AN OBSERVATIONAL STUDY OF THE FORMATION AND EVOLUTION OF SUNSPOTS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ASTRONOMY

DECEMBER, 2011

By
Sarah A. Jaeggli

Dissertation Committee:

H. Lin, Chairperson
S. Businger
S. Habbal
R. Joseph
I. Roussev
H. Uitenbroek

We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation for the degree of Doctor of Philosophy in Astronomy.

DISSERTATION COMMITTEE

_____

Chairperson

_____

_____

_____

_____

_____

There once was a girl named Sarah
Who studied the Sun for an era
She studied so long
Drank beer and played pong
Now she's got a new line for her vitae

Acknowledgements

The success of this dissertation is due to two individuals to whom I am profoundly thankful, Haosheng Lin, for his years of hard work, mango salsa, dedication, and encouragement to this dissertation and the FIRS project; and Han Uitenbroek, for his guidance and advice on all issues, theoretical and culinary. Additional thanks are necessary for the other members of my committee and those who advised my past research, Shadia Habbal, and Bob Joseph, Jeff Kuhn, Ilia Roussev, and Steve Businger.

This work would have gone faster but I would be a much less interesting person without the help of my family, friends, office mates, and house mates, especially Jeff Rich, who was a great help, keeping me fed and sane during the writing of this dissertation.

The FIRS project was funded by the National Science Foundation Major Research Instrument program, grant number ATM-0421582, and was completed through a collaboration between the Institute for Astronomy and the National Solar Observatory. The observations for this research were obtained using the Dunn Solar Telescope at the National Solar Observatory at Sacramento Peak, New Mexico, which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA), under cooperative agreement with the National Science Foundation. I would like to express my profound gratitude to the NSO for all of their assistance and would especially like to thank the DST observers, Doug Gilliam, Joe Elrod, and Mike Bradford for their patience and ingenuity

during the commissioning of FIRS and Tom Schad and Phil Judge for being the first of hopefully many to use FIRS.

Hinode is a Japanese mission developed and launched by ISAS/JAXA, with NAOJ as domestic partner and NASA and STFC (UK) as international partners. It is operated by these agencies in co-operation with ESA and NSC (Norway). Hinode SOT/SP Inversions were conducted at NCAR under the framework of the Community Spectro-polarimtetric Analysis Center (CSAC; http://www.csac.hao.ucar.edu/). I would especially like to thank Alfred de Wijn for his help with the Hinode data.

SOLIS data used here were produced cooperatively by NSF/NSO and NASA/LWS. HMI data used here is courtesy of NASA/SDO and the HMI science team.

I would also like to thank the Vector Magnetogram Comparison Group and Andres Asensio Ramos, and Hector Socas Navarro for their essential advice on inversion techniques, Huw Morgan and Ali Tristchler for input on the text of the upcoming ApJ paper.

Abstract

This dissertation focuses on the problem of molecules and the horizontal balance of forces in sunspots. Sunspots are quasi-static features on the solar surface and can be considered to be in a state of equilibrium. The weaker gas pressure of the cool sunspot interior is horizontally supported against the higher pressure of the hotter quiet-Sun by a strong vertical magnetic field. However, some sunspots show a rapid increase in magnetic pressure relative to the temperature of the gas in the coolest regions of the sunspot, implying that an isothermal decrease in the gas pressure must have occurred. The current model of sunspots is unable to describe this deviation from the assumed equilibrium state of the magnetic field and thermal gas pressure observed in these sunspots. Another method of altering the pressure of the gas must be occurring. The formation of molecules in sunspots may be the key to solving this puzzle. The sunspot interior provides a cool environment where molecules can form in abundance. As atoms become bound into molecules the total particle number of the gas is decreased. A sufficiently large molecular fraction could significantly alter the properties of the sunspot plasma, and specifically provide a mechanism for concentrating the magnetic field by non-thermally lowering the gas pressure.

I have investigated the equilibrium condition of sunspots of different sizes and in a variety of evolutionary phases through a Milne-Eddington inversion of spectropolarimetric observations of the Zeeman-split Fe I lines at 6302 and 15650 Å to obtain their thermal and magnetic topology. I carried out a calculation of the detailed radiative transfer and chemical equilibrium of model sunspot atmospheres to determine the molecular gas fraction. Several sunspots show unambiguous cases of isothermal magnetic field intensification, which

can only be explained by the formation or destruction of a large molecular population. All sunspots with magnetic fields stronger than 2500 G and temperatures cooler than 5800 K consistently show a signature of magnetic field over-concentration, consistent with molecular hydrogen formation of a few percent of the total gas fraction. The formation of this large molecular population has widespread implications for sunspot physics.

Table of Contents

List of Tables

List of Figures

Chapter 1

Perspective

1.1 Overview of this Dissertation

In the current chapter I present an overview of past observations and discoveries which have led to our current understanding of the Sun and fill in some necessary background for those unfamiliar with the field.

In Chapter 2 I discuss the specific problem with our current view of sunspot magnetohydrostatic equilibrium and identify molecule formation as a process which can alter the sunspot equilibrium, resulting in an isothermal increase in the magnetic field. I discuss the possible effects of molecule formation in detail and outline the observational sunspot survey which I have used to approach this problem.

In Chapter 3 I talk about the design and properties of the instruments which were used to conduct the sunspot survey observations, the properties of the observations themselves, and the instruments from which I obtained supporting observations.

In Chapter 4 I talk about the reduction of the survey data, including the correction of instrumental polarization and the unusual response properties of the infrared detector. In addition I present a direct comparison between my survey observations and those from the spectropolarimeter onboard Hinode.

In Chapter 5 I describe the Milne-Eddington inversion code that I have developed to analyze the spectropolarimetric data, and discuss the effect of systematic errors on the

retrieval of the inverted magnetic field parameters. I compare the results of this new code to other similar codes to promote understanding of the benefits and pitfalls of inversion techniques in general.

In Chapter 6 I describe the atmospheres that were selected and the radiative transfer and chemical equilibrium code that was used to retrieve the relation between H_2 fraction and the OH equivalent width, in addition to many other interesting results that have been determined from the models. The effect of stray light on the measured equivalent width of spectral lines is discussed in detail.

In Chapter 7 I present the results from the sunspot survey and discuss the history and results for each active region individually. Following this I discuss the entire sample as a whole and present the general results.

In Chapter 8 I speculate on the role that H_2 plays in sunspot evolution and discuss other effects which may change the observed thermal-magnetic relation. In addition I discuss the significance of this survey in light of the current solar activity cycle.

Finally in Chapter 9 I summarize this work and provide some concluding remarks. In addition I outline the prospects for future work on this problem and other work which may be possible with the survey data.

1.2 History of Solar Astronomy

The history of solar astronomy is deeply entangled in the development of understanding of the physical universe and many of the significant scientific discoveries in the last 1000 years of civilization. These first simple truths were gained at great personal and philosophical cost, providing a standard method of inquiry and setting the stage for the discoveries that would follow.

1.2.1 Early Solar Observations

The Sun is the most obvious of the astronomical bodies and its changing position in the sky throughout the day, and year, were first recorded to keep track of time. In nearly every ancient culture we find the remains of astronomical calendars. Babylonian, Greek, Arab, and Chinese astronomers measured and recorded the solar position, although there were a few who recognized that the Earth must be rotating, and orbiting around the Sun, many could not explain the wandering of the Sun's daily path in the sky. The idea of heliocentrism was not generally accepted until the efforts of Nicholas Copernicus in the 16th century.

Even early solar observers were able to challenge the view that the heavens were perfectly created, immutable, and unchanging. Sufficiently large features on the solar surface can be resolved by the human eye when the Sun becomes partially obscured by thin clouds or haze at the horizon. Early Chinese astronomers during the Han Dynasty (206 BCE - 220 CE) were the first to recognize sunspots as a recurring phenomenon and keep records of sunspots which were apparent to them. It was not known what these solar "blemishes" were until the invention of the telescope and its repurposing for astronomical observations in 1609. Galileo confirmed that sunspots were in fact features on the solar surface due to their apparent projection as they reached the edge of the Sun and not objects (or planets) passing between the Earth and the Sun. Following this critical technological advancement, early solar astronomers were able to observe the motion of sunspots across the disk of the Sun and quickly realized that the Sun was a spherical, rotating body. From observations of sunspot motion many early observers, including Galileo and Christoph Scheiner, were quickly able to quantify the solar rotation period of 27 days.

1.2.2 The Beginning of Modern Solar Astronomy

As improvements were made to instrumentation, so did the study of the Sun progress. Newton came up with the idea of dispersing sunlight with a prism, but it was Joseph Fraunhofer who built the first scientific apparatus for this purpose in the early 1800's. With his new spectroscope he analyzed dispersed sunlight and recorded hundreds of dark lines that

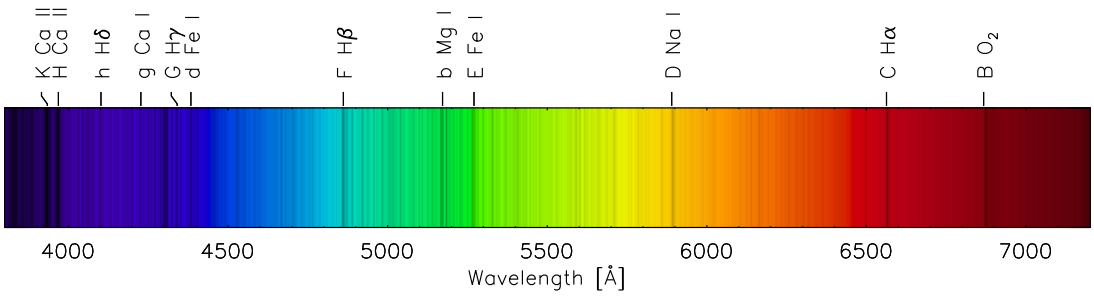

Figure 1.1 The solar spectrum over the visible range from 3800 to 7200 Å with the major Fraunhofer lines identified.

he saw in the solar spectrum. Figure 1.1 shows the solar spectrum with Fraunhofer's lettered dark lines and their identified responsible species. The first insight into the composition and processes that make the Sun shine came as a result of the laboratory work of Robert Bunsen and Gustav Kirchhoff, who confirmed that Fraunhofer's gaps in the solar spectrum corresponded to the emission wavelengths of specific elements. Indeed, they found that the solar surface was composed of hot gasses containing the same elements of which the Earth was made.

Joseph Stefan and Ludwig Boltzmann independently determined the relation between the temperature and luminosity of a radiating blackbody in the 1880s, and Stephan was able to determine a surface temperature of 5700 K for the Sun. But the source of the Sun's enormous energy was still a significant problem. It was not until the early 20th century that nuclear physics took leaps and bounds ahead, and a more reasonable explanation of the solar energy source was proposed. The application of Bohr's model of the atom and Saha's understanding of statistical equilibrium of ionized states in a gas showed that the Sun must be composed almost entirely of hydrogen. Sir Arthur Eddington defended the idea that all stars, including the Sun, must be enormously hot at the core based on his mass-luminosity relation (Eddington 1924), and could only be powered by nuclear fusion. Finally, in the 1930s Hans Bethe theorized correctly that the Sun must form hydrogen

into helium through a series of fusion reactions called the proton-proton chain which are responsible for the majority of the Sun's energy (Bethe 1939).

1.2.3 Discovery of the Magnetic Sun

Prior to the 20th century the study of sunspots was no more than a science of measuring sizes, positions, intensities, and their change with time. Dedicated observers found that sunspots were highly structured and dynamic, changing on hour timescales. In the 1840s Heinrich Schwabe recognized that there was an 11 year cycle of solar activity demonstrated by fluctuations in the number of sunspots (Schwabe 1843). It was then found that geomagnetic storms (i.e. aurorae) in the Earth's atmosphere closely followed the Sun's cycle of activity (Sabine 1852; Wolf 1852). Though it was obvious that the Sun was sometimes emitting particles or radiation which buffeted the Earth's atmosphere, the source of these storms was made abundantly clear in 1859. A gigantic sunspot group near the center of the solar disk experienced a series of flares observed by Richard Carrington (Carrington 1859). His sketch of the sunspot group is shown in Figure 1.2. During the following two days telegraph stations caught fire from arcing electricity induced in electrical lines and an unprecedented aurora was visible worldwide.

The magnetic nature of sunspots was long suspected due to the similarity in their appearance to iron filings arrayed in a magnetic field, but it could not be confirmed until the discovery of the Zeeman Effect in 1896. Pieter Zeeman had been experimenting with emission line sources and discovered that some of the lines became split into separate lines when the source was exposed to a strong magnetic field. Following Zeeman's discovery, the effect was explained by Lorentz through his theory of electromagnetic fields, and was later fully characterized using the quantum mechanical model of the atom. This laboratory discovery was of paramount importance to the solar physics that would follow.

In 1904 George E. Hale founded a new observatory at Mt. Wilson above the city of Los Angeles, and in 1907 the 60-ft solar tower was nearly complete. This was a new design of telescope, with a mirror positioned high above the ground to avoid the effects of

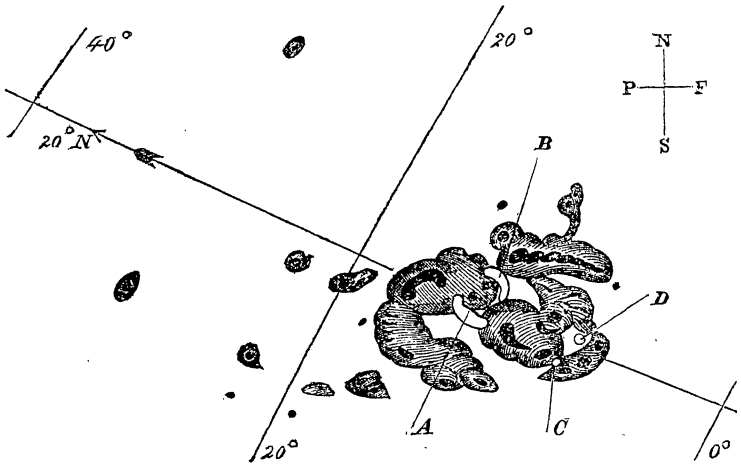

Figure 1.2 A drawing of the sunspot group responsible for the massive solar activity of September 1, 1859 [Figure 1 from Carrington (1859)].

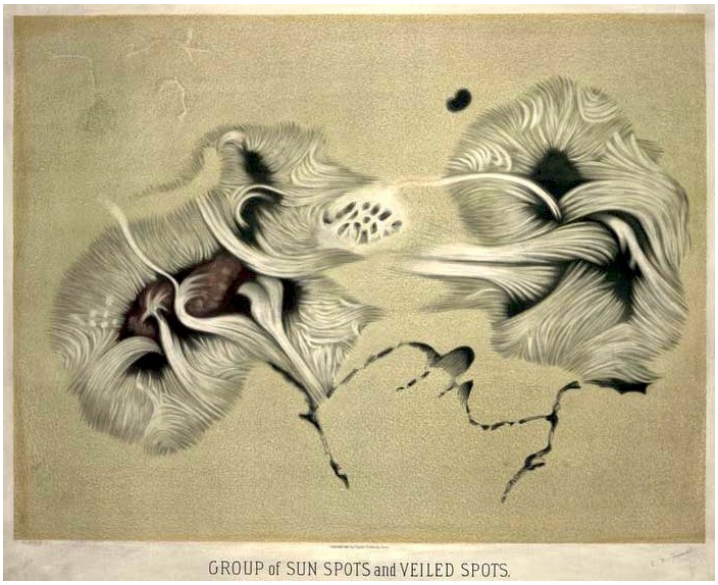

Figure 1.3 A lithograph of sunspots observed in 1875 by Étienne Trouvelot from the New York Public Library digital gallery.

turbulence from rising hot air (or "seeing"). An accompanying 30-foot grating spectrograph was located deep in the ground for thermal stability. With it Hale was able to achieve the large dispersion required to resolve the Zeeman splitting of iron and titanium lines in sunspots and definitively confirm the presence of kilo-Gauss magnetic fields in sunspots (Hale 1908). Hale was also able to measure the circular polarization of the magnetic fields which show whether the field is pointed towards or away from the observer. This was the birth of a new field, solar spectropolarimetry.

In the century that has followed since Hale's discovery the field of solar physics has developed in leaps and bounds, from the basic understanding of the global solar magnetic field, to the application of magnetohydrodynamics to understand the behavior of a huge variety of solar magnetic phenomena. In the last 20 years the development and application of digital detectors and adaptive optics has greatly increased the speed and clarity with which ground-based observations can be carried out, and every new satellite revolutionizes our understanding. In spite of the rich and lengthy history of solar observations we are continuing to learn more about the gloriously complicated and dynamic star which is our Sun.

1.3 Our Star, the Sun

A General Description of the Sun

Our Sun is a slightly larger than average star in the Milky Way galaxy. With a mass of $M_\odot = 1.989 \times 10^{33}$ g, the Sun maintains a radius of $R_\odot = 6.955 \times 10^{10}$ cm with a total luminosity of $L_\odot = 3.839 \times 10^{33}$ erg/s. In stellar terms, it is classified as a G2 dwarf with an absolute magnitude of 4.83. At a distance of 1 AU ($= 1.496 \times 10^{13}$ cm) the Sun subtends roughly 0.5° on the sky, or approximately 1,918", implying a scale of ∼730 km/arcsecond at the solar surface. Although is Sun is considered to be rich in heavy elements relative to many other stars in the galaxy, it is composed predominantly of hydrogen and helium

Table 1.1 Most Abundant Elements in the Solar Atmosphere

Element	Number [%]	Mass[%]
H	92.0	73.4
He	7.8	25.0
C	0.02	0.20
N	0.008	0.09
O	0.06	0.8
Ne	0.01	0.16
Mg	0.003	0.06
Si	0.004	0.09
S	0.002	0.05
Fe	0.003	0.14

with a few traces of heavier elements. A table of the most abundant elements in the solar atmosphere is given in Table 1.1

Lifecycle

The Sun and solar system formed out of the collapse of an enriched giant molecular cloud approximately 4.5 billion years ago. The Sun will last approximately another 5 billion years, gradually becoming larger and brighter as the hydrogen in the core is depleted. The Sun then passes through a brief phase as a red giant. With the energy source entirely depleted the core will begin to collapse under its own gravity. The increased pressure in the core causes the outer layers of the Sun's atmosphere to puff up to 250 times its current size. When a high enough density is reached, a short phase of the fusion of helium into carbon commences. As the helium is also depleted violent pulsations occur as the core once again collapses, throwing off the outer layers of the Sun's atmosphere, leaving the hot core bare, a white dwarf. The discarded atmosphere expands into a planetary nebula and gradually dissipates as the remains of the core slowly radiate the remaining energy over the billions of year to follow.

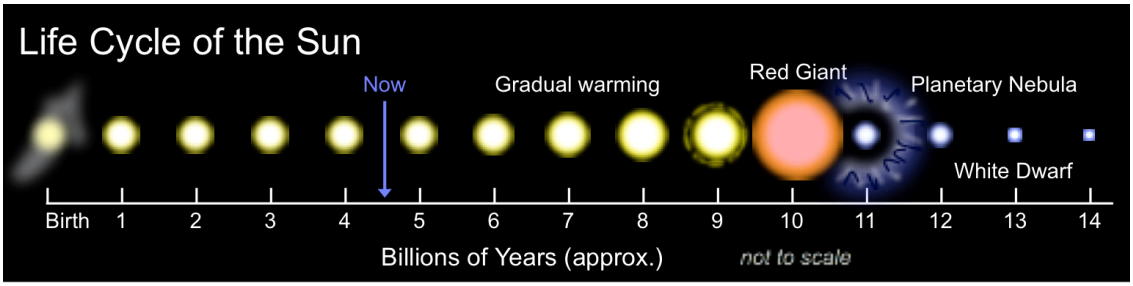

Figure 1.4 Diagram of the Sun's lifecycle.

1.3.1 Solar Structure

The Sun is composed of many parts. While only the outer layers are visible to us, there is strong evidence to support the current theoretical of the structure of the interior. A diagram of solar structure as we currently understand it is shown in Figure 1.5.

Core

The core of the Sun encompasses a region from the very center out to 0.2 $R_\odot$ which is responsible for most of the energy production in the Sun. The core is very dense (2.5×10^{17} dynes/cm^2) and hot (1.6×10^7 K) which facilitates the fusion of hydrogen into helium and energy. During the proton-proton chain proposed by Bethe, hydrogen is formed into deuterium, and then into helium.

$$\begin{aligned} {}^1_1H + {}^1_1H &\longrightarrow {}^2_1H + e^+ + \nu_e \\ {}^2_1H + {}^1_1H &\longrightarrow {}^3_2He + \gamma \end{aligned} \tag{1.1}$$

But the process does not stop there. The temperature and pressure in the core control the probabilities with which the next reactions occur. In the Sun's core 69% of the time 2 ${}^3_2He$ nuclei fuse to produce ${}^4_2He$ and hydrogen (PP I):

$${}^3_2He + {}^3_2He \longrightarrow {}^4_2He + 2{}^1_1H \tag{1.2}$$

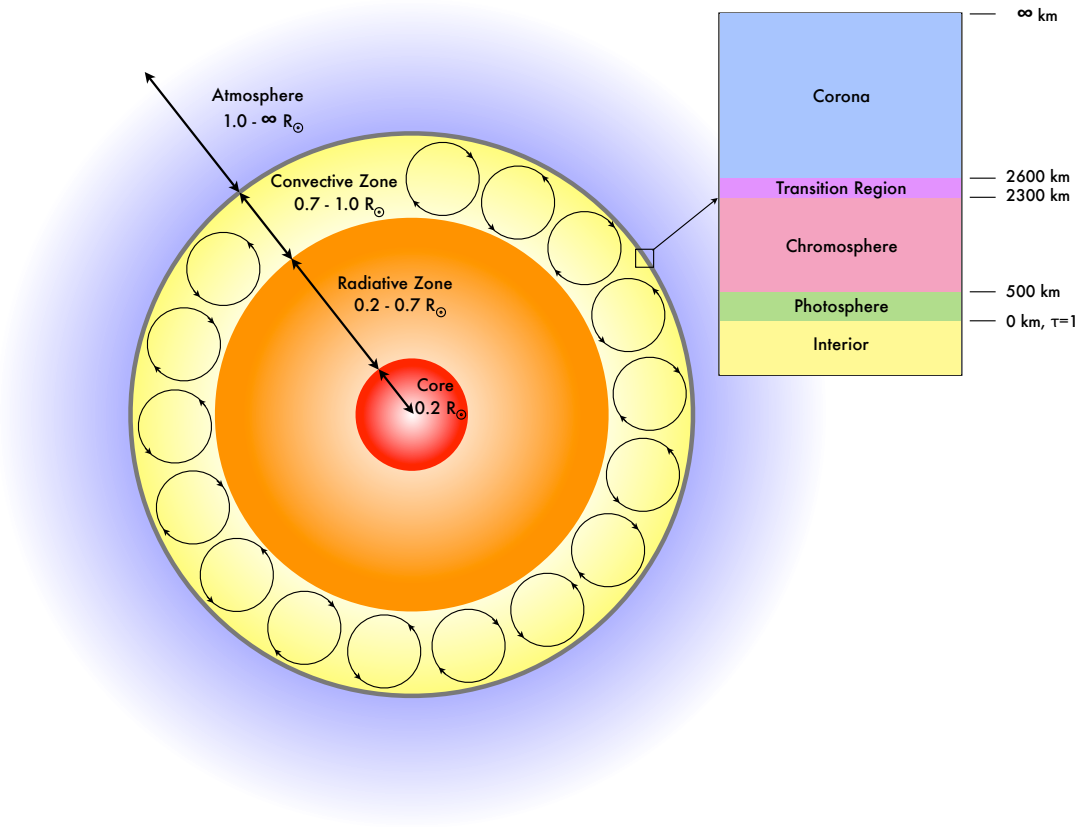

Figure 1.5 A cartoon of the regions in the Sun's atmosphere.

The other 21% of the time the resulting ${}^4_2He$ from PP I goes on to produce the light element berillium:

$$ {}^3_2He + {}^4_2 He \longrightarrow {}^7_4 Be + \gamma \tag{1.3} $$

After the production of Be, 99.7% of the time it combines with an electron and hydrogen atom, resulting in ${}^4_2He$ (PP II):

$$ \begin{aligned} {}^7_4Be + e^- &\longrightarrow {}^7_3Li + \nu_e \\ {}^7_3Li + {}^1_1 H &\longrightarrow 2{}^4_2He \end{aligned} \tag{1.4} $$

The other 0.3% of the time Be combines with a proton to form boron which decays to back to Be and finally to ${}^4_2He$ (PP III):

$$ \begin{aligned} {}^7_4Be + {}^1_1 H &\longrightarrow {}^8_5B + \gamma \\ {}^8_5B &\longrightarrow {}^8_4Be + e^+ + \nu_e \\ {}^8_4Be &\longrightarrow 2{}^4_2He \end{aligned} \tag{1.5} $$

While we cannot see directly into the solar core to know what occurs there, the fusion processes produce neutrinos at almost every step of the proton-proton chain. Because neutrinos are weakly interacting particles, they pass through the solar atmosphere unhindered, while photons from the core cannot reach us directly and must spend many thousands of years interacting with matter before they can be radiated into space. Experiments to detect solar neutrinos began in the 60s, and at the time it appeared that there were far too few neutrinos seen by detectors on Earth based on predictions from standard models of the solar interior (Davis & Cox 1991). Only recently has this problem been resolved by the discovery of neutrino oscillation from one flavor to another (Ahmad et al. 2002).

Radiative Zone

The radiative zone of the Sun stretches from the edge of the core at 0.2 to 0.7 $R_{\odot}$. In this region the material's ability to block light, or opacity, is low, therefore the energy produced in the core is most efficiently lost by radiation. However because the gas density is high photons quickly absorbed by atoms in the gas and emitted in a random direction. Due to this random-walk process, photos are trapped for thousands of years before actually reaching the edge of the radiative zone and escaping the interior.

Convection Zone

As the temperature decreases near the top of the radiative zone, the opacity in the solar atmosphere rapidly increases due to the recombination of metals, in particular iron and oxygen. Radiation is no longer an efficient means of energy transport and instability in the gas causes the formation of convective cells which carry hot plasma to the surface where the energy can be radiated into space.

Photosphere

At the top of the convection zone the optical depth of the solar atmosphere rapidly decreases over a few hundred kilometers. Optical depth (τ) is the wavelength-dependent measure of a material's transparency, and is related to the fraction of light not scattered or absorbed by $I/I_0 = e^{-\tau}$. For an optical depth greater than 1 the material is considered to be"optically thick" and most photons cannot pass through without being scattered or absorbed. When the optical depth is less than one, the material is considered to be "optically thin" and most photons are able to escape without interaction. The$\tau = 1$ surface layer where the photons we observe originate is called the photosphere. The surface of the Sun is described by an effective temperature of 5776 K, but the average temperature of the photosphere ranges from 6500 K at the bottom near the convection zone to 4400 K at the top. A graph of the temperature structure in the solar atmosphere is shown in Figure 1.6. By virtue of being visible to observation, the photosphere displays a wide variety of features: the tops of the

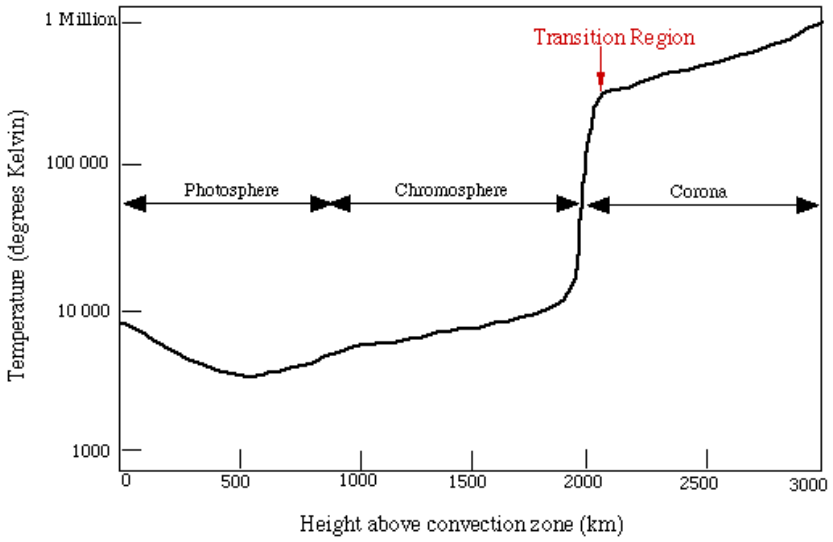

Figure 1.6 The temperature structure with height through the photosphere, chromosphere, transition region, and low corona [credit: MSU, adapted by M.B. Larson from Sun, Earth, Sky by Kenneth Lang].

convective cells create a pattern of granulation, the signature of oscillation modes in the solar interior, and magnetic features of all shapes and sizes.

Chromosphere

Immediately above the photosphere is the chromosphere, so called because of the characteristic pink color from the combination of 6563 Å Hα and 4861 Å Hβ emission observed during eclipses. The gas density drops as the distance from the photosphere increases and the temperature rises steadily from 4400 K to over 10,000 K. The chromosphere is highly structured on small scales due to the ubiquitous solar magnetic fields.

Transition Region

In a very small distance at the top of the chromosphere the solar atmosphere becomes very hot, rising to 10^6 K in a span of 200-300 km. This sudden increase in the temperature is

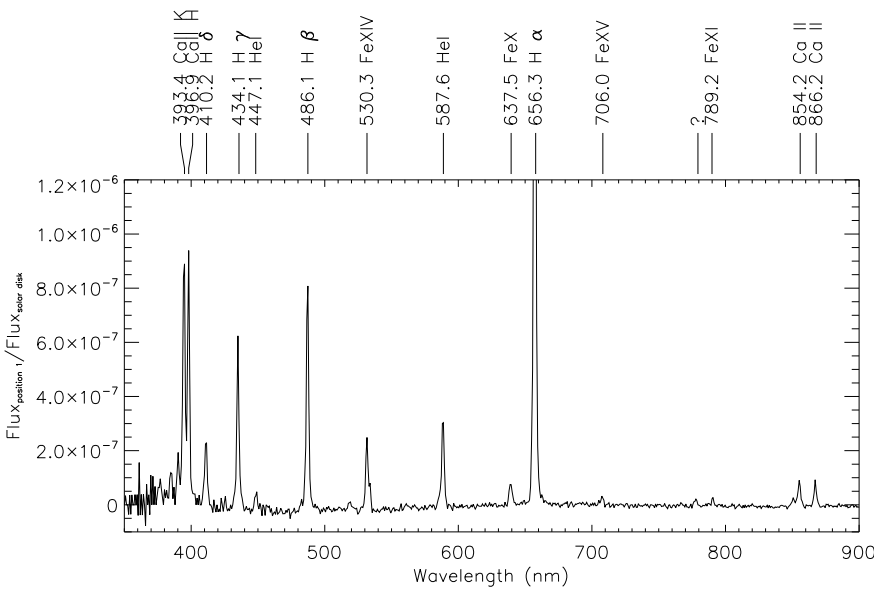

Figure 1.7 Emission spectrum from the inner solar corona at 0.25 $R_\odot$ above the limb, taken by S. Jaeggli during the total solar eclipse of March 29, 2006.

far more than one might expect from the radiative heating of the chromospheric gas, this is one of the most important outstanding problems in solar physics today.

Corona

Following the transition region the temperature increases more slowly with distance up to several million Kelvin and the gas becomes very tenuous in the thin outer atmosphere of the Sun, or corona. The corona is incredibly dim compared to the brightness of the solar disk, even the brightest inner portions are about 10^{-6} or roughly the same brightness as the daytime sky or full moon. The faintness of the corona makes it difficult to study with both space- and ground-based observations, except during total solar eclipses when the moon blocks the bright solar disk. An emission spectrum of the corona taken during a total solar eclipse is shown in Figure 1.7. The forbidden lines of highly ionized atomic species indicate a temperature of 2×10^6 K. In the corona close to the Sun a blurry solar disk spectrum can be seen scattering off of fast-moving electrons at many millions of degrees, while further

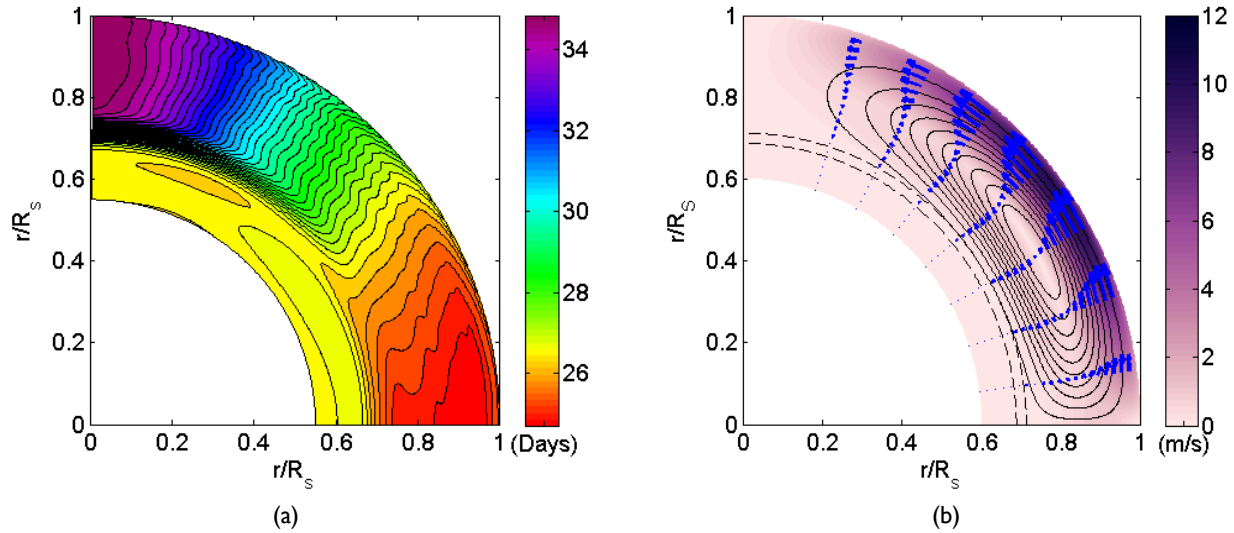

Figure 1.8 (a) Differential rotation in the solar convection zone with depth and latitude, the color scale indicates the rotation speed in days. (b) The meridional flow of the convection zone, showing strong poleward motion in the upper layers of the convection zone and weaker equatorward flow in the bottom layers [Figure 7 from Muñoz-Jaramillo (2011), data taken by GONG].

out the reddened sunlight scattered from dust particles combines with the zodiacal light at great distances from the Sun.

1.4 Solar Magnetic Fields

Everywhere we look, the processes of the active Sun are dominated by magnetic fields. The solar magnetic fields are a consequence of the fact that the convection zone rotates and circulates separately from the radiative zone, known as the solar dynamo (see Figure 1.8). The differential rotation with depth of the highly conductive solar plasma in the convection zone around the solid-body rotation of the radiative zone and core generates the dipolar magnetic field due to the presence of the highly sheared layer at the bottom of the convection zone. The Sun is also rotates differentially by latitude. It rotates faster at the equator than it does at the poles which causes the simple dipolar field to twist up

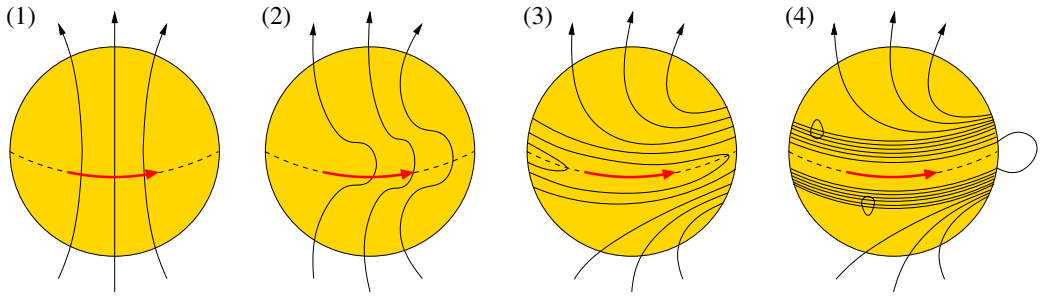

Figure 1.9 A figure showing the Babcock model (Babcock 1961) of winding up of the solar magnetic field through differential rotation [adapted from Freedman & Kauffman (2008)].

around the equator. The slower meridional flow at the bottom of the convection zone drags the magnetic activity belts to to equator. These complex motions cause the magnetic field lines to be forced through the surface of the photosphere in complicated patterns, causing magnetic activity in regions which drift from the poles to the equator as the field winds up, cancels out, and winds up again. A cartoon version of this winding process is shown in Figure 1.9.

