005). The comparisons in this chapter are made over the period that SIM and SATIRE-S data are available, i.e. from 2004 April 21 to 2009 October 31, with data available for 1764 days of this 2019-day period. This period is long enough to investigate variations on rotational and cycle-length timescales. After October 2009, large thermal effects in SIM data and glitches make it difficult to discern trends in the ascending phase of cycle 24, so no consideration is given to this period.

SIM, along with SCIAMACHY (Skupin et al., 2005), is one of the first instruments purpose-built to continually monitor solar UV, visible and IR radiation. The precision and measurement drift stability of SIM is at the 100 ppm level and it has an absolute accuracy of 2% (Harder et al., 2010). We consider SIM data observed by three of the four photodiode detectors: the UV photodiode observes 200-308 nm, 'vis1' covers 310-1000 nm and the IR 994-1655 nm, though we only use data up to 1630 nm. The integrated-flux over this range accounts for ~90% of TSI. Relative accuracy is wavelength dependent: it is worst with 0.5% at 310 nm, but above 500 nm it is better than 300 ppm (Harder et al., 2005a). Here we employ version 17, level 2 data with twice daily observations. For a space-based satellite, degradation of the instrument will occur on a long-term basis and SIM uses an onboard correction system that compares two spectrometers, SIM A and SIM B, to correct for long-term systematic degradation of the instrument (see appendix in Harder et al. (2009) for more detail).

5.1.1 Data preparation

Before comparing the data to SATIRE-S we process it as follows. First, dates when SIM data are known to contain instrument glitches are removed from the dataset. At some wavelengths severe glitches remain. Binomial smoothing (Marchand and Marmet, 1983) is used to produce an effective mean by smoothing the data, and dates with flux values outside 7σ of the instrument noise from the mean are identified. Dates with glitches in more than 200 wavelength elements and wavelengths with more than 300 days of glitches are identified and removed. Remaining individual elements are interpolated over from dates on either side. The process is then repeated at 5σ to capture smaller but similarly erroneous outliers, more easily identified now the larger ones have been removed.

These processes remove 0.6% of the 3623 time series elements and 1.6% of the 1829 SIM wavelength elements. Finally, all outliers above or below 3σ are interpolated over, so that ~7.0% of the remaining dataset are interpolated over. At this stage no smoothing of the data has occured, other than interpolation over glitches.

SIM oversamples by 6 wavelength elements per resolution element, so a function simulating the resolution is applied to the spectrum. This reduces the number of spectral elements to 226. The same window is applied to SATIRE-S with the centre wavelength matching SIM. Binomial smoothing with a [1,2,1] filter is then performed on each remaining SIM time series to dampen random fluctuations.

Finally, the SATIRE-S and SIM data sets need to be put on the same time-grid. As the SATIRE-S dataset has a lower cadence it is used as a basis. SIM data are averaged over 24 hour periods and interpolated onto the SATIRE-S time-grid over the full six-year period available.

5.2 Comparison between SIM and SATIRE-S

We now perform an analysis of SIM and SATIRE-S where lightcurves of the data over various wavelength ranges are presented. These are produced by integrating over the desired wavelengths and are broken down into original, smoothed and detrended lightcurves.

5.2.1 Integrated Total Solar Irradiance

Before breaking the solar spectrum down into small wavelength bands, we first compare 'pseudo-TSI' measurements of SIM and SATIRE-S with TIM. When the SIM data are integrated over the 200-1630 nm range ('Int-SIM') they account for ~90% of TSI with a flux level of ~1218 Wm^{-2}. This is in reasonable agreement with SATIRE-S integrated over this region ('Int-SATIRE') at 1204 Wm^{-2}, though SATIRE-S is not designed to reproduce the correct absolute values. The remaining ~140 Wm^{-2} are mainly emitted in the IR.

To compare Int-SIM and Int-SATIRE with TIM, these two datasets need to be multiplied (normalised) to TIM or shifted by a constant. Normalisation should provide an upper limit for the variability as it assumes that the same additional variability should be attributed

Table 5.1: Comparison of SORCE/TIM, Int-SIM and Int-SATIRE over the time period 2004 April 21 to 2009 October 31, with Int-SIM and Int-SATIRE normalised or shifted by addition to SORCE/TIM.

Normalisation by Multiplication				
Set 1	Set 2	Multip. factor	$r[r^2]$	Slope
TIM	Int-SIM	1.117	0.829 [0.688]	1.233
TIM	Int-SATIRE	1.130	0.968 [0.937]	1.096
Int-SIM	Int-SATIRE	-	0.813 [0.660]	0.885
Shift by Addition				
Set 1	Set 2	Add. [Wm^{-2}]	$r[r^2]$	Slope
TIM	Int-SIM	142.55	0.829 [0.688]	1.104
TIM	Int-SATIRE	156.42	0.968 [0.937]	0.970
Int-SIM	Int-SATIRE	-	0.813 [0.660]	0.875

to TSI from wavelengths outside the 200-1630 nm integrated region, while addition by a constant will assume no overall variability outside the 200-1630 nm region. In reality there is likely to be some additional variability and offsetting by different wavelengths. In Table 5.1 the effect of normalisation by multiplication or addition of a constant is shown with values derived from normalising or shifting to the mean of a one-year period centred on December 2008. The results indicate that the addition of a constant provides a significantly better result, i.e. a slope closer to 1.00 and therefore a better approximation for the missing variability. This is supported by the measurements from the SIM electrical substitution radiometer (ESR) instrument, covering 1630-2400 nm, which shows no long-term trend over the period to solar minimum ($\sim$7x10^{-4} Wm^{-2}) and is in line with the ever decreasing temperature sensitivity of the Planck function with increasing wavelength (Solanki and Unruh, 1998). Since the addition of a constant better represents the overall missing variability than normalisation, comparisons hereafter are made considering an addition of a constant to align flux levels with TIM.

In Table 5.2 we present correlation coefficients and regression slopes for the three datasets separately for the original, detrended and smoothed data. Along with the lightcurve plots of Fig. 5.1, these provide an additional comparison of short- and long-term behaviour.

Unruh et al. (2008) found that the short-term period May - July 2004 showed a high correlation of 0.97 (r^2 = 0.94) between Int-SIM and TIM, indicating very good short-term agreement. This period displays very little long-term variability and the majority of the correlation is captured by the short-term rotational variation from active region passages.

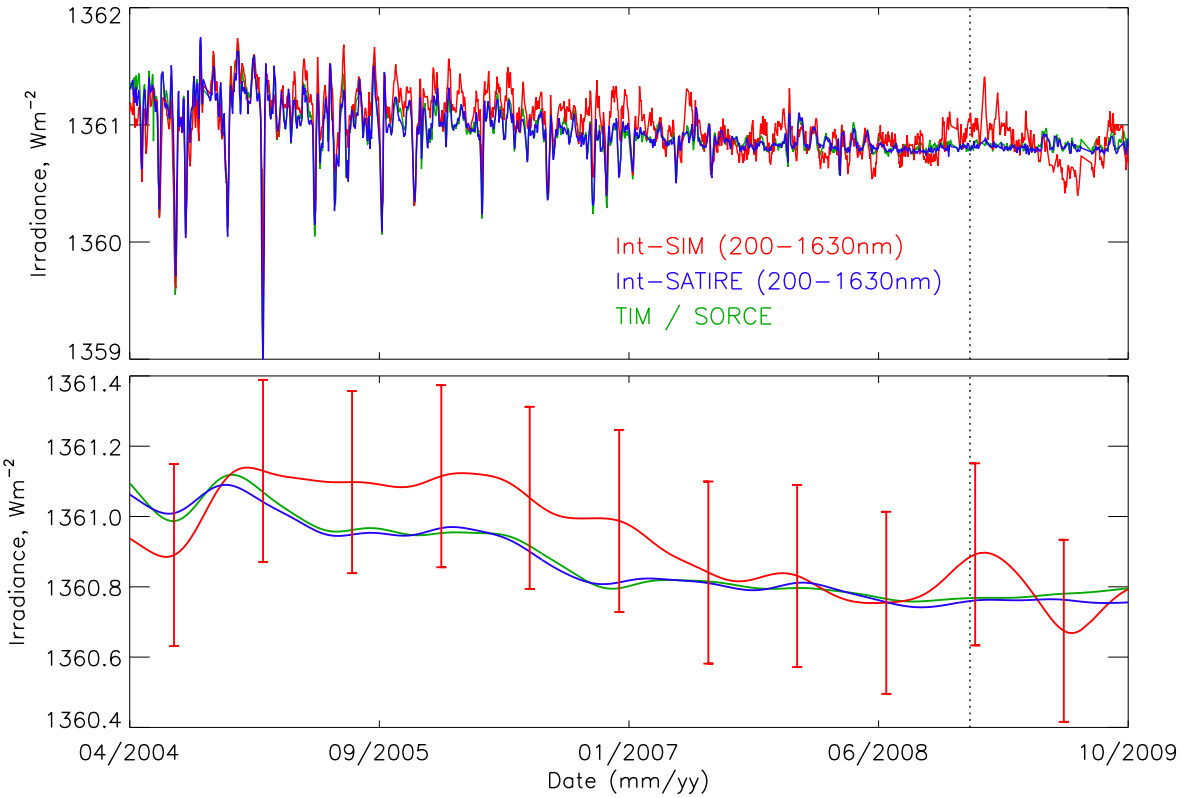


Figure 5.1: Lightcurves of Int-SIM (red, 200-1630 nm), Int-SATIRE (blue, 200-1630 nm) and SORCE/TIM (green) for (top) original data and (bottom) smoothed or long-term data. A constant shift (see Table 5.1) has been added to Int-SIM and Int-SATIRE. Note that the y-axes are on different scales: the scale for the bottom panel has been expanded by a factor of three compared to the top plot. Error bars represent one standard deviation in the long-term stability of Int-SIM and are 0.259 Wm^{-2} or 212 ppm. The vertical dotted line represents the time of solar minimum, December 2008.

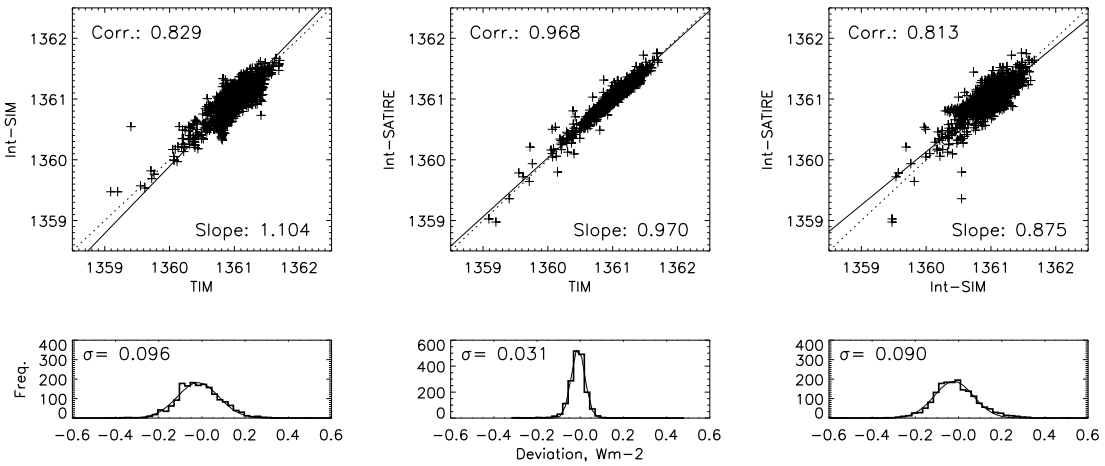


Figure 5.2: Regression plots for left: Int-SIM vs. SORCE/TIM, centre: SORCE/TIM vs. Int-SATIRE and right: Int-SIM vs Int-SATIRE. Here Int-SIM and Int-SATIRE have had an additive constant applied to bring their irradiance levels into line with TIM (see Table 5.1). Axes in the regression plots all have units of Wm^{-2}. Below the regression plots are histograms of the spread about the regression lines. A Gaussian is fit to the histograms and the standard deviation of the Gaussian is presented in the top left.

Table 5.2: Comparison between Int-SIM, Int-SATIRE and TIM for original, detrended and smoothed data for the period as in Tab. 5.1.

Original					
Set 1	Set 2	$r[r^2]$	Slope_{FITEXY}	$\text{Slope}_{regress}$	$\text{Slope}_{regress}$ error
TIM	Int-SIM	0.829 [0.688]	1.104	0.912	0.100
TIM	Int-SATIRE	0.968 [0.937]	0.970	0.939	0.100
Int-SIM	Int-SATIRE	0.813 [0.660]	0.875	-	-
Detrended / Short-term					
Set 1	Set 2	$r[r^2]$	Slope_{FITEXY}		
TIM	Int-SIM	0.861 [0.742]	1.025	-	-
TIM	Int-SATIRE	0.963 [0.927]	0.969	-	-
Int-SIM	Int-SATIRE	0.832 [0.693]	0.944	-	-
Smoothed / Long-term					
Set 1	Set 2	$r[r^2]$	Slope_{FITEXY}		
TIM	Int-SIM	0.791 [0.626]	1.343	-	-
TIM	Int-SATIRE	0.994 [0.988]	0.992	-	-
Int-SIM	Int-SATIRE	0.804 [0.647]	0.742	-	-

In Table 5.2 we find that the correlation between Int-SIM and TIM of the original data is reduced to 0.83 ($r^2 = 0.69$) over the six-year period compared to the three month period considered by Unruh et al. (2008). This is the result of the variability dropping below the instrument noise level during the recent extended minimum (see below).

Fig. 5.1 shows both the original (top) and smoothed (bottom) Int-SIM (red) and Int-SATIRE (blue) shifted by a constant as in Table 5.1 to bring them in line with TIM (green). The y-axis of the smoothed plot is smaller than the original data to show the long-term changes more clearly. The error bars of 0.26 Wm^{-2} (212 ppm) in the lower panel represent a one-standard deviation in the estimated long-term stability of Int-SIM. This instrument stability error is derived by computing the difference between the two onboard spectrometers, SIM A and SIM B, at as similar wavelengths as is achievable. The level is comparable to the noise equivalent irradiance of the detector. There are approximately annual oscillations in the smoothed plot of SIM with large deviations in the early period and around the solar minimum (December 2008, dotted line). These result from inexact degradation corrections made more difficult by slowly varying shifts in the wavelength scale and two spacecraft anomalies in 2009. The degradation corrections used here are the best that could be achieved at the time of writing.

Table 5.2 shows that agreement between original Int-SIM and TIM is reasonable with a correlation coefficient of 0.83, although the regression slope (assuming equal errors on either axis using FITEXY) of 1.10 indicates that the variability of SIM is larger. Indeed, separating the original data into detrended and smoothed lightcurves shows that this larger amplitude is due to the larger long-term trend of the Int-SIM data. The slope of 0.69 for the Int-SIM to TIM smoothed lightcurve indicates the larger variation in the long-term trend, as also shown in the lower half of Fig. 5.1. In contrast, Int-SATIRE shows very good agreement in both original and smoothed cases. It should be noted that disagreement captured by the lower correlation coefficient between Int-SIM and TIM (and Int-SATIRE) is mainly due to the discrepancies clearly seen in the early period and annual oscillations. The long-term trend remains within the error bars. Unfortunately, these are so large that the above statement is of limited value, as it is possible that Int-SIM actually increases over the time period considered instead of declining in line with TIM while remaining within the error bars.

The upper row of Fig. 5.2 shows regression plots comparing Int-SIM, Int-SATIRE and TIM with the regression slope and correlation quoted within. The lower panels display histograms of the scatter about the regression lines of the upper plots. Overplotted on

Table 5.3: Comparison of SORCE/SIM with SATIRE-S for selected inetgrated bands.

	(1) $\overline{\text{SIM}}$ [Wm^{-2}]	(2) ΔSIM [Wm^{-2}]	(3) $\overline{\text{SATIRE} - \text{S}}$ [Wm^{-2}]	(4) ΔSATIRE-S [Wm^{-2}]	(5) $r[r^2]$	(6) Slope
Integrated UV: 201-300 nm						
Original	15.06	0.34	15.05	0.06	0.772 [0.595]	0.195
Detrended	-	-	-	-	0.770 [0.593]	1.001
Smoothed	-	0.31	-	0.06	0.851 [0.724]	0.189
Integrated visible: 400-691 nm						
Original	521.27	-0.71	521.25	0.09	0.320 [0.102]	0.210
Detrended	-	-	-	-	0.884 [0.781]	0.972
Smoothed	-	-0.69	-	0.06	-0.576 [0.332]	-0.079
Integrated IR: 972-1630 nm						
Original	295.96	-0.24	295.95	0.00	0.402 [0.161]	0.159
Detrended	-	-	-	-	0.738 [0.545]	0.823
Smoothed	-	-0.24	-	-0.01	0.731 [0.534]	0.049
Integration over 200-1630 nm						
Original	1223.54	0.03	1223.46	0.34	0.810 [0.657]	0.877
Detrended	-	-	-	-	0.839 [0.702]	0.948
Smoothed	-	0.02	-	0.29	0.781 [0.610]	0.730

Columns (1) and (3) show the mean flux level of SIM and SATIRE-S, respectively, in December 2008; (2) and (4) show the flux change in SIM and SATIRE-S; (5) is the correlation coefficient and its square; (6) is the regression slope.

each histogram is a Gaussian fit with the standard deviation, in Wm^{-2}, quoted in the upper-left of the plot. All three plots show good agreement with the gaussian distribution. The scatter is a factor of three smaller for Int-SATIRE and TIM though there is a slight skew to the right owing to some facular overestimation in the model. The spread for the other two comparisons here is dominated by the internal variability of Int-SIM which is much larger than that of TIM or Int-SATIRE and hence the distributions are much more spread out. Fig. 5.1 suggests that the quiet, low S/N period during solar minimum is the main reason for the degraded correlation between Int-SIM and TIM. The lack of signal in Int-SATIRE during this minimum period (see section 2.1) leads to a further decrease of the correlation coefficient for Int-SIM and Int-SATIRE. Indeed, increased correlations are found when the period is restricted to more active times, e.g. ending in February 2007.

The results presented in Fig. 5.2 assumed errors calculated from the standard deviation of the data, which were approximately the same in all three datasets. However, given the error bars in Int-SIM (see Fig. 5.1) are significantly larger than the quoted uncertainty in TIM, a quick check is performed here assuming only errors in Int-SIM and Int-SATIRE. As in Fig. 5.2, Fig. 5.3 shows the regression of Int-SIM and Int-SATIRE against SORCE/TIM. The dashed lines are 1σ errors on the solid line of regression. The slopes are presented along with the 1σ errors in Tab. 5.2 in the fifth and sixth columns. The errors when calculating the slope in column four using FITEXY assumed that the uncertainty was equal in both datasets. Given the smaller uncertainty in TIM, one is justified in assuming that all the error is within the y-axis data and that TIM is the true answer. It is then found that both Int-SIM and Int-SATIRE agree with TIM within the formal uncertainty, though Int-SATIRE still has a better agreement from the slope. However, it is worth noting that the large deviations resulting from sunspot transits, producing the low values of integrated-flux, hide the smaller longer-term general deviations between Int-SIM and TIM during the earlier period, as discussed earlier. From hereon, we return to assuming almost equal errors on both axes as derived from the data.

Since long-term changes in TSI are small compared to rotational variability in the period considered, the short-term variability dominates the correlation coefficient results of the original datasets. This can be seen by comparing the original correlation coefficients with the detrended results in Table 5.2. The results also indicate that, even without IR longward of 1630 nm, Int-SATIRE still recreates the TSI accurately over the six year period and provides a good substitute. Short-term variability (Table 5.2, middle) of Int-SIM with TIM gives a good result. But, as can be seen from the upper plot of Fig. 5.1, on a short timescale Int-SIM tends to overestimate the short-term variations in TSI, in particular in

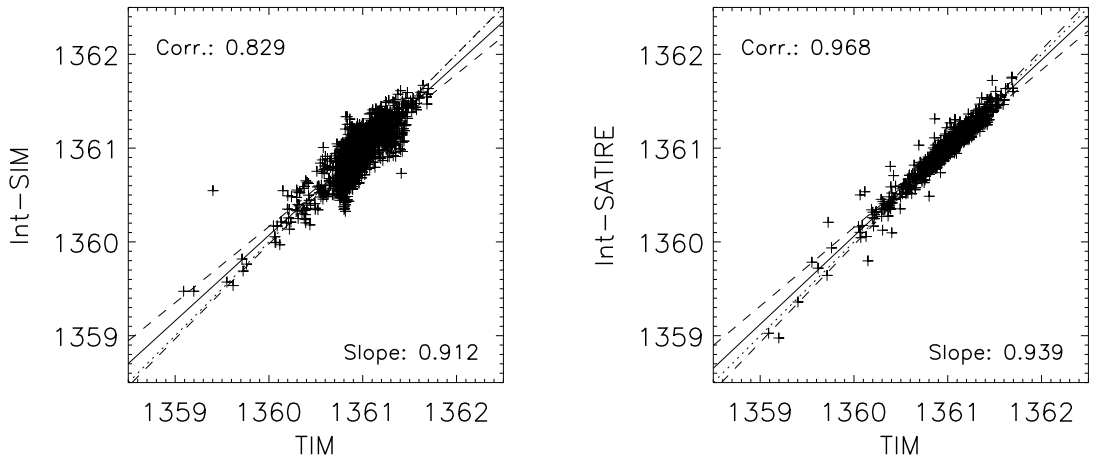


Figure 5.3: As for Fig. 5.2, regression plots for (**left**) Int-SIM vs. SORCE/TIM and (**right**) Int-SATIRE vs. SORCE/TIM calculated assuming only an error in the y-axis dataset. The solid line is the formal line of regression and the dashed lines are 1σ errors on the regression.

the later phases of the cycle.

Of most interest are the smoothed results. The long-term variation of some spectral regions were unknown before the launch of SIM. We find that Int-SIM correlates less well with both TIM and Int-SATIRE on a long-term basis. As seen in the long-term trend of the bottom panel of Fig. 5.1 this is at least partly caused by the oscillation visible in 2008 and 2009, not always in phase with TIM and Int-SATIRE. However, more significantly there is a difference of ∼0.2 Wm^{-2} between Int-SIM and TIM during April/May 2004. Although the two original datasets differ by only ∼150 ppm, this difference means that while TIM sees a decrease in irradiance of ∼0.3 Wm^{-2} between April 2004 and the solar minimum[2], Int-SIM shows a decrease of only ∼0.1 Wm^{-2}. By mid-2005 all three datasets have reached a similar level of irradiance variation and SIM follows the trend in TSI within errors. It turns out that the main part of the difference between TIM and Int-SIM occurs during the period of largest change in all spectral regions, but it should be noted that the difference is just within one standard deviation of the long-term stability. We will consider the early difference further in the following sections while discussing the spectral variation and trends.

[2]The value by which the datasets are shifted is taken from an average over a year centred on December 2008. This differs from the value in Table 5.3, but is done to smooth out the large bump during that period and make a fairer comparison.

5.2.2 Spectral Solar Irradiance

We now consider three of the spectral bands highlighted in Harder et al. (2009) where surprising SIM trends were first presented. Then our analysis will be expanded to narrower regions over the whole 200-1630 nm spectrum. In Figs 5.4, 5.5 and 5.6 the original, detrended and smoothed lightcurves are shown for the spectral regions 201-300 nm (UV), 400-691 nm (visible) and 972-1630 nm (IR), respectively, with SATIRE-S, in blue, normalised to SIM, in red, over a one year period centred on December 2008 (dotted line). Given in table 5.3 are the correlation coefficients, regression slopes and a measure of the change in irradiance between the average flux in the first month of the period and during the solar minimum month. Error bars represent the estimated long-term stability of the instrument over the integrated regions and are derived in the same way as for Int-SIM (see section 5.2.1).

In the following discussion focus should be given to the correlations when considering the short-term and regressions should be considered for discussion of the long-term variability. To highlight this, the example of a high correlation coefficient but poorly matching long-term trend in the UV 200-300 nm region explains nothing about how well they agree. The magnitude of the long-term change in this example is very much in disagreement and only the regression picks this up. A good regression for any detrended spectral bandwidth should be expected if the long-term trend is sufficiently removed and short-term variability is reasonably well recreated.

5.2.2.1 UV, 201-300 nm

In the UV, the most striking difference between SIM and SATIRE-S is the difference in gradient seen in the original (top panel) and smoothed (bottom) lightcurves of Fig. 5.4, in particular prior to the middle of 2006. The change in irradiance between 2004 and the reference date is 0.34 Wm^{-2} for SIM, around five times larger than SATIRE-S's 0.06 Wm^{-2}. This is also reflected in the low regression slope of 0.20 for both the original and smoothed UV data. Note that the very high correlation in the smoothed data is the result of consistent agreement in the direction of change in the UV. However, although not clear in the smoothed UV lightcurve, a minimum is reached in SIM more than a full year prior to SATIRE-S and TSI measurements, in August 2007 and October 2008 respectively. The error bars establish that the disagreement between SIM and SATIRE-S exceeds the estimated long-term uncertainty of the instrument. It should be noted that given that the decline in TSI over this period is approximately a third of the maximum-to-

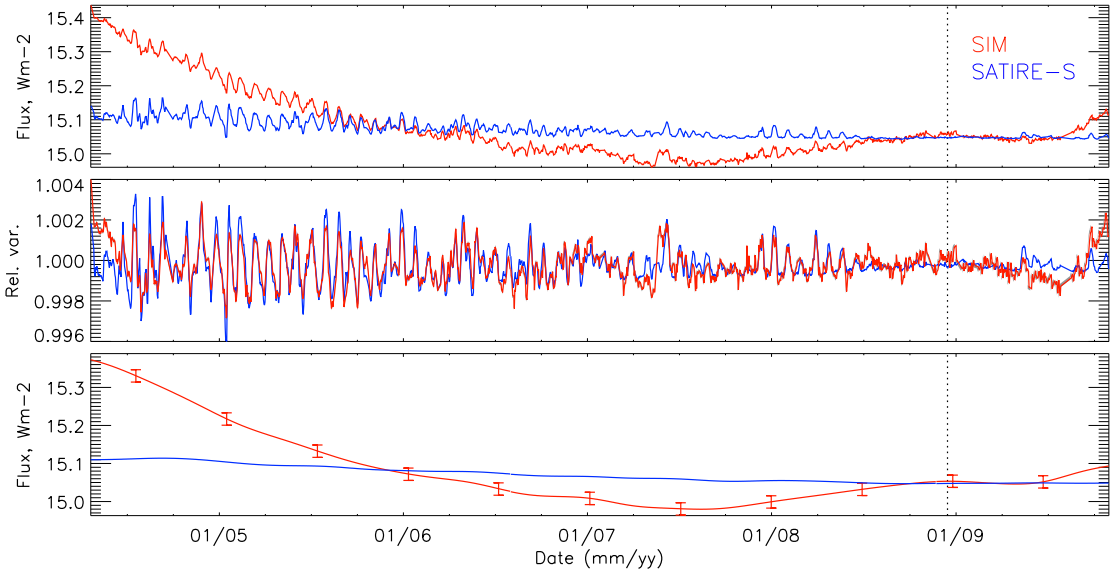


Figure 5.4: Lightcurves of the integrated UV region 201-300 nm for SIM (red) and SATIRE-S (blue). The top panel shows the original data, with SATIRE-S normalised by multiplication to SIM, while the middle and bottom panels show the detrended (short-term variations) and smoothed (long-term trends) curves. The one standard deviation error bars are for the long-term stability of the instrument over these integrated wavelengths is 0.016 Wm^{-2}. The dotted line is at the solar minimum at December 2008.

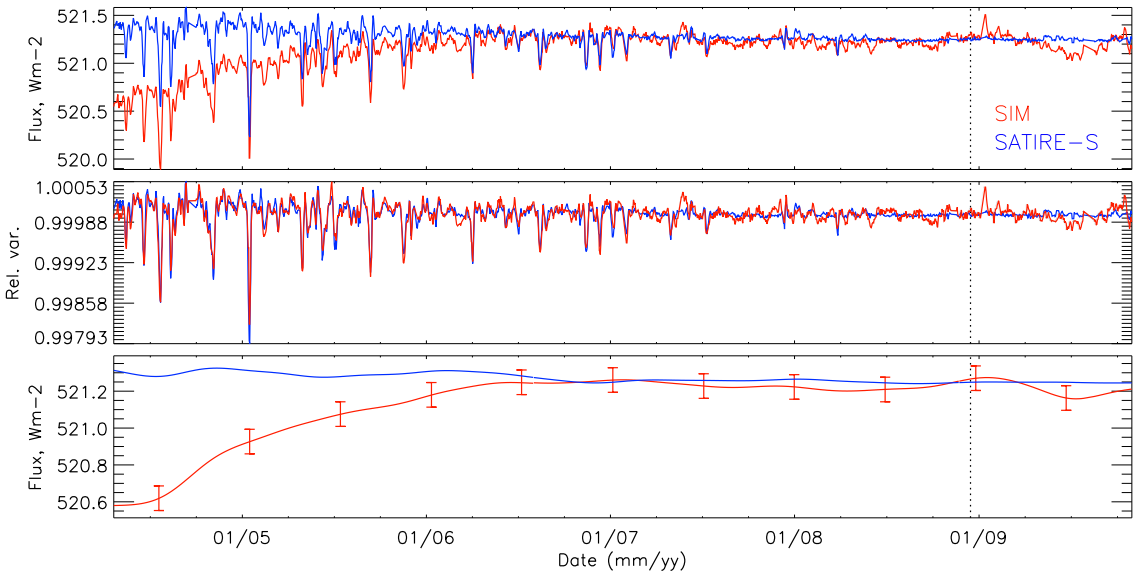


Figure 5.5: As Fig. 5.4 but for the integrated visible region of 400-691 nm. Error bars are 0.067 Wm^{-2}.

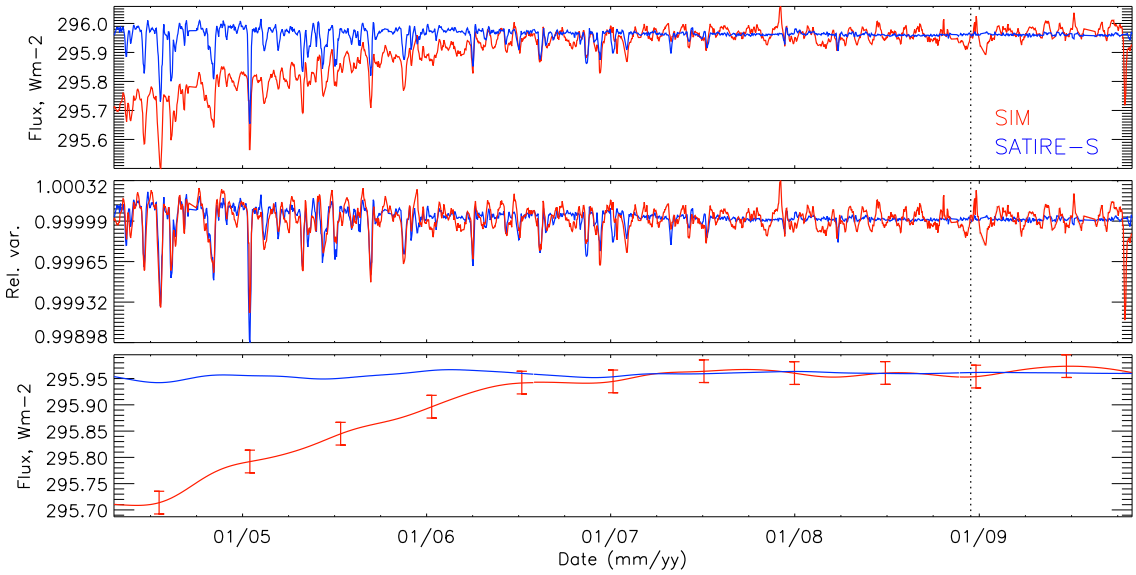


Figure 5.6: As Fig. 5.4 but for the integrated IR region of 972-1630 nm. Error bars are 0.022 Wm^{-2}.

minimum variation, this sets the cycle variation a significant factor above that observed by UARS/SUSIM and UARS/SOLSTICE (Floyd et al., 2003; Krivova et al., 2006). Over a similar period in the previous cycle, August 1993 to May 1996 (solar minimum), the 200-300 nm region declined by 0.06 Wm^{-2} from 14.51 to 14.45 Wm^{-2} according to level 3BS V22 SUSIM data on UARS. The decline from maximum (November 1991) to minimum in that cycle was 0.11 Wm^{-2} from 14.56 Wm^{-2}. Although the recent unusual minimum should be kept in mind, the change of TSI between the last two minima has been very small, if it has changed at all (Willson, 1997; Dewitte et al., 2004; Fröhlich, 2006, 2009). The reference date for the minimum is taken from sunspot number and it is interesting to note that according to SIM the UV minimum occured a year earlier, in 2007.

It should be noted that by varying the free parameter, B_{sat}, different gradients in the UV can also be achieved with SATIRE-S. A lower B_{sat} results in higher UV gradients (and also an incorrect TSI reconstruction), but even setting the value to a minimum the gradient seen in SIM cannot be reached and the resulting rotational variability becomes exceptionally large and unrealistic. In particular, SATIRE-S cannot match the huge drop in UV irradiance seen in SIM during the first two years. This early period also displays the largest change, larger than any other period considered, at all wavelengths and when integrated over the full range. This may be because early spacecraft degradation is larger during early phases of space-based missions and hard to correct for. In fact, the vast majority of cycle-length gradient and flux changes in SIM occur during this inital two-year

period.

The rotational variation is already known to agree reasonably well when considered over a period of several months (Unruh et al., 2008) and we also find that, over the six-year period considered here, rotational variation is well reproduced (Fig. 5.4). The UV region reflects facular variation and we can see that SATIRE-S tends to overestimate that contribution, predominately during periods of higher activity.

Facular overestimation occurs for two main reasons: the magnetic flux close to the limb is excessively amplified at locations where it is not vertical when applying a line-of-sight correction to the magnetograms, assuming all flux to be radial (see section 2.1). Furthermore, the contrast applied to flux elements near the limb is also overestimated due to the use of one-dimensional model atmospheres (Solanki and Unruh, 1998, and see also the discussion in section 2.2.1). Unruh et al. (2008) discuss the rotational variability in more depth and in comparison with the UARS/SUSIM instrument. They reported that below 240 nm and in some strong UV lines there is further disagreement since the intensties used to derive facular and spot contrasts are calculated on the assumption of LTE even where non-LTE would be more appropriate and so a difference is to be expected (Unruh et al., 1999; Krivova and Solanki, 2005; Haberreiter et al., 2005; Krivova et al., 2006). It was also shown by Shapiro et al. (2010), using an updated version of the COSI radiative transfer code which performs calculations under non-LTE conditions (Haberreiter et al., 2008), that LTE calculations can overestimate the facular contrast. Note that good agreement between SATIRE-S and TSI measured with SORCE/TIM on rotational timescales is found (see Fig. 5.1), although compensation by an underestimate at other wavelengths cannot be ruled out.

5.2.2.2 Visible, 400-691 nm

Moving to the visible region at 400-691 nm (Fig. 5.5 and middle section of Table 5.3), a difference between observation and model is keenly seen in the long-term trend of original and smoothed lightcurves in the first two years. As noted in Harder et al. (2009), there is an offset and steeply increasing trend in SIM for the early period in 2004 and 2005, in opposition to TSI, before flattening out towards the minimum. Again, the error bars in the SIM trend are much smaller than the trend itself. Over this period TIM TSI declines at a steady rate. SIM shows an increase in irradiance of 0.71 Wm^{-2}, almost an order of magnitude larger and in the opposite direction to SATIRE-S's 0.09 Wm^{-2} decrease. The

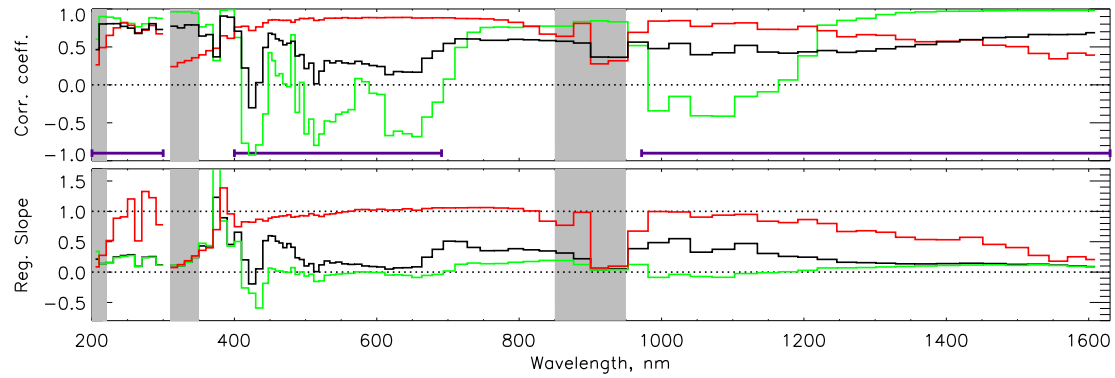


Figure 5.7: Correlation coefficents (top) and regression slopes (bottom) between SIM and SATIRE-S as a function of wavelength for the spectrum from 200-1630 nm. For $\lambda < 450$ nm wavelengths are compared in 10 nm bins; for $\lambda > 450$ nm bins correspond to resolution elements of SIM. There are three types of comparison made for each bin over the time period of 2004 April 21 to 2009 October 31: (black) original data, (red) detrended, short-term and (green) smoothed or long-term. Greyed-out regions highlight the detector edges where the signal is comparable to the instrument noise. Horizontal dotted lines are to aid the reader. The three horizontal purple bars in the upper plot highlight the integrated regions shown in Figs. 5.4, 5.5, 5.6.

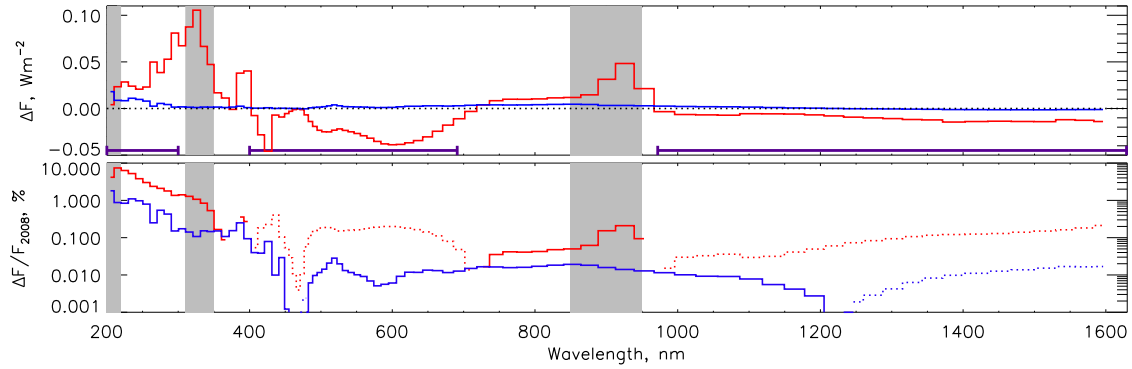


Figure 5.8: **top** Change in flux in the same bins as in Fig. 5.7 of (red) SIM and (blue) SATIRE-S between the mean level of the first month of SIM's data and solar minimum in December 2008. Negative values imply an increase in irradiance over the period. The dotted line marks zero flux change and the greyed-out regions indicate the detector edges. The purple horizontal bars are as in Fig. 5.7. **bottom** The same as the top plot, but the percentage change relative to December 2008. Dotted lines are negative changes.

regression slope reflects this. Note that after mid-2006 the SIM trend and the model agree within the error bars.

In stark contrast to the long-term trends is the agreement seen for the detrended, rotational variability, indicated by the correlation coefficient of 0.88. This means that rotational variation and especially the spot and facular contribution is being well reproduced in the model. This can be seen in the detrended lightcurve where the magnitude of variation is clearly simulated well. The relative lack of signal detected in MDI images leads to a lack of rotational variability in SATIRE-S between April 2008 and April 2009, so that SATIRE-S underestimates the observed rotational variabiltiy over this time interval; this is also apparent in the UV, but is most prominent in the IR (see below). It is the combination of clearly opposing early long-term trends and well matching short-term trends that results in the low correlation for the original data and the small regression slope seen in Table 5.3.

5.2.2.3 IR, 972-1630 nm

A similar result is found in the integrated-IR region of 972-1630 nm as in the visible, for both short-term agreement and the longer-term trends (Fig. 5.6). It should be noted that the facular model employed by SATIRE-S results in faculae that become dark for λ >1300 nm. As the region considered here captures wavelengths at which faculae are bright and dark, some cancellation will result in SATIRE-S over this region. Over the first three years, SIM observes an increase of 0.24 Wm^{-2} and SATIRE-S registers no change. Inspection of the lightcurve during active times shows that the effect of spot passages are well reproduced, though there is uncertainty about whether facular effects are too. Apart from a couple of spot passages in 2008, from July 2007 to the end of the period, SATIRE-S is unable to reproduce the rotational variability seen with SIM, as could be seen to some extent in the visible, resulting in a lower correlation coefficient of 0.76. As in the UV, the low magnetic flux during this later period means that only a very small rotational amplitude is produced, too small to be seen on this scale.

5.2.3 Comparison over all wavelength regions

The approach taken in the previous section is now extended to narrow bands across the entire spectrum. Fig. 5.7 shows the correlation and regression between SIM and SATIRE-S

at all wavelengths from 200 to 1630 nm, in bins of 10 nm below 450 nm and at the instrument resolution of SIM above this wavelength for the period 2004 April 21 to 2009 October 31.

In the upper and lower panels of Fig. 5.7 we plot correlation coefficients and regression slopes as a function of wavelength, respectively. These are split into the original (black lines), detrended (red) and smoothed (green) datasets. Regression slopes greater than 1.0 indicate wavelengths where the SATIRE-S model shows larger amplitudes than the SIM observations. Fig. 5.7 is therefore a general overview of the spectral comparison of SIM and SATIRE-S. The grey regions highlight the detector edges where instrument noise is large and masks real signal.

In the UV we see that short-term variation below 230 nm is poorly recreated, identified by a correlation coefficient below 0.5. SATIRE-S assumes LTE across all wavelengths while this region is known to be dominated by non-LTE effects. This is also a region where the instrument noise is comparable to variability in SIM. As wavelength increases, non-LTE effects become less important and LTE becomes a more realistic assumption. In parallel, the magnitude of solar variability increases relative to the instrument noise and this is reflected in an improved short-term correlation. The short-term regression slopes above 1.0 highlight the regions of the UV where SATIRE-S overestimates the rotational variability during active periods, namely the 250-260 nm and 270-290 nm regions. Up to wavelengths of 350 nm the magnitude of the gradient of the long-term trends are much larger in SIM and are reflected by the regression slopes close to zero.