While the specific processes which produce the solar dynamo are hidden from view in the solar interior, the dynamo produces simple observational consequences. Sunspot numbers, and other indicators of activity follow a regular 11 year cycle of solar activity during which the active regions of the Sun drift from the poles to the equator. Magnetic fields in the northern and southern hemispheres of the Sun swap polarity during the lulls in solar activity. Features of the photosphere, chromosphere, and corona, such as the rate of sunspot appearance, the brightness of chromospheric emission, the shape of the corona, the speed of the solar wind at the equator, the size and position of regions of fast solar wind (or coronal holes), alter with the magnetic cycle. While it is apparent that there exist long-term variations which alter the strength and frequency of solar activity from cycle to cycle, a clear pattern is not yet apparent.

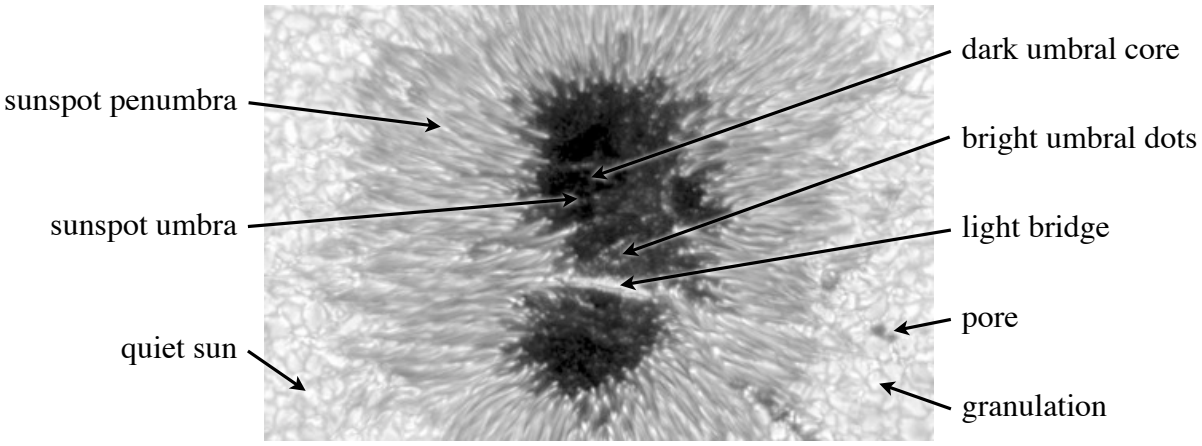

Figure 1.10 An image of a sunspot from the Swedish Solar Telescope which has been enhanced to show umbral features more clearly.

1.5 Sunspots

Sunspots are the most obvious manifestations of the solar magnetic field. They are created when strong magnetic fields, about 1000 G or larger, emerge through the solar surface. From Maxwell's equation $\nabla \times \vec{B} = \mu_0 \vec{j}$ plasma motions along magnetic field lines are permitted, but motions perpendicular to the magnetic field are suppressed. Therefore, the vertical magnetic field in a sunspot suppresses the cross-field overturning motion of convection at the top and bottom of the convection zone. The reduction in energy transport from below causes the sunspot atmosphere to be dimmer and cooler than the surrounding quiet-Sun by several thousand degrees. Figure 1.10 shows an example of a large sunspot. Sunspots are rich with features but are composed of two basic parts: the umbra and the penumbra (see excellent general reviews by Solanki (2003) and Moore & Rabin (1985), on fine structure in sunspots (Thomas & Weiss 2004), and on umbral fine structure (Socas-Navarro 2003)).

Umbra

The umbra is the darkest part of a sunspot. At first glance the structure of the umbra looks fairly dark and uniform compared to its surroundings, but further inspection shows a variety

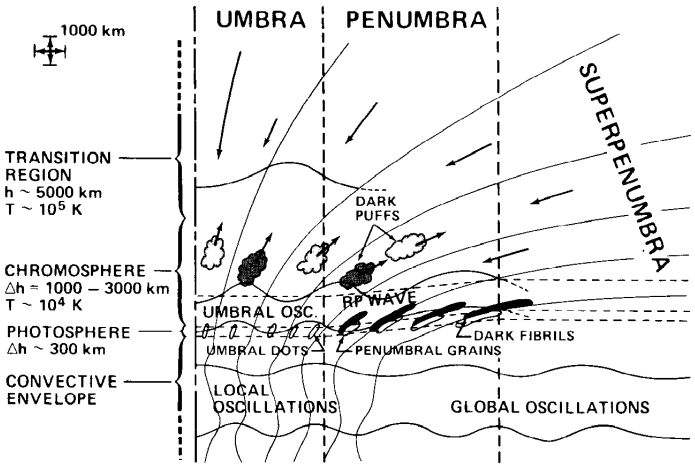

Figure 1.11 A sketch of the magnetic field in a sunspot (solid lines) and various features observed in the photosphere and chromosphere. Arrows indicate the qualitative magnitude and direction of the velocity of features at different heights and radii [Figure 9 from Moore & Rabin (1985)].

of structures that can be seen in Figure 1.10, bright dots, filamentary structure, dark cores, and even the signature of suppressed convective granulation. In the umbra the magnetic fields are mostly perpendicular to the surface of the Sun and typical umbral magnetic fields strengths range from 1500 to 3500 Gauss (Moore & Rabin 1985). The temperatures in the umbra range from 1000 to 2000 K cooler than the surrounding quiet-Sun, providing an environment cool enough where simple molecules can form, such as CN, CO, OH, H_2, and even H_2O.

Penumbra

While small sunspots may appear as a bare umbra, called a pore, in larger sunspots a radial filamentary structure can be seen surrounding part or all of the sunspot reaching from the quiet-Sun and into the umbra. The penumbra is a region where the magnetic field turns over from vertical to horizontal, and is characterized by magnetic field strengths of 700-1000 G and temperatures of 250 to 400 K cooler than the quiet-Sun (Moore & Rabin 1985). While it is not fully known what the physical structure of the penumbra is, recent modeling

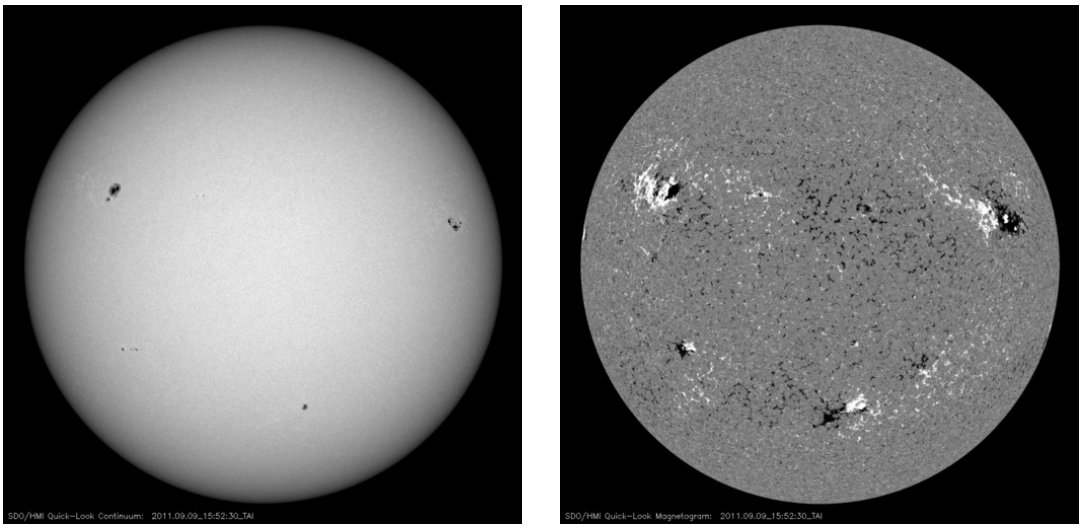

Figure 1.12 Full-disk intensity and radial magnetic flux on 9 Sept. 2011 from the Solar Dynamics Observatory - Helioseismic and Magnetic Imager.

efforts may reveal that we are seeing the granules of the quiet-Sun distended by the strong horizontal component of the magnetic field (Rempel et al. 2009).

Characterization of Sunspots

Sunspots appear in magnetically active regions where loops of the magnetic field emerge from the photosphere. Like the magnetic field from a horseshoe (U) magnet, the foot-points of each loop must have opposite magnetic polarity. Sunspots appear at the foot-points of the loops in pairs, groups, or alone and in a wide range of shapes and sizes. In a magnetic region the dominant polarity in the hemisphere often leads the region in the direction of the Sun's rotation, and the largest sunspots are often associated with the leading polarity. For example, Figure 1.12 shows the continuum intensity in the visible and radial magnetic flux of the Sun on September 9, 2011. In the northern hemisphere the strongest magnetic fields (in sunspots) are associated with the black regions which precede the white regions in the rotation direction (left to right). In the southern hemisphere this situation is reversed, with the white regions leading in the rotational direction. Therefore during this solar cycle, black

is the dominant polarity in the northern hemisphere and white is the dominant polarity in the southern hemisphere.

There are various terms to describe sunspot regions based on their magnetic field configurations. The Mt. Wilson classification used in this work to describe sunspots is defined as follows:

α: An α spot is a single monolithic umbra, usually of leading polarity with a following plage (or a region where a diffuse magnetic field is present).

β: A β-region contains two clear spots or spot groups with opposite magnetic polarity, one leading and one following.

γ: A γ-regions shows a complex distribution of the magnetic polarities.

$\beta\gamma$: A $\beta\gamma$ region is a bipolar group where one or both the leading and following groups show complex polarity distributions.

δ: In δ-spots the umbra of the spot is seen to have two different polarities contained in the same penumbra. This kind of configuration often leads to flare activity.

Evolution of Sunspots

Sunspots may last for hours, days, or weeks, but their longevity usually depends on the size of the sunspot and on the strength of its magnetic field. The evolution of a large sunspot begins in the quiet-Sun. Sunspots emerge rapidly from the background granulation as pores that grow as their magnetic fields strengthen. When a pore exceeds about 3.5 Mm in size a penumbra begins to grow around the spot (Solanki 2003). In a full size sunspot magnetic features can be seen flowing away through the penumbra and out through the quiet-Sun, and as magnetic flux is lost the sunspot magnetic field begins to weaken, the umbra warms up, and complex bright structures, such as light bridges may be seen. The sunspot will decay back into a pore, and then disappears completely, leaving a ghost of its magnetic signature on the quiet-Sun.

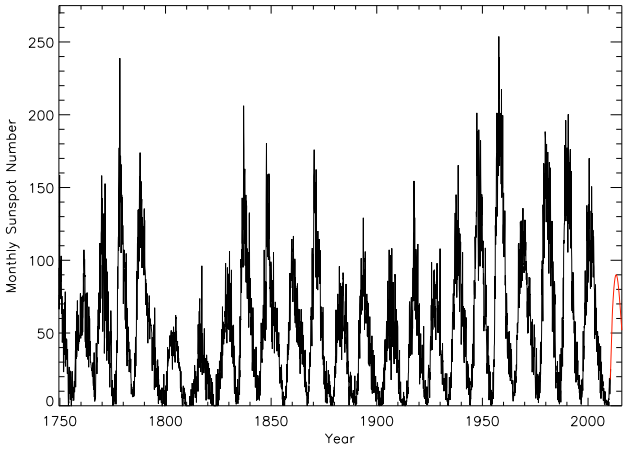

Figure 1.13 Monthly sunspot numbers from 1750 to 2011 (black) and the prediction for cycle 24 (red). Data taken from the SIDC, RWC Belgium, World Data Center for the Sunspot Index, Royal Observatory of Belgium.

1.5.1 Sunspots and Solar Activity

Reliable records of sunspot numbers have been kept from the 1750s to present day (Figure 1.13) and show a clear 11 year periodicity corresponding to the magnetic activity cycle discussed in the previous section and longer term variations for which a sufficient baseline has not yet been well established. The strength and number of sunspots is a direct measure of the strength of the solar magnetic cycle. Although it remains to be seen, our current solar cycle has been significantly weaker, and had a later onset than any cycle in the past 70 years, which may indicate lower levels of solar magnetic activity lie ahead of us.

As Carrington saw during the 1959 flare, sunspots are sites of violent solar activity. As the magnetic flux in complex regions is cancelled through magnetic reconnection, a large amount of magnetic energy is released as the magnetic fields relax into lower energy configurations in the form of a sudden brightening or flare. The release of energy during a flare causes the heating and acceleration of material above sunspots, and the ejection of high energy particles into space, sometimes toward the Earth, where they become trapped in the magnetic field and cause aurorae. This rearranging of the magnetic field during flares

is also associated with coronal mass ejections, which may remain bound to the Sun, as a prominence, or be entirely ejected into space.

1.6 Measuring Solar Magnetic Fields

A wide variety of solar activity is associated with the magnetic field. Therefore directly measuring magnetic fields on the Sun is of principle importance to understanding both the quiescent and dynamic behavior of the Sun which intimately affect the Earth. In the following section I describe spectroscopic methods of measuring the solar magnetic field which are an integral part of this work.

1.6.1 The Zeeman Effect

The Zeeman effect allows magnetic fields in sunspots to be measured directly from the magnitude of splitting in magnetically sensitive spectral lines. The phenomenon occurs when an externally applied magnetic field partially lifts the degeneracy of an energy level in an atom. The interaction between the combined spin and orbital angular momentum of the electron and the magnetic field causes a splitting in the energy levels (and wavelengths, $\Delta\lambda$) which depends on the magnetic field strength (B), the wavelength of the transition (λ), and the Landé g-factor (g) (Foukal 1990).

$$\Delta\lambda = \frac{\pi e}{m_e c} g \lambda^2 B \tag{1.6}$$

In this equation, e is the electron charge, m_e is the electron mass, and c is the speed of light.

The Zeeman effect occurs in a regime where the external magnetic field is much less than the internal magnetic field of the atom and the spin and orbital angular momentum remain coupled. Under stronger magnetic fields comparable to the internal magnetic field of the atom, the electron spin becomes entirely coupled to the external magnetic field which leads to additional splitting of the energy levels, and is known as the Paschen-Back effect. In the

case of the normal Zeeman Effect, in the presence of the magnetic field three components result from the original degenerate level, a higher energy level (blue-shifted) and lower energy level (red-shifted), the $+\sigma$ and $-\sigma$ components, and an unshifted component, called π, at the original wavelength.

Light is preferentially emitted or absorbed by the Zeeman-split atomic levels based on the geometry of the magnetic field, therefore the Zeeman effect also produces a polarization signature. The observed polarization of the π and σ components depends on the geometry between the magnetic field vector and the observed line of sight. For this reason a Zeeman-split spectral line is able to give information on both the magnetic field strength proportional to the line splitting, and the direction of the magnetic field vector, depending on the appearance of the polarized signal.

For most atomic transitions the Zeeman splitting of spectral lines is smaller than the line width due to the random thermal motions in the solar plasma (considering the g-factor for most lines, the low magnetic field strengths, and the poorly resolved granulation pattern of the quiet-Sun) but in sunspots the magnetic fields are stronger and certain lines of hydrogen, calcium, titanium, and iron that have large Landé g values are useful as magnetic field diagnostics. A large g-factor means that the splitting in wavelength of a line can be measured at a lower magnetic field strength, and magnetic field measurements can be made more accurately. From Equation 1.6 we can see that spectral lines at longer wavelengths should have much greater sensitivity, however the line width due to doppler broadening increases with wavelength as a factor of $1/\lambda$. In going to longer wavelengths the minimum splitting that can be resolved only increases by a factor of λ.

1.6.2 The Polarization of Light

Polarization arises from a variety of different optical phenomenon, for example by reflection or by retardance by a medium. Recent research indicates that some cephalopods are capable of both detecting and modulating the polarization of light (Mäthger & Hanlon 2006). Even the human eye is able to perceive linear polarization with limited sensitivity due to the

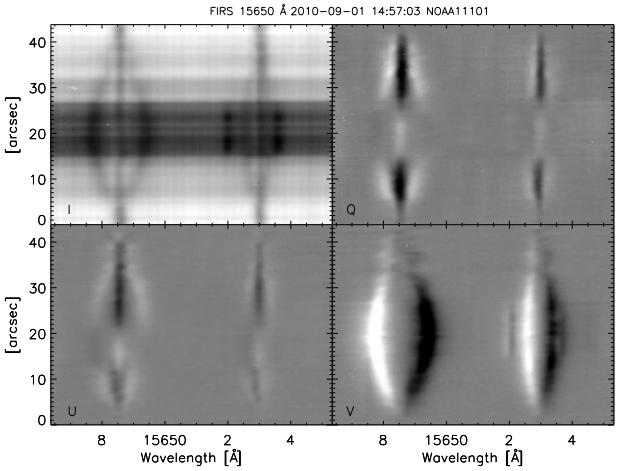

Figure 1.14 An example of the Stokes spectrum in a sunspot.

Table 1.2 Properties of Magnetic Diagnostic Lines

Species	λ_0 [Å]	χ_l [eV]	log gf [dex]	g_{eff}	Region	Reference
Fe I	6301.501	3.654	-0.718	1.67	upper photosphere	1
Fe I	6302.499	3.686	-1.235	2.49	upper photosphere	1
Si I	10827.089	4.954	0.363	1.50	photosphere	2
He I a	10829.09	...	...	2.0	chromosphere	3
He I b	10830.25	...	...	1.75	chromosphere	3
He I c	10830.34	...	...	0.85	chromosphere	3
Fe I	15648.515	5.426	-0.675	2.98	lower photosphere	2
Fe I	15652.874	6.246	-0.043	1.53	lower photosphere	2

(1)Vienna Atomic Line Database

(2)Borrero et al. (2003)

(3)Collados et al. (2003)

Figure 1.15 An exaggerated example of the appearance of Haidinger's brush pattern which can be seen by the human eye against a uniform polarized background.

orientation of the molecule xanthophyll, the pigment in blue cones. The yellow and blue pattern that can be percieved against strongly polarized uniform backgrounds is known as Haidnger's brush. An example is shown in Figure 1.15. Therefore it is crucial to have a standard method for measuring and describing the polarization of light.

1.6.3 Stokes Formalism and the Müller Matrix

The Stokes vector is a convenient way to fully describe the polarization state of an electromagnetic wave. It is defined as a four component vector:

$$\mathbf{S} = [I, Q, U, V] \tag{1.7}$$

where I is the total intensity of all light; $+Q$ is the light which is linearly polarized at 0° and $-Q$ is the light polarized linearly polarized at 90°; $+U$ is the light linearly polarized at 45° and $-U$ is the light linearly polarized at -45°; and $+V$ is the light which is right hand circularly polarized and $-V$ is the light which is left hand circularly polarized. By definition the net polarization should always be less than or equal to the total light,

$$I \leq \sqrt{Q^2 + U^2 + V^2} \tag{1.8}$$

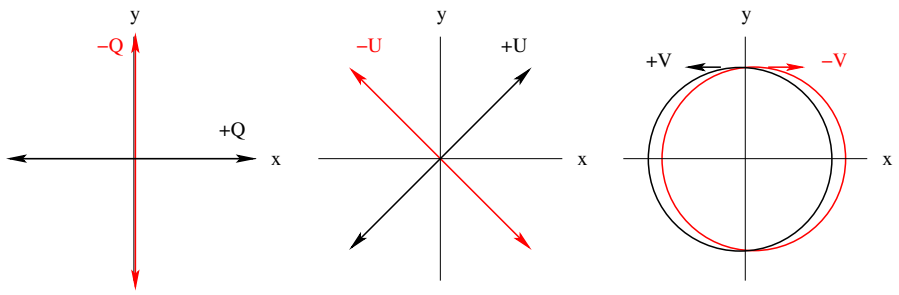

Figure 1.16 A visual representation of the polarized components of the Stokes vector.

It is possible to fully describe the polarization effects of an optical system with the Müller matrix. The transformation of the Stokes vector by a system ($\mathbf{S}'$) is written as the product of the input Stokes vector ($\mathbf{S}$) and the Müller matrix of the optical system ($\mathbf{M}$):

$$\mathbf{S}' = \mathbf{MS} \tag{1.9}$$

Some useful polarizing elements can be described in terms of their Müller matrix. A linear polarizer oriented at angle θ with respect to the input Stokes vector can be written:

$$\mathbf{M}_{lp}(\theta) = \begin{pmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \sin 2\theta \cos 2\theta & 0 \\ \sin 2\theta & \sin 2\theta \cos 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.10}$$

or a retarder with the fast axis oriented at angle θ and retardance ϕ is:

$$\mathbf{M}_{vr}(\theta, \phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + sin^2 2\theta \cos\phi & \sin 2\theta cos 2\theta(1 - \cos\phi) & -\sin 2\theta \sin\phi \\ 0 & \sin 2\theta \cos 2\theta(1 - \cos\phi) & \sin^2 2\theta + \cos^2 2\theta \cos\phi & \cos 2\theta \sin\phi \\ 0 & \sin 2\theta \sin\phi & -\cos 2\theta \sin\phi & \cos\phi \end{pmatrix} \tag{1.11}$$

Systems of linear polarizers and retarders can be combined to produce an optical system capable of measuring the Stokes vector in an efficient and accurate way. Such systems make

it possible to measure both the strength (from intensity) and orientation (from polarization) of Zeeman-split lines in solar magnetic fields. They system used in this research is fully described in Chapter 3.

1.7 Motivation for Studying Solar Magnetic Fields

While the Sun is important to the existence of life on Earth for obvious reasons, it is also the only star that can be investigated in detail, making understanding of its processes of principal importance in terms for understanding the energy budget of the Earth, the near-Earth environment, other stars, and the universe. The Sun is an astrophysical laboratory where the problems in fields such as relativity and quantum mechanics can be investigated. The study of solar magnetic fields, and sunspots in particular, is important for understanding the mechanisms involved in flares and coronal mass ejections. Sunspots are the most fascinating features of the visible Sun, their apparent stability against the pattern of granulation and their short lifetimes on a human timescale make their complicated evolution fascinating to observe. However many mysteries regarding their emergence, formation, and evolution, the formation of the penumbra, and the nature of features in the umbra, are still unresolved. In particular the research in this dissertation is focused on the equilibrium of gas and magnetic forces in the darkest regions of the sunspot umbra.

Chapter 2

Introduction

Abstract

In this chapter I identify the fundamental problem with the concept of magnetohydrostatic (MHS) equilibrium in a sunspot and address previous work that has been done on this subject. I define the best methodology for addressing this problem and summarize the work that will be presented in this dissertation.

2.1 Magnetohydrostatic Equilibrium in Sunspots

Sunspots are quasi-stable structures that can be considered to be in near magnetohydrostatic (MHS) equilibrium due to their long lifetimes relative to the hour-long dynamical timescale of the convective granules in the quiet-Sun (Meyer et al. 1977). Hale's historic observation of the Zeeman effect in sunspots revealed them as concentrations of strong magnetic fields (Hale 1908), and Biermann and Alfvén provided us with the first physically plausible explanation of the sunspot phenomenon (Biermann 1941; Alfvén 1943). Although the sunspot atmosphere has a low degree of ionization, its conductivity is high and the loose charges are able to flow freely through the plasma, therefore it can be assumed that the magnetic field is "frozen in" to the gas (i.e. the gas and magnetic field move together in the direction perpendicular to the magnetic field). In Alfvén's theory the strong vertical magnetic field in a sunspot inhibits the cross-field motion of the gas,

suppressing the convective transport of energy from below the sunspot, and maintaining the cool temperatures in the sunspot atmosphere. At any given height in the sunspot atmosphere, the magnetic field must support low gas pressure of the cool sunspot interior against the higher pressure of the hotter surrounding quiet-Sun (Deinzer 1965). For a sunspot with circular symmetry the MHS equilibrium state at any radius (r) and height (z) in the sunspot (for c.g.s. units) can be written as:

$$\begin{aligned} P_{qs}(z) &= n_{qs}(z)kT_{qs}(z) \\ &= n_s(z,s)kT_s(z,r) + \frac{1}{8\pi}(B_z^2(z,r) + F_c(z,r)) \end{aligned} \tag{2.1}$$

where the gas pressure in the quiet-Sun, P_{qs}, and the gas pressure in the sunspot are given by the product of the number density of the gas (n_{qs}, n_s), Boltzmann's constant (k), and temperature (T_{qs}, T_s). The lower gas pressure of the cooler sunspot is counteracted by the higher magnetic field which contributes a pressure term ($B_z^2/8\pi$) from the vertical component of the field (B_z) and a tension or curvature term (F_c) which is given by:

$$F_c(z,r) = 2\int_r^a B_z(z,r')\ \frac{\partial B_r(z,r')}{\partial z}\ dr' \tag{2.2}$$

where a is the maximum radius of the sunspot and B_r is the radial component of the magnetic field. The balance of forces in a sunspot is illustrated in Figure 2.1. The full derivation of these equations following the arguments in Jackson (1962) and Cowling (1976) is given in Appendix A.

Although this theory provides us with the basic description of the MHS equilibrium condition for sunspots, observational tests are difficult to carry out, even for the form of circular symmetry assumed in Equation 2.1. Because of radiative transfer (RT) effects, the magnetic and thermodynamic properties of a sunspot observed at different spatial locations may not originate from the same geometrical height in the atmosphere (i.e. sunspots are not flat), resulting in what is commonly known as the Wilson Depression. Figure 2.2 demonstrates the problem with continuum and magnetic field observations in

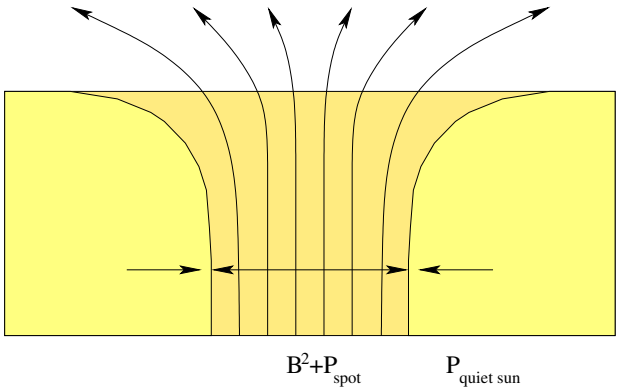

Figure 2.1 A diagram of the magnetic field in a simple magnetic flux tube.

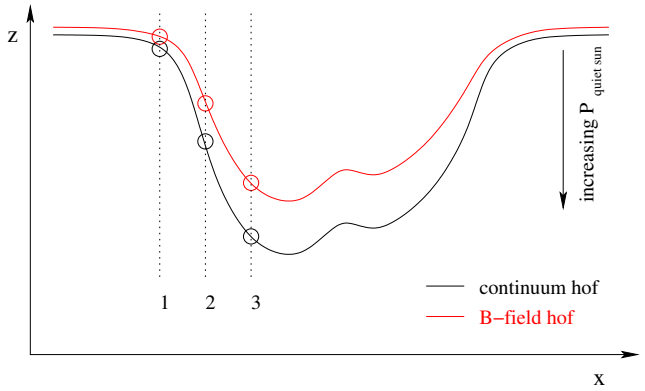

Figure 2.2 A figure demonstrating the problems in determining sunspot equilibrium caused by the RT effect. The different heights of formation for the continuum intensity and magnetic field are indicated by the black and red solid lines respectively.

the photosphere. The heights of formation of the continuum and magnetic field information vary with position in the sunspot atmosphere. Even observations from apparent positions which are relatively close together can originate at vastly different heights in the sunspot atmosphere. The equilibrium condition for each height is different because of the increase of the quiet-Sun and sunspot gas pressures at greater depth. Figure 2.3 shows an example of the estimated height of the $\tau = 1$ surface for the infrared continuum at 1.6 μm in a small, regular sunspot.

The height of a measurement in the sunspot atmosphere cannot be known, it changes from spot to spot depending on size, temperature, and magnetic field strength (see Figure

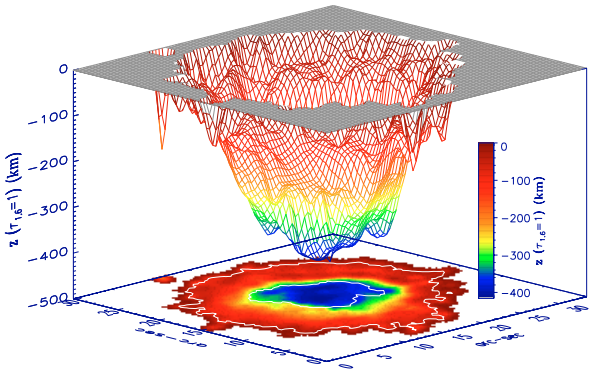

Figure 2.3 Depth of the Wilson Depression in a sunspot retrieved from IR observations. The vertical scale is roughly 50 times larger than the horizontal scale [Figure 7 from Mathew et al. (2004)].

2.2). Although it is possible to observe the foreshortening of sunspots as the Sun rotates them toward the limb, the height cannot be measured on a per-sunspot basis in a reliable way (e.g. see Watson et al. (2009)). Due to the RT effect, the thermal and magnetic pressures determined by observation do not necessarily reflect the true equilibrium state of a sunspot, and information about the vertical structure in a sunspot cannot be derived easily. Even observations covering multiple heights cannot be brought together in a meaningful way without preconceptions on how sunspots should work. Therefore the contribution to equilibrium from the curvature force term cannot be determined based on observations alone.

Sunspots cannot be in perfect equilibrium due to the fact that they evolve. In fact, they are the sites of many kinds of dynamic phenomena, such as the Evershed effect, which is a horizontal outflow seen in the penumbral photosphere (Evershed 1909), and at chromospheric heights the inverse Evershed effect is seen to flow into sunspots (Athay et al. 1982). Changes in the velocity and longitudal magnetic field strength corresponding to the 5 and 3 minute oscillations seen in the quiet-Sun can also be seen in sunspots (Kupke et al. 2000).

Under the restrictive conditions present in the sunspot umbra it is possible to make some simplifying assumptions. Evolution in the umbra occurs on much longer timescales than in the dynamic penumbra. The magnetic fields are mostly vertical, therefore the contribution from the curvature force term should be small and horizontal flows are prohibited. In addition, the radial gradient of temperature is small and the observations of magnetic field and temperature can be considered to originate at a single height. Therefore if we make the assumption that sunspots are vertical magnetic flux tubes with constant magnetic field strength, and have an ideal gas atmosphere with constant density and that information about the magnetic field and temperature come from a single height, then Equation 2.1 can be simplified to the thermal-magnetic relation:

$$B^2(r) \propto T_{qs} - T_s(r). \tag{2.3}$$

Although Equation 2.3 is elegantly simple it has not been verified observationally. There have been many investigations of the empirical relation between B^2 and T determined using observations of magnetically sensitive Zeeman-split lines in sunspots. While some of the relations derived are best described by a linear relationship (Gurman & House 1981; Livingston 2002; Penn et al. 2002, 2003a), the majority of modern observations show that B^2 is actually a highly non-linear function of T (Balthasar & Schmidt 1993; Kopp & Rabin 1993; Lites et al. 1993; Martínez Pillet & Vázquez 1993; Solanki et al. 1993; Stanchfield et al. 1997; Westendorp Plaza et al. 2001; Penn et al. 2003b; Mathew et al. 2004). Figure 2.4 gives an example of the thermal-magnetic relation for the main sunspot in NOAA 11130 that I have recently derived from spectropolarimetric observations of the Fe I lines at 15650 Å. The non-linearities in the B^2 vs. T relation are evident: the spot shows an unexpected sharp increase in magnetic field strength at relatively constant temperature in the darkest part of the sunspot. This result is puzzling. With the simplifying assumptions leading to Equation 2.3, Figure 2.4 implies the existence of a heretofore unidentified mechanism

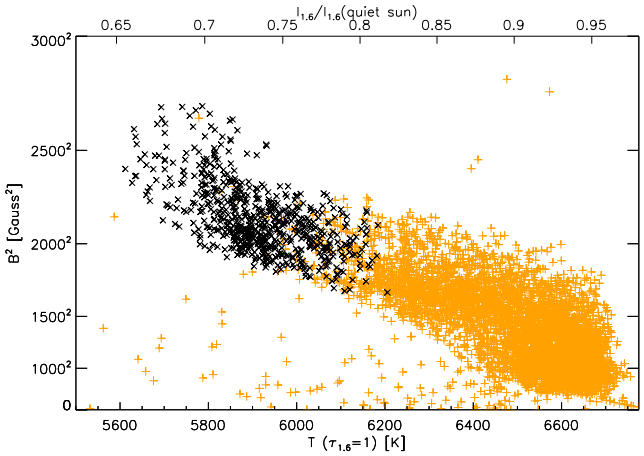

Figure 2.4 The B^2 vs T relation from NOAA 11130, which I observed on December 2, 2010 (see also Chapter 7).

at work in the coolest part of a sunspot that reduces the thermal pressure of the sunspot atmosphere to accommodate the higher magnetic pressure in these regions.

2.2 The Role of Molecular Hydrogen

In deriving Equation 2.3 we have made the assumptions that: 1) sunspot umbral magnetic fields are nearly vertical with respect to the local solar surface and that the curvature force has a negligible effect on the force balance of the sunspots, 2) the geometrical depth of the $\tau = 1$ level is roughly constant in the umbra, and that the temperature and magnetic field observed in the umbra of sunspots originate from roughly the same height in the sunspot atmosphere, and 3) the density of the solar atmosphere is a constant that does not change with temperature. While the validity of the first two assumptions is difficult to verify due to radiative transfer effects, they may be justified under restrictive conditions such as those encountered in the darkest part of the sunspot. On the other hand, the third assumption is obviously not warranted, especially under the temperature regime expected in the observable layer of the sunspots.

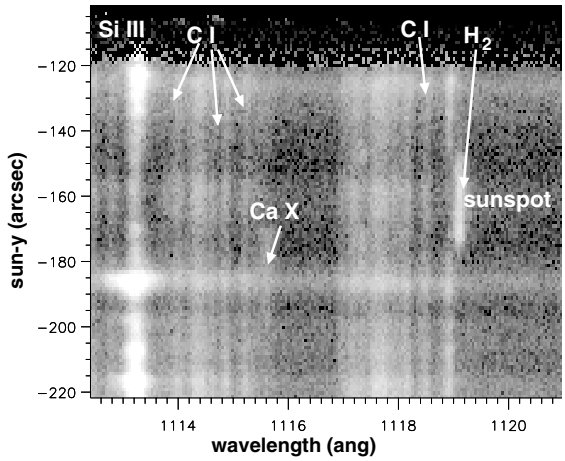

Figure 2.5 UV spectrum from the SUMER spectrograph with the slit placed across a sunspot [Figure 1 from Innes (2008)].

At an effective temperature of 5750 K the quiet-Sun photosphere is largely neutral and already harbors many heavy element molecular species with high dissociation energies (Grevesse & Sauval 1994). At even lower temperatures in the sunspot umbra many more molecular species, in particular molecular hydrogen (H_2), are able to form. Because the abundance of heavy elements (i.e. anything with an atomic number greater than 2) in the solar photosphere is very low compared to hydrogen, the number of molecules formed containing heavy elements is correspondingly low with respect to the total particle number density. Therefore the formation of heavy element molecules cannot have a significant effect on the thermodynamic properties of the solar atmosphere. However, in the umbra the atmosphere is very cool, making it possible to form a substantial fraction of H_2 molecules. The recombination of H_2 occurs most efficiently via a 3-body process, where two hydrogen atoms combine and a third particle escapes with the excess kinetic energy and bond energy released during recombination (Herzberg 1950). Observations of the fluorescent H_2 lines in the ultraviolet, such as the observation in Figure 2.5, confirm its presence in the chromosphere above sunspots (Jordan et al. 1978; Bartoe et al. 1979; Innes 2008), while

atmospheric models of the sunspot umbra predict a molecular hydrogen population of up to 10%, peaking near the height of continuum formation (Maltby et al. 1986).