For the vis1 photodiode above 400 nm, short-term correlation coefficients become better than 0.90 and the regression slope near to 1.0 across the entire visible region showing that rotational variability is well reproduced by SATIRE-S. Long-term regression slopes are negative and close to zero, this being a result of opposing trends between SATIRE-S and SIM prior to 2006 (see Fig. 5.5 and section 5.2.2.2) and only a very small long-term trend in SATIRE-S visible regions. The long-term results are quite wavelength dependent, with negative trends in SIM over wavelengths of 480-560 and 610-720 nm.

In the IR there is a gradual decline in correlation of the short-term variation, dropping from a correlation coefficient above 0.8 to below 0.4 near the edge of the detector at 1630 nm. We note that although spot passages are well identified in this spectral region, rotational variation due to faculae is not always. During periods of higher activity, and therefore higher magnetic flux, faculae in the IR can appear dark (Unruh et al., 2008;

Fontenla et al., 2004). But there appears to be a switchover to positive contrast in faculae of lower field strength, especially during lower activity. We see evidence for bright faculae in the IR from SIM during several rotations in 2008 that vary in phase with TSI and during which there were no spot passages. Over the IR domain there is a gradual change in the smoothed regression slope and correlation coefficient from negative to positive. This is because SATIRE-S assumes that faculae are all dark for $\lambda > 1300$ nm except near to the limb. Facular disk transits therefore result in reduced variability and ultimately produce a rotational variation in opposition to the flux measured in SIM in the later, quieter period. Integration over a wide IR spectral band, as in Fig. 5.6, combines these regions of positive and negative contrast which cancel each other to produce a flat smoothed lightcurve.

In many regions throughout the entire spectrum, the effect of conflicting short- and long-term trends is highlighted by the divergence of both green (smoothed) and red (detrended) lines from the black (original) line. While SATIRE-S generally matches the short-term variability, it cannot reproduce the magnitude or shape of the long-term trend in almost all spectral regions, though the accuracy of the observed long-term trends in SIM is still to be validated.

The top plot of Fig. 5.8 shows the contribution by each bin to the total change in Int-SATIRE and Int-SIM by taking the difference of the averaged 30-day period from 21 April 2004 to 21 May 2004 and the month of December 2008 at solar minimum. The lower plot shows the percentage change in flux relative to December 2008. Fig. 5.8 demonstrates that the offset that occurs between each spectral region is significantly less in SATIRE-S than with SIM. Although the integral over 200-1630 nm is not fully in agreement, the disagreement is within the error shown in Fig. 5.1. The missing 10% of irradiance >1630nm does not vary enough, suggested by the ESR detector, to make up this shortfall. Looking at Fig. 5.1, the missing flux must therefore be from the disagreement in the 2004 and the large oscillation around December 2008.

5.3 Discussion and conclusions

In this chapter we have made comparisons of total and spectral solar irradiance observations from the SORCE satellite, namely the TIM and SIM instruments (Kopp and Lawrence, 2005; Harder et al., 2005a), with the SATIRE-S model (Fligge et al., 2000; Krivova et al., 2003). We have broken down the broad spectrum from 200-1630 nm into

small regions and compared short-term rotational variability and long-term trends over a period of 6 years from moderate solar activity in 2004 to the solar minimum in 2008 and 2009.

Our investigation shows that although SATIRE-S is highly successful at recreating total solar irradiance over both rotational and cycle length periods, there are large differences between observation and model in a number of spectral regions, specifically regarding the gradients of the long-term trends between 2004 and 2006.

Over the period between 2004 and 2009 considered in this chapter SATIRE-S reproduces 95% of the TSI variation seen in the SORCE/TIM TSI time series, while Int-SIM, i.e. SIM integrated from 200 to 1630 nm, accounts for 69%. We find that Int-SIM shows excellent short-term agreement with TIM. The longer-term trends tend to agree reasonably well, though there is substantial diagreement in the first 12 months where there are rapid changes in all spectral regions and the largest variation in flux occurs.

On short timescales the SATIRE-S UV shows good agreement in capturing rotational variation, but overestimates the amplitude in some spectral regions during years of higher activity, e.g. at 250-260 nm and 270-290 nm. This results from overestimating the contrast for faculae close to the limb and the use of LTE approximation in a region where this is inadequate (Krivova et al., 2006). It should be noted that the limb- and flux-dependent contrasts applied in SATIRE-S are based on one-dimensional models and are therefore not likely to model the physics fully.

The rotational variability in the visible is the region best recreated by SATIRE-S, while short-term fluctuations are well matched with SIM in the IR during active periods. During the solar minimum in 2008-9 for $\lambda > 1300$ nm a clear signal is present in the observational data while almost no variation is seen in SATIRE-S at any wavelengths. This highlights that when using MDI 5-minute integrated magnetograms the signal of magnetic flux during the 2008-9 period is extremely weak and the resulting modelled lightcurves are almost completely devoid of rotational variability, whereas TIM shows small but clear variation (see section 2.1).

In SATIRE-S, during active periods, the short-term over-enhancement of facular regions in the UV more than compensates for the negative contrast applied in the IR to these regions. Magneto-convective simulations suggest that the IR contrast for low magnetic flux levels remains positive for $\lambda > 1300$ nm in agreement with the detrended SIM re-

sults (N. Afram, private communication). This evidence highlights that not all of the wavelength-dependent contrasts applied in SATIRE-S to facular pixels detected in MDI magnetograms are correct and that a transition from positive to negative contrast in the IR for $\lambda > 1300$ nm results in disagreement with the observations from SIM. Contrasts are dependent on wavelength, limb-angle and magnetic flux level (Foukal and Moran, 1994, and discussed in section 2.2.1). These variables change over the solar cycle and result in a different contribution to TSI. It is therefore important not only to improve the model rotational variability, but also to investigate the inter-cycle variation which is currently of great interest given the unusually quiet period the Sun recently went through.

The long-term UV trends observed in SIM and computed by SATIRE-S disagree strongly. The magnitude of variation over the first two-year period, 2004-2006, in the UV 200-300 nm region is five times larger in SIM than SATIRE-S. There is some agreement of this trend in SIM with the SORCE/SOLSTICE[3] (McClintock et al., 2005a) UV instrument, though this depends on the period of overlap considered. The short- and long-term trends of UARS/SUSIM agree with SATIRE-S in 2004. But, between 2005 and 2007 SORCE/SOLSTICE disagrees with SATIRE-S while having a reasonable agreement with SIM. Post-2007, SIM starts to increase again while SORCE/SOLSTICE continues to decline. Therefore, no clear picture emerges as yet from a comparison with the SORCE/SOLSTICE data even though there is a decline of 0.27 Wm^{-2} observed between April 2004 and December 2008, in reasonable agreement with SIM.

In the visible region, SIM observes an increase in irradiance as TSI decreases, large enough to offset the significant reduction in irradiance produced by the UV region. An increase in irradiance in this region as detailed in Harder et al. (2009) is in opposition to previously modelled spectral variation (Lean, 1991; Krivova et al., 2006) and to that used in global climate models (Haigh et al., 2010).

In the IR, except for the 750-900 nm region, long-term changes displayed by SIM are all in opposition to TSI and SATIRE-S for $\lambda < 1300$ nm. As a result of most of the spectrum in SATIRE-S decreasing with TSI, the contribution from each region to the overall change is, in some regions, more than an order of magnitude less than the trends observed in SIM.

The model assumes that variations in irradiance are directly related to the evolution of surface magnetic flux. Given this, if the assumption is made that the long-term UV results from SIM are indeed correct then there are physical implications. Either the long-term

[3]See next chapter for further discussion on SORCE/SOLSTICE data

change being observed is dominated by otherwise unknown non-magnetic processes or it is due to changes in small-scale flux which cannot be detected in 5-min MDI magnetograms[4]. It would also imply a physical change in the Sun during the decline of this solar cycle as compared to the previous one, and even compared with the rising phase of cycle 23, if the results from UARS/SUSIM and UARS/SOLSTICE are correct. Recall that SATIRE-S reproduced the data from these instruments for the same value of the free parameter as needed for TSI. Hence, although the declining phase of cycle 22 and rising phase of cycle 23 UV variations are consistent with surface magnetism as a source, this appears no longer to be the case in the declining phase of cycle 23. Such a sudden change in the physical mechanism seems rather unlikely and implies that there might be a problem in the calibration or stability of either the UARS or the SORCE UV measurements. It should be noted that the UARS instruments covered both the declining and rising solar cycle phases, whereas SORCE has so far mainly covered the declining phase. Also, the discrepancy of SIM to SATIRE-S is largely restricted to the first two years of SIM's operation. Many radiometric instruments studying the Sun in space have had problems maintaining the required level of stability in the first couple of years of observation. Recent work by Lean and DeLand (2011) and DeLand and Cebula (2011) have also reached similar conclusions suggesting that undercorrection of the degradation and sensitivity of the instrument has led to the trends observed in SIM.

Unfortunately, periods considered in both Harder et al. (2009) and this paper are only during a declining period in TSI. An inspection of the publically available SIM data online shows that, at the time of writing (January 2012), no conclusive trend reversals can be seen in SIM as online data are only available in the visible and IR regions up to the end of 2010. Therefore, any reversal in these individual spectral regions, which might provide weight to the plausibility of the spectral trends observed, cannot be considered for now. SOLSPEC (Thuillier et al., 2009), on the International Space Station, is also monitoring the region observed by SIM, but it only became active during the solar minimum and results are still pending.

The investigation in this chapter has highlighted areas of improvement needed for the model. Primarily, the contrasts used in SATIRE-S must be addressed in a physical way to incorporate dependence not just on limb-angle but also on the level of magnetic flux and wavelength. Work is currently underway to address this using results from magneto-

[4]Although it should be noted that, at least for TSI, the previous chapter has shown that SATIRE-S can reproduce TSI variability on all time scales, even when using a threshold that cuts out magnetic flux below 30 G

convective simulations (see chapter 8). It appears very unlikely, however, that such changes will be able to overcome the large difference in the trends between SIM and SATIRE-S at most wavelengths (Afram et al., 2009) and more observations are also needed to establish if the trends observed by SIM are indeed accurate.

Chapter 6

Spectral solar irradiance reconstruction of solar cycles 21-23 using SATIRE-S

Overview

In this chapter, the modelled spectral irradiance data as described in chapter 3, between 1974-2009, are prepared for use by the solar, climate and atmospheric communities. Two main steps are taken to prepare the data. First, the correction of Krivova et al. (2006) is applied to compensate for the LTE assumptions in the model atmopsheres below 270 nm. Then, all dates that lacked the continuum intensity images and magnetograms are filled using proxy indices of solar variability. The spectral reconstruction is then compared with the National Research Laboratory SSI (NRLSSI) model, which is commonly used in climate and atmospheric studies. Finally a comparison is made with the most current version of SORCE/SOLSTICE data, version 11, which have had the latest degradation corrections applied. This paves the way for an investigation, in the next chapter, of the cyclical solar forcings from SATIRE-S and SORCE/SOLSTICE on atmospheric ozone concentrations in the stratosphere.

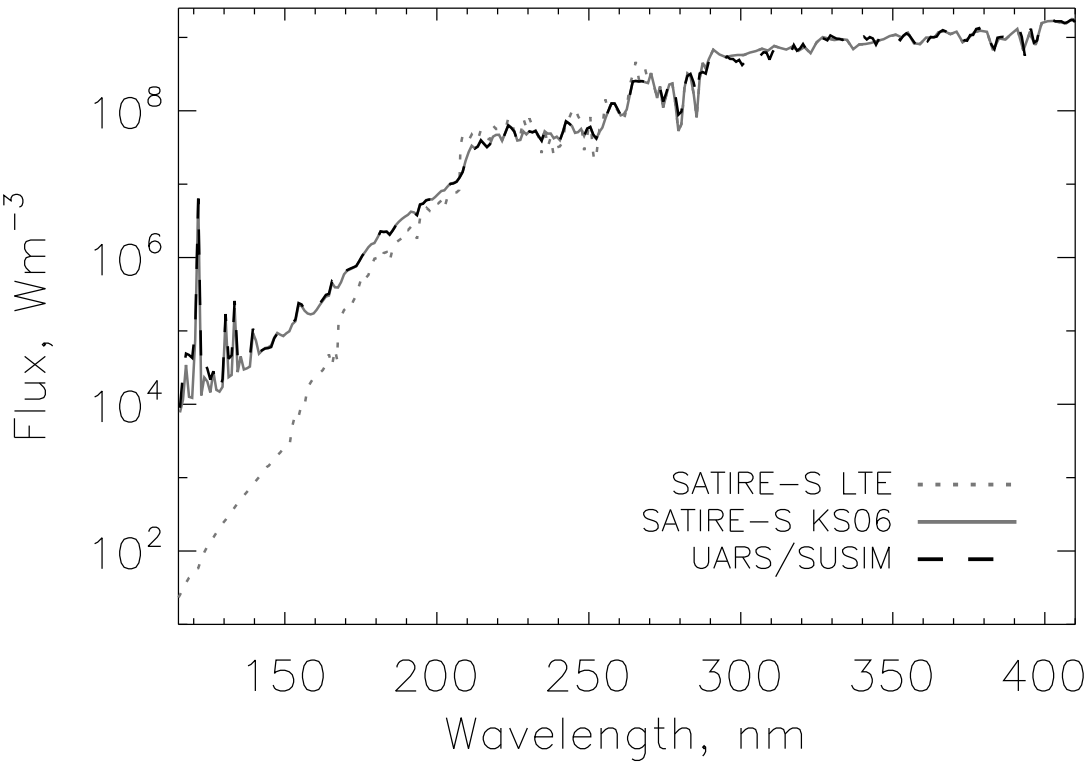


Figure 6.1: Flux as a function of wavelength on 2005 February 30. Shown are SATIRE-S employing LTE-only model atmospheres (grey, dotted line) and the resulting agreement with UARS/SUSIM (black, dashed) after applying the empirical modifications of Krivova et al. (2006) (grey, solid). Note that this plot is based on Fig. 1 of Krivova et al. (2006).

6.1 Preparation of SATIRE-S Dataset, 1974-2009

The previous chapter considered the modelled spectral data in comparison to SORCE/SIM. Here, the data need to be prepared for use in a climate model in the next chapter, and so adjustments need to be made to compensate for LTE approximations below 300 nm. This section descibes the corrections applied and the gap filling procedure used, to fill data gaps on just over 50% of dates between 1974 and 2009, and as a result improve the accuracy and usefulness of the spectral dataset.

6.1.1 Correction of wavelengths below 270 nm

The SATIRE-S model atmospheres assume LTE conditions at all wavelengths, an approximation known to become less certain below ~400 nm in all model atmospheres used in SATIRE-S including the modelled quiet Sun. A lot of this radiation is emitted from the optically thin region of the chromosphere and the variation is better reproduced in some

wavelengths of the model than the absolute fluxes. Either an understanding (and therefore accurate modelling) of how radiative transfer operates in non-local thermodynamic equilibrium, or a way of relating modelled to observed variations, is required.

KS06 found an empirical solution by using data from the UARS/SUSIM instrument (Brueckner et al., 1993) to extrapolate the model to regions below 270 nm and better reflect the Sun's spectral irradiance in these regions. They found that the integrated spectral region 220-240 nm, in SATIRE-S, correlated highly with the same region in SUSIM ($r_c = 0.97$) and showed good long-term agreement. This region in SUSIM, in turn, correlated well ($r_c > 0.8$) with around 80% of wavelengths between 115 and 410 nm. Through a regression analysis they calculated the variability and fluxes at other wavelengths using the derived regressions between the integrated 220-240 nm region and SUSIM. As a result, these wavelengths still depend on the variability determined by the underlying model of SATIRE-S from the 220-240 nm region, but now with better estimated absolute flux and variability than by using the model atmospheres in SATIRE-S under the assumption of LTE alone.

The uncertainty of SUSIM becomes comparable to the solar cycle variation at 250 nm, and higher at 300 nm and above. The model is therefore considered to be at least as reliable at these wavelengths. Artefacts appear in the SUSIM time series above 270 nm, meaning that replacing these wavelengths is less accurate than for wavelengths shorter than 270 nm. As a result, we choose to replace only wavelengths below 270 nm in SATIRE-S.

The resulting modified SATIRE-S spectral intensity can be seen in Fig. 6.1 (grey, solid line), along with SUSIM (black, dashed) and the original SATIRE-S that uses LTE alone (grey, dotted). This figure shows that the modification gives much better agreement with SUSIM. We apply the empirical corrections of KS06 to the new three-cycle spectral dataset.

6.1.2 Normalisation of KP/SPM reconstruction to SoHO/MDI

In chapter 3, the overlap period of KP/SPM and SoHO/MDI of 895 days between 1999 and 2003 was used to fix the free parameter in the KP reconstructions. The KP reconstruction was normalised to the level of the MDI reconstruction to create a homogeneous time series. In chapter 4, it was shown that the value of the magnetic flux registered in one instrument does not necessarily scale linearly, or in the same way at different positions,

with the other. Due to these differences, it was not possible to map the flux response of one instrument onto the other. As a result there is a slight difference in the relative distribution of magnetic flux as a function of limb angle. This leads to slightly different SSI trends and a different free parameter is needed for each reconstruction.

In Fig. 6.2a the ratio of the mean flux level for each wavelength in SPM and MDI is shown. This is the normalisation factor required to bring SPM into agreement with MDI between 115 and 500 nm. Above 500 nm the normalisation factor is very close to 1.0. The grey horizontal line is the normalisation factor for all wavelengths below 270 nm, which is constant because all wavelengths are derived from the 220-240 nm region (as is the case in Figs. 6.2b and 6.2c). Fig. 6.2b shows that the correlation coefficient between SPM and MDI is consistently high (>0.92) for all wavelengths, including above 2400 nm (not shown).

Fig. 6.2c shows the regression slope between MDI and KP/SPM after the normalisation in Fig. 6.2a has been applied to each wavelength. The regression results suggest that some wavelengths should display slightly different gradients prior to 1999. However, these regressions do not reflect long-term trends only. Due to the relatively low cycle variability for the 1999-2003 overlap period, the difference in short-term variability will also be important in producing the regressions shown. It is not possible to accurately correct for the different responses in KP/SPM and MDI magnetograms because of the variable response in the MDI instrument as a function of disk position. When each wavelength reconstructed using KP/SPM data is separately regressed onto MDI using the coefficents in Fig. 6.2c, the integrated-TSI remains within the uncertainty derived in chapter 3, and the change in TSI between cycle 23 minima reduces by only ~0.05 Wm^2. To maintain the integrated-TSI result, and because short-term variability makes the result of Fig. 6.2c less reliable for long-term reconstructions, we use the trends as they stand after normalising the KP dataset to MDI.

By integrating over the full SATIRE-S SSI, the resulting TSI is 0.47% lower than SORCE/TIM at the solar minimum of December 2008. We therefore increase all wavelengths by this amount, through normalisation, so that the integrated-value agrees with TSI from SORCE/TIM. This also brings the integrated-variability into agreement with SATIRE-S_{Ind} TSI calculated in chapter 3.

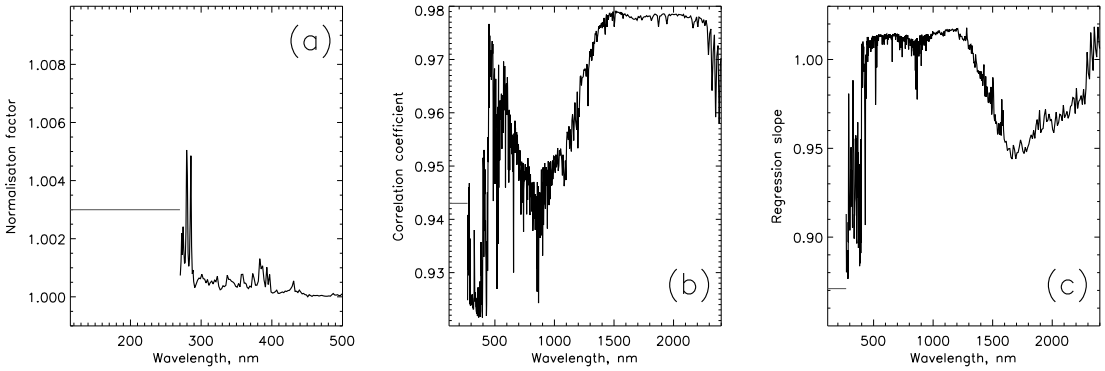


Figure 6.2: **(a)** The mean normalisation for each wavelength to bring flux levels in the SPM reconstruction into agreement with MDI during the overlap period between 1999 and 2003. Here, only normalisation is shown between 115 and 500 nm. **(b)** the correlation coefficient between SPM and MDI over the same period as (a) but for wavelengths 115 - 2400 nm. **(c)** as for (b) but showing the regression slope between SPM and MDI after normalisation to the same mean flux level.

6.1.3 Filling data gaps

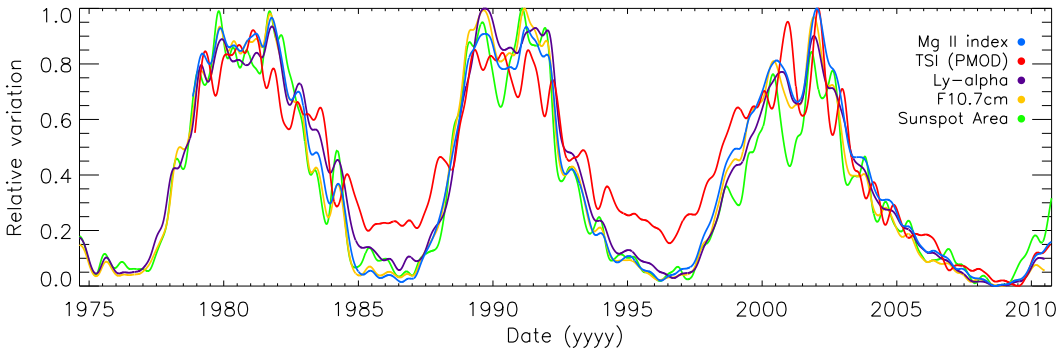


Figure 6.3: Smoothed time series of five indicators of solar variability between 1974 and 2009. The indices are indicated in the legend.

Of the 12853 days between 1974 December 10 and 2009 October 31, the period currently covered by the reconstruction, there are 6475 dates without available full-disk images from KP or MDI. Most of these data gaps are in cycles 21 and 22. We use the following datasets, collectively referred to here as *indices*, to fill gaps in SATIRE-S: the NOAA and LASP Mg II indices (Viereck et al., 2004; Snow et al., 2005), combined through linear regression; the Ly-α composite by Woods et al. (2000); the Penticton F10.7 cm radio flux (data available through the National Geophysical Data Center at http://www.ngdc.noaa.gov/); the TSI from version d41_62_1003 of the PMOD composite (Fröhlich, 2000) and the sunspot area (SSA) composite record by Balmaceda et al. (2009).

Each index is indicative of the behaviour of some feature (or combination of features) or source of radiative flux in the solar atmosphere, which in turn should correlate with activity and therefore variability in solar spectral irradiance. The variability of each index depends on the time scale being considered. In Fig. 6.3, each index is smoothed to removed rotational variability and the relative variation is plotted on a scale of 0 to 1, representing the minimum and maximum values of the time series, between 1974-2009; the legend gives the colours of each index. All time-series are smoothed using a Gaussian window with a full-width half-maximum of 127-days centred on each date. The indices show slightly different behaviour over medium-term periods of 1-2 years and on long-term, cycle-length periods. Although all indices are representative of some form of solar variability, it is not clear exactly how they relate to SSI (Dudok de Wit et al., 2008). Fig. 6.3 shows that not one single index can be used to represent all wavelengths of SSI on time scales longer than the solar rotation (and also compare the difference between spectral behaviours in the UV, visible and IR in section 5.2.2). The differences may also relate to the fact that, excluding the F10.7 radio flux, all of the indices are composites with data from different sources fitted together using various corrections and regressions.

In the following, we make no assumption about the long-term behaviour of SSI at each wavelength. To fill the gaps in the SATIRE-S reconstruction, we decompose each wavelength of SATIRE-S SSI into long and short-term components. This is done to avoid any bias in the long-term trends of the SSI reconstruction that might be introduced by the indices used to fill gaps. The detrended time series are produced by subtracting the smoothed time series from the original.

Gaps in the smoothed time-series are filled by linear interpolation. Most gaps are short: 37% of dates have a neighbouring data point; 70% of dates are within 6 days of another; 90% are no more than 27 days apart. Therefore, the long-term trend over the vast majority of the reconstruction is well approximated by a linear interpolation. Only five periods without data exceed two solar rotations, the longest of which is a 282 day period around the solar minimum of 1976. It is likely that a linear interpolation over this length of time is not appropriate to recreate the actual medium-term variation. However, the advantage is that no assumption is made about the expected behaviour of each wavelength or that each wavelength should vary in the same way as the best agreeing index.

To account for the rotational variability and fill gaps in the detrended SSI time series, we compare each detrended index with each detrended SATIRE-S SSI wavelength. We do

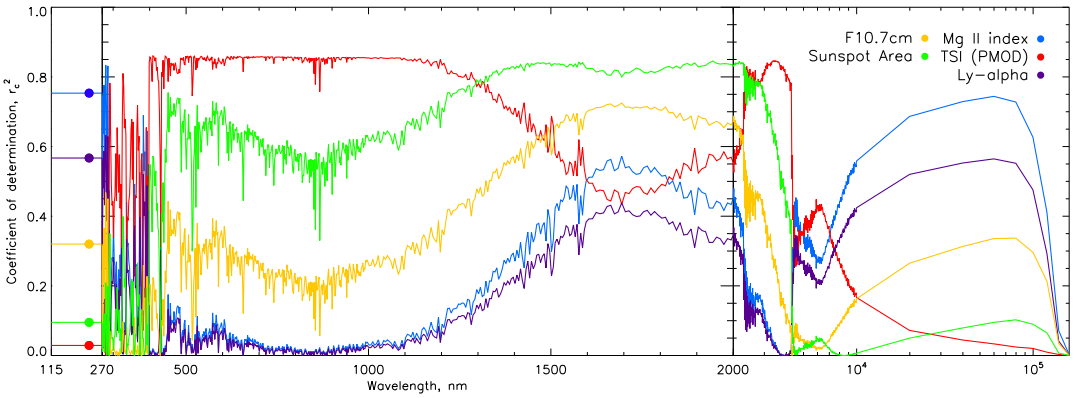


Figure 6.4: Correlation of determination, r^2, for each wavelength of SATIRE-S SSI compared with five measures of solar variability. See legend for detail of indices in plot.

this for all wavelengths above 270 nm and for the integrated region between 220-240 nm. Below 270 nm, all time series are determined from the integrated 220-240 nm region so only this one time series needs to be filled to account for all wavelengths in this wavelength region. In Fig. 6.4 the coefficient of determination, r_c^2 (the square of the correlation coefficient, r_c), shows how well each detrended index can account for the rotational variability at each wavelength. Colours are the same as in Fig. 6.3 and given in the legend. The scale on the x-axis is linear between 115 and 2000 nm and logarithmic above. The constant values of r_c^2 below 270 nm reflect the use of the 220-240 nm region, where the dots are located, to determine the wavelengths in this region.

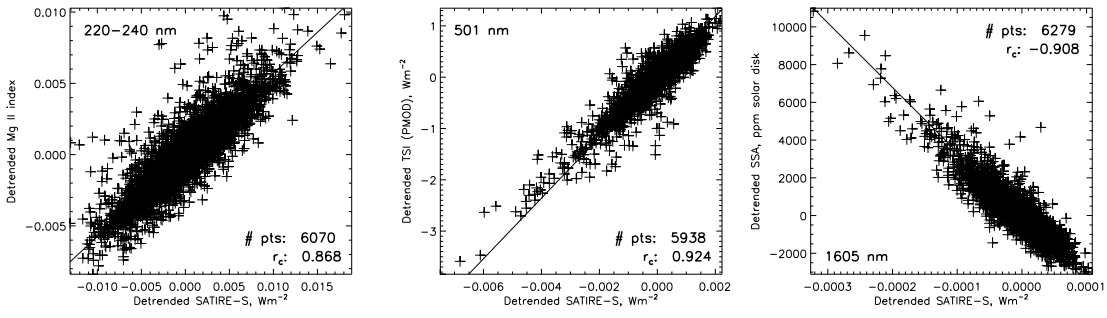


Figure 6.5: Regressions of detrended SATIRE-S wavelengths with the most common detrended index used to fill gaps in three spectral regions: (**left**) the integrated region 220-240 nm against the Mg II index; (**middle**) 501 nm against TSI; (**right**) 1605 nm against the SSA. The solid line is the line of regression. The number of dates (# pts) compared and the correlation coefficent, r_c are also shown.

Below 300 nm, the Mg II index agrees best with SATIRE-S SSI wavelengths, accounting for around 70 - 80% of the rotational variability. The Mg II index is a measure of the ratio of the emission reversal in the line core of the Mg II line (at 280 nm) and the Mg II wings. The line core is formed higher up in the solar atmosphere than the wings and is associated with regions of strong magnetic flux, as found in faculae and plages. Faculae dominate spectral variability below 300 nm, so the agreement of the Mg II index with these wavelengths is to be expected.

Between 300 and 400 nm, there is a transition between the dominance of faculae and of sunspots on spectral variability. This is reflected in the behaviour of the indices over this spectral interval where correlation coefficients are highly variable and quite poor at some wavelengths. The lower rotational agreement is not so much of a concern for cycle-length changes in these wavelengths, which is more important in the analysis of the long-term impact of SSI on, e.g., stratospheric chemistry as in chapter 7.

The regions between 400 and 1300 nm and 2300 and 4000 nm are best represented by TSI, which accounts for more than 80% of the variability at almost all of these wavelengths. TSI is replaced by the SSA and F10.7 cm radio flux between 1400 and 2200 nm as being the best representative indices of rotational variability; this region contains the opacity minimum and represents the deepest directly observable layers of the photosphere where sunspots and pores dominate and faculae have little influence. The F10.7 cm radio flux is formed from plasma trapped in coronal loops anchored to sunspots, so, although it is not formed in the photosphere, the F10.7 cm radio flux is highly correlated with the SSA on rotational time scales and hence the modelled irradiance at these wavelengths. Between 4000 and 10 000 nm none of the indices agree well with rotational variability in SATIRE-S SSI. As in the 300-400 nm region, there is a transition between a stronger influence from pores and sunspots below 4500 nm and faculae above; the contrast of faculae increases monotonically from here while penumbral and umbral contrasts remain relatively constant.

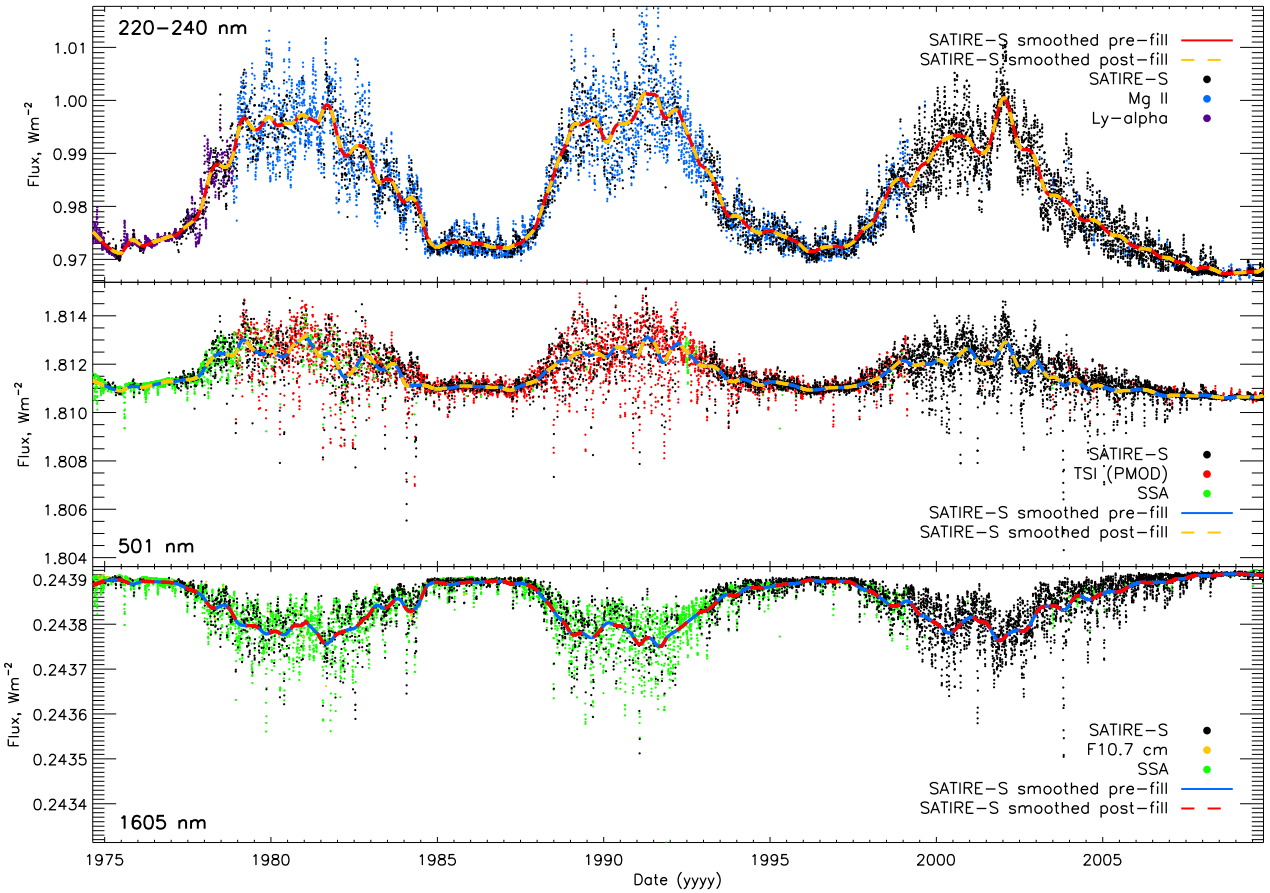


Figure 6.6: Three examples of the gap filling of different SATIRE-S spectral regions: (**top**) the integrated region 220-240 nm filled with the Mg II index (blue) and Ly-α (purple); (**middle**) 501 nm filled with the PMOD composite TSI (red) and SSA (green); (**bottom**) 1605 nm filled with the SSA and 10.7 cm radio flux (yellow). Each plot also shows the smoothed trend in the pre-gap filled (dashed, coloured lines) and post-gap filled time series.

Detrended gaps are filled in the following way. The regression between each index and every wavelength in SATIRE-S is calculated. Then on each date, the index with the highest value of r_c^2 is used to calculate the missing detrended SSI value for the gap. If data for that index does not exist on that particular date, the index with the next highest r_c^2 is used, and so on until all gaps are filled. Usually this only requires the two highest rated indices. In Fig. 6.5, three examples of the regressions are given. The integrated region of 220 - 240 nm and the individual wavelengths of 501 and 1605 nm are plotted against their highest rated index, Mg II, TSI and SSA, respectively. These plots show the good, linear agreement found at these wavelengths with their respective indices.

Finally, the detrended and smoothed time series are combined by addition to produce a spectral dataset that reflects the long-term variability of SATIRE-S while retaining rotational consistency with SATIRE-S at most wavelengths, i.e. the short-term variability at each wavelength is deduced from the most representative index, but scaled to the amplitude of SATIRE-S. The filled SSI time series for the same wavelengths as in Fig. 6.5 are shown in Fig. 6.6. In these plots the black dots represent spectral data produced directly by SATIRE-S, while coloured dots represent the indices used to fill the gaps. The colours for each index are as described in the plots. The smoothed time series before gaps are filled are shown as solid curves (red, top; blue, middle and bottom), and as dashed curves after the gaps have been filled (yellow, top and middle; red, bottom). These smoothed time series show that the long-term trend is almost completely unaffected by the filled gaps and supports the method of filling gaps by decomposition to retain the long-term behaviour of each wavelength as suggested by SATIRE-S.

According to Fig. 6.4, most wavelengths have at least one index accounting for over 70% of the variability. The TSI and Mg II index begin in 1978 and do not cover the period from 1974. The other indices with lower values of r_c^2 are used to fill this earlier period and, therefore, reduce the certainity of the period prior to 1977 in addition to the largest gaps in the long-term trends also being during this period. However, after 1992, the spectral reconstruction is relatively free of gaps and reflects the SATIRE-S model very well.

6.2 Comparison of SATIRE-S with the NRLSSI model

In this section we compare the, now complete, SATIRE-S SSI dataset with the National Research Laboratory Solar Spectral Irradiance (NRLSSI) dataset, the most widely used,

empirically derived, model of solar spectral irradiance (Lean et al., 1997; Lean, 2000; Lean et al., 2005).

NRLSSI covers the wavelength range 0-100 000 nm over the time period from 1950 to the present day. For the purposes of this comparison, we consider only wavelengths above 120 nm. NRLSSI uses the sunspot blocking function (PSI) and Mg II index as proxies for sunspots and faculae, respectively. Each wavelength is parameterized through multiple regression: the 120-300 nm region uses detrended UARS/SOLSTICE (Woods et al., 1993) time series in consort with the long-term behaviour of Mg II and PSI indices; above 300 nm, facular and sunspot contrasts from Solanki and Unruh (1998) are used. The multiple regression using detrended observed datasets, below 300 nm, is done to avoid long-term degradation and uncertainty in the UARS/SOLSTICE instrument introducing inaccurate long-term trends; the variability in wavelengths calculated using observations therefore assume that long-term variation changes in the same way as rotational.

Due to the wide spectral bandwidth and long period of coverage, the NRLSSI model has been used to investigate the impact of solar radiative forcing on, for example, stratospheric temperatures and ozone concentration, in e.g. recent studies by Gray et al. (2009); Meehl et al. (2009); Haigh et al. (2010) and Merkel et al. (2011).

6.2.1 TSI

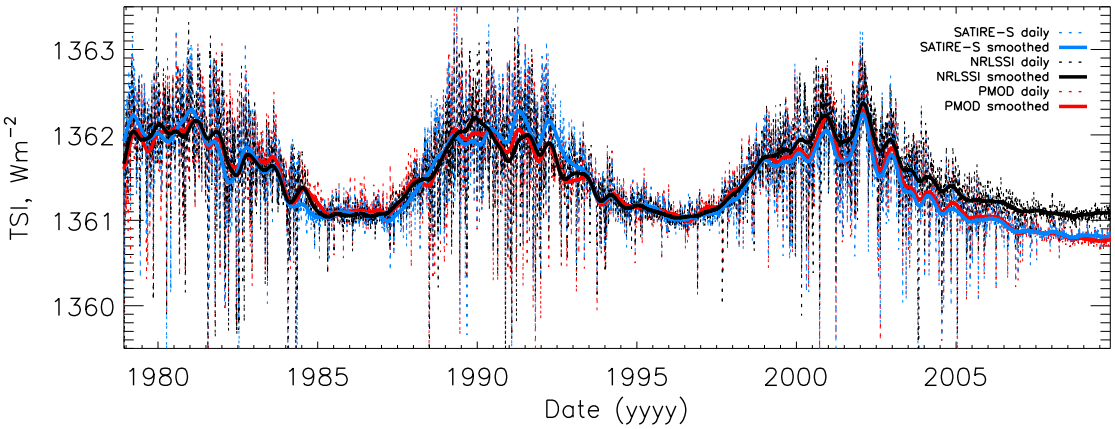


Figure 6.7: Daily (dashed lines) and smoothed (solid) TSI time series of SATIRE-S (blue), NRLSSI (black) and the PMOD composite (red) between 1978 and 2010. Only data that exists in all three time series have been plotted.

It is worth briefly considering how the NRLSSI TSI dataset differs from SATIRE-S TSI to give an indication of the comparative variability over all wavelengths between the two models. Here we take the SSI datasets from both models and integrate over the full range of wavelengths. In Fig. 6.7, daily (dashed lines) and smoothed (solid) time series are plotted for SATIRE-S (blue), NRLSSI (black) and the PMOD composite (red) between 1978 and 2009. Data are only compared when they exist in all three time series, i.e. for 10618 days. NRLSSI and PMOD have been normalised to the mean of SATIRE-S prior to 2000. Both SATIRE-S and NRLSSI agree well with the PMOD composite (Fröhlich, 2003).

In Tab. 6.1, the correlation coefficients and the standard deviation of the differences are given for original and detrended timeseries. SATIRE-S correlates better with PMOD than NRLSSI. This is the case for the original data as well as on a rotational basis. The standard deviation of the differences is a measure of how large deviations are, on average, between the compared datasets. Again, SATIRE-S shows a better agreement in both original and detrended comparisons, conveyed by the lower standard deviations.

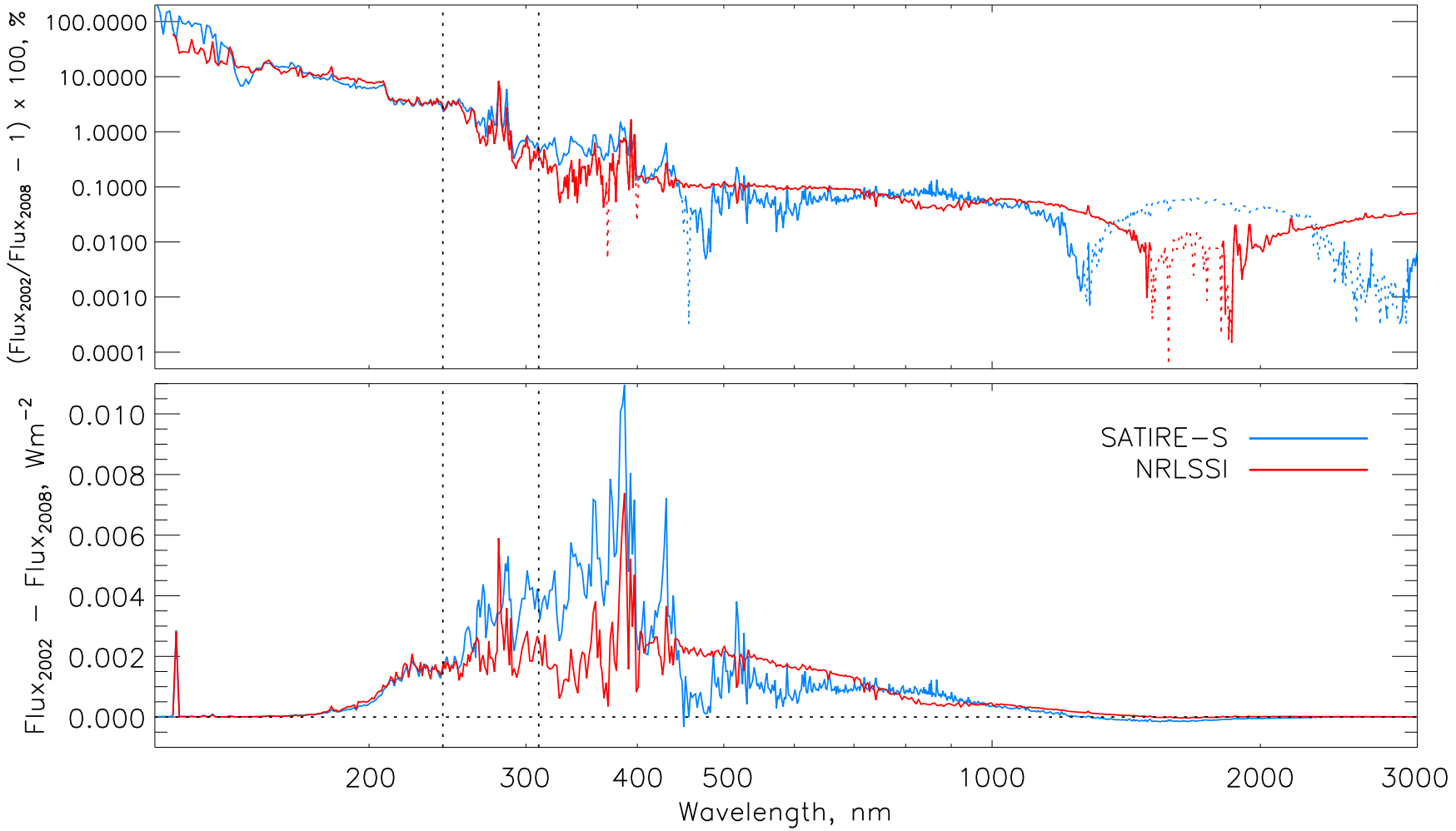


Figure 6.8: Changes between cycle 23 maximum and minimum for the SATIRE-S (blue) and NRLSSI models (red). (**top**) The percentage change in flux relative to the 2008 minimum; solid lines represent a decrease in flux while dotted lines represent an increase. (**bottom**) The absolute change in flux in units of Wm^{-2}; the horizontal dotted zero line acts as a guide. In both plots the vertical dotted lines are at 242 and 310 nm.