The formation of a large fraction of H_2 may have important effects on the thermodynamic properties of the solar atmosphere and the physics of sunspots. For example, as free atoms form into molecules the energy of recombination is released (up to a maximum of the molecular dissociation energy). This may increase the local thermal energy content of the atmosphere if it cannot be dissipated rapidly enough. On the other hand molecules have the ability to store energy in rotational and vibrational degrees of freedom which do not contribute to the thermal signature of the gas, resulting in an increased heat capacity. Therefore if a significant fraction of the hydrogen atoms in the photosphere of the sunspot exist in molecular form, then the thermodynamic properties of the sunspot will be significantly different from that of the quiet-Sun photosphere. Finally, the combination of free atoms into molecules decreases the total particle number density and pressure of the gas. This effect is of particular importance for the problem of MHS equilibrium in sunspots since it provides a mechanism for changing the gas pressure of the sunspot atmosphere without a corresponding change in the temperature and may explain the isothermal intensification of the magnetic field strength in the darkest regions of the sunspot seen in Figure 2.4.

Even a small alteration in the balance of gas pressures in a sunspot may translate into large changes in the sunspot magnetic field. Consider, for example, the pressure vs. height relation for a standard model of the quiet-Sun, like the combined radiative Harvard-Smithsonian Radiative Atmosphere and convective Spruit atmosphere (Gingerich et al. 1971; Spruit 1974) shown in Figure 2.7. The solar surface, at 0 km, is defined at the position where $\tau_{500} = 1$, with negative values indicating greater depth in the solar atmosphere. On the right side of the figure the magnetic field strength equivalent to the pressure ($P = B^2/8\pi$) is provided. In a very large sunspot the maximum magnetic field strength in the umbra is perhaps 3500 G, and the Wilson Depression is about 600 km. Assuming the magnetic field is entirely vertical, the magnetic pressure must be supporting roughly half of the horizontal gas pressure of the quiet-Sun at this depth ($\sim 1 \times 10^6$

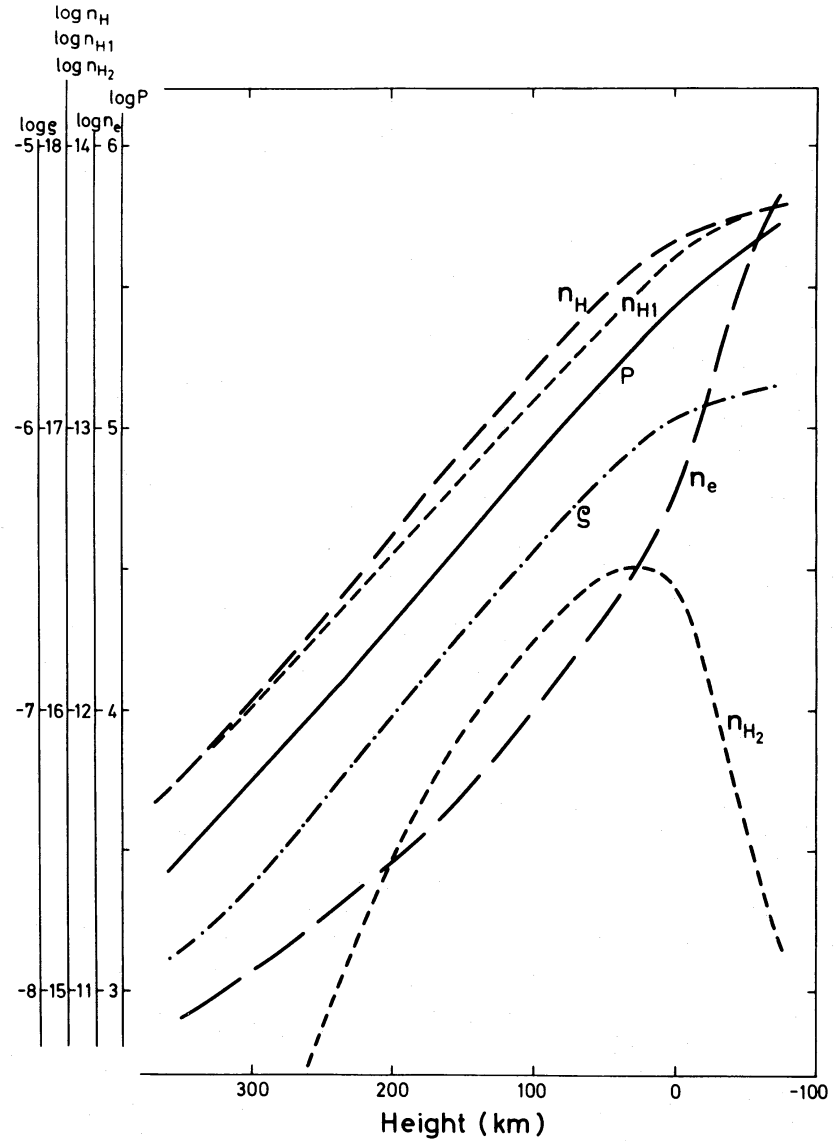

Figure 2.6 The stratification of various atmospheric properties with height (gas pressure P, density ρ, total hydrogen number density n_H, neutral hydrogen number density n_{H1}, molecular hydrogen number density n_{H_2}, and electron number density n_e) from a sunspot umbra model atmosphere [model M, Figure 9 from Maltby et al. (1986)].

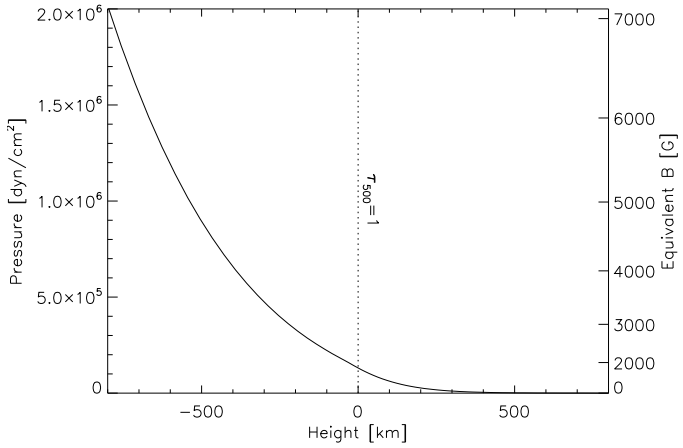

Figure 2.7 Pressure-height relation for the combined HSRA/Spruit quiet-Sun atmospheric model [data from Gingerich et al. (1971) and Spruit (1974)].

dyn/cm^2). If the atmosphere forms a 10% molecular hydrogen fraction, the gas pressure will be decreased by a factor 0.9 from what it originally was. Therefore a 10% increase in the magnetic pressure is necessary to maintain equilibrium of the sunspot, which translates into a 200 G increase in the magnetic field strength, which is of the order of magnitude seen in Figure 2.4. In smaller sunspots with higher temperatures the gas pressure term should be larger than the magnetic pressure, although less H_2 would be formed at higher temperatures, the effect on the magnetic field could be of the same order.

2.3 Objectives of This Research

Given the critical role that H_2 may potentially play in the physics of sunspots, it is the objective of this dissertation to prove that a significant fraction of H_2 exists in the photosphere of sunspots. I will also demonstrate that H_2 formation occurs at a temperature coinciding with the sharp rise of B^2 seen in Figure 2.4 to support the argument that this feature is due to the formation of H_2.

2.4 Methodology

To establish the existence of H_2 in the sunspot photosphere and its link with the isothermal intensification of the umbral magnetic fields, it is necessary to characterize the 2D magnetic field, temperature, and H_2 fraction of many sunspots of different sizes and at different stages of evolution. The past studies of the magnetohydrostatic force balance in sunspots listed in Section 2.1 were limited in spatial resolution, or included only a single sunspot. To this end, I have conducted a comprehensive survey of sunspots using a new high spectral and spatial resolution spectropolarimeter, obtaining observations of the infrared Fe I line pair at 15650 Å. I have derived the vector magnetic field configuration of the sunspots in the sample using an original Milne-Eddington inversion code. The direct detection of H_2 is exceptionally difficult, as will be described in Chapter 6, therefore a similar molecule, hydroxide (OH), is established as a proxy for H_2 based on spectral synthesis derived from atmospheric models.

Chapter 3

Instruments and Observations

Abstract

For each sunspot in this survey I have obtained spectropolarimetric observations of the Fe I lines near 15650 Å in the H-band and at 6302 Å in the visible. The infrared bandpass includes several strong OH lines. Observations were taken with the Horizontal Spectrograph (2001, 2005) and using the new Facility Infrared Spectropolarimeter (2009-2010) on the Dunn Solar Telescope which was commissioned by Haosheng Lin an myself during the course of this dissertation. Observations from additional instruments were used for the purposes of comparison and to provide context for the survey observations.

3.1 Observations

The original data presented in this dissertation were obtained using the Dunn Solar Telescope (DST) at the National Solar Observatory (NSO) located in Sunspot, New Mexico. Preliminary observations were taken with the Horizontal Spectrograph (HSG) in 2001 and 2005, and the main survey observations were taken over a period from July 2009 to December 2010 using the newly constructed Facility IR Spectropolarimeter (FIRS). For every sunspot in the survey I made spectropolarimetric observations of a 10 Å bandpass at 15650 Å covering the Fe I 15648.5 Å g=3 and Fe I 15652.9 Å g=1.7 lines, a blend of two OH lines at 15650.8 Å, and two OH lines on either side of the Fe I g=1.67 line at 15651.9 and 15653.7

Table 3.1 Summary of Sunspot Survey Observations

No.	Date	UT Time	Target	Config.	Lat [°]	Lon [°]	μ	IR S/N	Vis S/N
1	2001-04-18	14:44:17	NOAA 9429	α	8.39	13.42	0.947	298	···
2	2005-03-10	13:05:00	NOAA 10743	α	-6.41	62.80	0.456	1113	···
3	2005-03-10	23:31:01	NOAA 10741	β	-4.21	38.73	0.947	1095	···
4	2005-03-12	13:18:00	NOAA 10743	α	-6.38	39.80	0.770	873	···
5	2005-03-12	23:23:00	NOAA 10742	β	-4.93	9.86	0.954	972	···
6	2009-07-06	14:31:37	NOAA 11024	β	-24.63	25.45	0.786	1707	1627
7	2009-07-07	14:35:23	NOAA 11024	β	-23.84	33.25	0.728	2272	1859
8	2009-07-08	16:23:28	NOAA 11024	β	-23.76	52.54	0.525	1123	1051
9	2009-07-09	18:21:50	NOAA 11024	β	-27.64	59.70	0.407	796	781
10	2009-12-11	17:07:04	NOAA 11034	β	19.15	-47.93	0.614	1257	1283
11	2009-12-12	20:16:30	NOAA 11034	β	18.19	-28.25	0.818	1969	···
12	2009-12-15	16:22:10	NOAA 11035	β	28.36	0.30	0.865	736	749
13	2009-12-16	16:33:44	NOAA 11035 F	β	29.69	11.12	0.830	2337	2226
14	2009-12-17	15:21:43	NOAA 11035 F	$\beta\delta$	29.89	21.39	0.782	2323	2094
15	2009-12-17	15:58:13	NOAA 11035 L	$\beta\delta$	27.73	28.83	0.754	1710	1603
16	2009-12-18	15:35:23	NOAA 11035 F	β	29.66	33.45	0.695	1291	1278
17	2009-12-18	15:55:39	NOAA 11035 L	β	26.82	42.93	0.629	1325	1329
18	2010-02-12	16:41:27	NOAA 11046 L	$\beta\gamma$	21.79	0.88	0.865	2250	2253
19	2010-02-12	17:17:57	NOAA 11045 F	β	24.41	47.38	0.552	1635	1621
20	2010-02-12	17:49:06	NOAA 11046 F	$\beta\gamma$	22.83	-9.31	0.848	2482	2397
21	2010-02-13	16:54:13	NOAA 11046 L	$\beta\gamma$	22.57	13.90	0.834	2088	2084
22	2010-02-13	17:30:53	NOAA 11046 F	$\beta\gamma$	22.88	3.41	0.859	1936	1883
23	2010-02-14	16:05:49	NOAA 11046 F	β	22.47	17.08	0.823	1699	1657
24	2010-02-14	16:35:27	NOAA 11046 L	β	22.15	26.95	0.765	1696	1686
25	2010-02-15	15:34:59	NOAA 11046 F	β	22.54	29.50	0.743	1695	1660

continued on the next page

Table 3.1, continued.

No.	Date	UT Time	Target	Config.	Lat [°]	Lon [°]	μ	IR S/N	Vis S/N
26	2010-02-15	16:10:09	NOAA 11046 L	β	22.53	40.00	0.645	1339	1308
27	2010-02-15	16:38:17	NOAA 11048	β	18.92	-51.27	0.538	1463	1464
28	2010-02-17	16:12:27	NOAA 11049	β	-19.30	-3.60	0.980	1657	1652
29	2010-02-18	16:34:12	NOAA 11049 L	β	-20.72	5.28	0.973	1486	1440
30	2010-02-19	15:27:08	NOAA 11049 L	β	-19.00	23.65	0.904	1761	1766
31	2010-02-19	15:57:02	NOAA 11049 F	β	-20.61	17.72	0.933	1677	1691
32	2010-02-20	17:14:03	NOAA 11049 L	β	-20.53	32.47	0.833	1448	1458
33	2010-05-21	14:38:37	NOAA 11072	β	-16.33	-21.11	0.906	1696	1678
34	2010-05-22	14:51:39	NOAA 11072	β	-16.22	-7.13	0.965	2114	2029
35	2010-05-23	16:16:33	NOAA 11072	β	-16.04	7.86	0.964	818	797
36	2010-08-29	14:32:27	NOAA 11101	α	11.46	-13.70	0.968	1523	1278
37	2010-08-30	15:09:08	NOAA 11101	α	11.57	-0.43	0.995	1401	1293
38	2010-09-01	15:08:57	NOAA 11101	α	11.38	26.14	0.896	2003	1862
39	2010-09-01	15:47:18	NOAA 11102	β	26.36	44.09	0.687	1968	1905
40	2010-09-01	16:28:29	NOAA 11103/4	β	19.90	4.85	0.972	2216	2202
41	2010-09-02	14:27:11	NOAA 11101	β	10.97	38.70	0.782	1721	1634
42	2010-09-02	15:06:12	NOAA 11102	β	26.40	56.76	0.548	1773	1720
43	2010-09-02	15:33:49	NOAA 11103/4	β	24.90	43.58	0.697	1881	1867
44	2010-09-02	16:05:28	NOAA 11105	β	17.69	5.33	0.975	1698	1678
45	2010-09-04	14:20:15	NOAA 11105	β	17.86	35.29	0.802	1932	1879
46	2010-09-04	15:37:47	NOAA 11101	β	11.26	69.25	0.381	1798	1644
47	2010-09-05	15:45:56	NOAA 11101	β	11.78	80.75	0.233	659	669
48	2010-09-12	14:39:57	Pore Group	$\cdots$	30.27	-64.29	0.387	1137	1142
49	2010-09-12	15:30:41	NOAA 11106	α	-20.37	-55.85	0.486	1668	1597
50	2010-11-28	17:44:22	NOAA 11130 F	β	11.19	-5.52	0.976	2551	2538

continued on the next page

Table 3.1, continued.

No.	Date	UT Time	Target	Config.	Lat [°]	Lon [°]	μ	IR S/N	Vis S/N
51	2010-11-30	15:20:05	NOAA 11130 L	β	12.06	24.19	0.892	2341	2208
52	2010-11-30	16:58:33	NOAA 11130 F	β	12.49	19.14	0.922	2104	2059
53	2010-12-01	15:39:36	NOAA 11130 F	β	11.86	31.04	0.837	1946	1908
54	2010-12-01	16:15:27	NOAA 11130 L	β	11.56	37.45	0.776	2338	2236
55	2010-12-02	15:34:58	NOAA 11130 L	β	11.03	49.91	0.624	1702	$\cdots$
56	2010-12-02	16:41:57	NOAA 11131	α	30.45	-68.13	0.301	2258	2266
57	2010-12-03	15:49:37	Pore Group	$\cdots$	34.00	14.85	0.788	1829	1813
58	2010-12-03	16:02:50	NOAA 11131	α	30.59	-56.34	0.459	2447	2390
59	2010-12-03	16:46:10	NOAA 11130	β	11.10	63.32	0.426	2627	2545
60	2010-12-04	15:27:55	NOAA 11132	β	9.78	-2.76	0.982	1916	1909
61	2010-12-04	16:25:41	NOAA 11131	α	30.69	-43.94	0.605	2241	2142
62	2010-12-05	15:09:55	NOAA 11132	β	10.41	6.06	0.974	2671	2641
63	2010-12-05	15:55:07	NOAA 11131	β	31.38	-36.56	0.686	2213	2071
64	2010-12-05	17:03:58	NOAA 11133	α	13.75	-61.63	0.455	2099	2022
65	2010-12-06	16:00:07	NOAA 11133	α	13.67	-49.19	0.628	1584	1592
66	2010-12-06	17:00:15	NOAA 11131	β	29.78	-18.45	0.815	1672	1610

Å. In most cases simultaneous visible wavelength observations were made with the visible wavelength channel of FIRS at 6302 Å which spans 5 Å and includes the Fe I g=1.5 line at 6301.5 Å and the Fe I g=2.5 line at 6302.5 Å, the bandpass also includes two telluric O_2 lines at 6302.0 and 6302.8 Å. The complete survey contains 66 observations of sunspots and pores associated with 23 different active regions. Table 3.1 summarizes the observed active regions, the date and time of observation, their latitude, longitude, heliocentric position ($\mu = \cos\theta$), and signal to noise of the observation in the quiet sun.

3.2 The Dunn Solar Telescope

Both of the instruments used in this survey were located at the Richard B. Dunn Solar Telescope (DST) which is a vacuum-coelostat solar tower operated by the National Solar Observatory and located at an elevation of 9,250 ft. on Sacramento Peak, New Mexico, USA. A diagram of the telescope is shown in Figure 3.1. To avoid near-ground seeing effects, the light enters the telescope 41 m above ground. Turbulence due to air currents and scattered light within the telescope are mitigated by keeping a majority of the light path contained under a low vacuum. Sunlight enters through the 0.76 m diameter quartz vacuum window at the top of the tower. The light is directed vertically down by a pair of flat, 1.1 m diameter mirrors to the 1.5 m diameter main mirror, which has a focal length of 55 m. The telescope has an exceptionally slow effective focal ratio of f/72 to avoid excessive heating of the air within the light path. The main mirror returns the focused beam to the observing floor at ground level where the light exits though another vacuum window and is directed to the instruments. A schematic of the instrument layout on the observing floor is shown in Figure 3.2. The full interior length of the telescope, including the observing floor, is suspended from a mercury float so it is able to freely rotate to keep the image orientation fixed relative to the instruments as the telescope tracks the Sun throughout the day.

Sunlight collected by the telescope is directed to different exit ports by tilting the main mirror. A majority of the instruments are located off of Port 4 where the calibration and

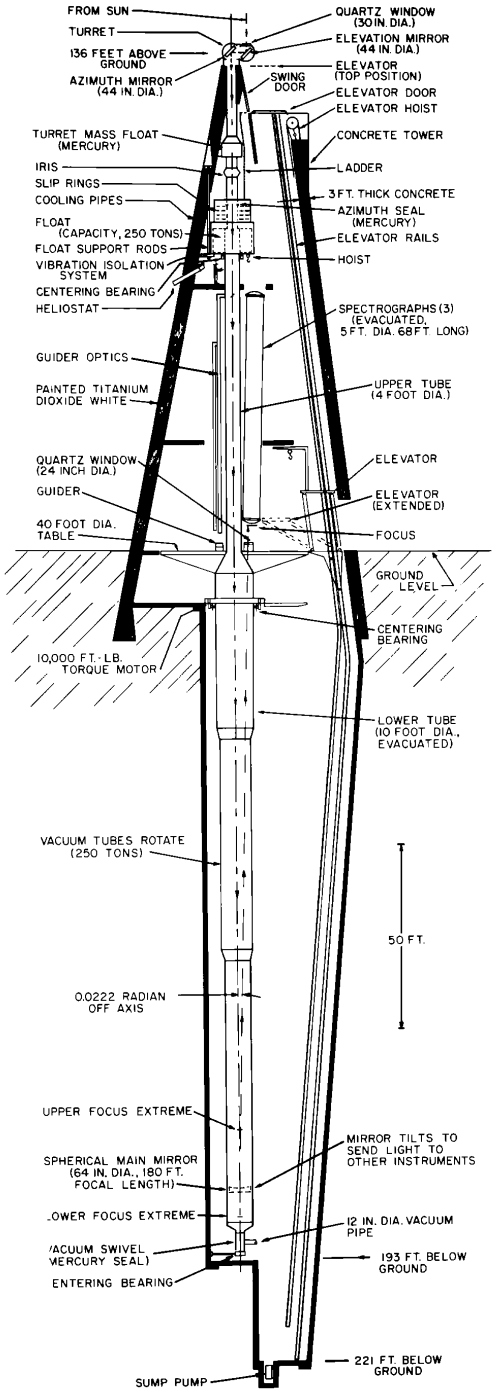

Figure 3.1 A diagram of the Richard B. Dunn Solar Telescope [credit: NSO].

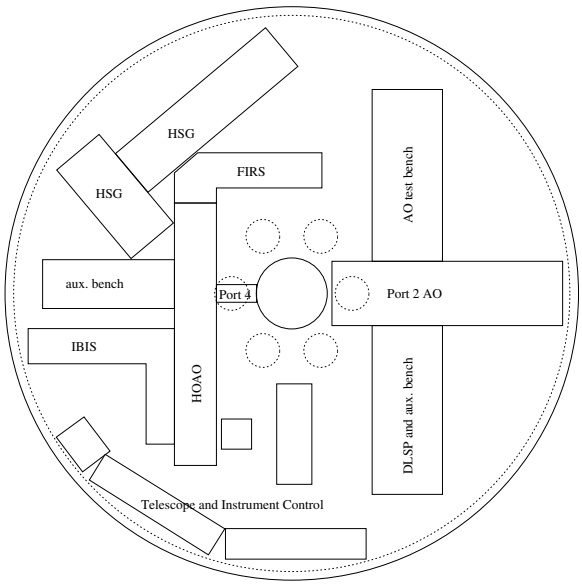

Figure 3.2 Layout of the DST observing floor.

modulation assembly described in Elmore et al. (1992), originally installed for the Advanced Stokes Polarimeter, resides. This assembly consists of an achromatic 1/4-waveplate and linear polarizer, each in rotation mounts, a pinhole, 150" × 150" (40 × 40 mm) field stop, and a wide variety of grids and targets for spatial calibration. From here the light is fed into the High Order Adaptive Optics (HOAO) System (Rimmele et al. 2004). The combined power of a correlation tracker and tip-tilt mirror, and an independently operating Shack-Hartmann sensor and 97 actuator deformable mirror, provides diffraction limited image quality at visible wavelengths in good daytime seeing conditions, and provides excellent image correction at infrared wavelengths even in mediocre seeing conditions.

3.3 The Horizontal Spectrograph

The Horizontal Spectrograph (HSG), used for the preliminary infrared sunspot observations, is a versatile and reconfigurable 3 m spectrograph arm with multiple available gratings. It consists of an adjustable rectangular slit, collimator lens, plane grating, camera lens and detector. An example of a setup on the HSG is shown in Figure 3.3 filling the two HSG

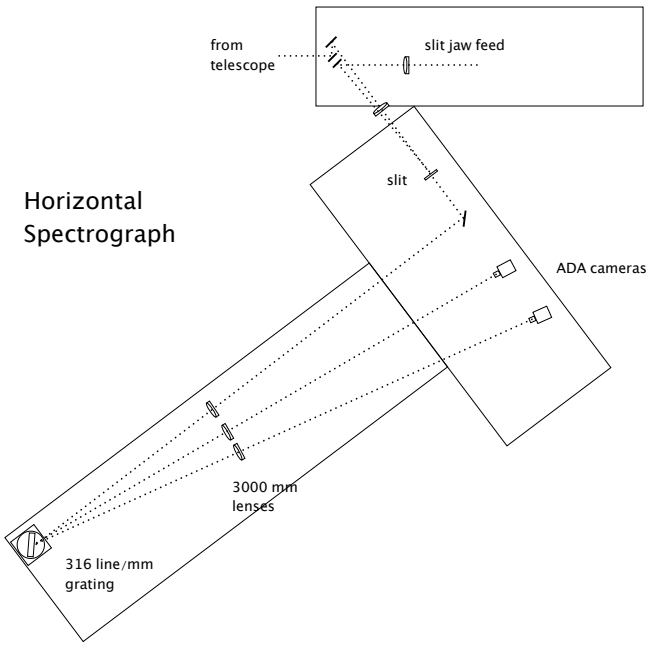

Figure 3.3 Example layout of the Horizontal Spectrograph (HSG) [credit: NSO].

optical benches and neighboring auxiliary bench. The experimental HSG setup for these sunspot observations was similar to those described in Lin (1995) and Lin et al. (1998). During the 2001 observations the 256×256 HgCdTe NICMOS 3 detector was used to record Stokes spectra with 47 mÅ/pixel dispersion, providing $0.3" \times 0.5"$(slit) spatial sampling. For the 2005 observations this detector was replaced with a 256×256 HgCdTe Rockwell TCM 2620 which was initially used by Kuhn et al. (1999), and in this setup the spectrograph provided 66 mÅ/pixel dispersion and $0.3" \times 0.3"$(slit) spatial pixels.

3.4 The Facility Infrared Spectopolarimeter

The Facility Infrared Spectropolarimeter (FIRS) was used for the majority of the survey observations and is a recently completed multi-wavelength spectropolarimeter built by the Institute for Astronomy, University of Hawai'i (Haosheng Lin and myself) in collaboration with the National Solar Observatory. FIRS is a versatile instrument which has been designed

to simultaneously observe two wavelength regions in the visible and infrared containing magnetic field diagnostics which span a range of heights in the solar photosphere and chromosphere, Fe I 6302 Å in the visible and Si I/He I 10830 or Fe I 15650 Å. FIRS incorporates several unique design elements to fulfill the need for a high cadence, multi-wavelength instrument with high spatial and spectral resolution, surpassing the abilities of other current instruments on the DST. The commissioning of FIRS is a critical part of this dissertation, and a large part of this work has already been presented in Jaeggli et al. (2010) and will soon be published in Lin et al. (in prep.). FIRS is a proof of concept for the next-generation Diffraction-Limited Near-Infrared Spectropolarimeter (DL-NIRSP) that will be built for the Advanced Technology Solar Telescope (ATST).

3.4.1 Layout of FIRS

Light enters FIRS directly through the end of the HOAO bench. A diagram of the light path through FIRS is shown in Figure 3.4. The beam first encounters a motorized mirror and 95/5 beam splitter which allow the beam to scan horizontally across the spectrograph slit unit downstream. Interchangeable motorized optics re-image the telescope beam and allow the user to select either a low-resolution f/36 beam (175" × 75") or a high-resolution f/108 beam (58" × 25"). A fold mirror redirects the beam onto the interchangeable slit unit, which is an optically flat glass element aluminized with the desired slit mask pattern. A variety of slit masks of different widths and configurations are available for FIRS (see Table 3.2) and allow the user to change the sampling resolution of the spectrograph. The beam reflected from the slit unit returns to a slit-jaw video camera which shows the observer where the spectrograph slits are located.

The FIRS spectrograph is in a near-Littrow configuration. Light passes through the slit unit and enters the spectrograph, traveling first to the off-axis parabolic mirror for collimation. The collimated beam then travels to a steeply inclined echelle grating and is spectrally dispersed. The light returns nearly along the same path to the parabolic mirror, which refocuses the beam. The pick-off mirrors then fold the beam down separate optical

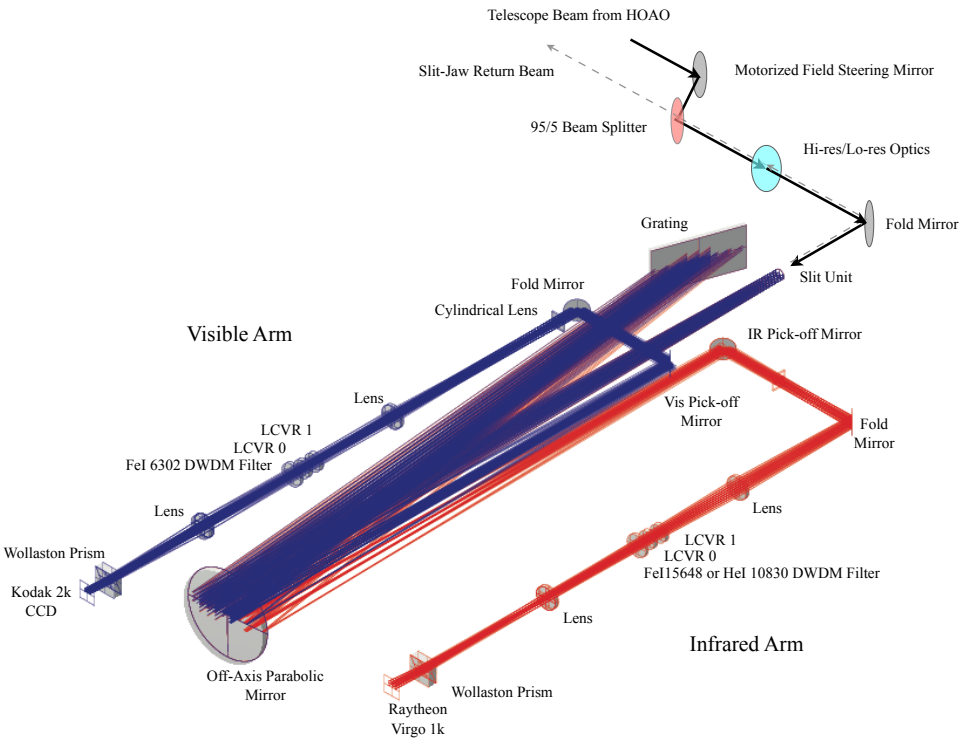

Figure 3.4 Diagram of the FIRS optical path following the Port 4 HOAO optics [credit: Don Mickey].

paths for the visible and infrared. Each light path has its own re-imaging optics, liquid crystal variable retarders (LCVRs), a narrow-band filter, a Wollaston prism (polarizing beam splitter), and a detector. The visible detector is a 2048×2048 Kodak CCD and the infrared detector is a 1024×1024 cryogenically cooled Raytheon Virgo HgCdTe array. The infrared arm can be easily converted from 10830 Å to 15648 Å modes simply by swapping the filters and LCVRs, and changing the position of the final lens between two preset positions.

To extend the chromospheric and photospheric wavelength coverage and increase the compatibility of the instruments on the DST, FIRS can operate simultaneously with the visible G-Band and white-light cameras, and the Interferometic Bi-dimensional Imaging Spectrometer (IBIS). A dichroic beamsplitter placed between HOAO and FIRS passes the Hα 6563 Å or Ca II 8542 Å bandpass to IBIS and sends the remaining visible and infrared

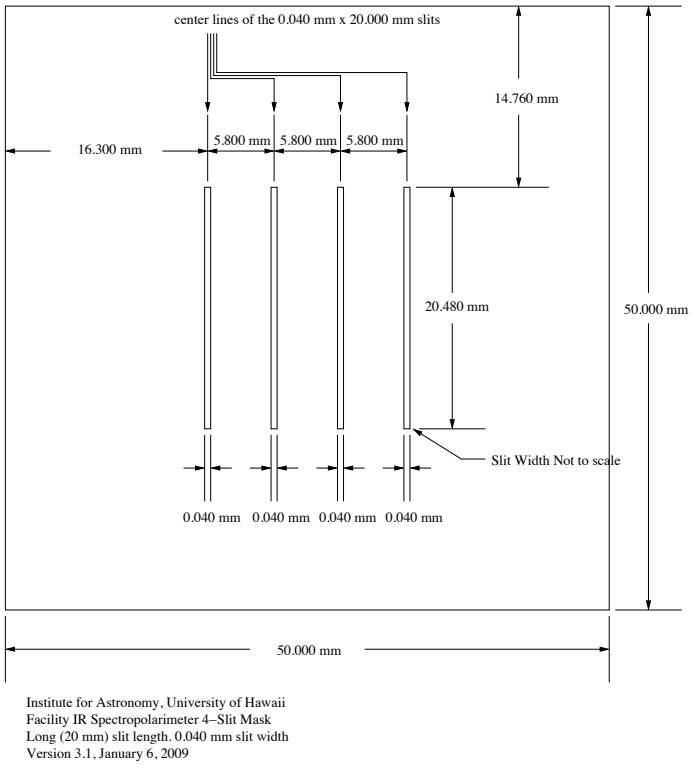

Figure 3.5 Schematic of the default FIRS slit unit containing four 40μm slits.

light to FIRS, however the FIRS visible bandpass cannot be observed simultaneously with IBIS Hα due to the proximity of the wavelengths.

3.4.2 Design Features

It should already be evident that FIRS incorporates many design features which are not typical to other spectrographs and warrant further explanation. In the following subsection I describe the unique features which allow FIRS to achieve high spectral and temporal resolution, and high polarization precision.

Slits and DWDMs

In order to build up an image of the solar surface the image plane is moved across the slit using the field scanning mirror, building up a scan (or raster) of spectral observations

at each position. In FIRS we have achieved this by stepping the motorized field scanning mirror in angle using a high precision motorized micrometer. A full-Stokes spectrum is obtained at each position as the micrometer scans through angles.

The magnetic diagnostics of interest cover a relatively small area of the detector while still achieving good spectral sampling. By filtering the spectral window to a narrow wavelength around the magnetically sensitive lines, and increasing the number of slits FIRS makes better use of the large area of the detector, and achieves an increased observing cadence over a traditional long-slit spectrograph. Figure 3.5 shows a schematic of the default 4-slit unit for FIRS. A number of slits with different widths and spacings are available, these are listed in Table 3.2. The multi-slit unit allows FIRS to scan a region four times larger than a single slit in the same amount of time.

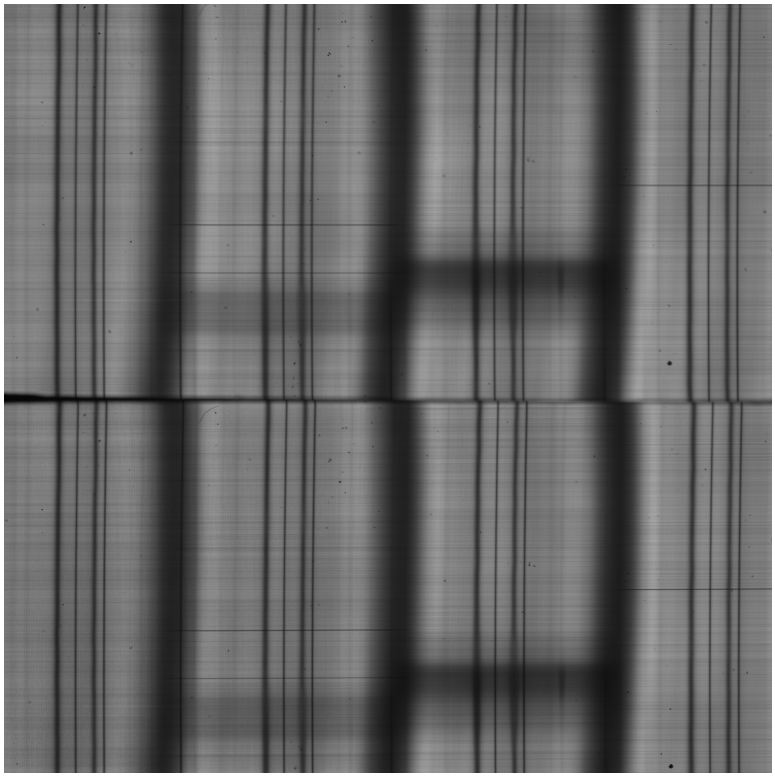

Figure 3.6 Example of a raw spectral intensity image in a sunspot obtained with FIRS 6302 Å.