Table 6.1: Comparison of original and detrended TSI between the PMOD composite and NRLSSI and SATIRE-S models for 10618 days between 1978 and 2009.

Original data			
Set 1 (S1)	Set 2 (S2)	$r[r^2]$	σ(S1-S2) [Wm$^-$2]
PMOD	NRLSSI	0.918 [0.843]	0.23
PMOD	SATIRE-S	0.959 [0.920]	0.17
Detrended data			
Set 1 (S1)	Set 2 (S2)	$r[r^2]$	σ(S1-S2) [Wm$^-$2]
PMOD	NRLSSI	0.872 [0.760]	0.20
PMOD	SATIRE-S	0.949 [0.901]	0.13

The most significant difference between NRLSSI and SATIRE-S TSI is that only the latter matches the decline seen in PMOD between the minima of 1996 and 2008. Fröhlich (2009) showed that the Mg II index and F10.7 cm radio flux, used in NRLSSI as a proxy for long-term variation, do not exhibit the same strong secular variation that may have occured in TSI between cycle 23 minima. This may be the reason why NRLSSI is only able to reproduce secular TSI and SSI variability in agreement with PSI and Mg II indices, though it should be noted that the uncertainty of the cycle minima in PMOD is also in agreement with no change between minima, as the NRLSSI model suggests.

6.2.2 SSI

We now consider the spectral differences between the two models, using the difference between two six-month periods centred on 2002 February 1, the second and highest peak of cycle 23 as in Fig. 6.7, and the cycle 23/24 minimum centred on 2008 December 15. This provides the largest range of change in cycle 23. For this comparison NRLSSI is normalised to SATIRE-S over the six-month minimum period centred on 2008 December 15.

In the upper plot of Fig. 6.8 the percentage change between cycle 23 maximum and minimum is shown with a logarithmic y-axis, while the lower plot shows the absolute change in flux on a linear y-axis. The dotted lines in the upper plot represent negative changes, i.e. an increase during the descending phase of cycle 23. The wavelength spacings in the two models are different, so NRLSSI has been interpolated onto the SATIRE-S wavelength grid using the astrolib *interpol* routine.

The regions below 242 nm and between 242 and 310 nm are important in ozone pro-

duction and destruction processes in the stratosphere (see next chapter), so these have been highlighted with vertical dotted lines. Below 250 nm, the two models roughly agree on the changes between maximum and minimum, with significant differences becoming apparent only below ~140 nm. NRLSSI flux integrated over 120-242 nm, a region important for the photodissociation of O_2 in the Earth's atmosphere, is ~6% larger between maximum and minimum than SATIRE-S. Larger differences in the total energy (lower plot) are apparent above 250 nm with almost twice the change in flux in SATIRE-S at wavelengths around 310 nm. The total change of the integrated flux over 242-310 nm is 50% larger in SATIRE-S than NRLSSI; this region is important in the photodissociation of O_3. Below 270 nm, SATIRE-S uses UARS/SUSIM to calculate the level of absolute flux, while retaining some of the variability calculated by SATIRE-S (see section 6.1.3). NRLSSI uses UARS/SOLSTICE and the Mg II index below 300 nm. Note that the uncertainty in the long-term trends reach cycle-variability below 300 nm in both instruments Woods et al. (1996). The divergence in the change in flux longward of 250 nm in the two models reflects these differences.

Around 370 nm, changes in spectral irradiance are three times larger in SATIRE-S, than in the NRLSSI model, and remain larger up to 440 nm. For most of the visible wavelength range, between 440 nm and 1250 nm, NRLSSI changes more than SATIRE-S. Both NRLSSI and SATIRE-S display negative variability in the IR, though this occurs over a wider range and with a higher variability in SATIRE-S than NRLSSI, the latter of which only shows negative variability between 1500 and 1850 nm. Above 300 nm, the two models reflect different responses because both models have different ways of calculating the spectral variability. NRLSSI uses the variability of the Mg II and PSI indices while SATIRE-S uses the limb-dependent contrasts derived from full disk images. Also, the two models use slightly different model atmospheres, with SATIRE-S employing those from Unruh et al. (1999) and NRLSSI using Solanki and Unruh (1998).

The differences discussed above will cause different responses in stratospheric ozone concentration, from wavelengths below 310 nm, and a difference in the heating of the atmosphere, from all wavelengths. The cycle length spectral variability is larger in the recent observations by the SORCE instrument. In the previous chapter, I compared the SATIRE-S model with SORCE/SIM Harder et al. (2005a,b) observations between 200 and 1630 nm. In the next section, I focus on a comparison between SATIRE-S and SORCE/SOLSTICE, which also monitors UV spectral variability during the descending phase of cycle 23.

6.3 SORCE/SOLSTICE Comparison

In this section we focus on a comparison between SORCE/SOLSTICE and the two models between the wavelengths of 120 and 310 nm. Another UV instrument, TIMED/SEE, has been observing spectral irradiance below 190 nm at the same time as SORCE/SOLSTICE, but SORCE/SOLSTICE is used to correct long-term degradation trends in TIMED/SEE. A comparison of the SORCE/SOLSTICE and TIMED/SEE trends, not shown here, showed very good long-term agreement over the declining phase of cycle 23 below 180 nm. Agreement is less good above 180 nm where SORCE/SOLSTICE changes between the FUV and MUV channels.

6.3.1 SORCE/SOLSTICE Dataset

The SOLar STellar IrradianCE (SOLSTICE) instrument (McClintock et al., 2005a) was launched in 2003 on board the SORCE satellite (Rottman, 2005), along with the SIM instrument. Observations cover the UV region 115-310 nm and daily data are available from 2003 May 14 to present day. Instead of using onboard calibration lamps as is the case for SORCE/SIM, the observed spectra are compared to 18 bright, blue calibration stars with low variability to allow degradation calibrations to be performed.

SOLSTICE has two identical instruments, called A and B. SOLSTICE A takes regular measurements and SOLSTICE B is used less often to allow cross-calibration corrections for the degradation of SOLSTICE A. Due to a problem with the SOLSTICE A entrance aperture, SOLSTICE A cannot observe the calibration stars, so this is performed by SOLSTICE B and used to help correct for any drift between the two instruments.

The instruments have two channels with the far-UV (FUV) channel covering 115-180 nm and the mid-UV (MUV) channel covering 170-320 nm. Version 11 data, released in January 2011, has improved degradation corrections over the previous version, most significantly in the MUV channel. The degradation correction provides a long-term uncertainty of 0.5% per year. The result of the improved degradation corrections is a change in the observed absolute levels, and relative changes, in spectral irradiance (private communication, Marty Snow). The focus is on version 11 data, but version 10 is plotted in Fig. 6.9 as it is used in the next chapter.

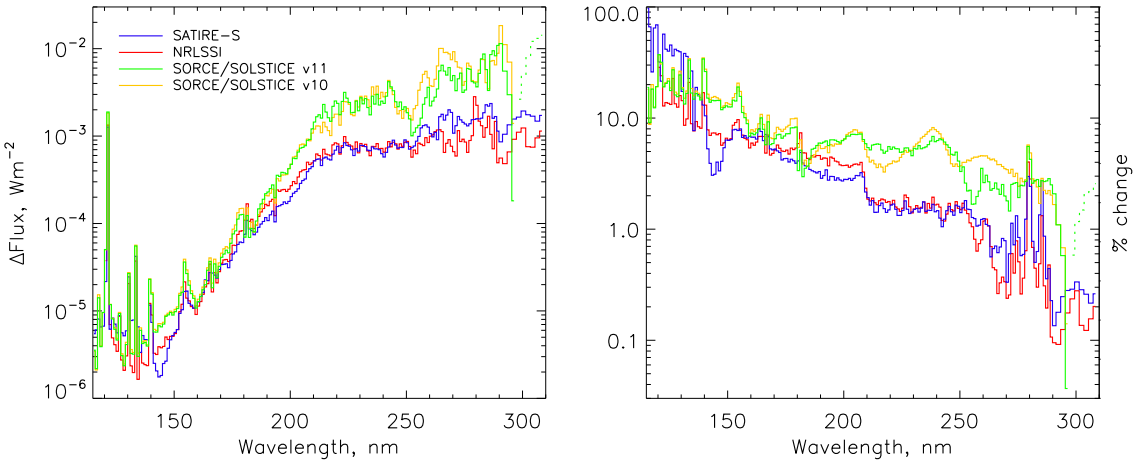


Figure 6.9: (**left**) The absolute change in flux between six month periods centred on July 2003 and December 2008 and (**right**) the percentage change, as a function of wavelength for the SORCE/SOLSTICE instrument (green) and the SATIRE-S (blue) and NRLSSI (red) models. Dotted SORCE/SOLSTICE lines above 298 nm indicate the change is negative. Wavelengths are 1 nm width bins below 290 nm and 2 nm bins above.

6.3.2 Comparison with models

In Fig. 6.9 the change in spectral irradiance between 115 and 310 nm is shown for SATIRE-S (blue), NRLSSI (red), SORCE/SOLSTICE version 11 (green) and version 10 (orange). The left plot gives the absolute difference in flux, ΔF, between the six month averaged periods centred on August 2003 and December 2008. The right gives the percentage change by dividing ΔF (left plot) by the mean absolute values from the December 2008 averaged-period. For the following discussion, focus is given to absolute changes in the left plot.

At 120 nm, SORCE/SOLSTICE shows roughly equal agreement with the models for ΔF. Except at around 140 nm, this agreement persists up to ~170 nm. By 190 nm SOLSTICE shows a two-fold change in flux relative to NRLSSI and SATIRE-S. Between 210 and 240 nm both models show approximately three times less change in flux than SOLSTICE, and between 240 and 250 nm this further decreases, with the models showing only a fourth of the change. The two models show different cycle variability above 250 nm, such that SOLSTICE shows, on average, nearly four times larger change than NRLSSI, and 2.5 times the change in SATIRE-S.

Solar cycle variability above 290 nm exceeds the long-term uncertainty of SOLSTICE (private communication, Marty Snow), so the long-term trends here cannot be relied upon for accuracy. Indeed, above 290 nm, SOLSTICE shows larger than five times the change

of either model and above 298 nm the change in flux for SOLSTICE becomes negative, so that it increases while TSI decreases during the declining phase of cycle 23. This negative response is very large (nine times larger than SATIRE-S with the opposite sign). Neither of the models, nor SORCE/SIM (see Fig. 5.8), display this negative variability, but it cannot be ruled out.

6.4 Conclusions

In this chapter a spectral dataset spanning 1974 to 2009 and covering 115 to 160 000 nm has been presented. It is based on the reconstruction procedure as documented in chapter 3. The preparation procedure included a correction for the LTE approximation below 270 nm using the work of KS06. The gaps in the data were filled by decomposing the time series at every wavelength into long and short-term variability. Long-term trends from SATIRE-S were retained by interpolating over gaps, while the detrended gaps were filled using the best-agreeing index of the five compared.

The model was then compared with the widely used NRLSSI model. First, the spectrally-integrated NRLSSI TSI was compared with SATIRE-S TSI and the PMOD composite. It was found that on all time scales PMOD agreed better with SATIRE-S than with NRLSSI. NRLSSI also did not show the decline between cycle 23 minima in 1996 and 2008 that both PMOD and SATIRE-S show.

I have presented spectral comparisons between the two models between 2002 and the minimum of 2008. The absolute and relative change in flux is in good agreement below 250 nm while they show different behaviour above. Between 250 and 430 nm, the change in flux of SATIRE-S is generally ~50% larger than NRLSSI, and at some wavelengths exceeds four times the change in flux. Above 430 nm NRLSSI and SATIRE-S repeatedly swap over from showing the largest change from maximum to minimum. They both show a negative response around the opacity minimum of 1600 nm, though SATIRE-S shows negative cycle changes over a larger range of wavelengths than NRLSSI.

Finally, the two models were put into context in the UV below 310 nm through a comparison with the newly updated version of SORCE/SOLSTICE data. This comparison showed that both models display significantly lower cycle variability than SOLSTICE for almost all wavelengths above 190 nm. The change in flux of SOLSTICE/SORCE is

similar to that of SIM between 200 and 310 nm and the given long-term uncertainty at 310 nm for SIM is the same as for SOLSTICE. However, it is also worth recalling that the instrumental noise in SIM is comparable to solar variability at and below 230 nm, so SOLSTICE is probably more accurate than SIM at capturing variations in this region (see section 5.2.3).

Only the cycle variation was considered when comparing SORCE/SOLSTICE because this puts the comparison of the three spectral datasets in the next chapter into context when considering their impact on stratospheric ozone.

In the near future the SATIRE-S spectral data, and an accompanying publication based on this and the following chapter, will be made available at http://www.mps.mpg.de/projects/sun-climate/data.html.

Chapter 7

Investigation into the effect of varying solar irradiance on ozone concentration in the stratosphere

Overview

The different spectral irradiance changes in SATIRE-S, relative to the NRLSSI model and the SORCE/SOLSTICE instrument, provide an opportunity to test the differing impacts that they have on Earth's atmosphere. While the influence of solar irradiance on the heating of the stratosphere is of importance, recent studies have focused on variations in observed ozone concentration, which have a direct influence on the stratospheric temperature and an indirect influence on stratospheric circulation, though the latter effect is not considered by the model and is outside the scope of this investigation. The primary aim of this chapter, then, is to investigate the sensitivity of changes in the concentration of ozone in the stratosphere to different changes in solar spectral irradiance. A two-dimensional, radiative-chemical transport atmospheric model by Harwood and Pyle (1975) is used to investigate this by keeping all parameters the same and only varying the cycle maximum to minimum change in flux, ΔF, at specific wavelengths in the UV. The work does not intend to make any conclusions about the accuracy of observations or the models, though these are discussed. Nor does it go into great detail regarding the physical processes underlying the atmospheric model as the focus remains on assessing the potential influence of varying the change in solar spectral irradiance.

7.1 2D Atmospheric model

The atmospheric model (referred to as HP model, henceforth) used in this chapter is a radiative-chemical transport atmospheric model. It has been used in many studies to investigate the dynamics and chemical interations of the Earth's atmosphere, ranging from the upper troposphere into the mesosphere, e.g. Bekki et al. (1996) and Warwick et al. (2004). It calculates the heat, momentum and chemical-constituent concentration in a time-dependent 2D zonal mean model. The model considers a spherical Earth with seasons, but with no inclusion of land topography or oceans. The model is resolved into 19 latitudes and 29 altitudes considered in pressure units, up to ~ 80 km.

The main inputs for the model are the solar spectral irradiance, sea surface radiative scheme and eddy momentum fluxes. These inputs allow the model to consider transport and feedback resulting from a change in the solar irradiance with the atmosphere through heating via absorption and chemical reactions. For example, the photochemical reactions involved in ozone production and destruction processes are responsible for the largest contribution to heating in the stratosphere through the absorption of short-wave solar flux below 310 nm, via photodissociation of O_2 and O_3 molecules and the subsequent release of energy through thermalisation. The absorption of this solar energy results in the heating of the atmosphere and leads to a modification of the zonal mean circulation, the movement of air and redistribution of heat around the planet, affecting the Earth's climate.

7.1.1 Variation in stratospheric ozone concentrations

This chapter is limited to considering the effect that solar radiation has on the ozone concentration in the stratosphere, or more specifically the change in ozone, ΔO_3. The work here uses that of Haigh et al. (2010) (H10) as a starting point. They investigated the impact that the significantly larger SORCE/SOLSTICE (used below 200 nm) and SORCE/SIM UV trends (used above 200 nm) had on ozone and temperature profiles in the stratosphere between ~25 and 60 km. Using SORCE/SIM, therefore, included the unusual variations in the visible that are out of phase with TSI during the declining period of the solar cycle (see chapter 5). H10 considered the period between 2004 and 2007; this period also shows the largest change in cycle length spectral variation within the SORCE record. With regards to stratospheric ozone, H10 found that when SORCE data were used as the input, counter to the previous view, ozone concentrations increased above 45 km and decreased more strongly at 35-40 km. This was in agreement with the suggested changes in ozone at

those heights given by ozone concentration data from the Earth Observing System Aura Microwave Limb Sounder (AURA-MLS) instrument (Lay et al., 2005) and contrasted significantly with the results when using NRLSSI spectra in the HP model, which showed nearly three times less change in concentration at 35-40 km, and no inverse in concentration trends at higher altitudes.

Merkel et al. (2011) (M11) also investigated ozone concentrations using the Whole Atmosphere Community Climate Model (WACCM), which is a full 3D global climate model and has a more complex and realistic consideration of the Earth's surface, including the coupled chemistry, radiation and dynamics over a larger range of altitudes extending up from the surface. They also employed NRLSSI and SORCE data as the solar input and found qualitatively similar results to H10. M11 compared their results to the ozone observations from the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument onboard TIMED (Rong et al., 2009). SABER results further support the SORCE SF results from both WACCM and from H10 in so much as the magnitude of the change in ΔO_3. However, SABER observations suggest that the sign change between increasing and decreasing ozone concentration occurs at ~1 hPa (or ~50 km) while the WACCM and HP models using SORCE SF suggest it is lower, around 5 hPa (~40 km) and 2 hPa (~46 km), respectively.

However, M11 also presented the equatorially-averaged (±15°) ozone percentage differences from SABER as a function of pressure (or altitude) over two time periods, 2004-2007 as in H10, and 2002/3 to 2008/9 covering almost the full period of available SABER data. These results showed that the differences, quantitatively, were similar when comparing the shorter and longer periods of time, except at night above 0.1 hPa (~60 km). This is interesting because the change in TSI, and by consequence the solar UV, between 2002 and 2008 is up to three times larger than between 2004 and 2007. This may suggest that other mechanisms may be involved in ozone modulation, or that the observations suffer from some problems.

The advantages of using a complex and more relatistic model, such as WACCM, is that the results should reflect the complexities of the real world and the observations drawn from it. The drawback is that it is more difficult to disentangle the competing and connected influences and feedbacks of the variable you are considering. By using a simple model, such as the HP model, it is easier to understand the mechanism being investigated, e.g. the effect of changes in ΔF, but with the restriction that the comparative power of the model results with observational data is reduced. Even so, the HP model produces

accurate reconstructions of variability in the stratosphere, but is less reliable below about 25 km due to approximations in the dynamical framework. The two types of model support each other, and the qualitative agreement of changes in ozone, between the HP and WACCM models in H10 and M11, respectively, reflect this. The agreement also shows that the HP model is sufficient for the additional investigations of SF that are performed here.

7.2 Changes in ozone concentration resulting from changes in SF

In this chapter we consider only the effect of changes in the spectral irradiance on ΔO_3 in the stratosphere and lower mesosphere. SORCE/SOLSTICE displays much larger spectral variability than the NRLSSI and SATIRE-S model fluxes, both of which also display sufficiently different spectral variability to produce clear differences in ΔO_3 as a function of latitude and altitude. Therefore, an exploration is performed in section 7.2.3 by varying the total flux in some UV bands to better understand the sensitivity of the HP model. This will also enable an investigation into whether different amounts of ΔF in some spectral regions can evoke the same ozone response and reproduce qualitatively similar results to those presented by H10 and M11.