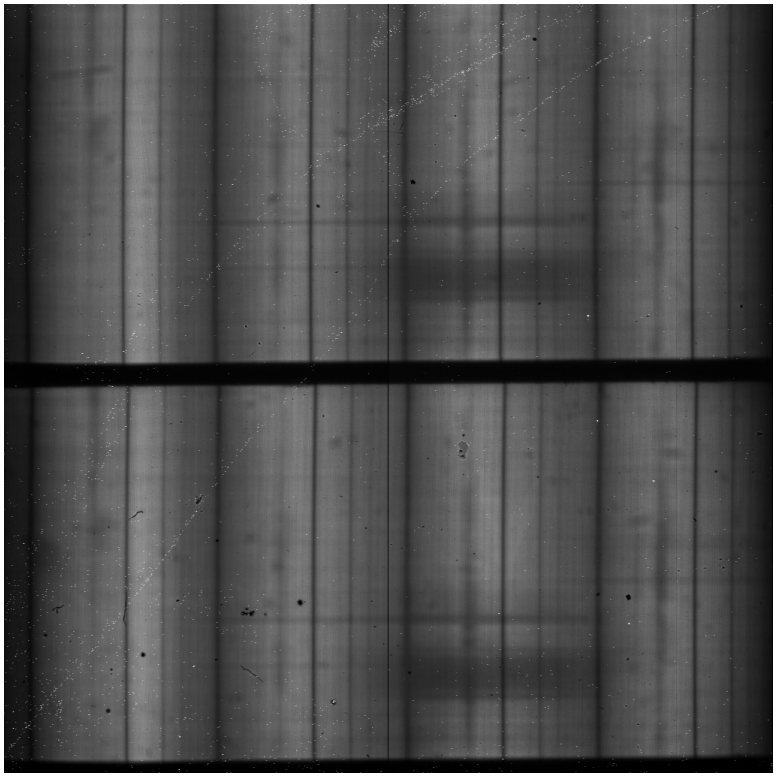

Figure 3.7 Example of a raw spectral intensity image in a sunspot obtained with FIRS 10830 Å.

Narrow-band filters for each wavelength are necessary to keep the spectra from multiple slits from overlapping on the detector and to prevent contamination from different wavelengths from overlapping orders. For this purpose narrow-band filters known as dense wavelength-division multiplexers (DWDMs) have been adapted from telecommunications applications. The effect of the filters can be seen in the intensity images produced by FIRS for each wavelength, such as those in Figures 3.6, 3.7, and 3.8. The spectra produced by each of the 4 slits are arrayed left to right across the detector and the three dark columns, most visible in the 6302 and 15650 Å data, are where the filter drops cuts off to the red and blue in each spectrum. The 10830 Å filter is slightly broader than the slit spacing, and the spectra from the 4 slits overlap, making it more difficult to disentangle the signal from each slit.

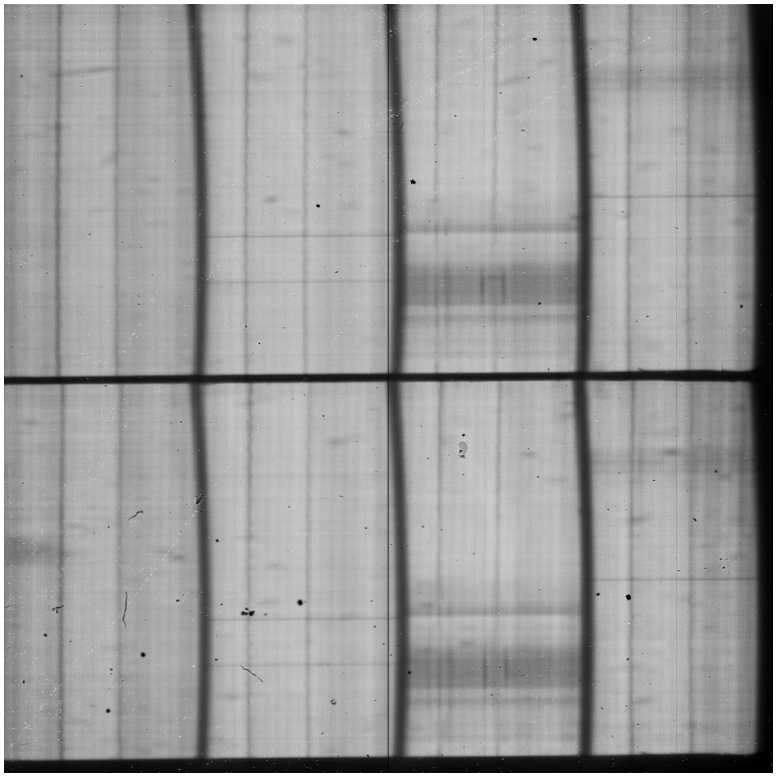

Figure 3.8 Example of a raw spectral intensity image in a sunspot obtained with FIRS 15650 Å.

Table 3.2 Available FIRS Slits

Slit Name	No. of Slits	Slit Width [μm]	Slit Length [mm]	Slit Spacing
15 μm short	4	15	10.0	5.70
15 μm long	4	15	20.48	5.70
30 μm short	4	30	10.0	5.70
30 μm long	4	30	20.48	5.70
30 μm double	2	30	20.48	11.4
40 μm single	1	40	20.48	$\cdots$
40 μm long	4	40	20.48	5.80
80 μm long	4	80	20.48	6.0

Grating

FIRS achieves high spatial resolution using a large echelle grating with a coarse 31.6 line/mm groove spacing and a blaze angle of 63.5°. The steep blaze angle of the grating provides

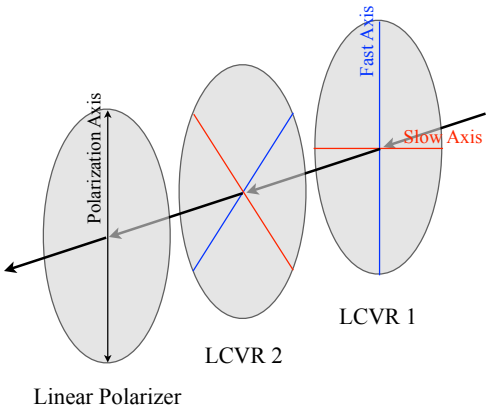

Figure 3.9 A cartoon of a full-Stokes polarimeter with LCVR modulators and a linear polarizer analyzer.

high efficiency at very spectrally dispersed orders. The spectrograph geometry has been optimized for high dispersion. The visible and infrared pick-off mirrors are positioned to select 6302 Å from the 90th order, 10830 Å from 34th order, and 15650 Å from 36th order. The spectrograph was designed to achieve a theoretical resolution of equal to 300,000 in the IR and 600,000 in the visible.

LCVRs and Wollaston Prism

Because FIRS runs at high cadence, it is necessary to measure the polarization states quickly, and with good precision. Knowledge of the full Stokes vector is necessary to recover the full geometry of the magnetic field. With the exception of radio wavelengths, it is not possible to directly measure the polarization of electromagnetic radiation. It is necessary to encode the polarization as an intensity signal before it can be measured. For FIRS we have chosen to use a full-Stokes polarimeter with two Meadowlark liquid crystal variable retarders (LCVRs) and a linear polarizer. A diagram for the light path in such a polarimeter is shown in Figure 3.9. The retardance of a liquid crystal changes based on the applied voltage, and they provide a good way to rapidly modulate the polarization of the beam. To obtain the Stokes vector, the LCVRs are operated in an "efficiency-

balanced" tuning scheme which yields the polarized components of the Stokes vector with equal efficiency. For this scheme four measurements must be made with the retardance of the first LCVR set to $[45°, 135°, 225°, 315°]$ and the retardance of the second LCVR set to $[54.7°, 126.3°, 54.7°, 126.3°]$ for the four images. The resulting combination of Stokes vector components (I, Q, U, V) encoded in the four intensity images (I_1 to I_4) are:

$$
\begin{aligned}
I_1 &= 1/2\ (I - Q/\sqrt{3} - U/\sqrt{3} + V/\sqrt{3}) \\
I_2 &= 1/2\ (I - Q/\sqrt{3} + U/\sqrt{3} - V/\sqrt{3}) \\
I_3 &= 1/2\ (I + Q/\sqrt{3} - U/\sqrt{3} - V/\sqrt{3}) \\
I_4 &= 1/2\ (I + Q/\sqrt{3} + U/\sqrt{3} + V/\sqrt{3})
\end{aligned}
\tag{3.1}
$$

The Stokes vector can be easily recovered from the observation by adding and subtracting images in post-processing.

The Wollaston prism or polarizing beam-splitter is the linear polarizer which analyzes the polarization modulated by the LCVRs, turning it into an intensity signal which can be measured by the detector, but it plays another essential role. The solar photosphere is richly structured on small scales and different regions often have different velocity signatures. Alterations in the seeing cause the light from slightly different regions, with slightly different velocities, to fall into the slit between spectral images taken even a short time apart, therefore the spectral lines are shifted. In retrieving the polarized components of the Stokes vector from a set of images for the efficiency-balanced scheme, we difference the images, and these small line shifts which change from image to image appear as an anti-symmetric polarization signal in Stokes Q, U, and V. This is known as the "spurious seeing" effect.

The Wollaston prism produces two parallel beams which are orthogonally polarized. In FIRS, the Wollaston prisms reside immediately in front of the detector for each beam to ensure there is no distortion or change in efficiency or polarization of the two separate beams due to non-uniformity of intervening optics. Because the two beams are orthogonally polarized, the real polarization signal in the demodulated Stokes vector will have the

Table 3.3 Properties of the FIRS Standard Configuration

Property	FIRS f/36	FIRS f/108	Hinode SOT/SP
Telescope	76.2 cm Coelostat	...	50 cm Gregorian
θ_R(6302 Å)	0.21"	...	0.32"
θ_R(10830 Å)	0.36"	...	...
θ_R(15650 Å)	0.52"	...	...
Field	174" × 75"	58" × 25"	160" × 151"
Spatial Sampling (Visible)	0.30" × 0.08"/pix	0.10" × 0.03"/pix	0.15" × 0.16"/pix
Spatial Sampling (Infrared)	0.30" × 0.15"/pix	0.10" × 0.05"/pix	...
Scan Time	18 min	25 min	83 min
Resolution (Sampling) 6302	0.01 (0.01) Å	...	0.03 (0.02) Å
Resolution (Sampling) 10830	... (0.04) Å	...	...
Resolution (Sampling) 15650	0.07 (0.05) Å	...	...

opposite sign in the different beams. The spurious seeing effect, however, results from the difference of the images and has the same sign between the two orthogonally polarized beams, so by subtracting the two beams the spurious seeing signal can be removed and we are able to achieve a higher accuracy in determining the real polarization signal and therefore, the vector magnetic field. In Figures 3.6, 3.7, and 3.8 the two beams are clearly visible in the top and bottom halves of the detector, and the same intensity features are reproduced in each beam. The two beams are cleanly divided by a dark row corresponding to the sharp top and bottom edges of the slit.

3.4.3 Instrument Properties

The general properties of FIRS are listed in Table 3.3. FIRS achieves a four times increase in cadence over the space-based Hinode Spectropolarimeter, for which I have also listed the properties.

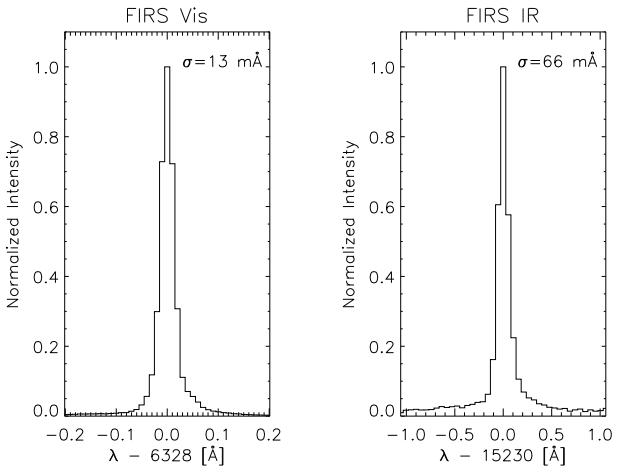

Figure 3.10 Averaged laser profiles measured with the visible and infrared for FIRS.

Laser Profiles

The spectral resolution of FIRS at both visible and infrared wavelengths was measured directly by taking laser profiles on 8 Sept. 2010. For our purposes the light from a laser can be considered to be entirely monochromatic, producing spectral lines of infinitesimal width. Light from a visible HeNe laser with a wavelength of 6328 Å was fed into a large core fiber, then the light from the fiber was collimated to illuminate the slit. It was necessary to remove the narrow band filter and refocus the spectral image by adjusting the last lens in front of the detector. The spectrograph grating was rotated such that the laser wavelengths fell onto the visible detector, and a set of laser observations were taken. Because it is not possible to see the infrared laser light, I made use of the existing alignment for the visible wavelength laser and directly replaced it with an infrared HeNe laser with a wavelength of 15230 Å. The spectrograph grating was rotated for this wavelength, the infrared filters were removed, the image was refocused, and laser observations were taken. The profiles from the observation were fit with a Gaussian function. The resulting average profiles for the visible and infrared are shown in Figure 3.10. The profiles had an average Gaussian width ($e^{-\lambda^2/2\sigma^2}$) of $\sigma = 13$ mÅ for the visible and $\sigma = 66$ mÅ for the infrared, implying the real spectral resolution is 340,000 for the visible and 160,000 for the infrared.

3.4.4 Future Upgrades

Several upgrades to extend the capabilities of FIRS are already being considered. With new Wollaston prisms for both the visible and infrared beams the spectrograph could be made more optically efficient. Increasing the beam separation at the camera and would increase the field of view and decrease the large factor of spatial oversampling. Additional narrow band filters would extend the simultaneous wavelength coverage of the spectrograph, and with additional detectors for both the optical and infrared FIRS would be able to simultaneously obtain spectra of the Fe I line pair at 6302 Å, Ca II line at 8542 Å, Si I and He I lines at 10830 Å, and Fe I line pair at 15648 Å with minimal changes to the current configuration.

3.5 Survey Observations

Observations with the HSG were taken by Haosheng Lin in 2001 and 2005. Data were taken with FIRS periodically throughout its commissioning, starting in April 2007. The instrument was in a continual state of change until July 2009 and data taken prior to this time were not consistent and of varying quality. Data considered for this dissertation began in July 2009 and continued until Dec 2010. Typical parameters for the survey observations are given in Table 3.5

3.6 Additional Observations from Other Instruments

I have made use of spectropolarimetric observations from additional instruments for comparison and to intensity observations to provide context for the survey observations obtained with FIRS and the HSG.

Hinode SOT/SP

The Hinode (Solar-B) satellite was launched on September 22, 2006, and consists of three main instruments: the Solar Optical Telescope (SOT), X-ray Telescope (XRT) and Extreme

Table 3.4 Summary of Observing Runs

Dates						Description
30	Apr	-	6	May	2007	FIRS Commissioning
11	Oct	-	21	Oct	2007	FIRS Commissioning
10	Dec	-	17	Dec	2007	FIRS Commissioning
26	Feb	-	9	Mar	2008	FIRS Commissioning
9	Jun	-	17	Jun	2008	FIRS Commissioning, Virgo 1K array in place
6	Sep	-	14	Sep	2008	FIRS Commissioning
8	Dec	-	16	Dec	2008	FIRS Commissioning, Virgo 1K upgraded to science array
17	Feb	-	1	Mar	2009	FIRS Commissioning
15	May	-	23	May	2009	FIRS Commissioning
7	Jul	-	11	Jul	2009	TO training, field scanning mirror replaced scanning slit mask
10	Sep	-	20	Sep	2009	FIRS Commissioning
9	Dec	-	18	Dec	2009	FIRS Commissioning
10	Feb	-	21	Feb	2010	SJ dissertation obs.
10	May	-	19	May	2010	SJ dissertation obs.
26	Aug	-	12	Sep	2010	SJ dissertation obs., measured laser profiles, Virgo 1K linearity
28	Nov	-	6	Dec	2010	SJ dissertation obs.

Table 3.5 Summary of Observations

Observation	IR Exp [msec]	IR Co-Add	Vis Exp [msec]	Vis Co-Add	Number
6302/15650 Å					
Scan	125	3	350	2	145
Dark	125	3	350	2	16
Flat	125	3	350	2	16
Pol. Cal.	125	3	350	2	36
6302/10830 Å					
Scan	250	2	350	2	145
Dark	250	2	350	2	16
Flat	250	2	350	2	16
Pol. Cal.	250	2	350	2	36

Ultraviolet Imaging Spectrometer (EIS) (Kosugi et al. 2007). The SOT (Tsuneta et al. 2008) is the largest optical solar telescope yet flown, and with an aperture of 50 cm it provides an unprecedented 0.2" resolution of the solar surface. The Spectro-Polarimeter (SP) on the SOT is, like FIRS, an off-axis Littrow echelle spectrograph which performs dual-beam full Stokes spectropolarimetry of the Fe I line pair at 6302 Å. The two instruments are very similar and achieve very comparable spatial and spectral resolution, so for this reason I have included the parameters of the SOT/SP in Table 3.3. However, the SP is limited by a single slit and does not achieve as high an observing cadence. Special care has been taken to accurately remove instrumental polarization from the reduced observations (Ichimoto et al. 2008), and preliminary Milne-Eddington inversion of the data has already been carried out. The SP is therefore valuable as a seeing-free reference and standard of comparison for FIRS and for the techniques I have applied in this dissertation. I have made use of the SP data products which are publicly available online through the Community Spectropolarimetric Analysis Center (CSAC). By serendipity the SOT/SP observation of the main sunspot in NOAA 11024 was taken almost simultaneously with the FIRS 6302/15650 Å observation on 2009-07-07, only about 4 minutes apart, and this dataset is heavily used for direct comparison of the data and analysis techniques.

SOLIS

The Synoptic Optical Long-term Investigations of the Sun (SOLIS) is a synoptic suite of instruments which was installed and began taking observations in 2003 at Kitt Peak, Arizona, that is intended to provide a long baseline for full disk solar observations. The SOLIS instrument suite consists of a Vector Spectromagnetograph (VSM), the Full Disk Patrol (FDP) imager, and the Integrated Sunlight Spectrometer (ISS). The VSM is another full-Stokes spectropolarimeter capable of operating at 6302 Å. Although it achieves only 1" sampling and slightly decreased spectral resolution, it is capable of covering the full disk of the Sun in less than 20 minutes. The SOLIS team is endeavoring to provide Milne-Eddington inverted data in addition to standard magnetograms. By coincidence SOLIS also

observed NOAA 11024 on 2009-07-07. I compare the method of analysis for this observation to the method applied to FIRS for further enlightenment.

MDI and HMI

Data from the Michelson Doppler Imager (MDI) on the Solar and Heliospheric Observatory (SOHO) and the Helioseismic and Magnetic Imager (HMI) on the Solar Dynamics Observatory (SDO) are used extensively to infer the evolutionary histories and configurations of the sunspots observed by FIRS. I have also made extensive use of the Solar Monitor currently operated by the Solar Physics Group at Trinity College, Dublin and the Solar Data Analysis Center.

Chapter 4

Data Reduction and Techniques

Abstract

In this section I describe the reduction of data for the FIRS, including critical corrections for the response of the infrared detector and instrumental polarization. The reduction of data from the HSG, Hinode, and SOLIS is also discussed. I directly compare the data products from FIRS and Hinode for verification of the FIRS data quality.

4.1 Data Reduction for FIRS Observations

4.1.1 Overview

The FIRS data reduction routines were written in IDL and are publicly available on the web at *http://kopiko.ifa.hawaii.edu/firs*. The procedures for data reduction are fully described in the manual (Jaeggli 2010). The reduction code transforms a set of raw data frames like those shown in Figures 3.6, 3.7, and 3.8, into a fully calibrated datacube. A complete set of observations consists of darks, flat-fields, science observations, and optional calibration images taken using the DST Port 4 Advanced Stokes Polarimeter calibration optics (Elmore et al. 1992). The reduction scripts are similar for all wavelengths however, due to the peculiarities of each wavelength, separate scripts are maintained for each, and we will note the differences in the summary of the data reduction that follows.

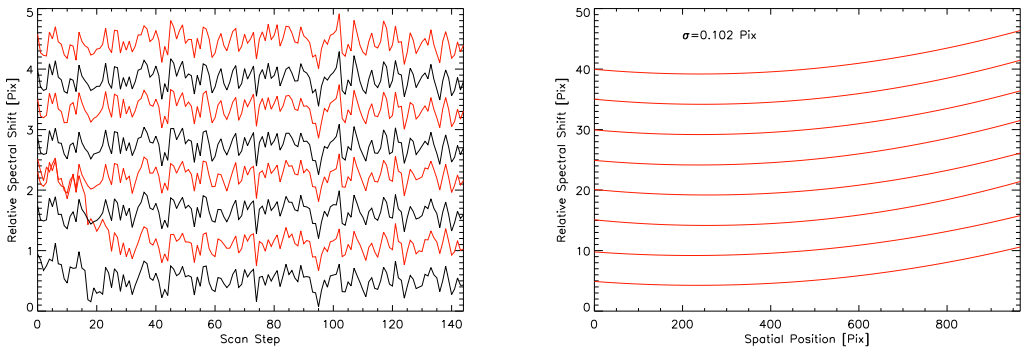

Figure 4.1 An example of spatial and spectral shifts measured from a telluric line in the visible data.

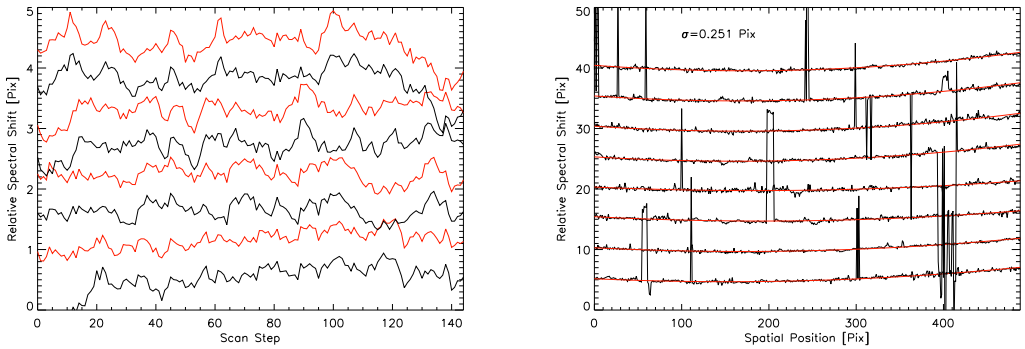

Figure 4.2 An example of spatial and spectral shifts measured from a Fe I line in the infrared data.

The data reduction begins by making master dark and flat fields by taking the average of 16 full-stokes dark and flat frames. The user interactively selects the position of the top and bottom beams in the image and the positions of the photospheric and telluric spectral lines, which are then fit with a Gaussian and removed from the flat-field (a Gaussian may not actually be suitable for Si I 10827 Å). The line centers fitted from the flat-field are used to determine the spectrograph curvature, and the line positions are used as a starting point to determine the spectrograph drift during the science observations (due to vibrations and thermal fluctuations in the instrument). Figures 4.1 and 4.2 show examples of the drift and spectral curvature measured from a scan. The telluric lines in the 6302 and 10830 Å data

allow the relative calibration of wavelength to a high accuracy, about a tenth of a pixel, but at 15650 Å the wavelength can only be calibrated to about 1/4 pixel from the quiet sun iron lines. It is interesting to note that the spectral drift from image to image during a scan is about 1/2 pixel for both the visible and infrared, or about 10 μm.

The science frames are corrected by the master dark and normalized by one of two methods. If we wish to determine the polarization crosstalk from Stokes I, the relative intensity between the observed images must be maintained, so the images are divided by the median of all four images together. If it is not necessary to know the crosstalk from Stokes I, each of the modulated frames may be divided by itself. Following this the science frames can be demodulated to obtain the Stokes I, Q, U, and V states. From the efficiency balanced-tuning scheme defined in Equation 3.1, the true Stokes vector can be found by:

$$\begin{aligned} I &= 1/2 \quad (\ \ I_1 + I_2 + I_3 + I_4) \\ Q &= \sqrt{3}/2\ (-I_1 - I_2 + I_3 + I_4) \\ U &= \sqrt{3}/2\ (-I_1 + I_2 - I_3 + I_4) \\ V &= \sqrt{3}/2\ (\ \ I_1 - I_2 - I_3 + I_4) \end{aligned} \tag{4.1}$$

where I_1 to I_4 are the observed spectral intensity in which the Stokes vector is encoded by the polarimeters. Note that there is a factor of 1/2 in front because the polarizer reduces the beam intensity by 1/2.

The flat field is dark subtracted, normalized in the same way as the science frames, demodulated, and then applied to the science frames. We find that the second method of normalization defined in the previous paragraph is preferable and greatly reduces the effect of fringes in the polarized states for the infrared data. The fringes cannot be removed entirely due to subtle differences in the illumination pattern on the detector between the flat field and the science observations, but the effect of fringes can be greatly reduced by taking flat fields in a nearby section of the quiet-sun immediately preceding or following science observations.

The science frames are sub-sectioned according to the selection of the beam areas from the flat field. Sub-pixel shifts are applied to the data using an FFT resampling method on a slit-by-slit and beam-by-beam basis according to the curvature and spectrograph drifts determined earlier. The orthogonally polarized beams are then combined to remove spurious seeing effects. The spatial sampling and distortion has been determined from scans of the Port 4 line grid. The Wollaston prisms do not introduce large distortions, and it is not necessary to transform the beams before combination for most scientific requirements. The processed spectra are spatially binned down from their original resolution to match the 0.29 arcsec/pixel slit width. The Stokes spectra from the combined beam are then normalized to the average local quiet-sun continuum in each slit to remove step-to-step intensity variations.

The instrumental polarization correction is applied to the combined beam using the Kuhn technique for calibration and removal of instrumental polarization, this discussed in detail in Section 4.1.3. The telluric lines in the science data are fitted and removed if desired and the resulting spectra are stacked into a data cube. Persistent streaking from the slit is removed from the Stokes spectrum and quick look maps are generated. The resulting output is saved, optionally as a FITS data cube or IDL structure which are suitable for further analysis. Examples of the spectra and scan maps which are the final results of the data reduction for the observation of the leading sunspot in NOAA 11035 taken on 2009-12-17 are shown in Figures 4.3 to 4.8

4.1.2 Response of the Infrared Detector

Linearity in the response of the detectors is of critical importance for the accurate determination of the contrast of solar features relative to the average quiet-sun intensity from which we determine the continuum intensity. During the course of this work it became obvious that the response of the Virgo 1k infrared detector was highly non-linear, and in addition the array has some unusual properties which further complicate the characterization of its response.

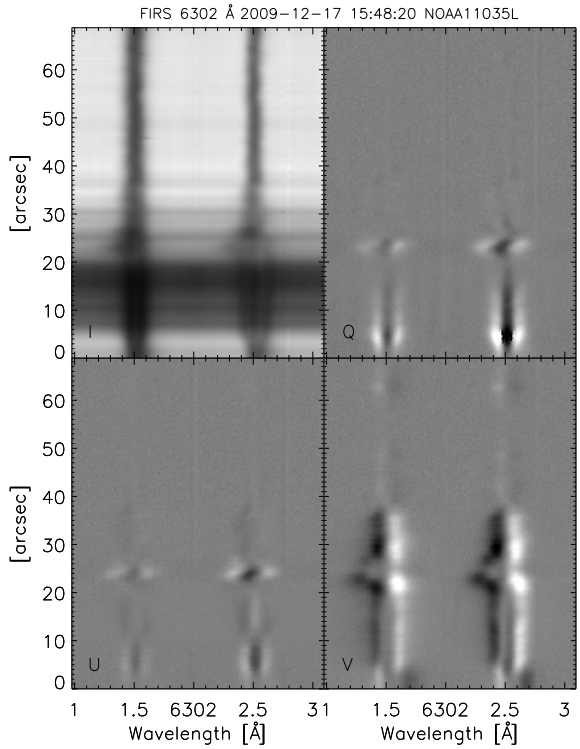

Figure 4.3 A full Stokes spectrum from the 6302 Å channel of FIRS.

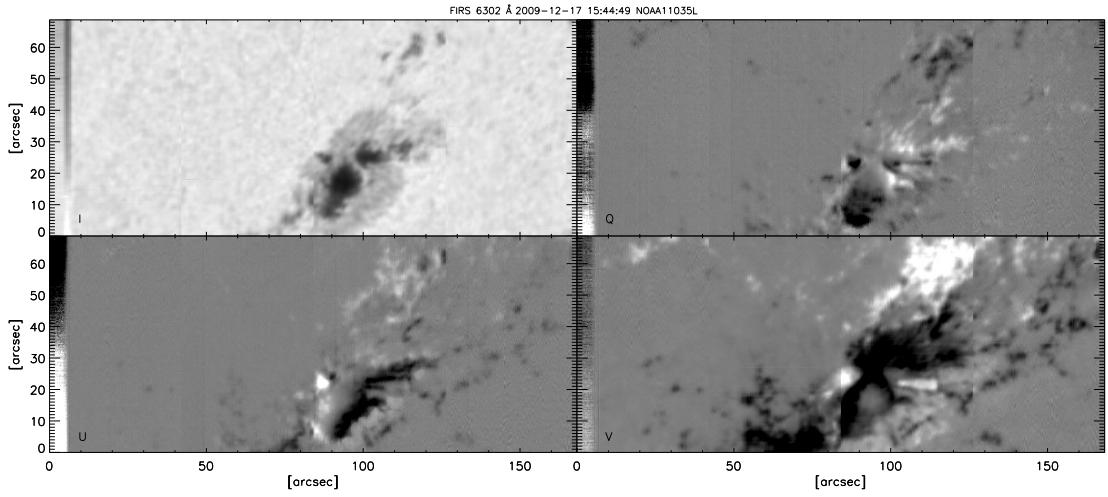

Figure 4.4 A quick-look map made from the Fe I Stokes profiles from the 6302 Å channel of FIRS.

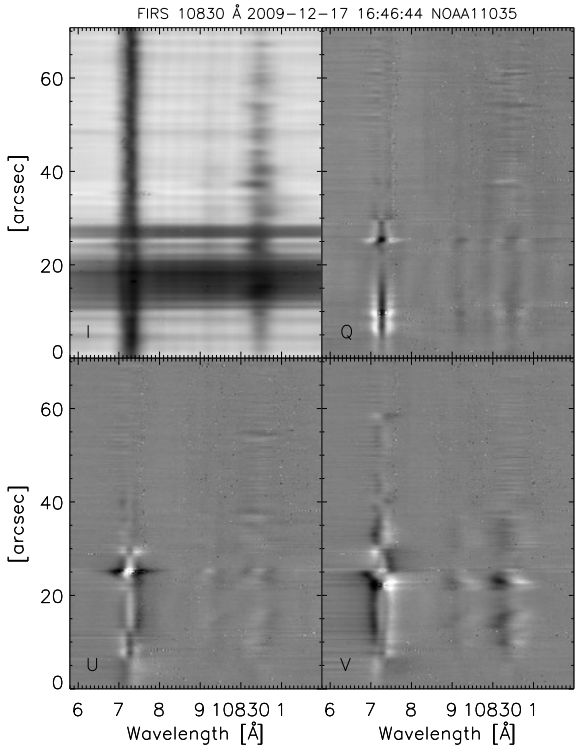

Figure 4.5 A full Stokes spectrum from the 18030 Å channel of FIRS.

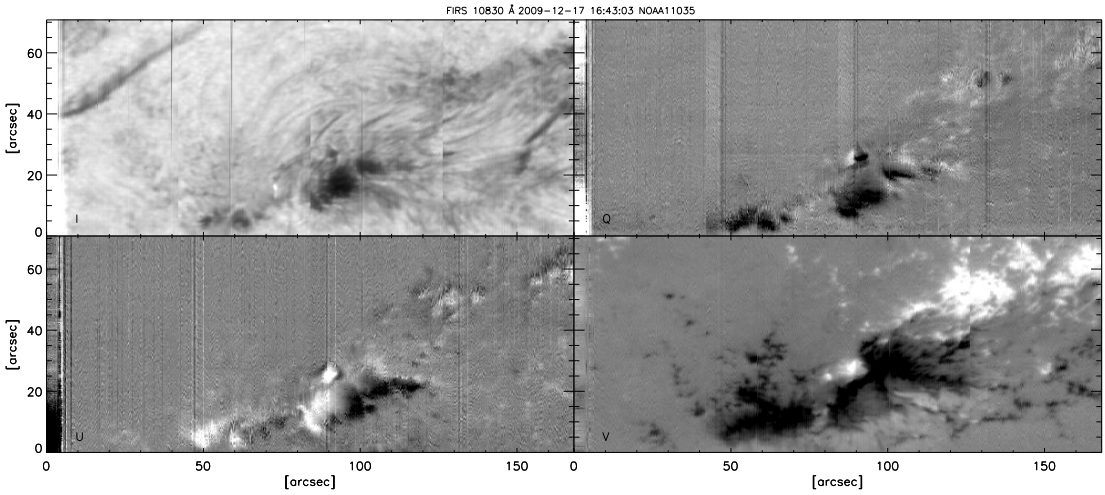

Figure 4.6 A quick-look map made from the Si I Stokes profiles from the 10830 Å channel of FIRS.

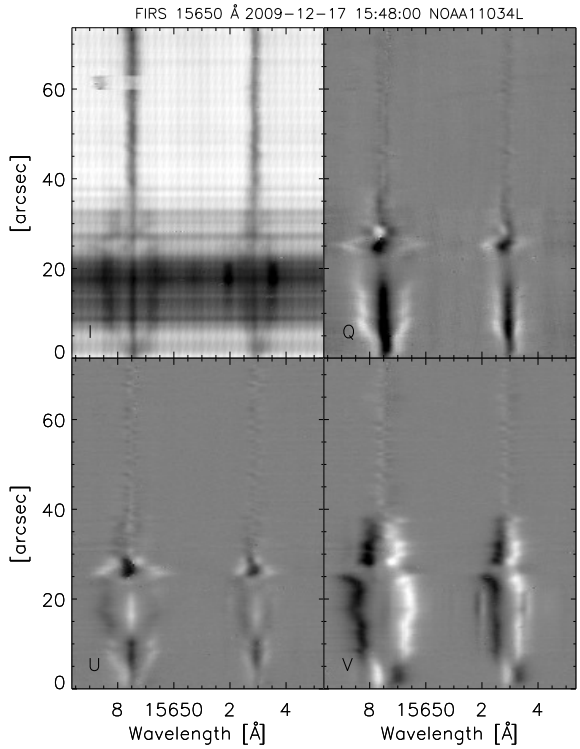

Figure 4.7 A full Stokes spectrum from the 15650 Å channel of FIRS.

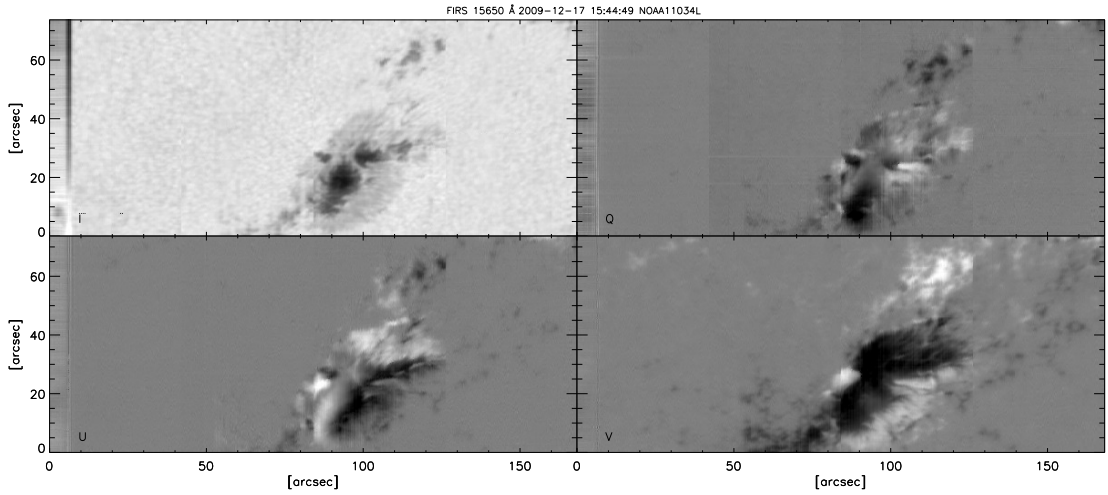

Figure 4.8 A quick-look map made from the Fe I Stokes profiles from the 15650 Å channel of FIRS.