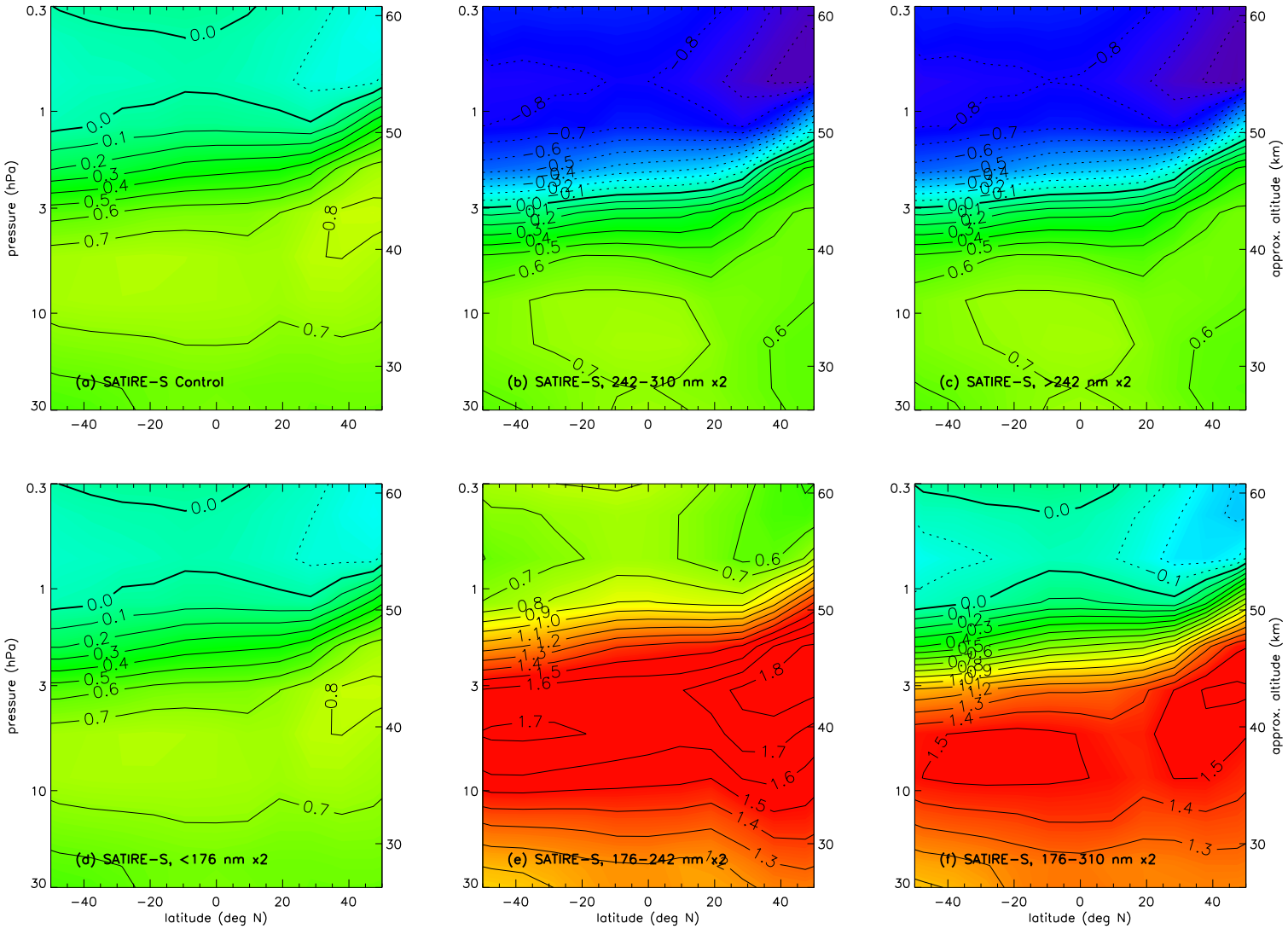


Figure 7.1: Contour plots of ΔO_3 in the middle atmosphere comparing the difference in ozone concentration bewteen 2003 and 2008. Plots (a) to (c) are on the top row and (d) to (f) on the bottom. **a** A reference run using the original spectral values from SATIRE-S; **b** as for (a) but with a doubling of the change in flux (ΔF) over 242-310 nm; **c** as for a but with a doubling of ΔF above 242 nm; **d** as for (a) but with ΔF below 176 nm doubled; **e** as for (a) but ΔF between 176 and 242 nm doubled; **f** as for (a) but ΔF between 176 and 310 nm doubled.

7.2.1 Investigating the effect of varying the spectral irradiance through control runs

First, the effect of spectral irradiance, with different ΔF in various spectral bands, on ΔO_3 must be established by investigating the regions of spectral irradiance important in O_2 and O_3 photodissociation which are key to modulating O_3 concentration. To identify and restrict the spectral regions that will be explored, tests are performed by comparing the change in ozone concentration between two six month periods centred on 2003 August 12, during an active solar period, and 2008 December 15, at solar minimum, i.e. the change in ozone concentration for a decline in solar irradiance. These periods will be used in all runs within this chapter. This period encompasses the full available time period and maximum forcing from SORCE/SOLSTICE. Figure 7.1a is the 'reference' run. This run uses SATIRE-S spectral irradiance with no modifications. The plots display latitude from -50 to +50 degrees North and height from 30 to 0.3 hPa, or approximately 25 to 60 km. Contours are given in 0.1% levels and negative changes (i.e. increases with decreasing solar activity) are shown with dotted contours. The thick solid line delineates between positive and negative changes. Colours support the contours with blue and purple reflecting negative changes and green, yellow and red being positive changes. The result of Fig. 7.1a is interesting in itself because there is a negative change in ozone concentration above ~55 km (as seen when using SORCE data, though of considerably lower magnitude), whereas using NRLSSI as an input into the model does not produce any negative changes over the altitude range presented, as in Fig. 2b of H10. All runs only consider the state of the atmosphere on December 25, i.e. a Northern hemisphere winter, which is why the lowest positive or largest negative, signal is seen in the Northern hemisphere, most notably at around 57 km.

Figures 7.1b and 7.1c consider the change in ozone concentration when doubling the change in flux, ΔF in all wavelengths between 242 and 310 nm and all wavelengths above 242 nm, respectively. The 242-310 nm flux causes the destruction of O_3 in the stratosphere through photodissociation by the reaction $O_3 + \frac{hc}{\lambda} \rightarrow O_2 + O$ (referred to as J_3). Shorter wavelengths, down to 176 nm, also contribute to O_3 destruction, but these photons also dissociate O_2 by the reaction $O_2 + \frac{hc}{\lambda} \rightarrow O + O$ (J_2). The release of the two oxygen atoms in this reaction enhances the production of O_3 by recombination with an oxygen molecule. Because only O_3 destruction is affected by modifying ΔF above 242 nm, this allows the effects of O_3 photodissociation to be separated from that of O_2 photodissociation. Figures 7.1b and 7.1c are almost identical in magnitude and spatial distribution. This is because only photons with wavelengths below 310 nm have enough energy to photodis-

sociate O_3 molecules. Spectral regions above 310 nm are involved in heating processes in the stratosphere involving O_3 (and other atmospheric constituents), but clearly these do not affect the concentration. This result means that only wavelengths below 310 nm need to be considered when exploring the sensitivity of the model over the range of latitudes and altitudes considered here.

In Fig. 7.1d, the change in flux in SATIRE-S below 176 nm, between 2003 and 2008, has been doubled. Once again, the change relative to Fig. 7.1a is barely noticable. This region is involved in O_2 destruction, mainly at altitudes exceeding 80 km which lies outside the altitude range considered here. Therefore, for this study, the region below 176 nm can be ignored.

In Fig. 7.1e, the ΔF over the spectral region that overlaps both O_2 and O_3 photodissociation processes is doubled, i.e. at 176-242 nm. The result is a larger change at all altitudes relative to the control run. At ~40 km, O_3 concentration has declined by an additional 1%, i.e. the forcings have evoked a stronger cycle reponse in ozone concentration, and the change above 50 km is now positive. This is because the higher flux in 2003 results in the production of more O_3, through the J_2 reaction and subsequent recombination of oxygen atoms with O_2, than is destroyed through photodissociation. The increase in O_3 has also been sufficient to offset the negative response in the reference run, above 50 km, in the mesosphere.

Figure 7.1f shows the effect of a doubling in the change of flux between 176 and 310 nm. It is similar to Fig. 7.1e in that the magntitude of the change in ozone is in agreement around 35-40 km, but above 50 km the ΔO_3 is once again negative as in the reference, though with a change in ozone around two to three times larger. Fig. 7.1f is the combination of forcings that result in Figs. 7.1b and 7.1e so this result confirms the processes discussed above.

The results of these runs reveal an important conclusion in the light of H10 and M11. These studies showed interesting results in observations supporting a negative response in the mesosphere (above 50 km) and strong positive response in the stratosphere. Figures 7.1a-f also reproduce changes in ozone concentration at similar altitudes and with similar magnitudes as in H10 and M11. However, when the ΔF is double in all wavelengths (Fig. 7.1f), they do not produce as strong a ΔO_3 in the stratosphere and mesosphere at the same time, showing that interactions between solar irradiance and O_2 and O_3 are not independent of each other. Although these test were performed by enhancing

fluxes in two wavelengths bands, the increase in ΔF was only by a factor of two, less than the flux difference between SATIRE-S and SORCE/SOLSTICE in these regions, as discussed in section 6.3.2 and shown in Fig. 6.9. The results in H10 were based on an earlier version of SORCE/SOLSTICE in combination with SORCE/SIM. It is, therefore, worth considering the effect that the newly recalibrated data from SORCE/SOLSTICE have on ΔO_3.

7.2.2 Change in ozone concentration using SORCE/SOLSTICE

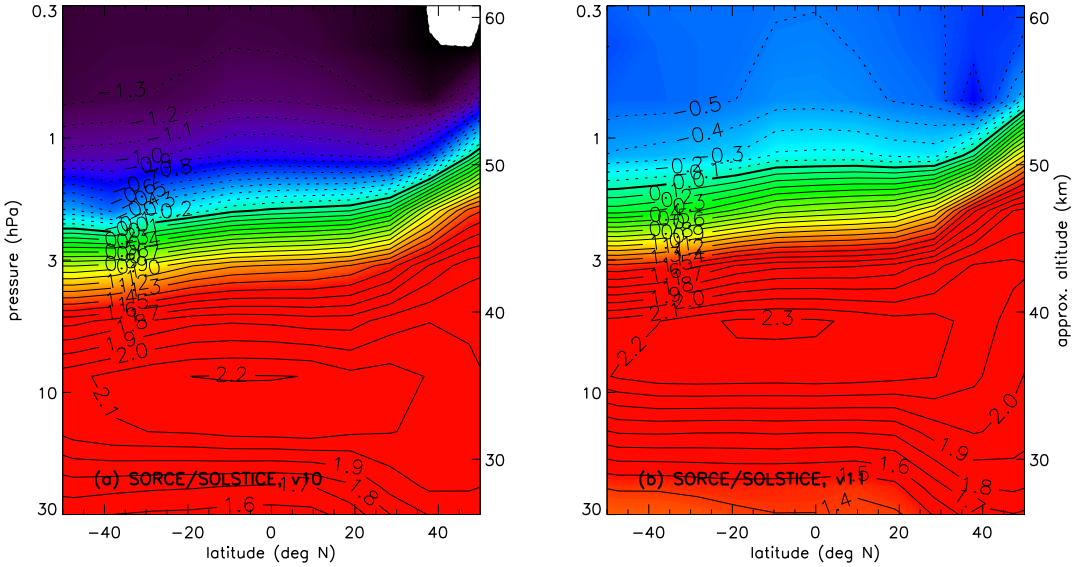


Figure 7.2: As for Fig. 7.1, ΔO_3 from the HP model using SORCE/SOLSTICE between 115 and 310 nm and SATIRE-S above as the solar input. The result of using SOLSTICE/SORCE version 10 data is shown in the left plot, and using version 11 is on the right.

Briefly, we consider the effect that SORCE/SOLSTICE has on the change in ozone concentration. The control runs of the previous subsection have shown that only wavelengths below 310 nm are important in modulating ozone concentration in the latitude-altitude region considered. Spectral observations are only available from SORCE/SOLSTICE between 115 and 310 nm, which is ideal for this investigation. The HP model requires input wavelengths up to 730 nm, but it has also been shown that the effect on ozone concentration above 310 nm is marginal. It is therefore possible to simply use SATIRE-S wavelengths above 310 nm.

In Fig. 7.2, results from a simulation using SORCE/SOLSTICE between 115 and 310 nm and SATIRE-S above are shown. Fig. 7.2a uses version 10 of SORCE/SOLSTICE data and 7.2b employs version 11 data. Exactly the same set up is used here as in the control run of Fig. 7.1a; colours are on the same scale (the white region at the top right of the Fig. 7.2a is due to the change in ozone exceeding the range of the contours).

The general spatial distribution is similar in both plots, though there are two key differences. Firstly, the magnitude in (negative) ΔO_3 in the mesosphere is significantly reduced when using level 11 data; the change at 55 km in Fig. 7.2b is less than half that (seen) in Fig. 7.2a. Secondly, the zero-line, the crossover from negative to positive changes, has also shifted up by 2-3 km and there is a slight enhancement at all levels below this. By considering the uncertainty in the long-term trends of 0.5% per year, the uncertainty in ΔO_3 is approximately $\pm 0.1\%$ at all altitudes. This uncertainty was considered in a simplistic way with only a blanket increase or decrease in the change in solar flux at all wavelengths below 310 nm.

This result reflects the reduced ΔF in version 11 relative to version 10, particularly in wavelengths above 250, as shown in Fig. 6.9, though the regions 207-225 and 242-250 nm show increases in ΔF. The SORCE/SOLSTICE data are still undergoing further corrections and so the results shown in Fig. 7.2 are likely to change somewhat in the future. The strength of change in ΔO_3 in the mesosphere in Fig. 7.2b is now of similar or lower magnitude than in the control run of Fig. 7.1b where flux in SATIRE-S between 242 and 310 nm was doubled.

The results highlights a key point, that similar ΔO_3 disribtuions can be reproduced when using SATIRE-S, even though its flux can be much lower at all wavelengths. This is because the O_3 concentration depends on the ratio of J_2 and J_3 photodissociation rates which are degenerate to certain combinations of changes in spectral irradiance. It may, therefore, be possible to recreate similar results to SORCE/SOLSTICE by varying the spectral variability in the SATIRE-S model artifically. To that end, we now consider how the spectral irradiance from the SATIRE-S model would have to be varied to produce the results seen in Fig. 7.2b from SORCE/SOLSTICE, though this is not an investigation into the validity of SORCE/SOLSTICE.

7.2.3 Exploration of changes in spectral irradiance using SATIRE-S

To investigate the potential range of ΔO_3 by varying ΔF in different spectral bands, some markers must be identified in the latitude-altitude plots and used as references while ΔF is varied. For this, the mean equatorial profile, averaged over ±14 degrees is used.

In Fig. 7.3, the mean equatorial profiles from SATIRE-S in Fig. 7.1a (SATIRE-S control, blue solid line), Fig. 7.1b (SATIRE-S with $\Delta F_{242-310nm}$, dotted blue line) and Fig. 7.1e (SATIRE-S with $\Delta F_{176-242nm}$, dotted blue line), as well SORCE/SOLSTICE versions 11 (solid black line) and 10 (dotted black line) from Fig. 7.2 are shown. These profiles make it somewhat easier to compare the general effect that changes in the modelled or observed ΔF have on ΔO_3.

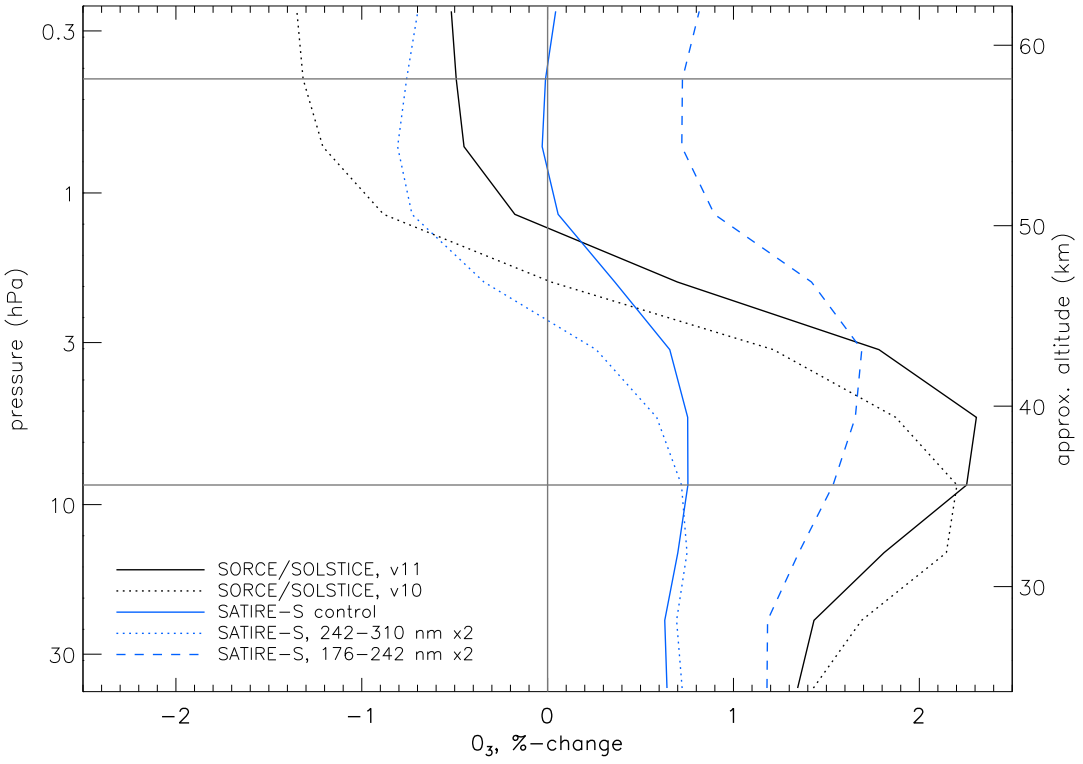


Figure 7.3: Equatorial (±14°) ΔO_3 profiles taken from Figs. 7.1a (blue, solid line), 7.1b (blue, dotted), 7.1e (blue, dashed) and from the left (black, dotted) and right plots (black, solid) of Fig. 7.2. The horizontal dotted lines indicate the marker levels used to produce the contour plots in Fig. 7.4.

From Fig. 7.3, the equatorial profiles allow markers of ΔO_3 at different altitudes to be chosen. Two are selected where the largest positive and negative values of ΔO_3 generally appear in all the plotted profiles, i.e. at 0.43 hPa (~ 58 km), where approximately the

largest negative changes are seen, and at 8.7 hPa (~ 36 km), where approximately the largest positive changes occur. These two altitudes are indicated by the horizontal grey lines. These two levels are also useful because, as indicated by the SATIRE-S profiles, an increase in $\Delta F_{242-310nm}$ produces large changes at 0.43 hPa without significantly affecting ΔO_3 at 8.7 hPa. This is not the case with $\Delta F_{176-242nm}$, which modulates the entire profile. The result of varying irradiance in both spectral bands should produce unique results that could potentially be used to identify changes in spectral irradiance are reasonably correct, if observations are accurate. This is not done in this study, but the sensitivity of the model to changes in solar irradiance can be investigated.

In Fig. 7.4, $\Delta F_{176-242nm}$ and $\Delta F_{242-310nm}$ are independently varied by multiplying each region by 0.2, 0.25, 0.33, 0.5, 1, 2, 3, 4 and 5 (data points are at the intersections of the dotted horizontal and vertical lines). Then, the value of ΔO_3 at 0.43 hPa (left plot) and 8.7 hPa (right) is plotted to produce the contour plots. Contours of ΔO_3 are given in levels of 1% and colours support the contours and aid understanding; colour represent the same level in both plots. The contours are smoothly varying in most regions in the plot. However, there are localised spikes in the plots at, e.g., coordinates (0.25, 1) and (2, 4). High resolution contour plots show that these spikes are usually very localised, but at the lower resolution of the plots in Fig. 7.4 they appear to be much larger. The source is not well understood, but could be the result of rounding errors within the code. These plots should be considered as illustrative only of the sensitivity of ozone to spectral variations (as determined by the HP model).

The contour plot for 0.43 hPa shows that the change in ozone concentration is dependent on ΔF in both spectral bands and the resulting value of ΔO_3 is more sensitive when ΔF is much stronger in all UV wavelengths, and less when the ΔF is much lower. At 8.5 hPa, the result is different. Varying $\Delta F_{242-310nm}$ only a marginal effect on the resulting ΔO_3, while almost all changes in ozone at that altitude result from varying $\Delta F_{176-242nm}$.