For a typical photoconductive device, the detection of photons is achieved when a photon of sufficient energy (E > band-gap energy) excites an electron from the valence band to the conduction band of the material. Once in the conduction band the electrons can be transferred, amplified, and counted by the array electronics. Two effects contribute to the non-linearity of the response of the array material itself. When the conduction band is empty more photons are needed to lift the first few electrons into the conduction band (activation energy). When the conduction band is nearly full, electrons raised from the valence band will no longer stay in the conduction band (the saturation limit of the detector). The amplifiers can also contribute to the non-linearity of photon detection.

For traditional methods of measuring the linearity of a detector, images are taken at different signal levels to fully sample the response of the pixels. The noise, characterized by the standard deviation (σ), at each signal level is determined from the resulting images is calculated on a pixel-by-pixel basis or for a small region of well-behaved pixels in the image. The noise should of course be related to the number of photons that are counted (by Poisson statistics) when the read noise contributed by the electronics is low. The relation between the noise (σ), detected signal (analog-to-digital units or N_{ADU}), and the gain (or ADU detected per photon) is: $gain * \sigma^2_{ADU} = N_{ADU}$. This method of characterization typically works well for CCDs and infrared arrays.

A thorough set of calibration images was taken on September 8, 2010. Images were taken with exposure times from 125 to 2000 msec in increments of 125 msec at a constant intensity level. The Virgo 1k shows an unusual characteristic. The noise initially increases with the measured intensity, but then decreases again at high signal levels. This result is so strange and unexpected that three different methods of measuring the noise were employed to reveal any possible systematic effects:

1. The signal and noise are determined directly from the dark-subtracted images.
2. The signal and noise are determined from dark-subtracted and flat-fielded images.
3. The signal is determined from the dark-subtracted images, the noise is determined from a pair of subtracted images at the same intensity level and scaled by $\sqrt{2}$.

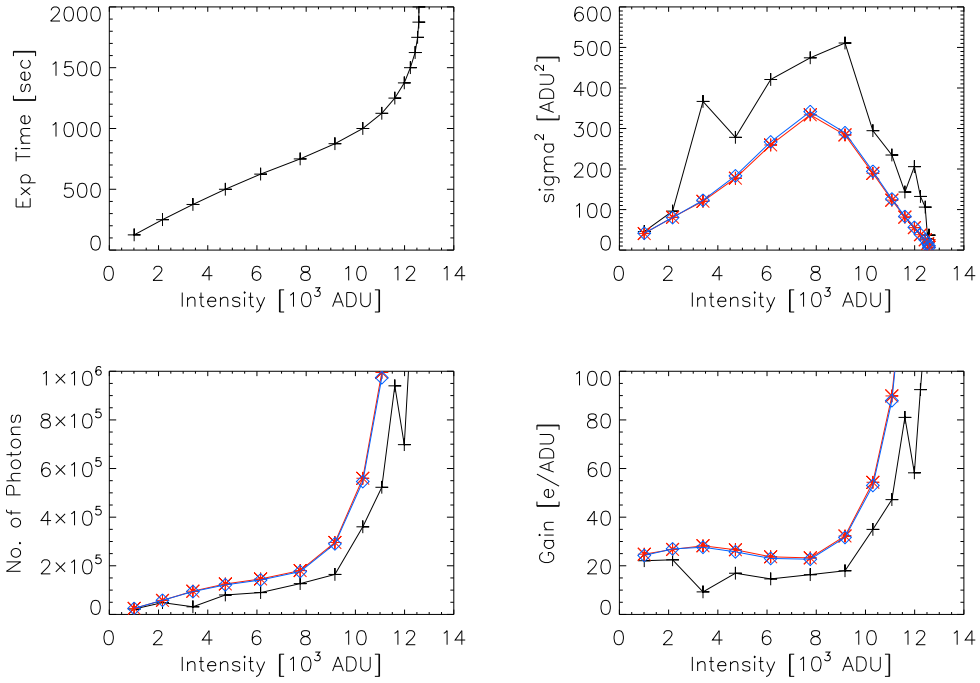

Figure 4.9 Calibration curves for the Virgo 1K array obtained using three methods to measure the image noise. Note that this data has been dark-subtracted.

The square of the noise plotted against measured signal level for each method is shown in the top right panel of Figure 4.9. Methods 2 (red) and 3 (blue) are in very good agreement. Method 1 (black) is sensitive to the macroscopic variation of intensity and bad pixels within the subsection of the array, which results in a higher noise measurement and is consistent with the other two methods. The lower two panels in Figure 4.9 show the resulting detector response and gain that can be inferred from the noise measurement. If the detector response (photons vs. ADU) were real it should be in good agreement with the plot of exposure time against intensity shown in the top left panel. Because these two are so different we can only assume that there is something funny about the noise characteristics of the detector, which may be intrinsic, or due to the readout electronics. Therefore the detector response (and other important characteristics) cannot be inferred from the noise measured from the image.

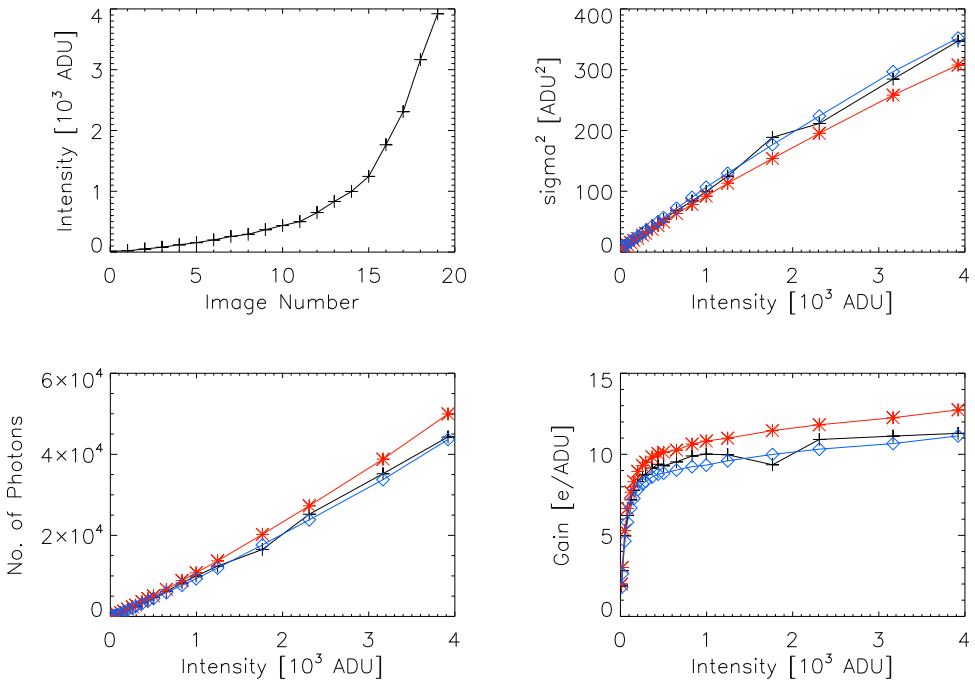

Figure 4.10 Calibration curves for the Kodak 2K CCD obtained using three methods to measure the image noise.

To determine the detector response we assume a constant flux of photons in time and use our knowledge of the camera exposure time. In the past, the variability of shutter speeds of infrared cameras made this measurement too imprecise, however the Virgo 1k has no shutter, and the array is clocked at a constant rate with signals supplied by a computer timing card. Images were taken with exposure times from 125 to 2000 msec in increments of 125 msec at four constant intensity levels (set using the incandescent overhead lights in the telescope on a dimmer switch) to fully sample the detector response. The four data sets were combined by scaling the exposure times using a factor determined from quadratic interpolation of the intensity for each of the 4 levels. The resulting data points were fit using a chi-squared minimized spline to produce a lookup table for every possible value of the detector output. The data and the resulting fit are shown in Figure 4.11. From this curve we have generated a look-up table for ADU values from 0 to 16,000. The correction is

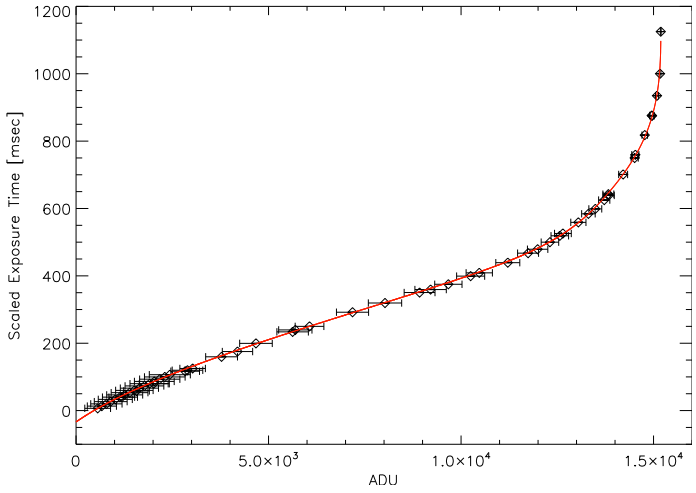

Figure 4.11 Response of the Virgo 1k detector as shown in exposure time vs. ADU. The error bars show the measured error from the intensity frame. Note that this data has not been dark-subtracted.

applied to every raw image used during the reduction of the infrared data before any other processing is preformed.

As a test of this correction I took observations of a sunspot with an adjustable iris placed at a pupil position just upstream of the field-steering mirror in Figure 3.4. The observations were taken on 2010-09-01 with the slits at a single spatial position across the umbra and penumbra of the sunspot in NOAA 11101. The observations with 5 different pupil apertures were taken in rapid succession so changes in the spatial structure in the sunspot and the effects of seeing were small between images. In Figure 4.12 I show the pixel-by-pixel scatterplot of a section of the spectra containing the sunspot, normalized to the local quiet-Sun continuum in each spectrum. The images are compared to the one with the iris opened to the largest aperture. If the detector response were linear, the scatter of points would fall mostly in a line, however we see that even small changes in the input intensity reveal the non-linear effect of the array. If we apply the correction determined above, the relations in Figure 4.13 result. The correction does a pretty good job, but there is still some residual non-linear behavior evident between the highest and lowest intensity levels.

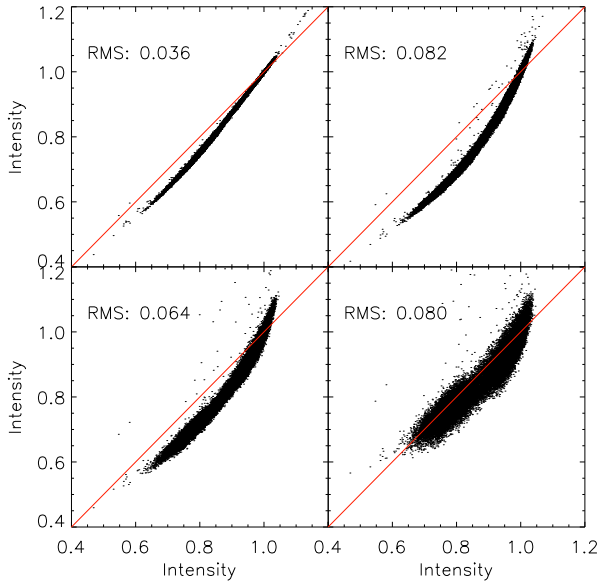

Figure 4.12 The distribution of continuum contrast in spectra which have not been corrected for the IR detector response in a sunspot for different light levels controlled with an adjustable iris, all referenced to the same spectrum with the iris fully open.

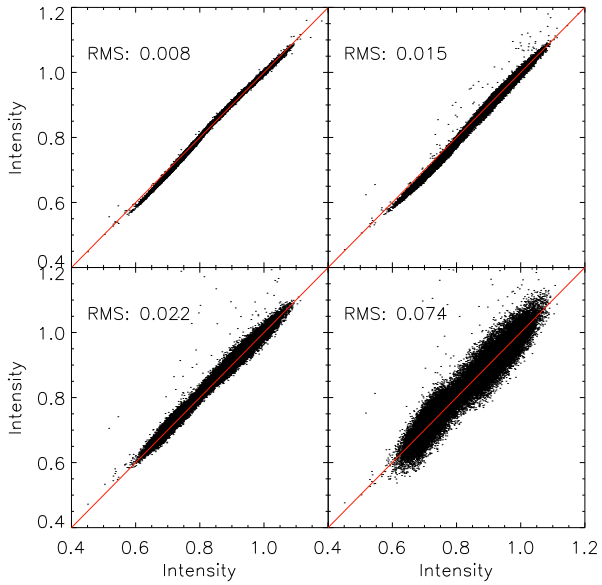

Figure 4.13 The distribution of continuum contrast in spectra which have been corrected for the detector response in a sunspot for different light levels controlled with an adjustable iris, all referenced to the same spectrum with the iris fully open.

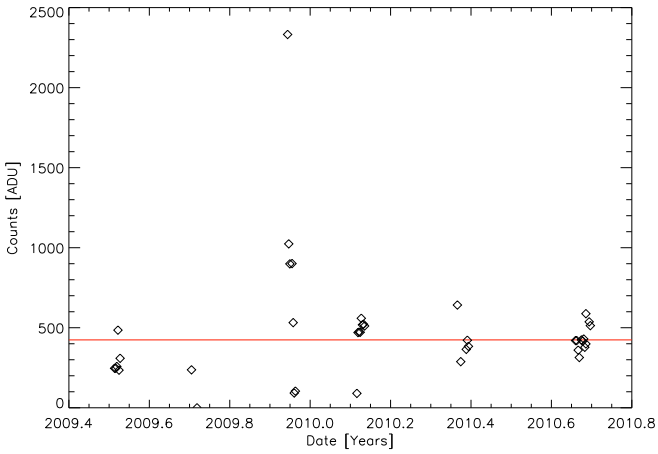

Figure 4.14 Median level of signal in the dark frames with time. The red line shows the level of the dark during the calibration of the detector response.

There is an additional source of error in the response correction to be aware of. The data taken prior to Feb 14, 2010 suffers from a wandering bias level from manual adjustment of the computerized settings for the detector. The extremely low humidity at the observatory site, aggravated by a continuously running fan, caused static charge to build up on the external camera electronics (A-to-D and amplifiers) during the night when they were not in use. After the camera system was again powered on for the day this static charge, which manifested as an excess of noisy pixels below a certain intensity threshold, took several hours to dissipate. Before the problem was properly identified (and the fan turned off at night) it was thought that exercising the bias level might fix this problem. The bias level in the older data must be adjusted to match the current bias level before applying the detector response. I have attempted to correct for this effect by comparing the average dark level to a reference dark level determined from the data taken after Feb 2010 (see Figure 4.14) and adjusting the pixel values before applying the response correction. Although other effects can alter the dark level (eg. parasitic light, array temperature) this correction is probably good to 200 counts.

4.1.3 Instrumental Polarization Calibration and Correction

For every spectropolarimeter the accurate correction of polarization cross-talk due to the instrument is critical for the determination of the vector magnetic field. Instrumental polarization cross-talk arises from the presence of non-ideal optical elements in the combined telescope and instrument system and is aggravated by geometry. The vacuum-coelostat design of the DST is highly non-ideal for the study of solar polarization. The thin films which form on the surfaces of reflective optics due to the accumulation of dust, vacuum oil, and the oxidation of the aluminum coatings can act as both retarders and as partial linear polarizers when they are highly inclined to the beam. The second effect is particularly important because the first two reflections in the DST are often about 45°. In addition the entrance and exit vacuum windows have non-negligible retardation due to the strain exerted on them. The polarization cross-talk introduced by the telescope is dependent on the observing geometry and changes throughout the course of the year and day. The field-scanning mirror and 95/5 beamsplitter for FIRS also have reflections at angles greater than 45°. Although the polarization from these mirrors is not considered to change significantly with time or scanning position, they do appear to introduce a considerable amount of polarization cross-talk in addition to the telescope. As a result, FIRS observations show a large amount of crosstalk between polarization states. The left panel in Figure 4.15 shows an example of the observed polarization in a reduced observation. The telescope primarily exchanges the polarization between Stokes U and V.

Attempts have been made to measure the polarization properties of many solar telescopes and generate a pointing-dependent Müller matrix, [for examples see Bernasconi (1997), Elmore et al. (2010), and Socas-Navarro et al. (2011)]. However, any pointing-dependent polarization model must be continually updated to keep up with the changing retardance of the mirror surfaces with time. The method of crosstalk correction in Kuhn et al. (1994) and a similar method in Collados (2003) have the advantage that they require no prior knowledge of the instrumental polarization cross-talk, they simply use the polarization of the regular Zeeman-split line profiles themselves, assuming symmetry of the linearly

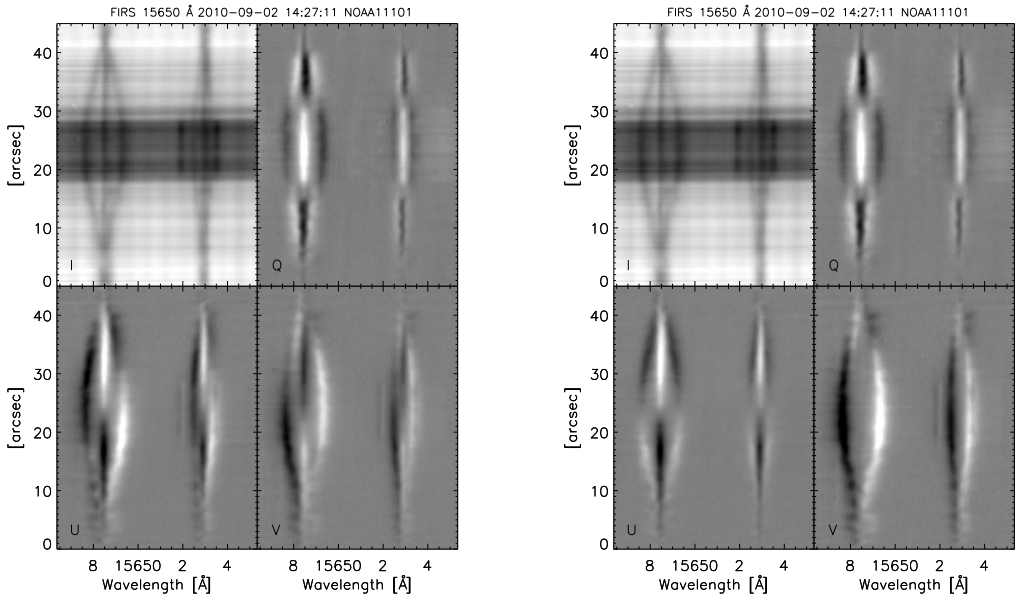

Figure 4.15 An example of the uncorrected Stokes vector in a sunspot showing large polarization cross-talk between U and V (left), and the Stokes vector corrected using the Kuhn method.

polarized Stokes Q and U components and anti-symmetry of the circularly polarized Stokes V component to find the coefficients that best meet these requirements.

Method for Polarization Correction

The Kuhn method for polarization correction is best suited to the large amount of sunspot data that was taken for this dissertation, and in this section I outline the arguments presented in Kuhn et al. (1994). Consider a set of observations of the Stokes vector (I', Q', U', V') in a sunspot. The observed vector suffers from polarization cross-talk due to the instrumental system. In Kuhn et al. (1994) the cross-talk from Stokes I into polarization states was considered to be small and set to zero. This may not always be the case, but the cross-talk from I is easily measured from the continuum of the polarized components and removed:

$$Q' \;=\; Q'' - eI$$

$$\begin{aligned} U' &= U'' - fI \\ V' &= V'' - gI \end{aligned} \tag{4.2}$$

or in matrix form:

$$M^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -e & 1 & 0 & 0 \\ -f & 0 & 1 & 0 \\ -g & 0 & 0 & 1 \end{pmatrix} \tag{4.3}$$

where the observed Stokes vector (S'') is related to the I cross-talk removed Stokes vector by (S') by $S' = M^{-1}S''$.

Cross-talk between linear and circular polarization states arises from non-ideal thin films on the surface of reflective optics and birefringence of transmissive optics among other things. Cross-talk between linear polarization states is the equivalent of a rotation and is not considered here. To first order the relation between the true and observed Stokes vector can be written as:

$$\begin{aligned} V &= V' - aQ' - bU' \\ Q &= Q' - cV \\ U &= U' - dV \end{aligned} \tag{4.4}$$

Considering cross-talk between Q, U, and V, the inverse Müller matrix of the optical system is given by:

$$M^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ e & 1+ac & bc & -c \\ f & ad & 1+bd & -d \\ g & -a & -b & 1 \end{pmatrix} \tag{4.5}$$

where the observed Stokes vector (S') is related to the true Stokes vector (S) by $S = M^{-1}S'$.

To determine the coefficients a and b, first consider taking the integral of the observed profiles over the entire wavelength range:

$$\begin{aligned} \int Q\ d\lambda &= \int (Q' - cV)\ d\lambda \\ \int U\ d\lambda &= \int (U' - dV)\ d\lambda \\ \int V\ d\lambda &= \int (V' - aQ' + bU')\ d\lambda \end{aligned} \tag{4.6}$$

For most cases, the true Stokes V line profile should be anti-symmetric, and the integral over the line will be zero, therefore we obtain the following relations:

$$\begin{aligned} \int Q\ d\lambda &= \int Q'\ d\lambda \\ \int U\ d\lambda &= \int U'\ d\lambda \\ \int V'\ d\lambda &= a \int Q'\ d\lambda\ +\ b \int U'\ d\lambda \end{aligned} \tag{4.7}$$

A linear regression supplied with the wavelength-integrated observed Stokes profiles will return the coefficients a and b, allowing us to correct the observed V profiles.

In order to get the coefficients c and d, consider the second order integrals (or correlations) of the Stokes profiles with wavelength:

$$\begin{aligned} \int QV\ d\lambda &= \int (Q'V - cV^2)\ d\lambda \\ \int UV\ d\lambda &= \int (U'V - dV^2)\ d\lambda \end{aligned} \tag{4.8}$$

The true Q and V, or U and V profiles should be uncorrelated, and the second order integrals will be zero, leaving us with the coefficients:

$$\begin{aligned} c &= \frac{\int Q'V\ d\lambda}{\int V^2\ d\lambda} \\ d &= \frac{\int U'V\ d\lambda}{\int V^2\ d\lambda} \end{aligned} \tag{4.9}$$

Now it is possible to correct the polarization cross-talk in the Stokes vector to first order.

Limitations

The right panel of Figure 4.15 shows the corrected Stokes vector for our example observation. Despite the seemingly good correction that the technique provides, this method is often called into question. The emergent line profiles, especially in high resolution data, are not always symmetric if, for example, multiple velocity components exist along a single line of sight. A detailed comparison of the FIRS data with the pristine polarization measurements made with Hinode is discussed in a following section. To first order this technique retrieves the Müller matrix of the entire optical train at the time of observation, and though the results are not discussed here, work is underway to reconcile the polarization calibration determined from FIRS observations with a pointing-dependent polarization model of the DST combining the Müller matrix determined here for the full telescope and instrument system with polarization cross-talk determined using another calibration method which is discussed in Appendix B.

4.2 Data Reduction for HSG Observations

Data reduction for the four observations from the experimental HSG setup was carried out by Haosheng Lin in much the same way as it was done in Lin (1995) and Lin et al. (1998), and was also similar to the FIRS data reduction method (except for a single slit). The observations taken in 2001 with the NICMOS 3 detector have an unrealistic umbral contrast and it is difficult to say whether this might be a result of the data reduction or perhaps the non-linear response of the detector. The intensity levels in the 2001 observation were adjusted by applying a linear offset to the spectra and renormalizing the datacube to the level of the quiet sun continuum.

4.3 Data Reduction for Other Observations

Minimal processing was necessary for the Hinode SOT/SP and SOLIS VSM data. Level 1 (calibrated data) fits files from the SP were obtained from the CSAC archive (http://www.csac.hao.ucar.edu/csac). Level 2 (inverted data) MERLIN Hinode inversion client. Relevant observing parameters in addition to the Stokes spectra, were loaded into the IDL data structures used for FIRS so the same analysis could be performed.

Level 1 data from SOLIS was obtained directly from the SOLIS team for the purposes of comparison and Level 2 data was obtained freely online from the SOLIS data archive.

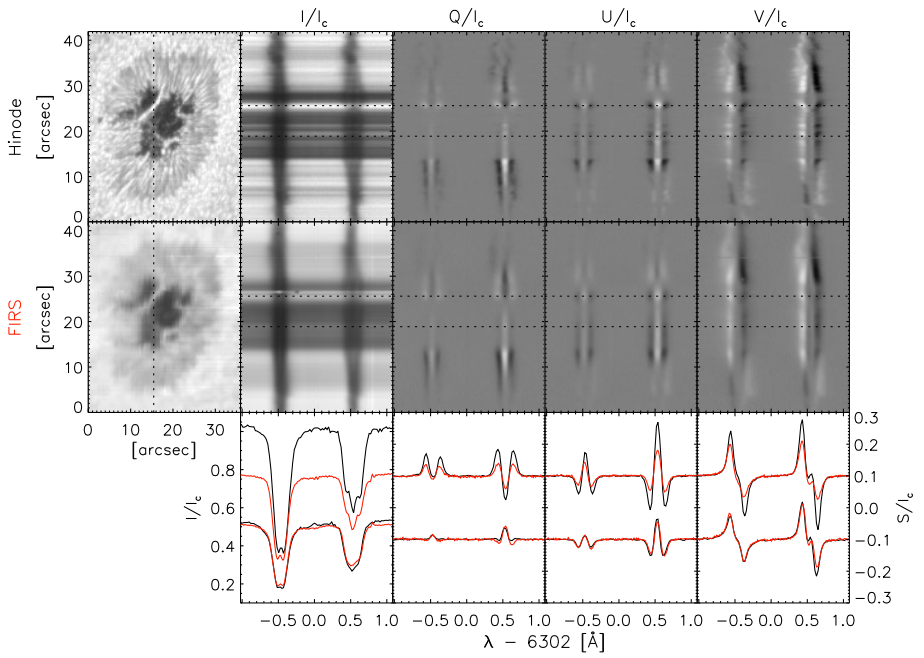

Figure 4.16 Direct comparison of the continuum maps (top left panels), 2D full-Stokes spectra (top right 8 panels) from the position indicated by the dotted line in the map, and 1D full Stokes line spectra (bottom 4 panels) from the position of the dotted line in the 2D spectra from FIRS (red) and Hinode SOT/SP (black) observations of the leading polarity sunspot in NOAA 11024 taken 4 minutes apart on 2009-07-07.

4.4 Direct Comparison of Spectropolarimetric Observations from FIRS and Hinode

Based on the assumptions made by the Kuhn method of polarization correction, one would think that the it might not work well for physical situations which produce non- symmetric and anti-symmetric Stokes profiles. Therefore, we have verified the appropriateness of the Kuhn method, and the data quality in general, by comparing corrected FIRS observations to reduced spectra from the Hinode Solar Optical Telescope Spectropolarimeter (SOT/SP) which were closely coincident in time (about 4 minutes apart). The polarization cross-talk in the SOT/SP has been very carefully measured and corrected as described by Ichimoto et al. (2008). The maps and example spectra from each observation are shown in Figure 4.16, and in Figure 4.17 we show a point-by-point comparison of all quiet sun and sunspot spatial and spectral positions from the aligned datacubes of Stokes spectra from FIRS and the SOT/SP. Most predominantly noticeable is the effect of seeing on the observations. Small scale features with extreme intensity, velocity, or polarization signatures are smoothed over in the FIRS spectrum, such as the upper line spectrum in Figure 4.16 which is from a position in the light bridge. However, the overall behavior of the spectra is similar, and non-symmetric profiles seen in the Hinode data are faithfully reproduced by FIRS. In the scatter plots the overall adherence to the 1:1 line (esp. at low signal levels in Stokes I) is encouraging. The flattening of Stokes Q, U, and V in Figure 4.17 indicates that the highest polarization signals seen by Hinode are not detected by FIRS, these signals are spatially discrete and become smoothed over by seeing. The high Q and U values in the Hinode data where corresponding FIRS values are zero, are the result of a bad scan-step on the penumbral/quiet sun boundary in the SOT/SP scan.

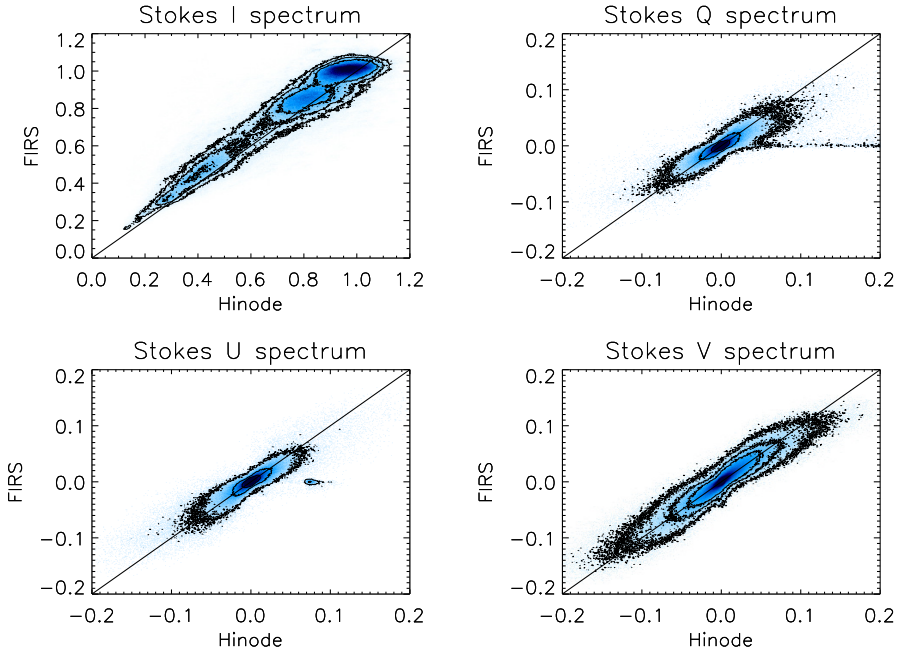

Figure 4.17 Point -by-point scatter plot comparison of the full Stokes spectra from Hinode SOT/SP and FIRS data cubes, using coincident sunspot observations taken on 2009-07-07. 50, 75, 95% contours are displayed for I and 90, 95, 99% contours for Q, U, and V. The solid line indicates the 1:1 correspondence.

Chapter 5

Inversion of Observations

Abstract

This chapter describes the techniques that were used to retrieve the magnetic field geometries and line parameters for the Fe I and OH lines using a Milne-Eddington inversion technique. There is an overabundance of inversion codes available for this purpose, but to better understand the mechanics of inversion techniques I have written my own code. Consequently it is necessary to compare this code with other inversion techniques to verify the results.

5.1 Introduction

The magnetic field strength, intensity, and molecular line parameters must be determined in some way from the spectropolarimetric observations. However, "nobody can measure physical quantities of the solar atmosphere" (del Toro Iniesta & Ruiz Cobo 1996), we can only hope to infer the average properties over the heights in the solar atmosphere at which the line diagnostics form through radiative transfer. The retrieval of unknown parameters of a given phenomenon by fitting a model to observations is known as an"inverse problem." In solar physics, the techniques used to retrieve atmospheric parameters and magnetic field geometries from spectropolarimetric observations are simply called "inversions." There are three basic approaches for inversion codes. The first approach simplifies the equations of

polarized radiative transfer by assuming the line forming region exists at a single height, and is known as a Milne-Eddington inversion. The second method is more computationally intensive and attempts to solve the equations of radiative transfer to obtain the vertical stratification of parameters given a basic atmospheric model. The third method matches a database of line profiles generated using known model parameters to data using principal component analysis (PCA) or trained artificial neural networks (ANN). This kind of method is appealing for large datasets due to its inherent speed, however it is less able to address unique situations or probe the detailed line physics. Wood & Fox (1995) present a very general approach to inverse problems in astronomical spectropolarimetry. More recent overviews of inversion techniques can be found in del Toro Iniesta et al. (1995); Socas-Navarro (2001); del Toro Iniesta (2003); Bellot Rubio (2006); Lites et al. (2007); Ruiz Cobo (2007).

Inversion techniques began with the analytical solutions derived by Unno (1956) from the equations for radiative transfer to determine magnetic parameters directly from the Stokes profiles, assuming that the line and continuum form in narrow layers over which the atmospheric properties change little, known as the Milne-Eddington approximation. Computationally more intensive techniques began with Auer et al. (1977) and their simplified Milne-Eddington approach where they assumed a constant magnetic field, ignored quantum mechanical and magneto-optical effects, with a 7 parameter fit to the polarized line profile. In the work that followed, the approach of Auer et al. (1977) was revised to retain a greater level of complexity which was made possible by high resolution spectropolarimetry (Landolfi & Landi degl'Innocenti 1982; Jefferies et al. 1989), and resulted in the standard Milne-Eddington inversion code which is still used today (Skumanich & Lites 1987).

While the methods of fitting Stokes profiles for a Milne-Eddington model have improved, the improvements on the physics themselves have not been made. Because of the assumption of a flat atmosphere, the ME codes are inherently limited and cannot describe physical situations that result in asymmetric polarized line profiles. To move beyond the Milne-Eddington approximation it is necessary to solve the equations of radiative transfer

themselves (Landi degl'Innocenti & Landi degl'Innocenti 1985), using knowledge of the atmospheric stratification. This allows the parameters for complicated line profiles with multiple components along the line of sight to be successfully inverted. There are many techniques (e.g. SIR Ruiz Cobo & del Toro Iniesta (1992)) that allow for the simultaneous inversion of multiple spectral regimes at once (Cabrera Solana et al. 2006; Martínez González et al. 2008). These are very interesting for the future of FIRS data, but have not been extensively applied in this work.

5.2 Two Component Magneto-Optical Inversion Code (2CMO)

There exists a variety of options when choosing a Milne-Eddington inversion technique. Inversion techniques are sophisticated pieces of software which are difficult to digest without prior knowledge of how they should work, and there is a tendency to treat their inner workings as black magic. I have written my own code to better understand the benefits and pitfalls of a simple Milne-Eddington inversion technique. The code has been implemented in IDL following the formalism of Jefferies et al. (1989).

5.2.1 Description

The 2CMO inversion fits the full-Stokes spectrum using Voigt and Faraday-Voigt profiles for the Zeeman-split atomic lines which are calculated to high accuracy for each of the three Zeeman-split components using the numerical approximation provided in Matta & Reichel (1971). The magnetic component is described by ten parameters: the magnetic filling fraction (f), the source function (B_0), source function gradient (B_1), velocity (λ), Doppler width ($\Delta\lambda$), damping parameter (a), ratio of line to continuum absorption (η_0), magnetic field strength (B), inclination (γ), and azimuth (χ). The non-magnetic component is allowed an independent velocity (λ_s), but it is assumed that the non-magnetic fill fraction is $1 - f$. For the 15650 Å observations the magnetic component also includes the Gaussian

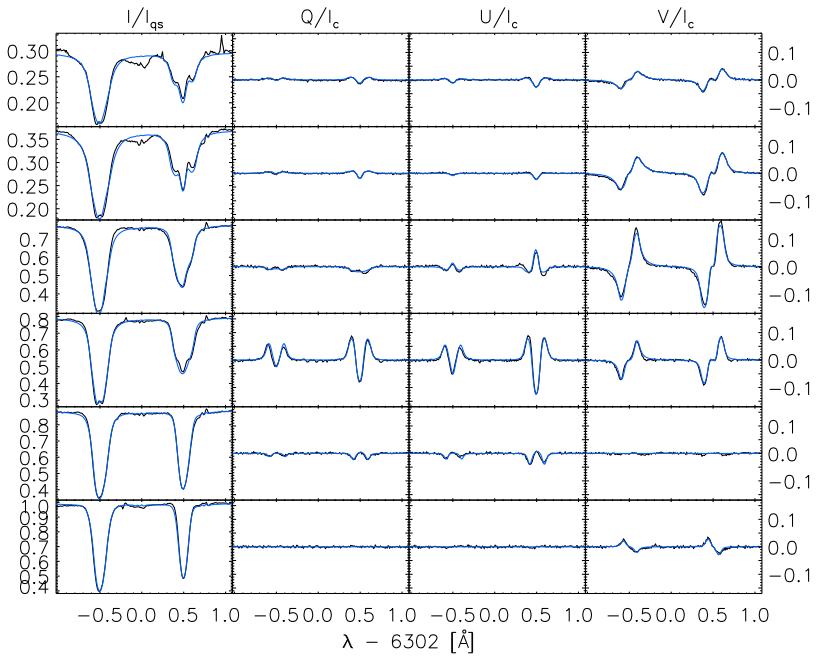

Figure 5.1 An example of inverted spectra from a 6302 Å observation of the sunspot in NOAA 11101. The observed spectra are shown in black and fits are shown in blue. Stokes I, Q, U, and V are spectra are shown in the panels from left to right, and the different positions in the sunspot are shown in the panels from top to bottom.

profiles used to fit the molecular lines (and molecular Zeeman effect). The width and depth of the OH lines are fit in addition to the other parameters. In addition to these free parameters it is necessary to know the rest wavelength, $log(gf)$, and g_{eff} for all of the fitted atomic lines. We have adopted the values from VALD (Vienna Atomic Line Database) for the 6302 Å lines and Borrero et al. (2003) for the 15650 Å Fe I lines and use the maximum g_{eff} (of 0.15) from Berdyugina et al. (2003) for the OH lines.

The 2CMO inversion module operates in a pre-existing IDL software framework for conducting inversions by H. Lin. Only slight modifications to the inversion code were necessary to implement the 2CMO module. During the inversion of a data cube, first the stray light profile is generated by taking the median of intensity profiles from identified quiet-Sun regions. For every spatial position a guess is generated for each parameter from a direct measurement of the Stokes profiles. The IDL *curvefit* function is supplied with

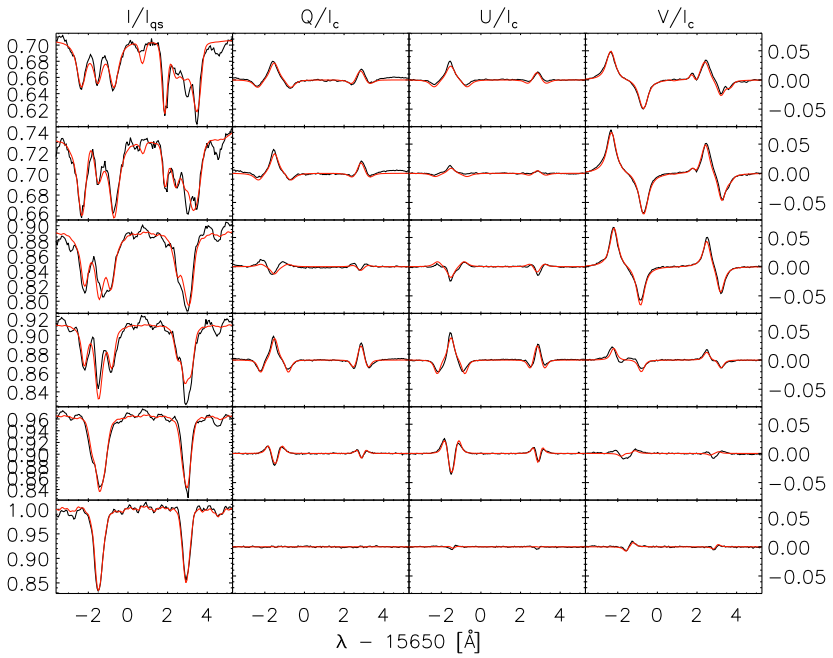

Figure 5.2 An example of inverted spectra from a 15650 Å observation of the sunspot in NOAA 11101. The observed spectra are shown in black and fits are shown in red. Stokes I, Q, U, and V are spectra are shown in the panels from left to right, and the different positions in the sunspot are shown in the panels from top to bottom.

this guess and a functional form containing only the magnetic component to refine the guess. Then *curvefit* is supplied with the functional form containing both the magnetic and non-magnetic components and the resulting parameters and synthetic spectra are saved. Following the fitting of the whole data cube, new guesses for poorly fitted profiles are generated from neighboring points and the profile fitting is executed a second time for these points. Examples of the resulting magnetic field geometry produced by 2CMO for the visible and infrared data are shown in Figures 5.3 and 5.4. The magnetic field strength is scaled 0-3000 G, the inclination runs 0-90°, and the azimuth runs 0-180°.

By simultaneously fitting all Zeeman-split lines in a spectrum we place additional constraints on the solution space for the magnetic field strength and filling factor in regimes where the lines are not fully split. The Fe I lines in the pair (at either visible or infrared wavelengths) have very similar line properties (within each pair), and only their Landé g_{eff}

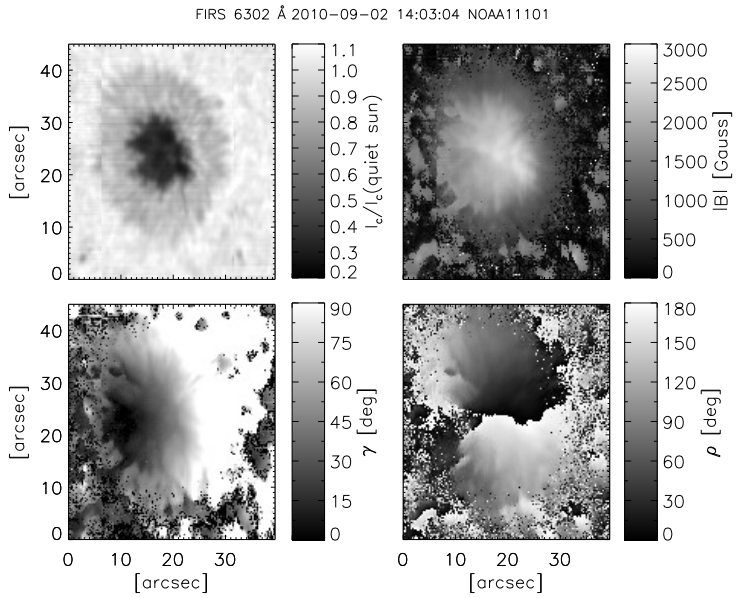

Figure 5.3 An example of the inverted magnetic field strength (B), inclination (γ), and azimuth (ρ) determined for a 6302 Å observation of the sunspot in NOAA 11101. The map of continuum intensity in the sunspot is shown for reference.

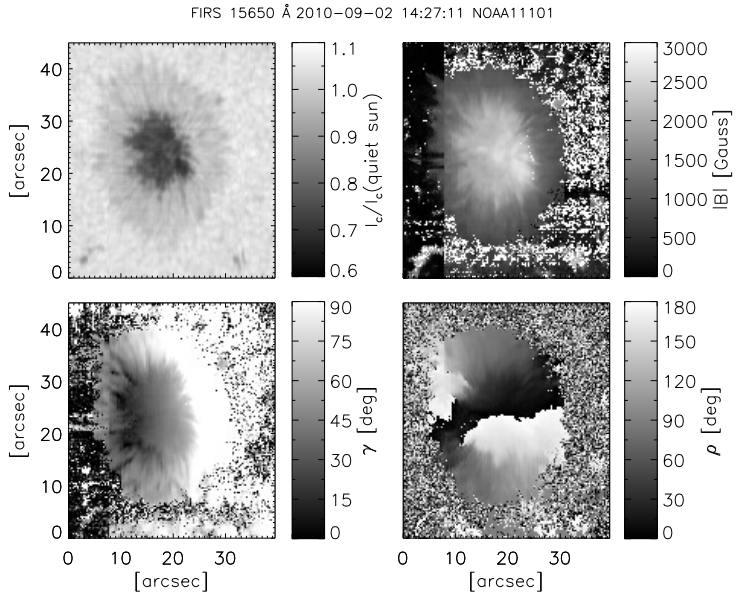

Figure 5.4 An example of the inverted magnetic field strength (B), inclination (γ), and azimuth (ρ) determined for a 15650 Å observation of the sunspot in NOAA 11101. The map of continuum intensity in the sunspot is shown for reference.

is different, allowing us to measure fields as low as 100 G (Solanki et al. 1992). However, the value of η_0 is different for each of the lines (because they also come from different depths) and must be scaled by a constant factor of $g_l f e^{-E_l/kT}$, where g_l is the degeneracy of the lower level, f is the oscillator strength of the transition, E_l is the energy of the lower level for the transition, and T is the gas temperature (we assume an average temperature of 5000 K). The justification for this scaling is given in Appendix D.

5.2.2 The Effect of Systematic Errors on Inversion Precision

With high resolution observations of sunspots, it is possible to produce highly asymmetric Stokes profiles in regions where, for example, multiple velocity components exist along a single sight line. In these cases a Milne-Eddington inversion will necessarily fail to accurately describe the physical situation in the atmosphere. The inversion of Hinode/SP data is a good test of the robustness of an inversion technique in this limit. Efforts are being made to understand the results of Milne-Eddington inversions in the face of such high spatial resolution data (Orozco Suárez et al. 2010). A good deal of effort is also being put into understanding the limitations of the parameter sets selected for Milne-Eddington inversion techniques. Asensio Ramos et al. (2007) have made a very thorough attempt to understand the errors and covariance of parameters arising from Milne-Eddington techniques and address the problem of uniqueness of the solutions.

These problems are universal to Milne-Eddington inversions, and in the context of this research I am most interested in investigating the effect of observational systematics on the fitting precision, i.e. how consistently a result is produced. To test the inversion code I have conducted Monte Carlo tests by constructing a very high signal-to-noise spectrum from sunspot data encompassing umbral, penumbral, and quiet-Sun regimes. The effect of different systematics were tested separately, varying the level of Poisson noise, fringing, and polarization cross-talk and comparing the results.

Figures 5.5 and 5.6 show the results from a test where I applied Poisson noise with a signal-to-noise ratio of 1000, randomized polarization cross-talk with an amplitude of up

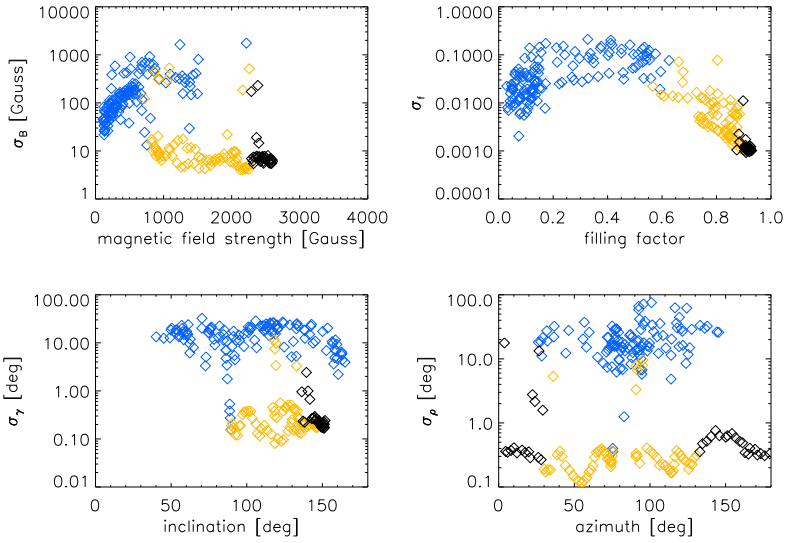

Figure 5.5 Results from Test 2 of the 2CMO Monte-Carlo test for the 6302 Å data showing the standard deviation (σ) for the magnetic field parameters as a function of their value in the quiet-Sun (blue), penumbra (gold), and umbra (black).

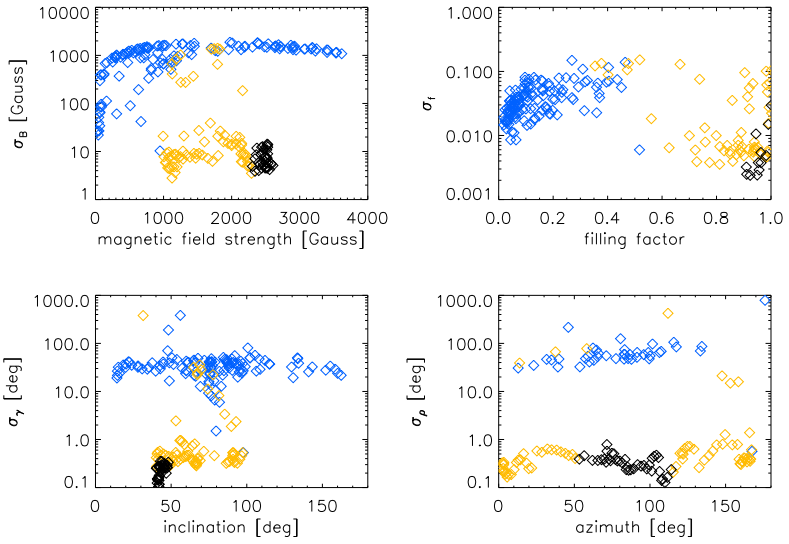

Figure 5.6 Results from Test 2 of the 2CMO Monte-Carlo test for the 15650 Å data.

to 10%, and for just the infrared data synthetic fringes with an amplitude of up to 10%. The plots for the additional tests are shown in the supplemental material for this chapter in Appendix E. The plots show the standard deviation of magnetic field parameters (magnetic field strength, filling factor, azimuth, inclination) produced by the inversion code plotted against the mean value determined at every spatial position. The black points indicate umbral positions, gold points indicate penumbral positions, and blue points indicate quiet-Sun positions. The code does not produce faithful results in the quiet-Sun regime. The magnetic field is retrieved to an accuracy of 10 G for magnetic field strengths larger than 750 G. The azimuth and inclination are faithfully reproduced to an accuracy of 0.5°. In the umbra the magnetic filling factor is precisely determined to 0.1%, while in the penumbra the uncertainty is 10% or larger, as a result of the degeneracy with the magnetic field strength in regimes where the line is not completely split. It is interesting to note that in spite of the difference in magnetic sensitivity of the Fe I lines at 6302 and 15650 Å the fitting precision is roughly the same for both wavelengths, but the good fits are produced down to 750 G for the visible and only about 1000 G for the infrared, which is contrary to what we would imagine. Fringes were not applied to the visible data because this effect is negligible, however in the IR the fringes greatly effect the strength of magnetic field that can be accurately detected. Greater effort must be put forth to reduce the fringes in the infrared observations, but for the purposes of this research 10 G precision in the umbra is sufficient.

5.2.3 Resolution of the 180° Ambiguity of the Magnetic Field Vector

Due to the symmetry of the Q and U signals produced by the Zeeman effect there is an ambiguity of 180° in the component of the magnetic field perpendicular to the line of sight. The resolution of the magnetic field vectors, often called "disambiguation," can be accomplished in many ways. Georgoulis (2005) presented a method for a non-potential magnetic field calculation to determine the most probable configuration of the magnetic field vectors by minimizing current density, but it is possible to carry out the disambiguation

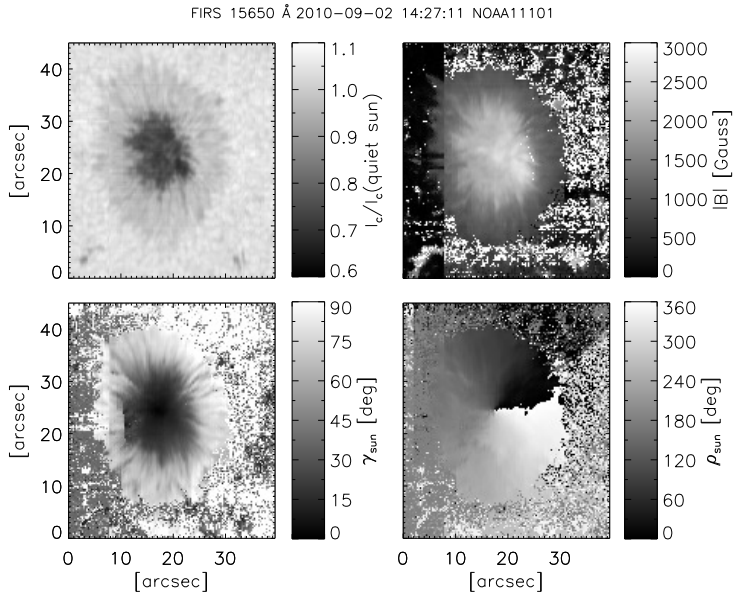

Figure 5.7 An example of the inverted magnetic field strength (B), and the disambiguated inclination (γ) and azimuth (ρ) in local solar coordinates determined for a 15650 Å observation of the sunspot in NOAA 11101. The map of continuum intensity in the sunspot is shown for reference.

more intuitively. The magnetic field of a fairly round sunspot should be symmetrically diverging in azimuth and pointing out of the solar surface. I have disambiguated the vector magnetic field maps by interactively picking the boundaries for the azimuth. Once this is accomplished it is straightforward to project the magnetic field vectors from the observed reference frame into a local reference frame normal to the solar surface using the information known about the telescope pointing and solar geometry at the time of observation. Figure 5.7 shows an example of the disambiguated map previously shown in Figure 5.4. Here the smallest magnetic field inclination is centered on the sunspot and the azimuth smoothly varies about the center from 0° to 360°. Discontinuities such as the one seen in the left hand penumbral inclination angle are unavoidable for this method in sunspots even modestly far from disk center.

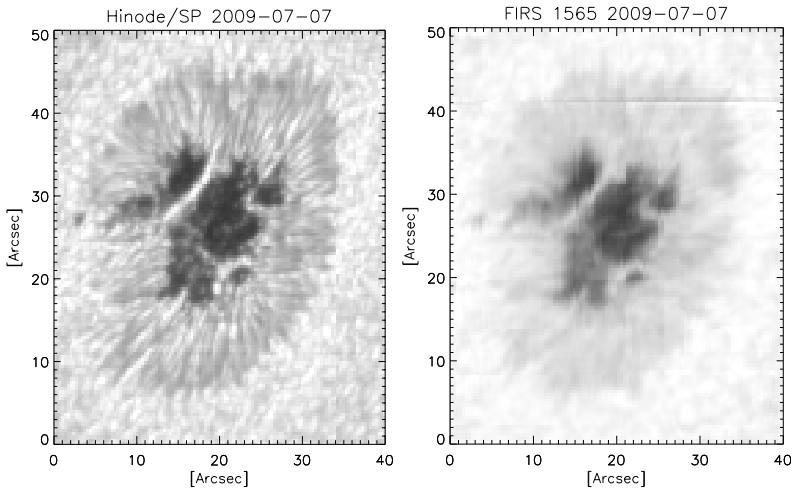

Figure 5.8 Maps of continuum intensity of the main sunspot in NOAA 11024 observed on 7 July 2009 with the SOT/SP on Hinode (6302 Å, left) and FIRS (15650 Å, right).

5.3 Comparison with Other Inversion Methods

In Chapter 4 I compared nearly simultaneous data from FIRS and the Hinode SOT/SP and found them to be consistent. To confirm the validity of the solutions from the 2CMO inversion code, and to contribute to the general understanding of inversion techniques in the community, I have conducted a detailed comparison of the results from four additional codes: MERLIN, NICOLE, SOLIS ME, and the original Gaussian-fitting code by H. Lin that 2CMO is an improvement upon. I have used observations of the main sunspot in NOAA 11024 taken on 6-7 July 2009 from the 15650 Å channel of FIRS, and from the 6302 Å Hinode SOT/SP and SOLIS VSM. Maps of the continuum intensity for these three observations are shown in Figures 5.8 and 5.9. The comparison of inversion codes and observations is summarized in Table 5.1.

5.3.1 Comparison with MERLIN

The Milne-Eddington gRid Linear Inversion Network (MERLIN) from the Community Spectropolarimetric Analysis Center at the National Center for Atmospheric Research is a modern implementation of the ME code originally written for observations from the

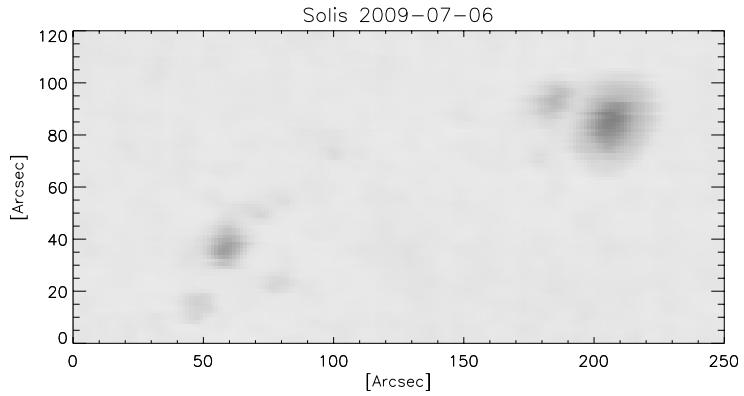

Figure 5.9 Map of the continuum intensity of the sunspot group NOAA 11024 observed on 6 July 2009 with the SOLIS/VSM (6302 Å).

Table 5.1 Summary of Inversion Technique Comparisons

Start Time (UT)	Target	Instrument	Wavelength [Å]	Method
2009-07-07 14:34:50	NOAA 11024	Hinode SOT/SP	6302	2CMO MERLIN NICOLE
2009-07-06 22:38:57	NOAA 11024	SOLIS VSM	6302	2CMO SOLIS ME
2009-07-07 14:33:43	NOAA 11024	FIRS	15650	2CMO Gaussian

Advanced Stokes Polarimeter (Skumanich & Lites 1987; Lites et al. 2007). This inversion code is regarded as the"industry standard" for the Milne-Eddington inversion of Zeeman-split photospheric lines and is used for the analysis of all spectropolarimetric observations from the SOT/SP on Hinode. MERLIN fits the same set of parameters as 2CMO, and although the subtleties of the implementation, initialization, and calculation are different, the two codes describe the same physics and should produce similar results when performed on identical data. Therefore, it is a very relevant and valuable standard for comparison.

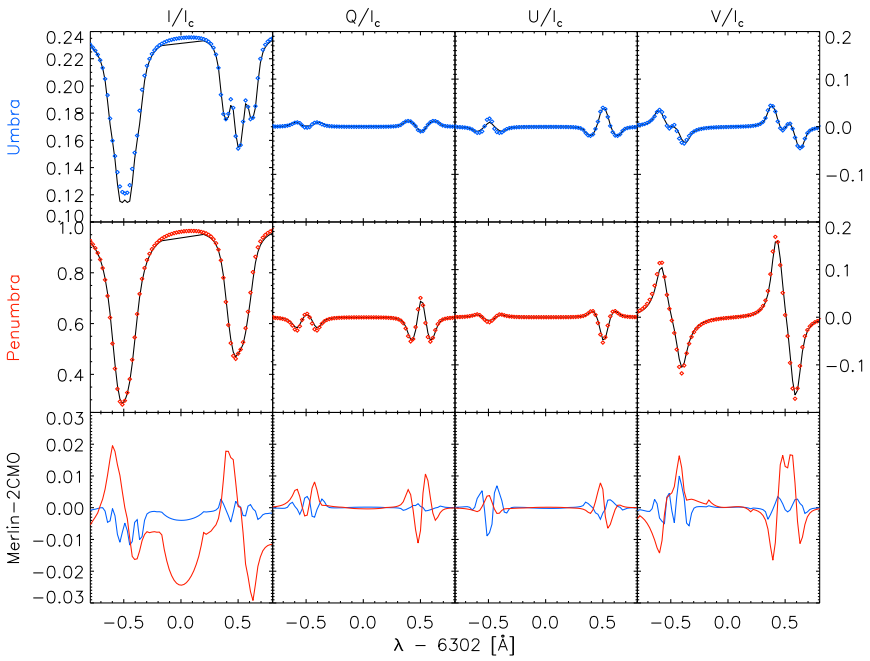

Figure 5.10 An example of synthetic Stokes spectra from MERLIN (black line) and 2CMO (colored points) for the same parameter set simulating the umbral (blue, top) and penumbral (red, middle) positions in NOAA 11024. The MERLIN-2CMO residuals are shown for each position in the bottom panels.

The MERLIN inversion of Hinode data is conducted at the CSAC HAO/NCAR, however the full inversion results and synthetic spectra are publicly available on the web through the CSAC Hinode inversion client (http://www.csac.hao.ucar.edu/).

Comparison of Forward Models

MERLIN and 2CMO fit identical parameter sets, and the forward model of the codes can be compared directly. In this example, I have taken the parameter output from the inverted observation of NOAA 11024 produced by MERLIN and supplied it directly to the 2CMO forward model to generate synthetic spectra. Figure 5.10 shows an example of the synthetic Stokes profiles from MERLN (black line) and 2CMO (colored points) for positions in the sunspot umbra and penumbra. The profiles produced by MERLIN and 2CMO are nearly identical, as they should be. The residuals from the subtracted profiles are near the 2% level.

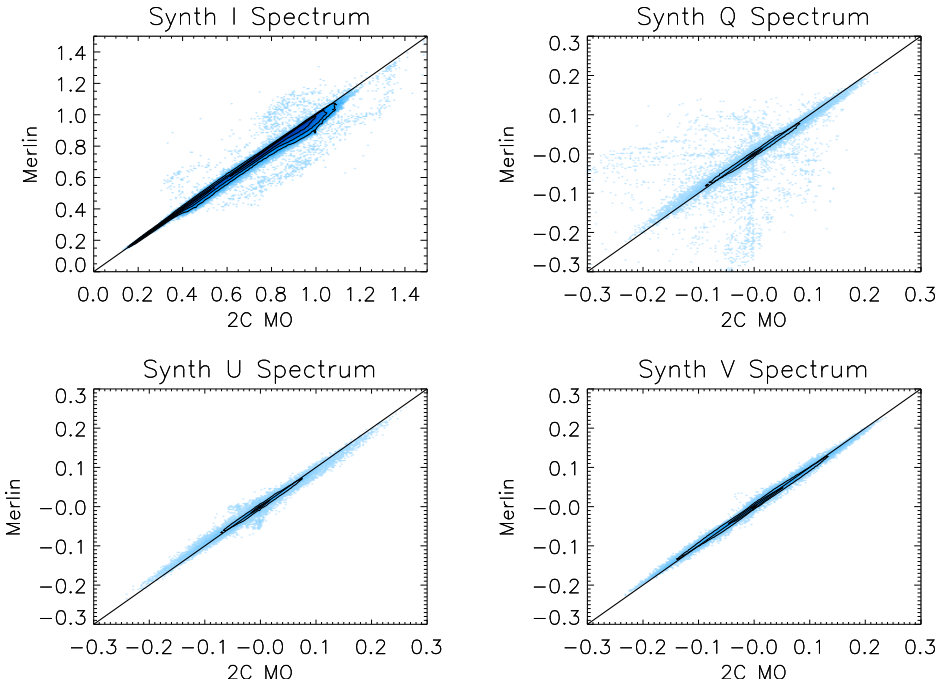

Figure 5.11 Comparison scatter plots showing the agreement between the synthetic Stokes profiles from MERLIN and 2CMO forward model for the Fe I 6302.5 Å line for all fitted points in the NOAA 11024.

The 2CMO Stokes V profile in the umbra shows inversion in the line core of this magnetically weaker line, we expect that this is because we use a different factor (as determined from line constants) to scale the value of η_0 between the two lines (or MERLIN may actually fit separate values of η_0 for each line).

In Figure 5.11 I give a point-by-point scatterplot comparison of every spatial and spectral point in the full-Stokes synthetic spectra produced by MERLIN and 2CMO for the Fe I 6302.5 Å line. In these plots values below 0.8 in Stokes I and greater than +/- 0.01 from the origin for Stokes Q, U, and V, correspond to the line profiles and are of greater interest. The profile values are in quite good agreement and we can conclude that the forward models used by MERLIN and 2CMO are effectively identical.

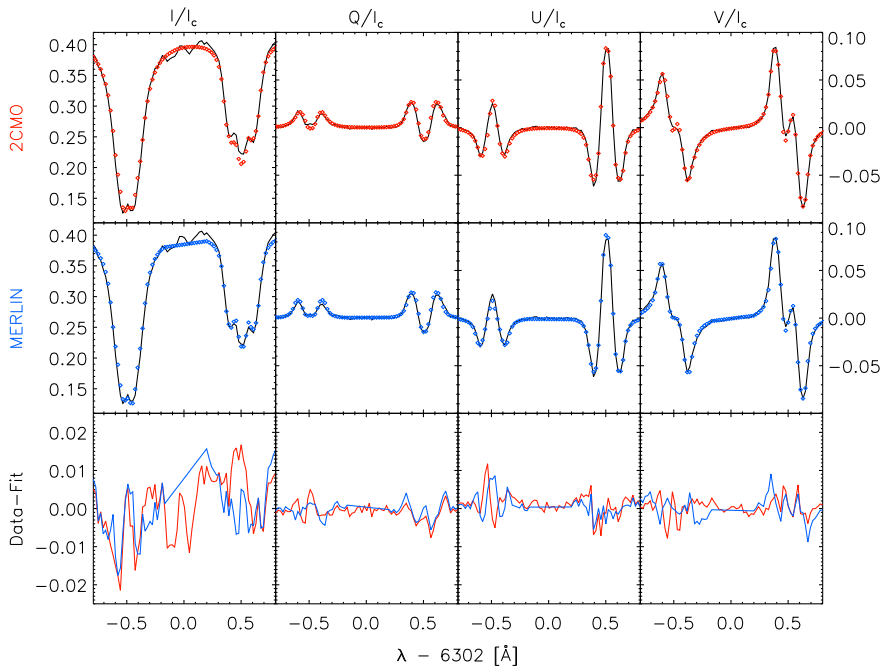

Figure 5.12 Fitted Stokes spectra from 2CMO (red) and MERLIN (blue) and data-fit residuals compared.

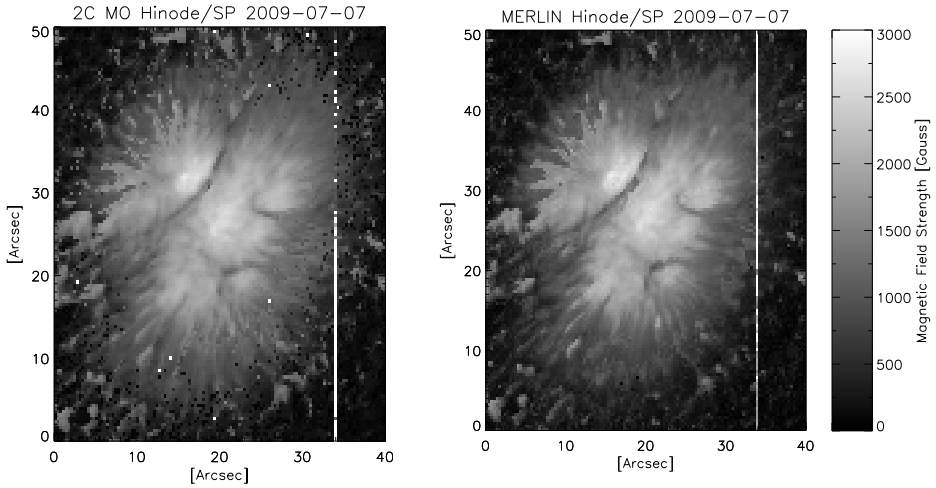

Figure 5.13 Maps of the magnetic field strength from the 2CMO and MERLIN inversions of the Hinode/SP 6302 Å observation of the main sunspot in NOAA 11024.

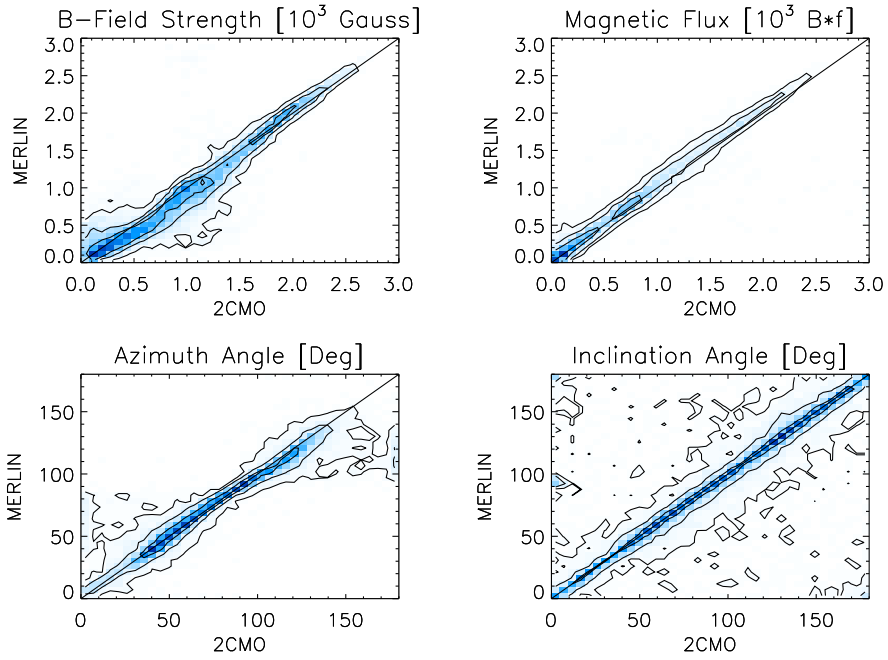

Figure 5.14 Comparison scatter plots for the magnetic field parameters determined from the MERLIN and 2CMO inversion of the two Fe I lines at 6302 Å.

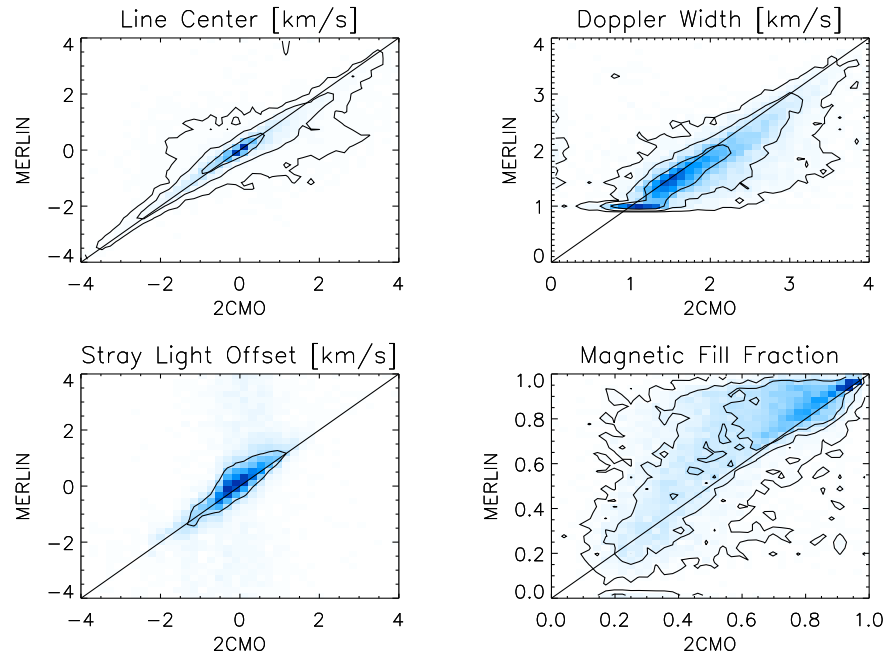

Figure 5.15 Comparison scatter plots for the line parameters determined from the MERLIN and 2CMO inversion of the two Fe I lines at 6302 Å.