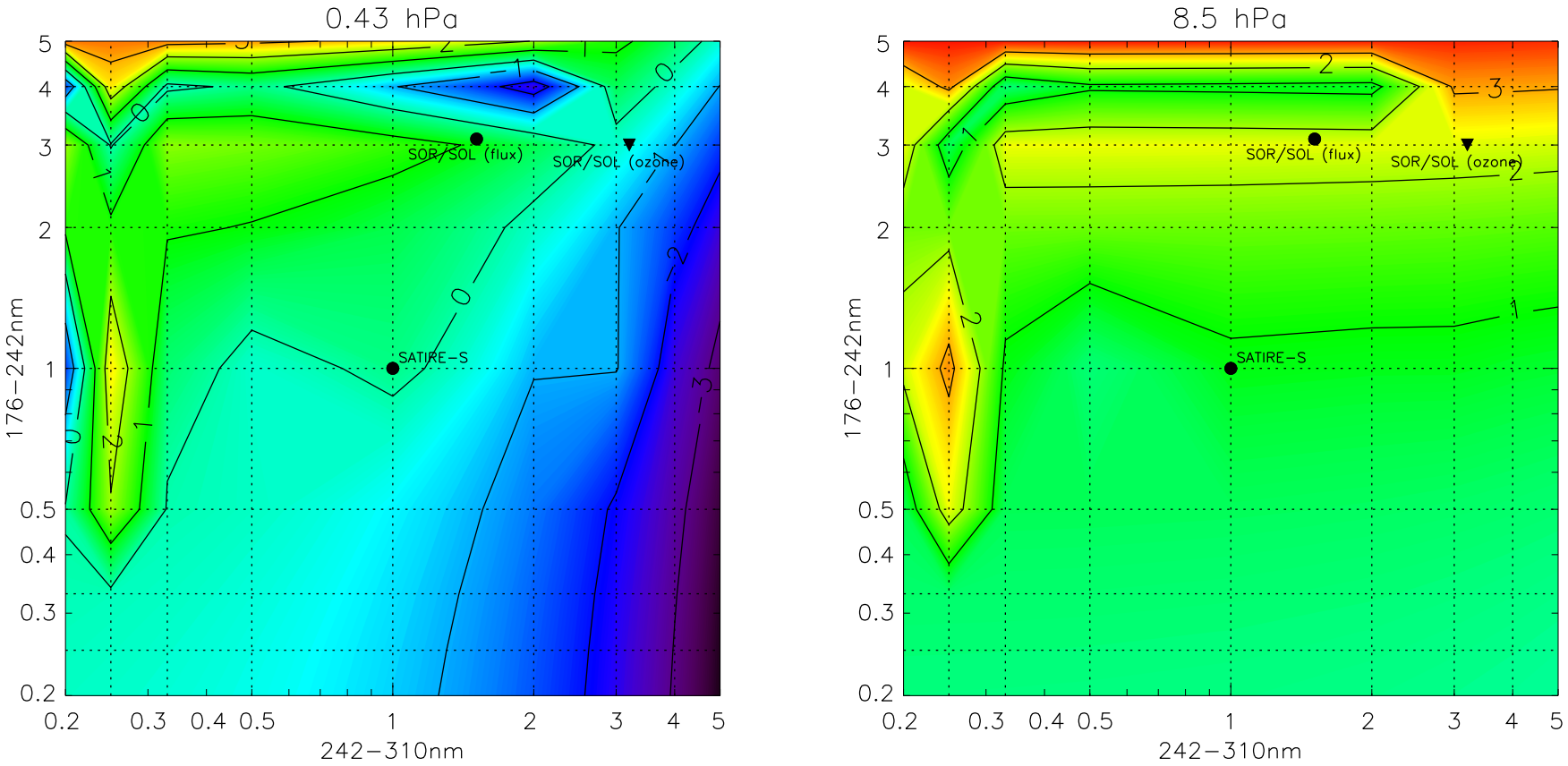


Figure 7.4: The ΔO_3 at two pressure levels (altitudes), 0.43 hPa (~60 km, left) and 8.5 hPa (37 km, right). The filled circle labelled 'SATIRE-S' marks the ΔO_3 with a ΔF from SATIRE-S with no enhancements. The filled circle labelled 'SOR/SOL (flux)' marks ΔO_3 that SATIRE-S would give if the integrated-fluxes were scaled to SORCE/SOLSTICE without consideration of spectral distribtion. The filled triangle labelled 'SOR/SOL (ozone)' marks the approximate position of ΔO_3 resulting from the SORCE/SOLSTICE irradiance distribution when used in HP in place of the SATIRE-S irradiance distribution.

By including the HP model results when using SORCE/SOLSTICE (version 11), the ozone feedback is revealed to be more complicated than a simple variation of ΔF in two broad spectral bands. The control run of SATIRE-S is marked (and labelled) on the contour plots with a filled black circle and labelled 'SOR/SOL (flux)'. The integrated-flux in SOLSTICE over 242-310 nm is 1.5 times larger than SATIRE-S and 3.1 times larger in the 176-242 nm band. This puts SORCE/SOLSTICE at the location near the top, middle of the contour plots and is also marked with a filled black circle. However, the ΔO_3 at this position is not what results from the HP model when the SOLSTICE fluxes are used in SATIRE-S, i.e. with the same change in integrated flux. Instead, the change in ozone concentration as marked by the filled black triangle (and labelled 'SOR/SOL (ozone)') reflects where SOLSTICE would appear on these contour plots relative to ΔO_3, when considering ΔO_3 in both contour plots together. The reason the filled circle and triangle are in different locations is possibly the result of ΔO_3 being dependent on narrower bands of ΔF than the ones used here or because the absolute levels are slightly different.

The contour plots show that there is a strong spectral dependence of ΔO_3 on ΔF which becomes more acute at higher levels of flux. It also shows, significantly, that using observations and atmospheric model results can not be considered at one altitude alone to support the trends seen in solar spectral irradiance observations. This is made difficult given that uncertainty increases to solar cycle levels in space-based UV satellite observations around 300 nm (Woods et al., 1996; McClintock et al., 2005b). The uncertainty in ozone observations should also be considered, as well as possible limitations on the atmospheric model.

7.3 Conclusions

In this chapter, the SATIRE-S model SSI and SORCE/SOLSTICE observational datasets have been used to test the sensitivity of changes in ozone concentration using the 2D atmospheric model of Harwood and Pyle (1975). Recent publications, e.g. Haigh et al. (2010) and Merkel et al. (2011), have shown interesting results using SORCE SSI in models that show some agreement with changes in atmospheric ozone concentrations. However, the results in chapter 5 show that SORCE/SIM displays very different cycle variability to previous UV observations and models. SORCE/SOLSTICE shows lower cycle variability than SIM, but this is still significantly larger than the SATIRE-S and NRLSSI models, as shown in the previous chapter.

The results of these comparisons suggest that caution is required in the use of stratospheric ozone observations to provide validation of any particular change in the solar spectrum. The work in this chapter is has attempted to better understand how sensitive atmospheric models can be to solar inputs in the light of the SORCE results.

It has been shown, as is already known, that the change of ozone between cycle maximum and minimum is dependent on the competing effects of O_2 and O_3 photodissociation in broad spectral bands. In addition to this, the work here indicates that the response is highly dependent on the spectral trends (possibly) in addition to the absolute flux level. This result emerges from experiments with the atmospheric model which shows that varying SATIRE-S to have the same change in flux as SORCE/SOLSTICE over two broad UV ranges does not reproduce the same result as actually replacing the individual spectral bands in the model solar input with the absolute levels from SORCE/SOLSTICE.

Further work is needed to understand better which spectral regions are most important and to what extent the absolute level is important in relation the the cycle changes in solar flux.

Chapter 8

Conclusions and Future Work

8.1 Thesis conclusions

In this thesis, I investigate the relationship between the solar surface magnetic fields and changes in the total and spectral irradiance of the Sun. This has been conducted through the SATIRE-S model and encompasses the period of solar observations from 1974 to 2009. The work presented within demonstrates that solar irradiance variations on all timescales over this period, including secular changes, can be understood in terms of the solar magnetic field without the need to invoke any other causes of irrradiance variation.

In chapter 2, the SATIRE-S model was explained and it was shown that despite using a simple approximation and relationship between the magnetic field strength and the area covered by faculae, the model reproduces TSI variations extremely well. In chapter 2, I explained how I reconsidered the way continuum intensity images and magnetograms were preprocessed and found that an improvement of 6% in the ability of the model to account for solar variability, on cycle-length timescales, could be made. While this may seem modest, the improvement brings TSI reconstructions to the level of agreement seen between independent measurements of TSI instruments.

In chapter 3, I reconsidered the input datasets in SATIRE-S previously used by Wenzler et al. (2006) to enable the reconstruction of TSI and SSI datasets that would be as independent of observations as possible. In particular, the reconstruction was extended to 2009 using SoHO/MDI data and the free parameter was set using SORCE/TIM data between 2003 and 2009, meaning that the free parameter was determined using data not used in any of the composites. The work, along with the choice of an appropriate method

of determining the free parameter from chapter 2, led to a reconstruction that was fully independent of the TSI composites after 1989, with an uncertainity range independent from observations at all times. The most important result is that the model recreates the change in TSI between the solar cycle minima of 1996 and 2008, in agreement with the change estimated by the PMOD composite. The small reconstructed change between the other two minima is also supportive of PMOD and less consistent with the other two. Because the reconstruction is independent of the composites during cycle 23, this is good evidence that changes in the photospheric magnetic field are responsible for solar irradiance changes on all timescales, including cycle and rotational variability. Due to the good agreement with, and relative independence from, observations, the model may now be used as a tool to assess the accuracy of the composites, especially over periods when individual TSI datasets are combined or an attempt to account for instrument degradation is made.

The result of chapter 3 showed that the SoHO/MDI magnetograms cannot be used to estimate the changes in the surface magnetic flux on timescales longer than the solar cycle. In chapter 4, I investigated how the magnetograms had changed over time. I found that changes in the focus setting and degradation in the early phase of the mission led to a change in the response of the instrument. I demonstrated that by taking into account just the deliberate focus changes, almost half of the change in the reported magnetic flux, between 1996 and 2008, could be accounted for. It, therefore, became clear that reconstructions of solar irradiance using MDI magnetograms need to take account of these changes, in addition to the degradation of the instrument, if conclusions are to be made about the Sun's behaviour over timescales exceeding the solar cycle.

The remaining chapters focused on the SSI dataset resulting from the TSI reconstruction in chapter 3. In chapter 5, a detailed comparison between SATIRE-S SSI and the SORCE/SIM instrument was performed. This is interesting because SORCE/SIM is the first instrument to provide calibrated, long-term trends in wavelengths above 400 nm, i.e. in the visible and IR. The work showed that SATIRE-S and SORCE/SIM rotational variability agrees very well at most wavelengths and especially in the visible. However, SORCE/SIM displays significantly larger cycle variability in UV than previous instruments and inverse trends, with respect to TSI, at longer wavelengths. This was unexpected and neither the SATIRE-S model, nor the NRLSSI model, can reproduce these trends. It does appear, however, that there may still be instrument degradation that has not been accounted for, as the largest changes occur during the first two years of operation, and the trend in the integrated-SSI does not agree with TSI from SORCE/TIM during that period.

Unfortunately, the uncertainty on integrated-SIM measurements are too large to allow any definate conclusions to made in regards to the accuracy of the trends.

In chapter 6, I prepared the SATIRE-S spectral irradiance dataset for publication and for use within the atmospheric and climate communities. This required using the correction of Krivova et al. (2006) to the spectral region below 270 nm so as to better account for the SATIRE-S LTE approximation. I also filled all data gaps, at all wavelengths, by decomposing the long and short-term variability to retain the model's long-term trends and rotational amplitudes. The data were then compared with the most widely used SSI model, the NRLSSI model. TSI agreement with PMOD was higher with SATIRE-S than NRLSSI on all timescales. The cycle variability was different in the two models at most wavelengths above 250 nm, which has implications when solar spectral irradiance is employed in atmospheric and climate models.

With that in mind, I tested the SSI from SATIRE-S in the simple 2D atmospheric model by Harwood and Pyle (1975) to compare how the spectral irradiance changes affected the stratospheric ozone concentration. In particular, the sensitivity of the atmospheric model was investigated. This is important because recent studies by Haigh et al. (2010) and Merkel et al. (2011) have shown clear signals in ΔO_3 in the descending phase of cycle 23 when using the large changes in SSI from SORCE/SIM and SORCE/SOLSTICE in agreement with some atmospheric satellite observations. SATIRE-S showed agreement in the direction of change in ΔO_3 as a function of altitude, but with a significantly lower magnitude of change. However, by artifically varying the change in spectral irradiance, I showed that qualitatively similar changes to those suggested by SORCE/SOLSTICE and SORCE/SIM could potentially be produced by different changes in spectral irradiance. A comparison of level 10 and 11 SORCE/SOLSTICE data in the Harwood and Pyle (1975) model also showed a large difference after additional calibration corrections have been applied. Therefore, the work from chapter 7 suggests that caution, along with a good understanding of the uncertainty and limitations of direct solar and atmospheric observations, is required before making firm, quantitative conclusions about the way the Sun influences the Earth's atmosphere.

In conclusion, the work in this thesis provides evidence that solar irradiance, especially TSI, can be explained on all timescales over the last 34 years. However, significant work still needs to be done, through both observations and modelling, to gain a better understanding of how solar spectral irradiance influences the Earth on climate timescales.

8.2 Future work

There are many avenues of research that follow from the results I have described in this thesis. In this final section, I outline potential, interesting directions this work could and should be taken in.

The model used Kitt Peak and SoHO continuum intensity images and magnetograms. There was some uncertainty in the KP/512 dataset that required a correction to all pre-1990 magnetograms Arge et al. (2002); Wenzler et al. (2006). It would be worthwhile to investigate whether Mount Wilson Observatory and Wilcox Solar Observatory magnetogram data can provide a better estimate of solar irradiance behaviour prior to 1989 if employed within SATIRE-S.

Magnetograms are not the only source of information relating to photospheric magnetic fields. Ca II K and H_α full-disk images also provide information through observations of plage regions in the solar chromosphere. For example, thousands of Ca II K spectroheliograms exist in overlapping historical datasets from the Arcetri Astrophysical Observatory and Kodaikanal Observatory between 1907 and 1995 (Makarov et al., 2004; Ermolli et al., 2009a). Other overlapping spectroheliogram timeseries also exist. If these images are well calibrated, it should be possible to relate the observed Ca II K, H_α intensities to facular brightness and use these to assess the secular variability of the Sun. However, it would need to be established that variations in low-level magnetic flux can be identified through these intensity observations. If so, then they could potentially provide an invaluable tool for assessing the climate length variability of the Sun.

Looking forward, there remain 18 months of MDI continuum intensity image and magnetogram data between the end of the reconstruction presented here and the cessation of MDI observations in April 2011. I have already done some work to prepare data up to October 2010, though there are problems with the data after October 2009, due to the age of the instrument, that still need to be resolved. Completing the dataset is vital to provide (almost a year of) overlap with the new Solar Dynamics Observatory (SDO) Helioseismic and Magnetic Imager (HMI) and extend the SATIRE-S reconstruction. HMI has 16 times the number of pixels and four times the resolution of MDI and will allow even higher precision reconstructions to be developed. Kok Leng Yeo (MPS) is already working on this, and I will work with him to complete the integration of the two datasets. Lockwood et al. (2011) have pointed out that the unusual solar minimum along with predictions for cycle 24 may mean that the Sun is heading into a period of much lower activity and, as a result,

potentially lower solar irradiance. The results from this thesis show that it is important to continue extending the SATIRE-S model into the future with HMI and continue to test if the evolution of solar magnetic flux continues to be the likely cause of changes in TSI on secular time scales. If the assumption holds, then this will allow stronger constraints on solar irradiance reconstructions back to the Maunder Minimum and beyond.

As for spectral irradiance variability, the trends seen by SORCE/SIM need independent validation. This will come when future spectral irradiance missions are launched. In addition, reanalysis of the data is shortly to begin on the SORCE/SIM data to assess if better estimates of the long-term trends can be achieved.

Work on SATIRE-S can and should continue. The following major enhancements to the model should be completed over the next few years. As has been made clear in this thesis, the model atmospheres and emergent intensities used in SATIRE-S have been calculated under the assumption of LTE, which has been demonstrated to be inaccurate, so exploratory calculations in some non-LTE regions below 400 nm will be performed to better estimate the variability in that spectral region. In addition, SATIRE-S relies on 1D atmospheric models that do not provide realistic geometric descriptions of faculae as a function of disk position and there is no prescription for the very small scale network contribution to long-term solar irradiance. These shortcomings, together, are being addressed by using the full three-dimensional MHD simulations of the photosphere using the MPS/University of Chicago Radiative MHD, or MURaM Vögler (2004); Vögler et al. (2005). The MuRAM results show similar magnetic flux and limb dependent contrast variation to those suggested by observations. Afram et al. (2011) have begun to develop spectral contrasts as a function of magnetic flux and limb angle. I will continue this work as a part of my RA. The results should fully constrain the model and remove the need for the free parameter. This will be done by relating the magnetic flux observed in a magnetogram pixel to the expected distribution of faculae and network components within. If successful, it will then lead to better estimates of solar irradiance variability and, therefore, a better understanding of influence that the Sun has on the Earth.

Bibliography

Afram, N., Unruh, Y. C., Solanki, S. K., Schssler, M., and Mathew, S. K.: 2009, *Proceedings of the International Astronomical Union* **5(Symposium S264)**, 63

Afram, N., Unruh, Y. C., Solanki, S. K., Schüssler, M., Lagg, A., and Vögler, A.: 2011, *Astronomy & Astrophysics* **526**, A120

Arge, C. N., Hildner, E., Pizzo, V. J., and Harvey, J. W.: 2002, *Journal of Geophysical Research (Space Physics)* **107**, 1319

Baliunas, S. and Jastrow, R.: 1990, Nature **348**, 520

Ball, W. T., Unruh, Y. C., Krivova, N. A., Solanki, S., and Harder, J. W.: 2011, *Astronomy & Astrophysics* **530**, A71

Balmaceda, L., Krivova, N. A., and Solanki, S. K.: 2007, *Adv. Sp. Res.* **40**, 986

Balmaceda, L., Solanki, S. K., and Krivova, N.: 2005, *Mem. S.A.It.* **76**, 929

Balmaceda, L. A., Solanki, S. K., Krivova, N. A., and Foster, S.: 2009, *Journal of Geophysical Research (Space Physics)* **114(A13)**, 7104

Beck, J. G.: 2000, *Solar Physics* **191**, 47

Beer, J., Blinov, A., Bonani, G., Hofmann, H. J., and Finkel, R. C.: 1990, Nature **347**, 164

Bekki, S., Pyle, J. A., Zhong, W., Toumi, R., Haigh, J. D., and Pyle, D. M.: 1996, *Geophysical Research Letters* **23**, 2669

Biermann, L.: 1941, *Vierteljahrschrift Astron.Ges.* **76**, 194

Bond, G., Kromer, B., Beer, J., Muscheler, R., Evans, M. N., Showers, W., Hoffmann, S., Lotti-Bond, R., Hajdas, I., and Bonani, G.: 2001, *Science* **294**, 2130

Brandt, P. N., Schmidt, W., and Steinegger, M.: 1990, *Solar Physics* **129**, 191

Browning, M. K., Miesch, M. S., Brun, A. S., and Toomre, J.: 2006, ApJL **648**, L157

Brueckner, G. E., Edlow, K. L., Floyd, IV, L. E., Lean, J. L., and Vanhoosier, M. E.: 1993, *Journal of Geophysical Research* **98**, 10695

Chapman, G. A., Cookson, A. M., and Dobias, J. J.: 1996, *Journal of Geophysical Research* **101**, 13541

Charbonneau, P.: 2010, *Living Rev. Solar Phys. 7, (2010), 3. URL (cited on 2011/11/11): http://www.livingreviews.org/lrsp-2010-3*

Crommelynck, D., Fichot, A., Lee, III, R. B., and Romero, J.: 1995, *Adv. Sp. Res.* **16**, 17

Danilovic, S., Gandorfer, A., Lagg, A., Schüssler, M., Solanki, S. K., Vögler, A., Katsukawa, Y., and Tsuneta, S.: 2008, *Astronomy & Astrophysics* **484**, L17

DeLand, M. T. and Cebula, R. P.: 2008, *Journal of Geophysical Research* **113**, 11103

DeLand, Matthew, T. and Cebula, R. P.: 2011, *Journal of Atmospheric and Solar-Terrestrial Physics, in press.*

Dessler, A. E., Burrage, M. D., Grooss, J.-U., Holton, J. R., Lean, J. L., Massie, S. T., Schoeberl, M. R., Douglass, A. R., and Jackman, C. H.: 1998, *Reviews of Geophysics* **36**, 183

Dewitte, S., Crommelynck, D., Mekaoui, S., and Joukoff, A.: 2004, *Solar Physics* **224**, 209

Domínguez Cerdeña, I., Sánchez Almeida, J., and Kneer, F.: 2006, *The Astrophysical Journal* **636**, 496

Dudok de Wit, T., Kretzschmar, M., Aboudarham, J., Amblard, P.-O., Auchère, F., and Lilensten, J.: 2008, *Advances in Space Research* **42**, 903

Ermolli, I., Criscuoli, S., Uitenbroek, H., Giorgi, F., Rast, M. P., and Solanki, S. K.: 2010, *Astronomy & Astrophysics* **523**, A55

Ermolli, I., Marchei, E., Centrone, M., Criscuoli, S., Giorgi, F., and Perna, C.: 2009a, A&A **499**, 627

Ermolli, I., Solanki, S. K., Tlatov, A. G., Krivova, N. A., Ulrich, R. K., and Singh, J.: 2009b, *The Astrophysical Journal* **698**, 1000

Feynman, J. and Ruzmaikin, A.: 2011, *Solar Physics* p. 335

Fisk, L. A. and Schwadron, N. A.: 2001, *The Astrophysical Journal* **560**, 425

Fligge, M., Solanki, S. K., and Unruh, Y. C.: 2000, *Astronomy & Astrophysics* **353**, 380

Fligge, M., Solanki, S. K., Unruh, Y. C., Froehlich, C., and Wehrli, C.: 1998, A&A **335**, 709

Floyd, L., Rottman, G., Deland, M., and Pap, J.: 2003, in A. Wilson (ed.), *Solar Variability as an Input to the Earth's Environment*, Vol. 535 of *ESA Special Publication*, pp 195–203

Fontenla, J., White, O. R., Fox, P. A., Avrett, E. H., and Kurucz, R. L.: 1999, *The Astrophysical Journal* **518**, 480

Fontenla, J. M., Avrett, E. H., and Loeser, R.: 1993, *The Astrophysical Journal* **406**, 319