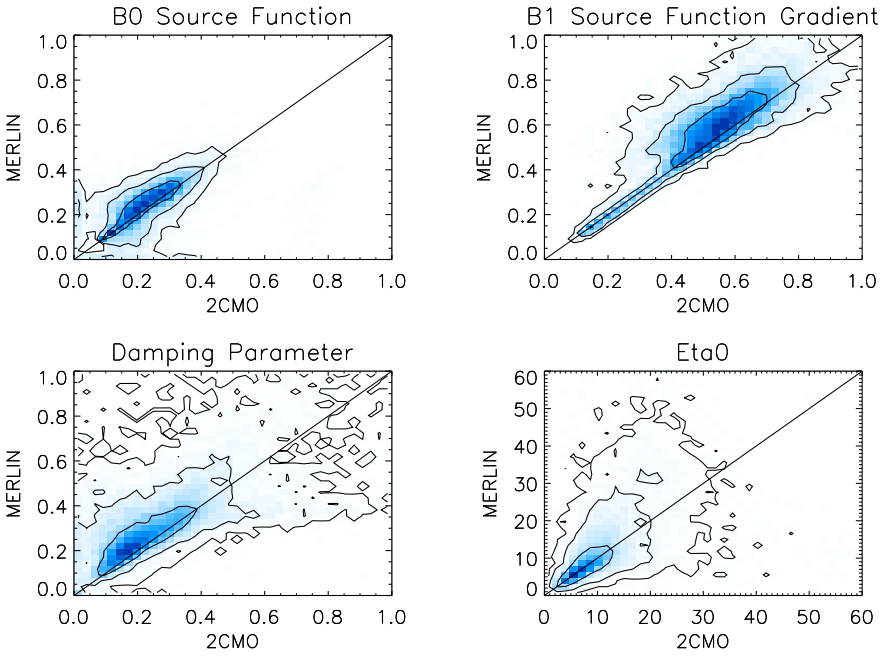

Figure 5.16 Comparison scatter plots for the magnetic field parameters determined from the MERLIN and 2CMO inversion of the two Fe I lines at 6302 Å.

Comparison of Results

Because the profiles produced by the forward model of each code are in such good agreement, any differences in their inverted results must be due to implementation, e.g. the initial guess generation and behavior of the non-linear least squares solver. The profiles fitted to the observed spectra in Figure 5.12 are similar and even produce similar data-fit residuals (bottom panels). The magnetic field maps produced by the two inversion techniques, shown in Figure 5.13, by eye appear to agree very well. Figures 5.14 to 5.16 show the point-by-point scatterplot comparison of the parameters produced by the two techniques. The agreement between magnetic field parameters is best in the umbra where the filling factor and magnetic field strength are high and at least one of the Zeeman-split lines can be fully resolved. The scatter in the magnetic filling fraction, magnetic field strength, and azimuth increases at lower magnetic field strengths, which is expected due to the degeneracy of these two parameters in regimes where the Zeeman-split components are unresolved. The

velocity determined from the line centers agree well, and again there is a little more scatter in the stray light component because it is relatively poorly constrained by the observations at low magnetic field strengths. The line parameters such as the Doppler width, damping parameter, and ratio of line to continuum absorption (η_0) have a larger solution space, and therefore the scatter is larger here, but they still fall symmetrically about the 1:1 line. The overall agreement in the source function and its gradient between the two codes is good, the increased scatter is likely due to low magnetic field strength regions.

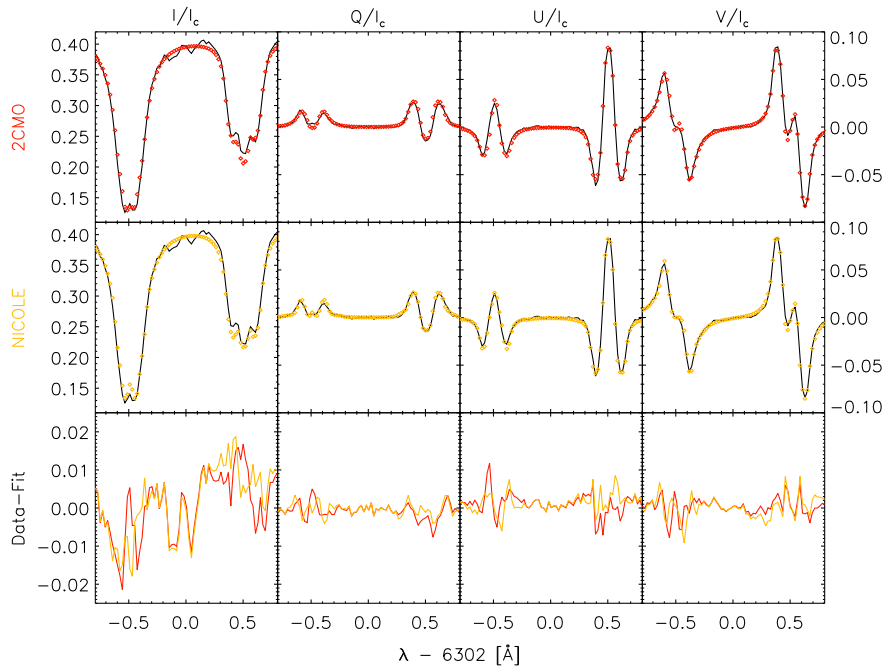

Figure 5.17 An example of fitted Stokes spectra from 2CMO (red) and NICOLE (gold) and data-fit residuals compared.

5.3.2 Comparison with NICOLE

The Non-LTE Inversions COde using the Lorien Engine (NICOLE) is a code by Hector Socas-Navarro which is still currently under development, but I have obtained and compiled a preliminary version of the code for testing purposes. NICOLE is not an ME inversion, and belongs to the second family of codes which does atmospheric fitting based on response

functions. In order to compare the vertically stratified atmospheric parameters produced by NICOLE with the single height parameters from 2CMO, the NICOLE atmospheres were averaged over the $\log \tau_{500}$ = -1 to -2.5 heights which correspond roughly to the 6302 Å line-forming region. Only the magnetic field and velocity parameters can be compared, other parameters which are fit by ME inversions are intrinsic to the model atmosphere and supplied line parameters in a RF inversion. NICOLE was run with just a magnetic component with no accounting for stray light.

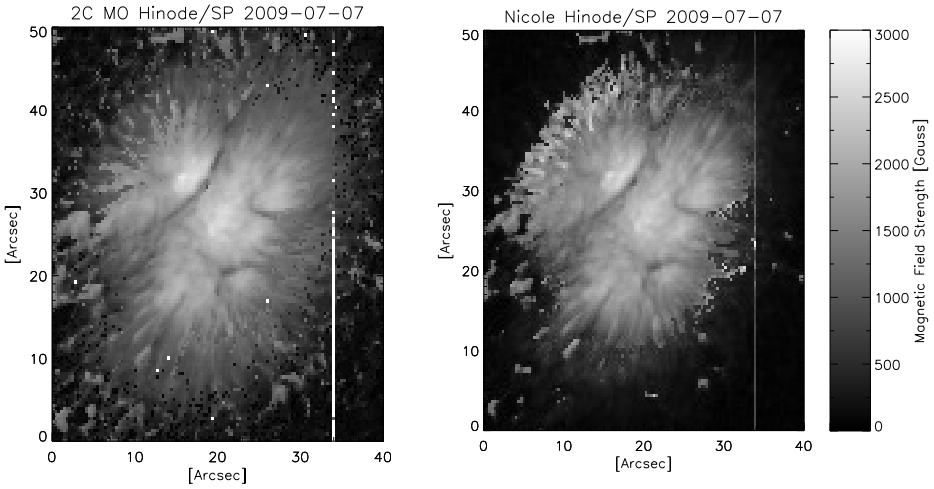

Figure 5.18 Maps of magnetic field strength from the 2CMO and NICOLE inversions for the main sunspot in NOAA 11024 observed on 7 July 2009 by the SOT/SP on Hinode.

The spectra fitted by NICOLE are compared with the 2CMO spectra in Figure 5.17. The synthetic spectra produced by NICOLE are even more similar to 2CMO than MERLIN, and the data-fit residuals are nearly identical. This is probably because we have used the same line constants in NICOLE as were supplied to 2CMO, in spite of the different assumptions made by these techniques, they can in fact describe the same physics. The differences in the magnetic field maps are most evident in the penumbra where the magnetic filling fraction in 2CMO is lower. NICOLE assumes a filling fraction of 1 in this region and produces a correspondingly lower estimate of the magnetic field which can be seen more clearly in the scatter plot in Figure 5.19. However, the umbral field strengths from NICOLE agree well

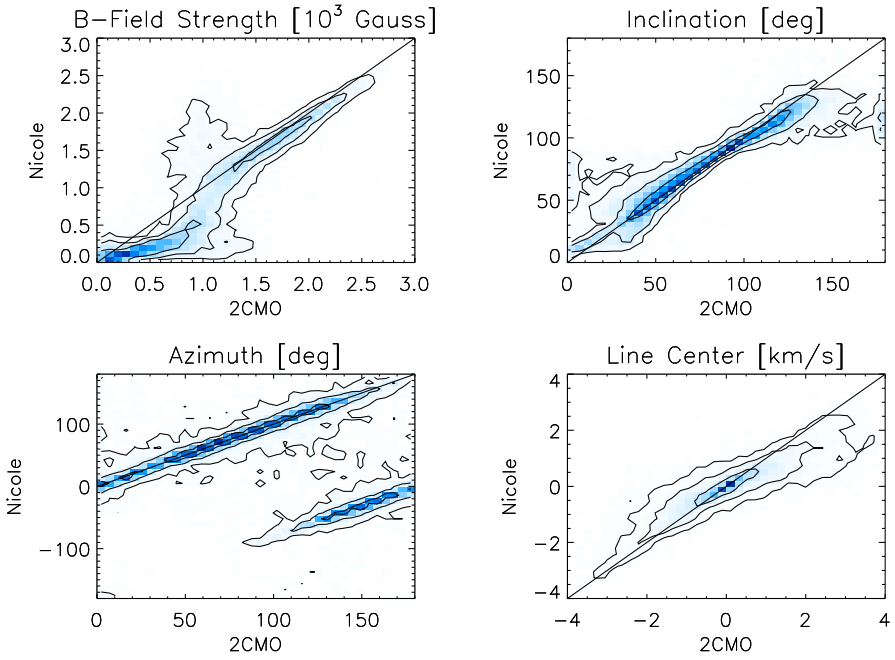

Figure 5.19 Comparison scatter plots for the magnetic field parameters determined from NICOLE and 2CMO inversions of the two Fe I lines at 6302 Å.

in addition to the magnetic field azimuth and inclination, and velocity determined form the atmosphere. The azimuth angle determined for NICOLE runs over a different range of angles but the scatter is still quite small.

5.3.3 Comparison with Gaussian Fitting

The 2CMO inversion code is an addition to an existing inversion code framework by Haosheng Lin, as was mentioned in a previous section. The original code further simplifies the ME assumption by fitting Gaussian profiles to the Stokes spectra. The use of Gaussian profiles greatly increases the speed at which a solution can be reached, but the Gaussian assumption is not always appropriate (e.g. for lines where the magneto-optical effect is significant, Fe I 6302 Å). Like 2CMO, the Gaussian-fitting code simultaneously fits both a magnetic and non-magnetic component using ten free parameters: three polynomial terms to describe the continuum shape; velocity and Doppler width of the line; independent

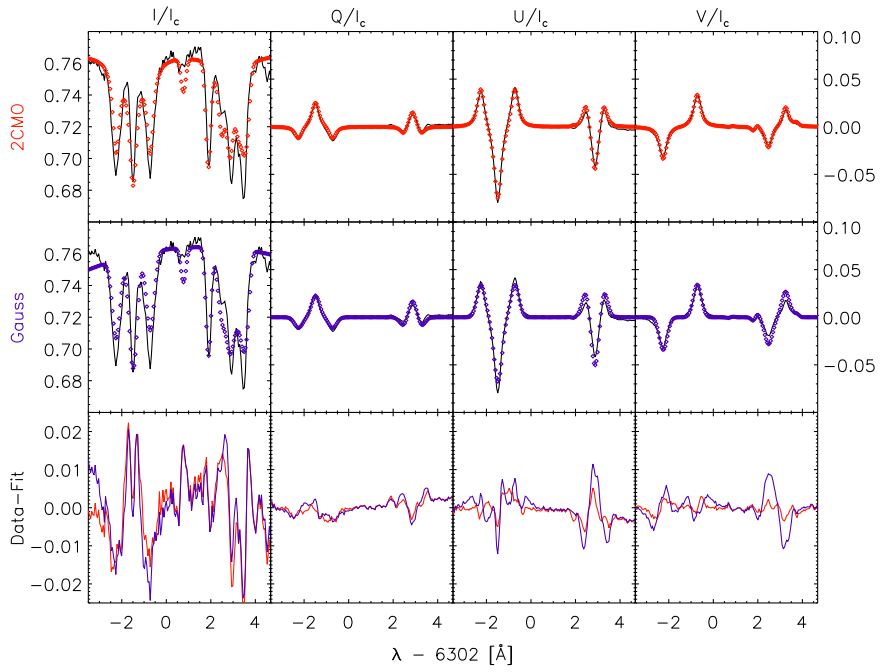

Figure 5.20 Fitted stokes spectra from 2CMO (red) and Gaussian-fitting (purple) inversions and data-fit residuals compared.

magnetic and non-magnetic filling factors; and the magnetic field strength, inclination, and azimuth. The width and depth of the OH lines are fit in addition for 15650 Å observations. The relative amplitude of the Fe I lines are determined beforehand from an average quiet-Sun intensity profile, it and additional line parameters are assumed to be constant over the region.

2CMO and the Gaussian fitting code use the same initial guess generator, however they differ in their treatment of stray light and the way the magnetic filling factor is applied. The synthetic Stokes spectra produced by each method are shown in Figure 5.20. Both codes actually fit the selected Stokes I spectrum relatively poorly. The spacing of the Zeeman split lines is fit accurately, but the magnitudes of the blue and red-shifted components of the line are too small. The differences between the codes become more apparent when looking at the residuals, especially for Stokes V. The Gaussian fitting residual is much larger for the g=1.7 line, and here 2CMO does a much better job, the residual is composed

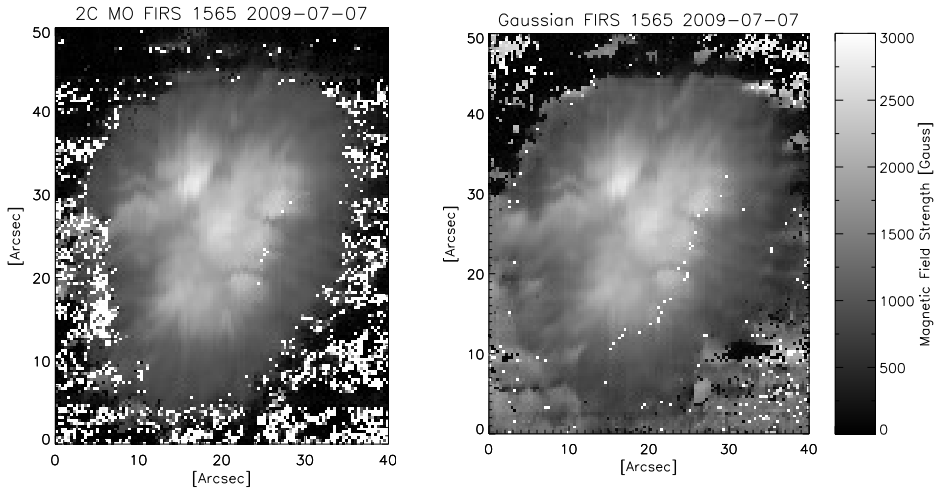

Figure 5.21 Maps of magnetic field strength from the 2CMO and Gaussian-fitting inversions for the main sunspot in NOAA 11024 observed on 7 July 2009 with the FIRS 15650 Å channel.

mostly of noise. This improvement may have more to do with the application of standard line parameters as opposed to the use of quiet-Sun line amplitudes, than the use of more sophisticated line profile calculations as the profiles themselves are of similar shape. The magnetic field maps produced by the inversions shown in Figure 5.21 are strikingly similar, and we can see from the comparison of the parameters in Figure 5.22 and 5.23 that there is virtually no difference. The large splitting of the Fe I g=3 line leaves little room for error all the way down to about 500 G. The azimuth and inclination are also determined similarly by each method, although the scatter is large and random for low magnetic field strengths. Agreement with line wavelength is excellent, the Doppler width is different because for 2CMO we also include the line damping parameter which alters the line shape.

5.3.4 Comparison with SOLIS ME

In this example I compare results from the inversion of SOLIS/VSM observations of NOAA 11024 which were taken on from 2009-07-06. The observations from this instrument are standardly analyzed with a Milne-Eddington inversion code (which we call here SOLIS ME). Like MERLIN, it is based on the original CSAC/ASP code Skumanich & Lites (1987). The

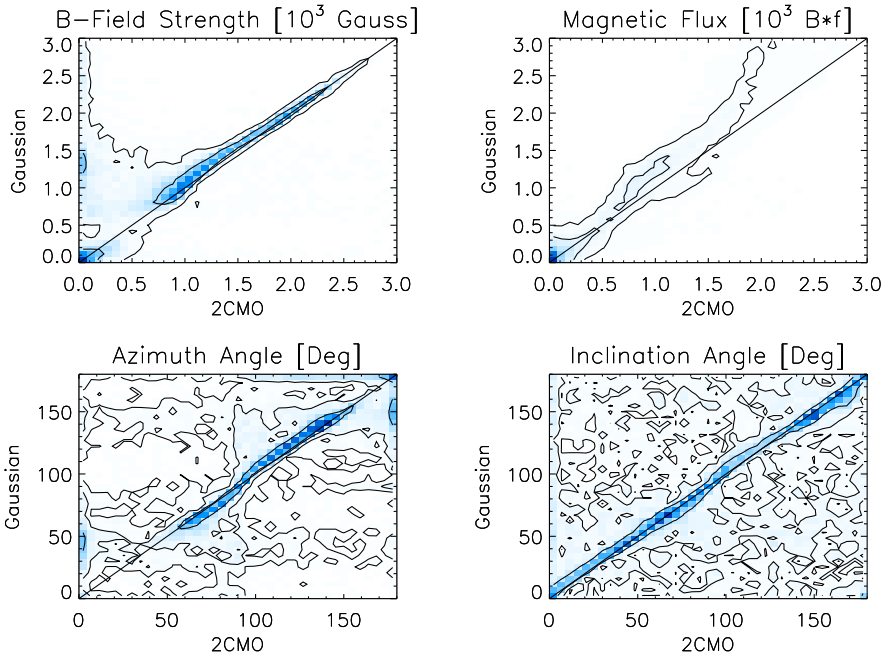

Figure 5.22 Comparison scatter plots for the magnetic field parameters determined from the Gaussian and 2CMO inversion of the two Fe I lines at 15650 Å.

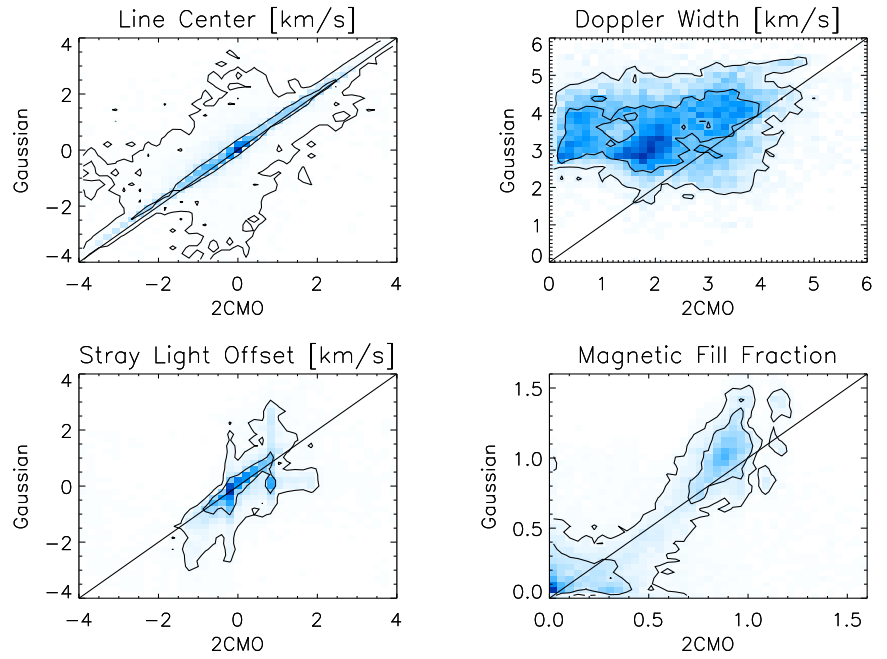

Figure 5.23 Comparison scatter plots for the line parameters determined from the Gaussian and 2CMO inversion of the two Fe I lines at 15650 Å.

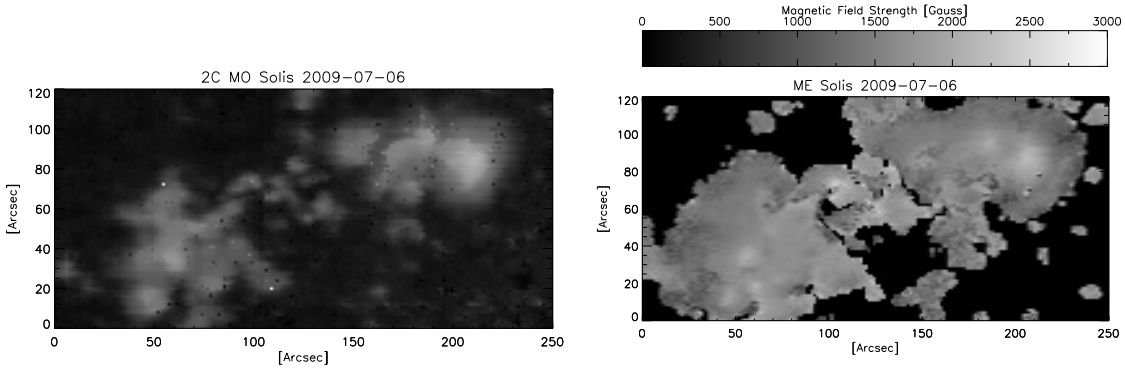

Figure 5.24 Maps of magnetic field strength from the 2CMO and SOLIS ME inversions for the main sunspot in NOAA 11024 observed on 6 July 2009 with SOLIS.

SOLIS ME code has been modified to use a simulated instrumental stray light profile for every position on the disk instead of treating stray light as a free parameter. The SOLIS ME inverted results are publicly available online. The original reduced line profiles were obtained from the SOLIS team and inverted with 2CMO. The SOLIS ME inverted parameter maps have been rotated from the observed orientation such that solar north is aligned with the up direction in the image, the parameter maps resulting from the 2CMO inversion have been aligned and rotated to match this orientation. Unfortunately the synthetic spectra from the inversion were not provided, so a direct comparison is not possible. The maps of the continuum and the magnetic field strength from SOLIS ME and 2CMO are shown in Figure 5.24. SOLIS ME only inverts spectra in the full disk map with polarization signal above a certain threshold to save time, resulting in the patches of signal in the SOLIS ME magnetic field strength map.

A comparison of the 2CMO results and the more limited results available from the SOLIS ME inversion is shown in Figure 5.25. Only the magnetic field geometry is provided by the SOLIS ME result, so we compare magnetic field strength, magnetic flux, inclination, and azimuth. The large fraction of points with zero values in the SOLIS ME maps indicate where the inversion was not carried out due to low signal strength. All parameters show extremely high scatter. The SOLIS ME yields unusually high values of the magnetic field

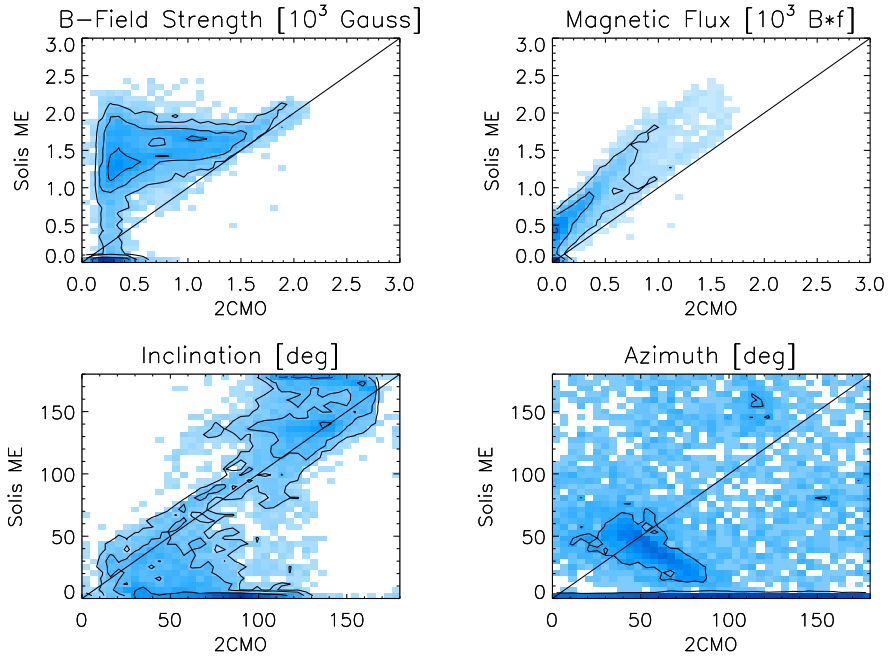

Figure 5.25 Comparison scatter plots for the magnetic field parameters determined from the SOLIS ME and 2CMO inversion of the two Fe I lines at 6302 Å.

strength. However, agreement seems to be reached in regions with the highest magnetic field strength. The magnetic flux determined by SOLIS ME is also about a factor of 1.5 to 2 higher than expected. It should be noted that 2CMO shows some discontinuities in the inverted B map. The SOLIS ME azimuth map has been solved for ambiguity using the method of Georgoulis (2005). The azimuth angles run in the opposite direction and show a 90° offset with respect to the magnetic vector azimuth from 2CMO, making them appear perpendicular to the 1:1 line in the comparison scatter plot.

The overestimation of the magnetic field may have something to do with the treatment of scattered light, the degeneracy of the magnetic field strength and filling factor for weak fields, and the low spectral resolution of the data. The resulting magnetic field map from 2CMO depends heavily on how the Stokes profiles are weighted when they are supplied to the least squares solver. This seems to indicate that the spectral resolution is not high enough for the method to produce unique results within the tolerance of the fit to the Stokes

profiles. They can only supply, at best, an upper limit to the magnetic field strength in all but the strongest magnetic fields. It is difficult to tell whether either method may be at fault without further investigation and detailed comparison of the synthetic line profiles and other fitted parameters which we do not have access to.

In conclusion, the results from SOLIS ME and 2CMO are not in good agreement, in spite of the fact that these two codes should be able to produce similar results. I suggest that further efforts should be made by the SOLIS team in order to increase the quality of their inverted results.

5.4 Discussion on Inversion Results

Based on the favorable comparison of 2CMO with another Milne-Eddington technique, MERLIN, the atmospheric model fitting based on response functions, NICOLE, and a more simple Gaussian-fitting code, I have concluded that 2CMO is able to accurately determine the magnetic field and atmospheric parameters for high spectral resolution observations within the limitations enforced by the Milne-Eddington approximation. From the analysis of the effects of systematic errors and by comparison of the inversion techniques we can assume ~ 70 G error in the umbral magnetic field.

Chapter 6

Atmospheric Modeling and Spectral Synthesis

Abstract

The H_2 abundance in the photosphere cannot be measured directly from H_2 line diagnostics and must be inferred for each sunspot observation based on other observed quantities, i.e. the measured equivalent width of the infrared OH lines. In this chapter I describe the method used to establish these diagnostics from atmospheric models that fully consider the radiative transfer and chemical equilibrium to produce synthetic spectra. These models are also useful as a sanity check and for determining the differences in atmospheric formation height of the magnetic field and temperature diagnostics. I also discuss the determination of OH equivalent width and the effect of stray light on the inferred temperature and H_2 abundance.

6.1 The Measurement of H_2

Obtaining an estimate of the H_2 density from direct observation of H_2 lines is an exceedingly difficult task. Rotational-vibrational transitions of H_2 are scattered throughout the near-infrared. However, these transitions are highly forbidden, and are unlikely to be seen in the solar atmosphere due to the high collisional frequency (private communication, Tom Ayers). The lines are produced by diffuse gas in low-excitation regions (Gautier III et al. 1976; Black & van Dishoeck 1987; Scoville et al. 1983) and have been seen in the outer envelopes of cool

K and M giant stars (Spinrad 1966) which share little in common with the solar atmosphere. We have no reason to believe that the Sun also provides an environment conducive to their formation. Their current non-detection means that they must exist at a level of at least 10^{-5} in intensity (Sonnett 2007). A measurement with such ambitiously high signal-to-noise has not yet been carried out and an instrument capable of such a measurement could easily be put to better purpose.

Transitions of the H_2 molecule can be excited via fluorescence, and in the UV H_2 lines excited by high energy photons produced mostly by the Lyman band and C II can be clearly identified in sunspots (Figure 2.5). These UV lines originate at chromospheric heights many hundreds of kilometers higher than the height of formation of the magnetic diagnostics. Because they are fluorescent transitions, they are highly dependent on excitation geometry which may not be uniform. In addition, the wavelengths at which these lines reside must necessarily be observed using space-based instrumentation, either by rocket (HRTS, Jordan et al. (1978); Bartoe et al. (1979)) or satellite (SUMER, Schühle et al. (1999); Curdt et al. (2004); Innes (2008)). Unfortunately there is no spectrograph currently operating that can observe these wavelengths, but there is hope that an upcoming space-based instrument, the Interface Region Imaging Spectrograph (IRIS), will provide coverage of this wavelength regime. With detailed observations and modeling, we may be able to interpret the UV H_2 lines, but at the moment this option is unsuitable for the purposes of this project.

The formation of H_2 molecules is directly controlled by the temperature of the gas in the solar atmosphere, therefore it is possible to infer the photospheric H_2 abundance from proxy measurements of other similar molecules with stronger spectral signatures. The dissociation energy of the H_2 molecule (4.48 eV) is very similar to the dissociation energy of the hydroxide (OH) molecule (4.39 eV). Barring any competition with the formation of other molecular species we can approximate the ratio of OH to H_2 through Saha-Boltzmann statistics. The ratio of the diatomic molecular species (AB) to its atomic constituents (A and B) is controlled by temperature (T), molecular mass (m_{AB}), and the dissociation energy

(D):

$$\frac{n_A n_B}{n_{AB}} = \left(\frac{2\pi m_{AB} kT}{h^2}\right)^{3/2} e^{-D/kT} \left[\frac{Q_A Q_B}{Q_{AB}}\right] \tag{6.1}$$

where k is Boltzmann's constant, h is Plank's constant, and Q_A, Q_B, and Q_{AB} are the partition functions for the separate atoms and the molecule. The ratio n_{OH}/n_{H_2} can therefore be written:

$$\frac{n_{OH}}{n_{H_2}} = \frac{n_O}{n_H} \left(\frac{m_{H_2}}{m_{OH}}\right)^{3/2} e^{(D_{OH}-D_{H_2})/kT} \left[\frac{Q_H Q_{OH}}{Q_O Q_{H_2}}\right] \tag{6.2}$$

The mass ratio of H_2 to OH is 2/17, and for solar abundances, the oxygen to hydrogen ratio by number is 4.57×10^{-4} ($8.66 = log(n_O/n_H) + 12$ from Asplund et al. (2004)). The tabulated partition function for many diatomic molecules including H_2 and OH can be found in Tatum (1966). If we pick a temperature of 5000 k from this table, the ratio of $Q_{OH}/Q_{H_2} \simeq 28$. Assuming the ratio of atomic partition functions is of order unity, the relation results in $n_{OH}/n_{H_2} \simeq 4 \times 10^{-4}$.

The OH molecule exhibits strong absorption features throughout the visible and infrared solar spectrum. Serendipitously, it has strong absorption lines inside the spectral window of the sunspot magnetic field measurement of the Fe I 15650 Å line pair. Because the direct detection of H_2 in the photosphere is not possible, I have established the equivalent width of the OH 15652 Å line as a proxy for the H_2 abundance through detailed calculation of the chemical equilibrium and OH spectral line synthesis for model solar atmospheres with different effective temperatures. The methods and results are described in the following sections.

6.2 Model Atmospheres

In order to characterize the average properties of sunspot atmospheres we examine 1-D atmospheric models with solar gravity and abundances, but having different effective temperatures to cover a variety of situations from quiet-Sun temperatures to the coolest

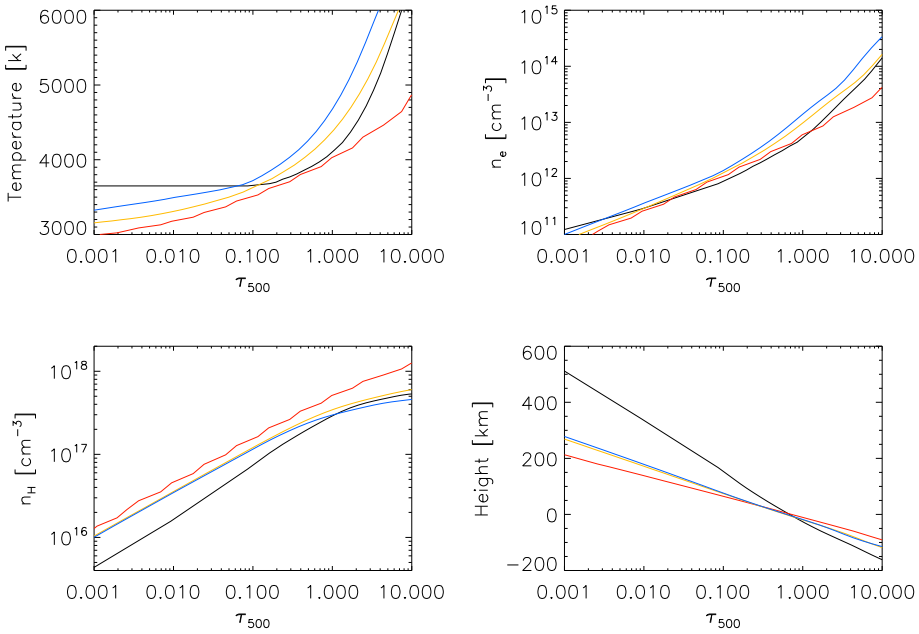

Figure 6.1 A comparison of the atmospheric stratification of temperature (upper left), electron number density (upper right), neutral hydrogen number density (lower left), and geometrical height (lower right) with optical depth for different atmospheric models. The semi-empirical medium umbra model from Maltby et al. (1986) (black) is compared with the Kurucz 4250 K (blue), 4000 K (gold), and Phoenix 3700 K (red) atmospheric models.

sunspot atmosphere. I used the Kurucz model atmosphere grids (T_{eff}=4000 to 7000 K in 250 K steps) (http://kurucz.harvard.edu/grids.html). To fully cover the range of temperatures in sunspot umbrae, cooler models from the updated Phoenix model atmosphere grids (T_{eff}=2600 to 3900 K in 100 K steps) were kindly provided by Peter Hauschildt (Hauschildt et al. 1997).

These models are stellar atmospheres and may not be appropriate to use as sunspot atmospheric models, therefore they must be compared with empirical sunspot atmospheric models for verification. Figure 6.1 shows a comparison of the atmospheric stratification of parameters for the semi-empirical medium temperature umbra model (M) from Maltby et al. (1986) with similar temperature Kurucz and Phoenix models. The overall behavior of the parameters from the different models is similar. However, the Maltby M model does not

agree particularly well with any one model over all heights, and values are in best agreement close to an optical depth of unity where the measurements for the semi-empirical model are made. From the plots in the following section it is apparent that there is some discontinuity between the low temperature Phoenix models and the high temperature Kurucz models in terms of how the atmospheric model effective temperature is determined, but this does not seem to affect the derived diagnostics very much.