Fontenla, J. M., Harder, J., Rottman, G., Woods, T. N., Lawrence, G. M., and Davis, S.: 2004, *The Astrophysical Journall* **605**, L85

Foukal, P.: 1994, *Science* **264**, 238

Foukal, P. and Lean, J.: 1986, *The Astrophysical Journal* **302**, 826

Foukal, P. and Lean, J.: 1988, *The Astrophysical Journal* **328**, 347

Foukal, P. and Moran, T.: 1994, in D. M. Rabin, J. T. Jefferies, & C. Lindsey (ed.), *Infrared Solar Physics*, Vol. 154 of *IAU Symposium*, p. p.23

Fröhlich, C.: 2000, *Space Science Reviews* **94**, 15

Fröhlich, C.: 2003, *Metrologia* **40**, 60

Fröhlich, C.: 2006, *Space Science Reviews* **125**, 53

Fröhlich, C.: 2009, *Astronomy & Astrophysics* **501**, L27

Fröhlich, C.: 2011, *Space Science Reviews* pp 133–+

Fröhlich, C. and Lean, J.: 1998, *Geophysical Research Letters* **25**, 4377

Gray, L. J., Rumbold, S. T., and Shine, K. P.: 2009, *Journal of Atmospheric Sciences* **66**, 2402

Grossmann-Doerth, U., Knoelker, M., Schuessler, M., and Solanki, S. K.: 1994, *Astronomy & Astrophysics* **285**, 648

Haberreiter, M., Krivova, N. A., Schmutz, W., and Wenzler, T.: 2005, *Adv. Sp. Res.* **35**, 365

Haberreiter, M., Schmutz, W., and Hubeny, I.: 2008, *Astronomy & Astrophysics* **492**, 833

Hagenaar, H. J., Schrijver, C. J., and Title, A. M.: 2003, *The Astrophysical Journal* **584**, 1107

Haigh, J. D.: 1994, Nature **370**, 544

Haigh, J. D.: 2007, *Living Reviews in Solar Physics, http://www.livingreviews.org/lrsp-2007-2* **4**, 2

Haigh, J. D., Lockwood, M., Giampapa, M. S., Rüedi, I., Güdel, M., and Schmutz, W.: 2005, in J. D. Haigh, M. Lockwood, M. S. Giampapa, I. Rüedi, M. Güdel, and W. Schmutz (eds.), *Saas-Fee Advanced Course 34: The Sun, Solar Analogs and the Climate*

Haigh, J. D., Winning, A. R., Toumi, R., and Harder, J. W.: 2010, Nature **467**, 696

Hale, G. E.: 1908, *Publications of the Astronomical Society of the Pacific* **20**, 287

Hall, J. C. and Lockwood, G. W.: 2004, in *Bulletin of the American Astronomical Society*, Vol. 36 of *Bulletin of the American Astronomical Society*, p. 671

Harder, J., Lawrence, G., Fontenla, J., Rottman, G., and Woods, T.: 2005a, *Solar Physics* **230**, 141

Harder, J. W., Fontenla, J., Lawrence, G., Woods, T., and Rottman, G.: 2005b, *Solar Physics* **230**, 169

Harder, J. W., Fontenla, J. M., Pilewskie, P., Richard, E. C., and Woods, T. N.: 2009, *Geophysical Research Letters* **36**, 7801

Harder, J. W., Thuillier, G., Richard, E. C., Brown, S. W., Lykke, K. R., Snow, M., McClintock, W. E., Fontenla, J. M., Woods, T. N., and Pilewskie, P.: 2010, *Solar Physics* **263**, 3

Harwood, R. S. and Pyle, J. A.: 1975, *Quarterly Journal of the Royal Meteorological Society* **101**, 723

Hickey, J. R., Alton, B. M., Kyle, H. L., and Hoyt, D.: 1988, *Space Science Reviews* **48**, 321

Hinteregger, H. E., Fukui, K., and Gilson, B. R.: 1981, *Geophysical Research Letters* **8**, 1147

Hoyt, D. V., Kyle, H. L., Hickey, J. R., and Maschhoff, R. H.: 1992, *Journal of Geophysical Research* **97**, 51

Hoyt, D. V. and Schatten, K. H.: 1998, *Solar Physics* **181**, 491

Huang, T. Y. W. and Brasseur, G. P.: 1993, *Journal of Geophysical Research* **98**, 20413

IPCC: 2007, *Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*, Cambridge University Press, Cambridge, UK and New York, NY, USA

Jones, H. P. and Ceja, J. A.: 2001, in M. Sigwarth (ed.), *Advanced Solar Polarimetry – Theory, Observation, and Instrumentation*, Vol. 236 of *Astronomical Society of the Pacific Conference Series*, p. 87

Jones, H. P., Duvall, Jr., T. L., Harvey, J. W., Mahaffey, C. T., Schwitters, J. D., and Simmons, J. E.: 1992, *Solar Physics* **139**, 211

Keller, C. U., Deubner, F.-L., Egger, U., Fleck, B., and Povel, H. P.: 1994, *Astronomy & Astrophysics* **286**, 626

Keller, C. U., Schüssler, M., Vögler, A., and Zakharov, V.: 2004, *The Astrophysical Journall* **607**, L59

Kiselman, D. and Löfdahl, M.: 2003, *Royal Swedish Academy of Sciences*

Kobel, P., Solanki, S. K., and Borrero, J. M.: 2011, *Astronomy & Astrophysics* **531**, A112

Kopp, G. and Lawrence, G.: 2005, *Solar Physics* **230**, 91

Kopp, G., Lawrence, G., and Rottman, G.: 2005, *Solar Physics* **230**, 129

Kopp, G. and Lean, J. L.: 2011, *Geophysical Research Letters* **38**, L01706

Kozyra, J. U., Crowley, G., Goncharenko, L. P., Hagan, M. E., Lu, G., Mlynczak, M. G., Paxton, L. J., Russell, J. M., Solomon, S. C., Talaat, E. R., and Yee, J.: 2004, *AGU Fall Meeting Abstracts* pp B3+

Krivova, N. A., Balmaceda, L., and Solanki, S. K.: 2007, A&A **467**, 335

Krivova, N. A. and Solanki, S. K.: 2004, *Astronomy & Astrophysics* **417**, 1125

Krivova, N. A. and Solanki, S. K.: 2005, *Adv. Sp. Res.* **35**, 361

Krivova, N. A. and Solanki, S. K.: 2008, *Journal of Astrophysics and Astronomy* **29**, 151

Krivova, N. A., Solanki, S. K., Fligge, M., and Unruh, Y. C.: 2003, *Astronomy & Astrophysics* **399**, L1

Krivova, N. A., Solanki, S. K., and Floyd, L.: 2006, *Astronomy & Astrophysics* **452**, 631

Krivova, N. A., Solanki, S. K., and Schmutz, W.: 2011a, *Astronomy & Astrophysics* **529**, A81

Krivova, N. A., Solanki, S. K., and Unruh, Y. C.: 2011b, *Journal of Atmospheric and Solar-Terrestrial Physics* **73**, 223

Krivova, N. A., Solanki, S. K., and Wenzler, T.: 2009, *Geophysical Research Letters* **362**, L20101

Krivova, N. A., Vieira, L. E. A., and Solanki, S. K.: 2010, *Journal of Geophysical Research (Space Physics)* **115(A14)**, 12112

Kundt, W.: 1992, *Astrophysics and Space Science* **187**, 75

Kurucz, R.: 1993, *ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993.* 13

Labitzke, K., Austin, J., Butchart, N., Knight, J., Takahashi, M., Nakamoto, M., Nagashima, T., Haigh, J., and Williams, V.: 2002, *J. Atmos. Solar-Terr. Phys.* **64**, 203

Labitzke, K. and van Loon, H.: 1995, *Tellus Series A* **47**, 275

Lay, R. R., Lee, K. A., Holden, J. R., Oswald, J. E., Jarnot, R. F., Pickett, H. M., Stek, P. C., Cofield, R. E., Flower, D. A., Schwartz, M. J., and Shoemaker, C. M.: 2005, *Proceedings of the SPIE* **5882**, 476

Lean, J.: 1991, *Reviews of Geophysics* **29**, 505

Lean, J.: 2000, *Geophysical Research Letters* **27**, 2425

Lean, J., Rottman, G., Harder, J., and Kopp, G.: 2005, *Journal of Solar Physics* **230**, 27

Lean, J. L., Cook, J., Marquette, W., and Johannesson, A.: 1998, *The Astrophysical Journal* **492**, 390

Lean, J. L., Rottman, G. J., Kyle, H. L., Woods, T. N., Hickey, J. R., and Puga, L. C.: 1997, *Journal of Geophysical Research* **102**, 29939

Lean, Judith, L. and DeLand, Matthew, T.: 2011, *NTRS, in press.*

Lites, B. W.: 2002, *The Astrophysical Journal* **573**, 431

Lites, B. W., Scharmer, G. B., Berger, T. E., and Title, A. M.: 2004, *Solar Physics* **221**, 65

Livingston, W. C., Harvey, J., Pierce, A. K., Schrage, D., Gillespie, B., Simmons, J., and Slaughter, C.: 1976, *Applied Optics* **15**, 33

Lockwood, G. W., Skiff, B. A., and Radick, R. R.: 1997, *The Astrophysical Journal* **485**, 789

Lockwood, M.: 2010, *Proc. R. Soc. A* **466**, 303

Lockwood, M.: 2011, *Journal of Geophysical Research (Atmospheres)* **116(D15)**, D16103

Lockwood, M., Harrison, R. G., Woollings, T., and Solanki, S. K.: 2010, *Environmental Research Letters* **5(2)**, 024001

Lockwood, M., Owens, M. J., Barnard, L., Davis, C. J., and Steinhilber, F.: 2011, *Geophysical Research Letters* **38**, 22105

Makarov, V. I., Tlatov, A. G., Singh, J., and Gupta, S. S.: 2004, in A. V. Stepanov, E. E. Benevolenskaya, & A. G. Kosovichev (ed.), *Multi-Wavelength Investigations of Solar Activity*, Vol. 223 of *IAU Symposium*, pp 125–126

Marchand, P. and Marmet, L.: 1983, *Rev. Sci. Instr.* **54**, 1034

Martinez Pillet, V., Lites, B. W., and Skumanich, A.: 1997, ApJ **474**, 810

Martinez Pillet, V. and Vazquez, M.: 1990, *Astrophysics and Space Science* **170**, 75

Matthes, K., Langematz, U., Gray, L. L., Kodera, K., and Labitzke, K.: 2004, *Journal of Geophysical Research (Atmospheres)* **109**, 6101

McClintock, W. E., Rottman, G. J., and Woods, T. N.: 2005a, *Solar Physics* **230**, 225

McClintock, W. E., Snow, M., and Woods, T. N.: 2005b, *Solar Physics* **230**, 259

Meehl, G. A., Arblaster, J. M., Matthes, K., Sassi, F., and van Loon, H.: 2009, *Science* **325**, 1114

Mendoza, B.: 1997, *The Astrophysical Journal* **483**, 523

Merkel, A. W., Harder, J. W., Marsh, D. R., Smith, A. K., Fontenla, J. M., and Woods, T. N.: 2011, *Geophysical Research Letters* **38**, 13802

Neckel, H. and Labs, D.: 1994, *Solar Physics* **153**, 91

Noël, S., Bovensmann, H., Skupin, J., Wuttke, M. W., Burrows, J. P., Gottwald, M., and Krieg, E.: 2003, *Adv. Sp. Res.* **32**, 2123

Nordlund, Å., Stein, R. F., and Asplund, M.: 2009, *Living Reviews in Solar Physics* **6**, 2

Norton, A. A. and Gilman, P. A.: 2004, ApJ **603**, 348

Olemskoy, S. V. and Mordvinov, A. V.: 2009, *Geomagnetism and Aeronomy Geomagnetizm i Aeronomiia* **49**, 179

Ortiz, A., Domingo, V., and Sanahuja, B.: 2006, *Astronomy & Astrophysics* **452**, 311

Ortiz, A., Solanki, S. K., Domingo, V., Fligge, M., and Sanahuja, B.: 2002, *Astronomy & Astrophysics* **388**, 1036

Owens, M. J., Crooker, N. U., and Lockwood, M.: 2011, *Journal of Geophysical Research (Space Physics)* **116(A15)**, 4111

Pagaran, J., Weber, M., and Burrows, J.: 2009, *The Astrophysical Journal* **700**, 1884

Pap, J., Anklin, M., Fröhlich, C., Wehrli, C., Varadi, F., and Floyd, L.: 1999, *Advances in Space Research* **24**, 215

Penn, M. J. and Livingston, W.: 2006, *The Astrophysical Journall* **649**, L45

Preminger, D. G., Walton, S. R., and Chapman, G. A.: 2002, *Journal of Geophysical Research* **107**, 1354

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: 1992, *Numerical recipes in C. The art of scientific computing*, Cambridge University Press

Radick, R. R., Lockwood, G. W., Skiff, B. A., and Baliunas, S. L.: 1998, *Astrophysical Journal Supplement Series* **118**, 239

Ribas, I.: 2010, *Proceedings of the International Astronomical Union* **264**, 3

Rong, P. P., Russell, J. M., Mlynczak, M. G., Remsberg, E. E., Marshall, B. T., Gordley, L. L., and López-Puertas, M.: 2009, *Journal of Geophysical Research (Atmospheres)* **114**, 4306

Rottman, G.: 2005, *Solar Physics* **230**, 7

Rottman, G. J., Woods, T. N., and Sparn, T. P.: 1993, *Journal of Geophysical Research* **98**, 10667

Rouppe van der Voort, L. H. M., Hansteen, V. H., Carlsson, M., Fossum, A., Marthinussen, E., van Noort, M. J., and Berger, T. E.: 2005, *Astronomy & Astrophysics* **435**, 327

Rozelot, J. P., Lefebvre, S., Pireaux, S., and Ajabshirizadeh, A.: 2004, *Solar Physics* **224**, 229

Scafetta, N. and Willson, R. C.: 2009, *Geophysical Research Letters* **36**, 5701

Schatten, K. H., Mayr, H. G., Omidvar, K., and Maier, E.: 1986, ApJ **311**, 460

Scherrer, P. H., Bogart, R. S., Bush, R. I., Hoeksema, J. T., Kosovichev, A. G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T. D., Title, A., Wolfson, C. J., Zayer, I., and MDI Engineering Team: 1995, *Solar Physics* **162**, 129

Schrijver, C. J., Livingston, W. C., Woods, T. N., and Mewaldt, R. A.: 2011, *Geophysical Research Letters* **380**, L06701

Shapiro, A. I., Schmutz, W., Rozanov, E., Schoell, M., Haberreiter, M., Shapiro, A. V., and Nyeki, S.: 2011, *Astronomy & Astrophysics* **529**, A67

Shapiro, A. I., Schmutz, W., Schoell, M., Haberreiter, M., and Rozanov, E.: 2010, *Astronomy & Astrophysics* **517**, A48

Shine, R. A. and Linsky, J. L.: 1972, *Solar Physics* **25**, 357

SIDC-team: 2011, *Monthly Report on the International Sunspot Number, online catalogue*

Skumanich, A., Smythe, C., and Frazier, E. N.: 1975, *The Astrophysical Journal* **200**, 747

Skupin, J., Noël, S., Wuttke, M. W., Gottwald, M., Bovensmann, H., Weber, M., and Burrows, J. P.: 2005, *Adv. Sp. Res.* **35**, 370

Snow, M., McClintock, W. E., Woods, T. N., White, O. R., Harder, J. W., and Rottman, G.: 2005, *Solar Physics* **230**, 325

Sobotka, M.: 2003, *Astronomische Nachrichten* **324**, 369

Solanki, S. K.: 2003, *Astronomy & Astrophysicsr* **11**, 153

Solanki, S. K. and Krivova, N. A.: 2011, *Science* **334**, 916

Solanki, S. K. and Unruh, Y. C.: 1998, *Astronomy & Astrophysics* **329**, 747

Solanki, S. K., Usoskin, I. G., Kromer, B., Schüssler, M., and Beer, J.: 2004, Nature **431**, 1084

Solanki, S. K., Walther, U., and Livingston, W.: 1993, *Astronomy & Astrophysics* **277**, 639

Spiegel, E. A. and Zahn, J.-P.: 1992, *Astronomy & Astrophysics* **265**, 106

Spruit, H. C.: 1976, *Solar Physics* **50**, 269

Steinhilber, F.: 2010, *Astronomy & Astrophysics* **523**, A39

Steinhilber, F., Abreu, J. A., and Beer, J.: 2008, *Astrophysics and Space Sciences Transactions* **4**, 1

Steinhilber, F., Beer, J., and Fröhlich, C.: 2009, *Geophysical Research Letters* **361**, L19704

Tapping, K. F., Boteler, D., Charbonneau, P., Crouch, A., Manson, A., and Paquette, H.: 2007, *Solar Physics* **246**, 309

Thomas, J. H. and Weiss, O. W.: 2008, *Sunspots and Starspots*, Cambridge University Press

Thuillier, G., Foujols, T., Bolsée, D., Gillotay, D., Hersé, M., Peetermans, W., Decuyper, W., Mandel, H., Sperfeld, P., Pape, S., Taubert, D. R., and Hartmann, J.: 2009, *Solar Physics* **257**, 185

Topka, K. P., Tarbell, T. D., and Title, A. M.: 1992, *The Astrophysical Journal* **396**, 351

Topka, K. P., Tarbell, T. D., and Title, A. M.: 1997, *The Astrophysical Journal* **484**, 479

Tran, T., Bertello, L., Ulrich, R. K., and Evans, S.: 2005, *The Astrophysical Journals* **156**, 295

Trujillo Bueno, J., Shchukina, N., and Asensio Ramos, A.: 2004, Nature **430**, 326

Unruh, Y. C., Krivova, N. A., Solanki, S. K., Harder, J. W., and Kopp, G.: 2008, *Astronomy & Astrophysics* **486**, 311

Unruh, Y. C., Solanki, S. K., and Fligge, M.: 1999, *Astronomy & Astrophysics* **345**, 635

Unruh, Y. C., Solanki, S. K., and Fligge, M.: 2000, *Space Science Reviews* **94**, 145

Usoskin, I. G., Solanki, S. K., and Kovaltsov, G. A.: 2007, *Astronomy & Astrophysics* **471**, 301

van Loon, H. and Shea, D. J.: 1999, *Geophysical Research Letters* **26**, 2893

Vieira, L. E. A. and Solanki, S. K.: 2010, *Astronomy & Astrophysics* **509**, A100

Viereck, R. A., Floyd, L. E., Crane, P. C., Woods, T. N., Knapp, B. G., Rottman, G., Weber, M., Puga, L. C., and DeLand, M. T.: 2004, *Space Weather* **2**, S10005

Vögler, A.: 2004, *Mem. S.A.It.* **75**, 282

Vögler, A., Shelyag, S., Schüssler, M., Cattaneo, F., Emonet, T., and Linde, T.: 2005, *Astronomy & Astrophysics* **429**, 335

Wang, H. and Zirin, H.: 1992, *Solar Physics* **140**, 41

Wang, Y.-M., Sheeley, Jr., N. R., and Rouillard, A. P.: 2006, *The Astrophysical Journal* **644**, 638

Warwick, N. J., Bekki, S., Nisbet, E. G., and Pyle, J. A.: 2004, *Geophysical Research Letters* **31**, 5107

Wenzler, T.: 2005, *Ph.D. thesis*, Eidgenoessische Technische Hochschule Zürich (Switzerland), Switzerland

Wenzler, T., Solanki, S. K., and Krivova, N. A.: 2005, *Astronomy & Astrophysics* **432**, 1057

Wenzler, T., Solanki, S. K., and Krivova, N. A.: 2009, *Geophysical Research Letters* **36**, L11102

Wenzler, T., Solanki, S. K., Krivova, N. A., and Fluri, D. M.: 2004, *Astronomy & Astrophysics* **427**, 1031

Wenzler, T., Solanki, S. K., Krivova, N. A., and Fröhlich, C.: 2006, *Astronomy & Astrophysics* **460**, 583

Willson, R. C.: 1997, *Science* **277**, 1963

Willson, R. C. and Mordvinov, A. V.: 2003, *Geophysical Research Letters* **30(5)**, 050000

Wilson, O. C.: 1978, *The Astrophysical Journal* **226**, 379

Wilson, R. M.: 1987, *Solar Physics* **111**, 255

Woods, T. N., Eparvier, F. G., Bailey, S. M., Chamberlin, P. C., Lean, J., Rottman, G. J., Solomon, S. C., Tobiska, W. K., and Woodraska, D. L.: 2005, *Journal of Geophysical Research (Space Physics)* **110**, 1312

Woods, T. N., Prinz, D. K., Rottman, G. J., London, J., Crane, P. C., Cebula, R. P., Hilsenrath, E., Brueckner, G. E., Andrews, M. D., White, O. R., VanHoosier, M. E., Floyd, L. E., Herring, L. C., Knapp, B. G., Pankratz, C. K., and Reiser, P. A.: 1996, *Journal of Geophysical Research* **101**, 9541

Woods, T. N., Rottman, G. J., and Ucker, G. J.: 1993, *Journal of Geophysical Research* **98**, 10679

Woods, T. N., Tobiska, W. K., Rottman, G. J., and Worden, J. R.: 2000, *Journal of Geophysical Research* **105**, 27195

Wuebbles, D. J., Wei, C.-F., and Patten, K. O.: 1998, *Geophysical Research Letters* **25**, 523

Zeilik, M. and Gregory, S.: 1998, *Introductory Astronomy and Astrophysics*, Harcourt College Publishers