6.3 Radiative Transfer and Chemical Equilibrium

To generate the necessary spectral diagnostics, the model atmospheres were supplied to the Rybicky-Hummer (RH) code developed by Han Uitenbroek, which has been demonstrated in Uitenbroek (2000a,b) where it was applied to the interpretation of the CO ro-vibrational bands. The code solves the equations of radiative transfer and statistical and chemical equilibrium for atomic and molecular species in LTE and non-LTE situations. All of the significant solar molecules are accounted for in the calculation of chemical equilibrium: H_2, $H_2{}^+$, C_2, N_2, O_2, CH, CO, CN, NH, NO, OH, and H_2O. I have used the LTE solutions from the code to generate synthetic spectra and obtain the necessary diagnostics for our 1-D plane parallel model geometry. In addition to calculating the molecular populations at every height in the model, I used the RH code to produce a synthetic spectrum of the 6302 and 15650 Å ranges including the Fe I and OH lines for each atmospheric model for ten heliocentric angles. In Figure 6.2 and 6.3 I show for each wavelength regime the continuum normalized spectra from disk center for each atmospheric model from an effective temperature of 2600 K (red) to 7000 K (violet).

The line equivalent widths and the continuum intensity (normalized to the quiet-Sun model T_{eff}=5750 K) were measured directly from the synthetic spectrum for each temperature model and heliocentric angle. The molecular abundances were determined for two heights in the atmosphere which we are concerned with, near $\tau = 1$ in the continuum where the intensity/temperature is measured, and near $\tau = 1$ in the Fe I line core where the

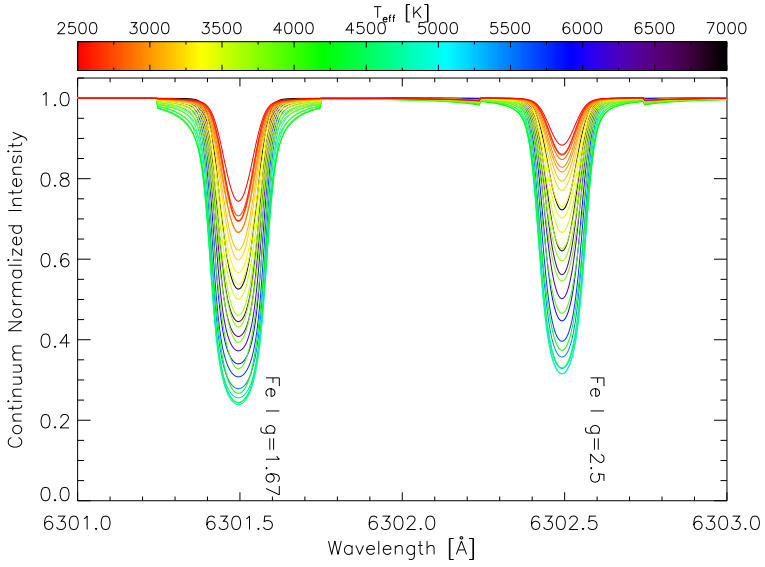

Figure 6.2 The synthetic spectra from RH for the 6302 Å bandpass. The model atmosphere effective temperature for each spectrum is indicated by the range of colors from violet/black (hottest) to red (coolest).

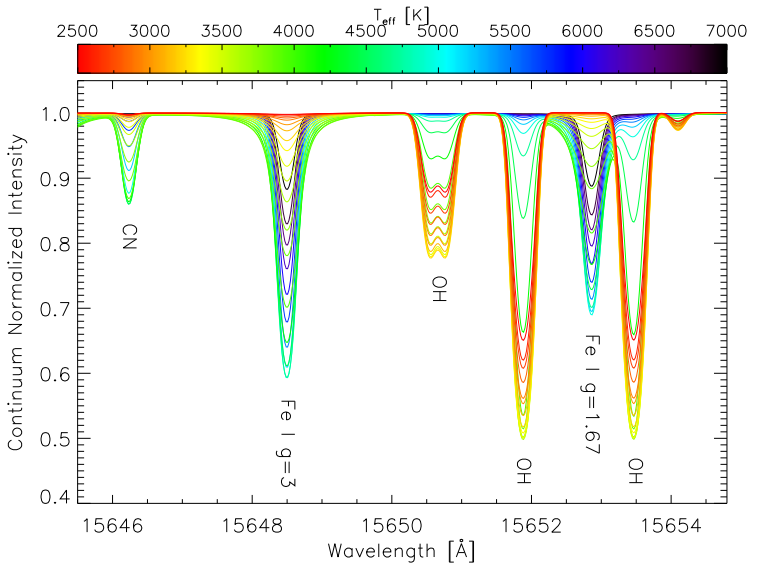

Figure 6.3 The synthetic spectra from RH for the 15650 Å bandpass. The model atmosphere effective temperature for each spectrum is indicated by the range of colors from violet/black (hottest) to red (coolest).

magnetic field strength measurement can be considered to originate (Sanchez Almeida et al. 1996). A linear interpolation was used to retrieve the parameters for arbitrary heliocentric angles and atmospheric temperatures.

6.4 Resulting Theoretical Relations

From the atmospheric parameters, the resulting equilibrium molecular and atomic populations, and the synthetic spectra generated by RH it is possible to extract a variety of useful information about the atmospheres of sunspots. Not only is it possible to investigate the relation of the concentration of H_2 to the OH equivalent width, but also the relation between temperatures and intensities at different wavelengths, the heights of the 15650 and 6302 Å lines and continua, and the possible energy increase resulting from molecule formation.

The use of multiple heliocentric angles is necessary in determining the correct relation between the equivalent width of the 15651.9 Å OH line and the concentration of H_2. The height of $\tau = 1$ changes as a function of heliocentric angle, therefore the OH line strength, intensity ratios, and atmospheric parameters change for sunspots at different positions. For sunspots within about 45° of disk center, the difference is expected to be less than the measurement errors, but is larger for spots close to the limb. To make use of observations at large heliocentric angle we apply a correction based on the atmospheric models. For the heliocentric angle of each sunspot observation we determine the relation of the intensity and OH equivalent width at the heliocentric angle to the equivalent disk center temperature at an optical depth of unity ($\tau_{1.6}$=1) in the IR continuum and H_2 fraction of the total particle number (including H^+, He, and electrons) respectively. In this way data from a wide range of angles can be compared directly. Unless otherwise indicated, the plots and values given are those for a normal viewing angle of the atmospheric model corresponding to disk center.

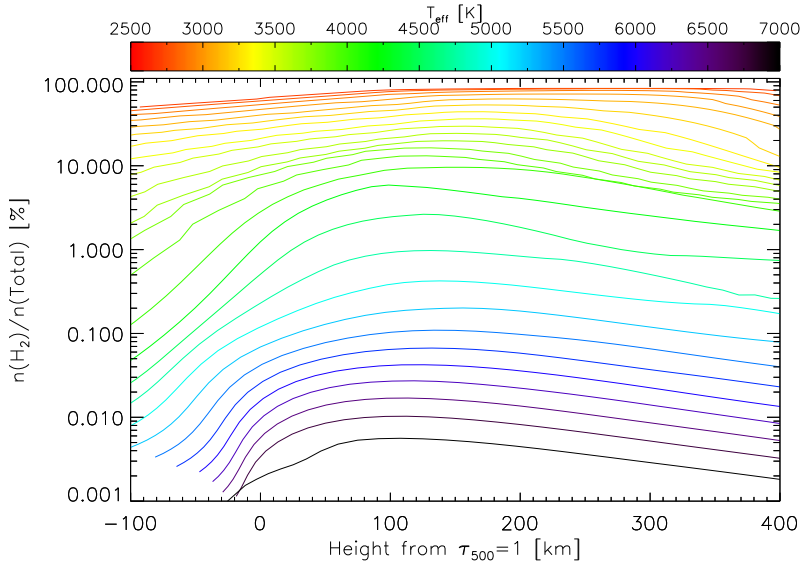

Figure 6.4 H_2 fraction of the total particle number density as a function of height for individual atmospheric models. The model atmosphere effective temperature is indicated by the range of colors from violet/black (hottest) to red (coolest).

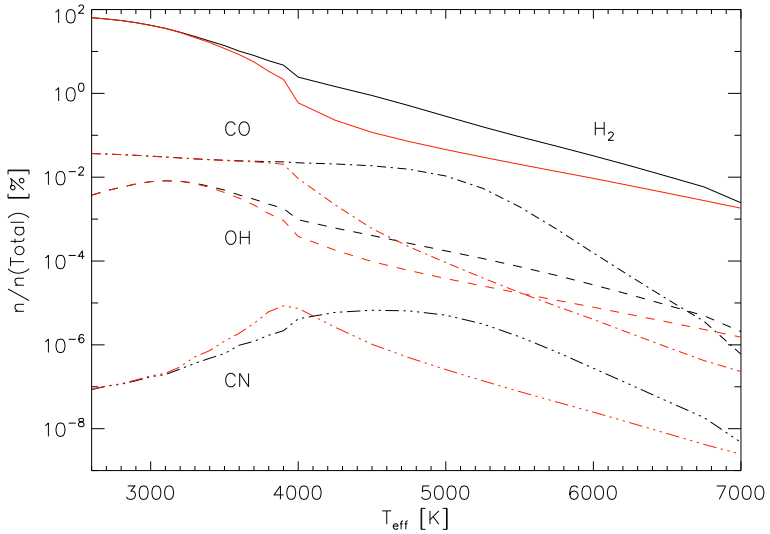

Figure 6.5 Molecular fraction of the total particle number for different molecular species as a function of the model atmosphere effective temperature. The fraction at the $\tau = 1$ height in the 15650 Fe I line core and continuum is indicated in black and red respectively.

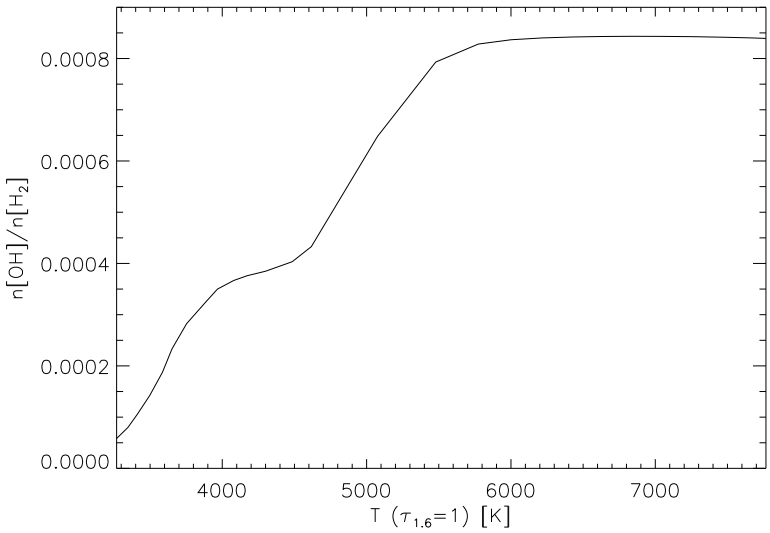

Figure 6.6 The fraction of OH to H_2 molecules at the $\tau = 1$ height in the 15650 Å continuum as a function of the temperature at the same height.

As we expected from their similar composition and dissociation energies, the selection of OH as the proxy of H_2 is very appropriate. The number density of OH very closely tracks the behavior of H_2 as can be seen in Figure 6.5 which shows the logarithmic number densities of the molecules H_2, OH, CO, and CN at τ=1 in the 15648.5 Å Fe I line core (black lines) and continuum (red lines) as a function of the atmospheric model effective temperature. The OH abundance closely tracks the increase of the H_2 abundance from about 3500 to 6500 K, but below 3500 K the OH abundance decreases due to competition with CO for atomic oxygen. The other molecular species do not follow the behavior of H_2 particularly well. The ratio of n_{OH}/n_{H_2} estimated in the first section using Saha-Boltzmann statistics (4×10^{-4}) matches well with the values from the calculation of chemical equilibrium for the atmospheric models shown in Figure 6.6.

Figure 6.7 shows the predicted equivalent widths of the OH 15651.9 Å line over the range of heliocentric angles spanned by the sample as a function of the 15650 Å continuum temperature and intensity at disk center (black shaded region). The fractional number density of H_2 for the line core (red dashed) and continuum (black dashed) are included in

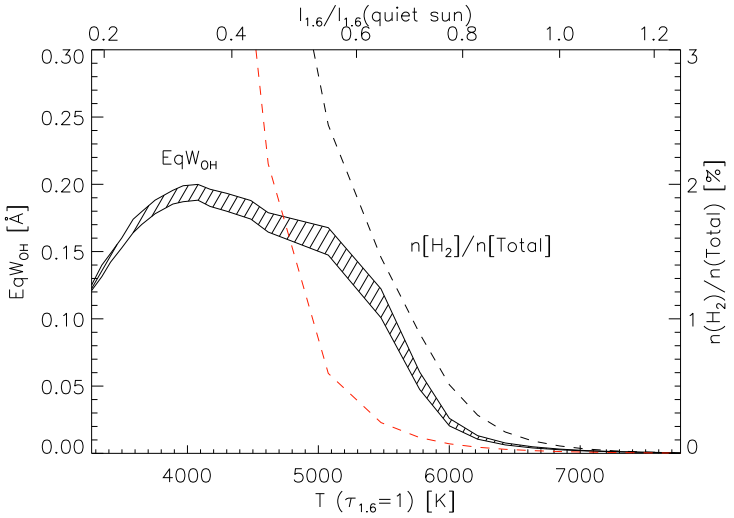

Figure 6.7 The equivalent width of the synthesized OH 15651.9 Å line plotted against the 15650 Å temperature at $\tau = 1$ and continuum intensity. The H_2 density fraction on a linear scale is included for reference and should be read from the right hand scale.

linear scale to show the good correspondence between OH equivalent width and H_2 density. Figure 6.8 shows the direct relation between OH equivalent width and the H_2 fraction at disk center.

It is interesting to note that the $\tau = 1$ levels of the 6302 Å Fe I lines and continuum sit considerably higher in the solar atmosphere than the infrared lines, by about 150 km as can be seen from the plotted heights of the different $\tau = 1$ levels relative to the standard $\tau_{500} = 1$ level for each model effective temperature in Figure 6.10. The line core and their continua come from significantly different heights, and the difference is greater for 6302 Å (i.e. the lines are deeper). The OH lines however originate close to the continuum for the majority of cases that we consider, meaning that it should be closely related to the temperature at $\tau_{1.6} = 1$ in the infrared continuum. Also, because the 6302 Å lines are formed higher in the atmosphere and the environment that produces the lines is cooler, the H_2 fraction present in the atmosphere at this height is more than a factor of 2 greater than at the height of the 15650 Å Fe I line. In Figure 6.11 I compare the H_2 fractions at the height of $\tau = 1$ in

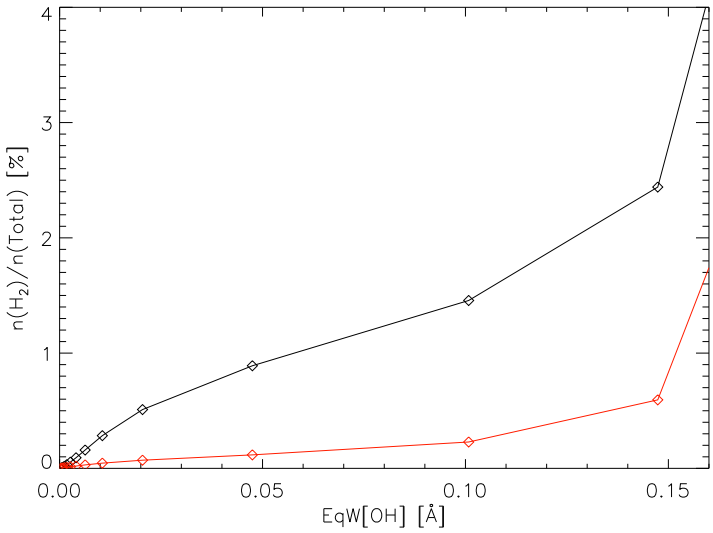

Figure 6.8 H_2 fraction of the total particle number density at the height of $\tau = 1$ in the 15650 Å Fe I line core (black) and continuum (red) as a function of the synthetic OH equivalent width for the range of observed values of EqW_{OH}.

the continuum and line core of the Fe I lines at 6302 and 15650 Å on a linear scale. Based on this, we might expect the formation of H_2 would result in a stronger signature in the thermal-magnetic relation.

The temperatures at height of the line and continuum formation ($\tau = 1$) and their relationship to spectral intensity are not immediately apparent from the data. However, these quantities are easily determined from the atmospheric models and synthesized spectra. In Figure 6.12 I show the relationship between the model atmosphere effective temperature and the various temperatures and intensities for the 6302 and 15650 Å regimes. The continuum temperature from the model is in good agreement with the continuum brightness temperature that can be calculated from the Plank function based on the continuum intensity if the wavelength-dependent continuum temperature at the quiet-Sun is known.

As a sanity check I have calculated the energy resulting from the formation of H_2 to verify that local heating in the atmosphere will not prevent rapid H_2 formation. During the formation of a molecule the maximum possible energy released is the dissociation energy

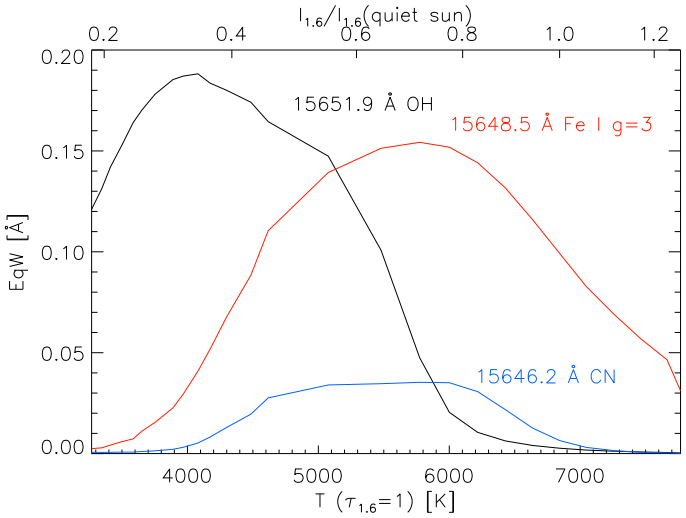

Figure 6.9 Equivalent widths of the OH (black), CN (blue), and Fe I g=3 (red) lines near 15650 Å as a function of the temperature at $\tau_{1.6} = 1$.

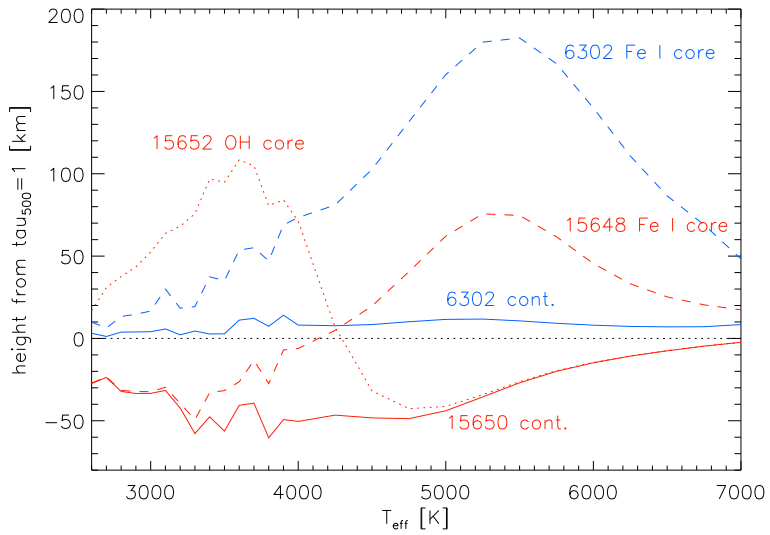

Figure 6.10 Height of the $\tau = 1$ layer in the line core (dashed lines) and continuum (solid lines) for the visible (blue) and infrared (red) diagnostic Fe I lines measured from $\tau_{500} = 1$, as a function of the model effective temperature.

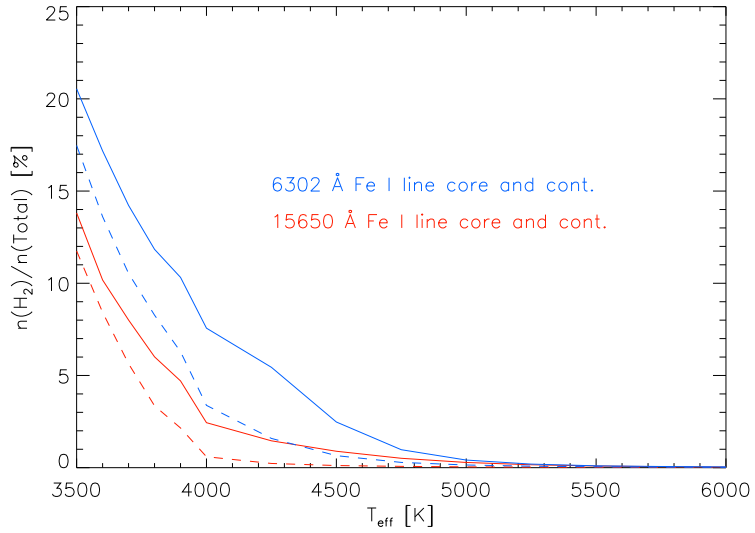

Figure 6.11 H_2 fraction of the total particle number density at the $\tau = 1$ heights of the Fe I 6302 Å and Fe I 15650 Å line cores and continua as a function of effective temperature of the atmospheric model.

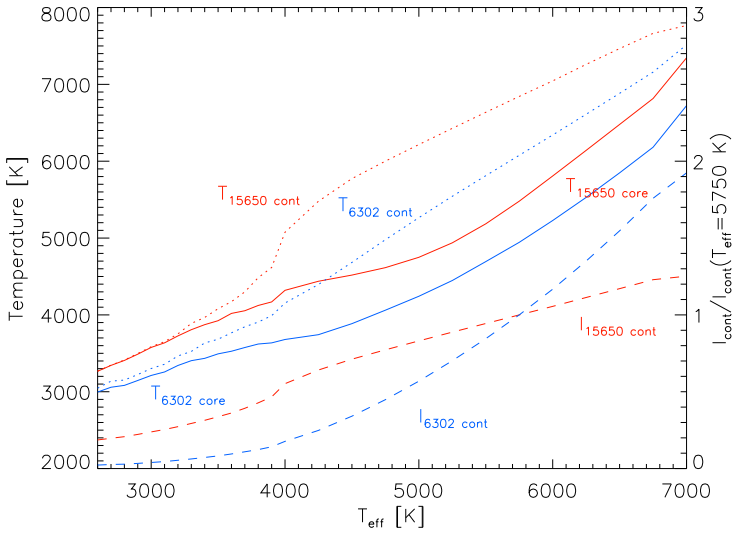

Figure 6.12 The various temperatures as a function of the model atmosphere effective temperature. The temperatures are: temperature at $\tau = 1$ in the 15650 Å Fe I line core (red dotted) and continuum (red solid), and in the 6302 Å Fe I line core (blue dotted) and continuum (blue solid). The 6302 (blue dashed) and 15650 (red dashed) Å continuum intensities have been normalized to the T_{eff}=5750 K model, values are given on the right.

(4.48 eV for H_2). If this energy is transferred directly into the thermal motion of the gas (and not into the rotational and vibrational modes of the molecule) there will be an increase in temperature if the energy cannot be radiated away rapidly enough. In Figure 6.13, for the height of $\tau = 1$ in the 15650 Å line core I plot the maximum possible energy density released by molecule formation (the dissociation energy), this is compared with the mean particle energy for the gas (H_2, H, He, protons, and electrons) on a log scale. For the most part, the energy released during H_2 formation is very very small compared to the thermal energy of the gas. At cooler temperatures (T < 4000 K) however the gas has less energy, the molecules are much more numerous, and the release of their energy of formation is highly significant. If we assume all of the molecules form instantaneously and release their maximum energy we can calculate a cooling time for the gas based on the rate at which energy is radiated away from the atmosphere assuming a blackbody with the atmospheric model effective temperature using the Stefan-Boltzmann equation (surface flux = σT^4), this is shown in Figure 6.14. Even for cool atmospheric temperatures the cooling time is very short, less than a second. We can conclude that release of the molecular energy of recombination in typical sunspot atmospheres would not have a significant effect on the temperature of the atmosphere.

6.5 Application to Observations

6.5.1 Treatment of Stray Light

Sunspots are dark, magnetic regions embedded in extended regions of bright quiet-Sun, which is on average non-magnetic. Unsurprisingly, the contamination of light from one region (brighter) into another (darker) region, or stray light, is one of the largest sources of error in sunspot temperature and magnetic field measurements. Stray light is produced by the scattering of sunlight, mainly from dust grains, in the Earth's atmosphere, and by the scattering of light off optical elements in the telescope and instrument. Seeing can introduce additional blurring and the correction applied by adaptive optics is limited. In

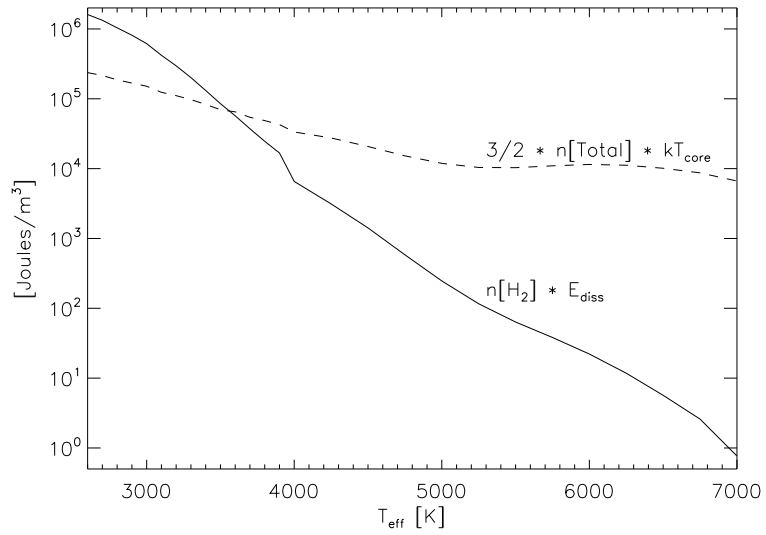

Figure 6.13 The maximum possible H_2 energy of formation released and the average particle energy at the height of the $\tau = 1$ in the 15650 Å Fe I line core for each atmospheric model T_{eff}.

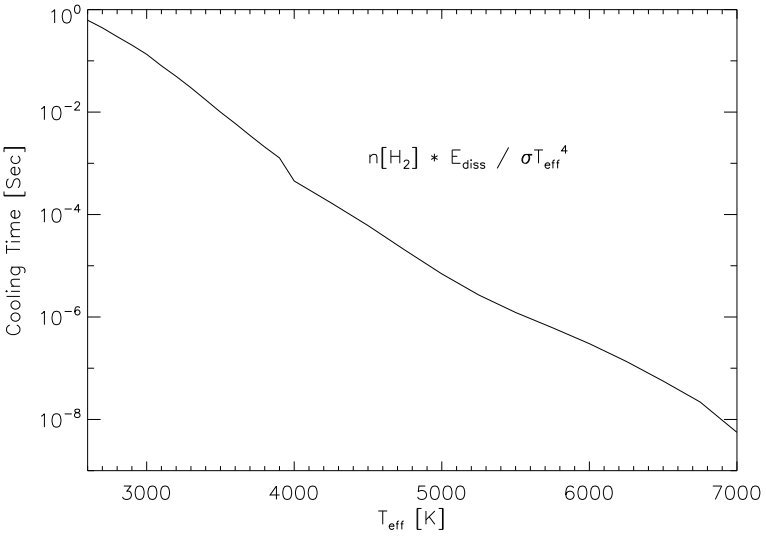

Figure 6.14 The cooling time after spontaneous H_2 formation based on the rate of energy loss due to radiation of the solar atmosphere for each atmospheric model T_{eff}.

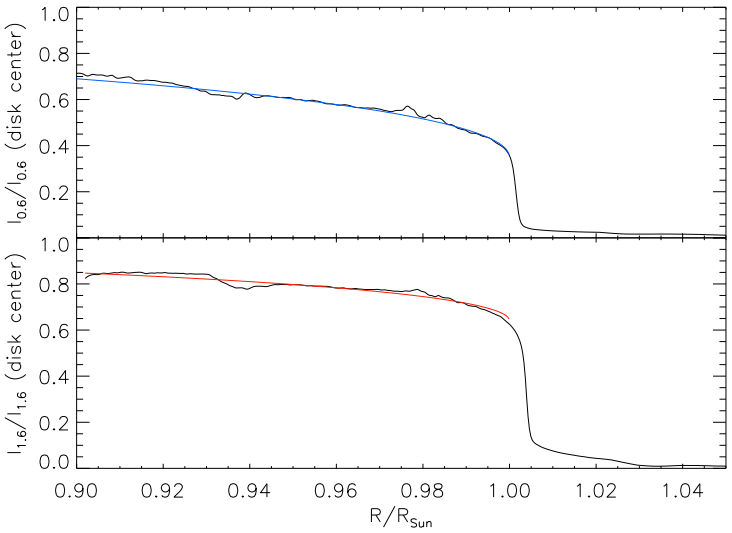

Figure 6.15 Profiles of the 6302 Å (top) and 15650 Å (bottom) continuum intensity at the solar limb as a function of distance from solar center. The colored line in each case is the theoretical limb darkening function from Cox (2000) for 0.6 and 1.6 μm respectively.

fact it can produce large lobes away from the central PSF of the telescope. These all cause light to stray in spatial directions before entering the spectrograph, but downstream from the spectrograph slit light can scatter from optics in both the spectral and spatial directions. Parasitic stray light from the surrounding environment, or scattered from the telescope beam may also contaminate the spectral measurements. Other wavelengths from overlapping orders may find their way onto the detector and, because FIRS has multiple slits, light from the edge of the filter may find its way into adjacent slits.

Detailed discussions of the effects of stray light and its correction have been carried out in Martínez Pillet & Vázquez (1993) and Solanki et al. (1993). In general, stray light results in reduced contrast of the continuum with respect to the quiet-Sun, the contamination of sunspot intensity profiles with a stray quiet-Sun signature, and reduced amplitudes of the polarized components. The correction of stray light is particularly important for the accurate determination of quantities which depend on intensity such as temperature and the equivalent width of spectral lines.

Spatial stray light may arise from both atmospheric and instrumental sources, but without proper characterization of the instrumental and atmospheric contribution to spatial stray light, the fraction due to the instrument and telescope cannot be distinguished from the unresolved mixing of adjacent magnetic and non-magnetic regions in a single resolution element. The degeneracy of magnetic field strength and its filling fraction in weak field regimes is an additional source of ambiguity. The filling fraction determined by the inversion can only place an upper limit on the spatial stray light. Detailed measurements to determine the stray light for FIRS have not been carried out. However in observations at the solar limb (shown in Figure 6.15) the stray fraction at 0.02 $R_\odot$ or about 20" above the limb is approximately 3% and 5% of the disk center brightness at 6302 and 15650 Å, or about 8% of the intensity at the limb. The theoretical limb darkening from Cox (2000) for 6000 and 16000 Å is plotted the data in color. The off-disk stray light fraction agrees with the filling fraction determined in the umbra of very large sunspots, where the photosphere should be almost entirely magnetic. Therefore we apply the inverted magnetic filling fraction to correct the spatial stray light in the intensity spectrum, assuming that all magnetic elements are resolved.

There is no line-free source of spatial stray light, therefore any spectral stray light must arise only following the slit of the instrument, from scattering of light off the imperfect surface of the grating, and optics, the overlap of spectral orders which pass through the filter, or the presence of parasitic light from the surroundings. It is possible to estimate the spectral stray light by measuring the equivalent width of spectral lines and comparing them to a measurement free of spectral stray light. The presence of a possible large spectral stray fraction in FIRS and its effect on the measurement of equivalent width is discussed further in the following section.

6.5.2 Measurement of OH Equivalent Width

The 15651.9 Å OH line is the best candidate as a diagnostic for H_2. It is stronger than the blend at 15651.0 Å and less affected by the Zeeman splitting of the Fe I g=1.7 line than the

15653.4 Å line. However, the g=1.7 Fe I line would still contaminate a direct measurement of the equivalent width of the OH line due to the Zeeman effect and significant changes in the equivalent width with temperature, seen in Figure 6.9. Observational systematics in FIRS, such as fringing and improper correction of the filter profile, also increase the uncertainty in the direct measurement of equivalent width from the integrated spectrum. Therefore in the umbra I have calculated the equivalent width from the inverted parameters for this OH line (the inversion fits are actually very good for this line, see Figure 5.20), and elsewhere in the sunspot the equivalent width is calculated directly by integrating over the Stokes I spectrum because it does not produce poor results when the OH lines are weak. The details of the equivalent width measurement are discussed in further detail in Appendix F.

A great deal of care has been put into ensuring that the equivalent widths for the OH line have been properly determined. Yet the empirical relation between the OH line equivalent width and temperature does not agree with the relation predicted by spectral synthesis of model atmospheres, as I will discuss further in this section. Figure 6.16 shows the compiled EqW_{OH} vs. continuum brightness temperature at 15650 Å for the entire sunspot survey. The data points have been arranged into 50 K bins and the mean and standard deviation of each bin were calculated, as indicated by the point and associated error bar. The solid red line indicates the theoretical OH EqW at disk center. Although the relations do not agree, the data and model qualitatively describe the same curve which appears to be different only by a factor of about 1.55 at temperatures below 6000 K where the OH lines are strong enough to be measured accurately.

This disagreement between the empirical and theoretical relation of OH equivalent width with temperature must be due to systematics in the observations or the realism of the model atmospheres and the RH code. A large amount of spectral stray light could account for the consistent underestimation of the OH equivalent width. If we assume a model where the spectral stray light level at any given position is a constant fraction of the quiet-Sun continuum ($I' = I + f * I_0$, e.g. from scattered or parasitic light), then the corrected

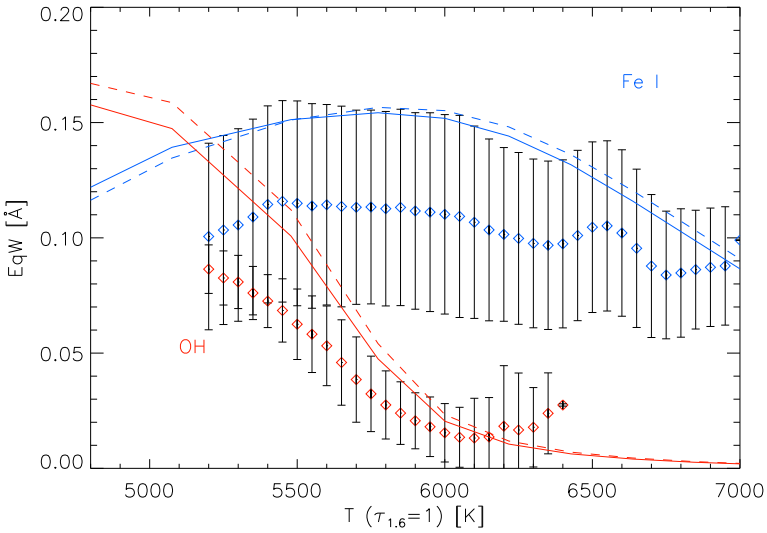

Figure 6.16 Equivalent width of the OH 15651.9 Å and Fe I 15654.8 g=3 line compiled from the sample vs. continuum temperature at 15650 Å.

equivalent width would be increased, but the continuum contrast would also be increased, and the resulting relation would move up and to the left in Figure 6.16. However, if we assume that the stray fraction at any given position is proportional to the continuum intensity at that position ($I' = I*(1+f)$, e.g. from overlapping orders), then the equivalent width increases by a multiple of the stray fraction, but the continuum contrast is preserved. The most likely situation for the spectral stray light is a combination of these two, where the stray fraction is a constant fraction of the intensity input through the slit, which for FIRS is very complicated to determine.

If there was a large fraction of spectral stray light it should be seen to affect the iron lines in the same way. Figure 6.16 also shows the equivalent width of the whole Fe I g=3 Zeeman triplet calculated for the sunspots and the surrounding regions of quiet-Sun. The data were binned and the mean and standard deviation were determined in the same way as for the OH line, and are shown in blue. The large scatter in the Fe I measurement makes the possible stray light contamination difficult to determine for these lines, but it appears

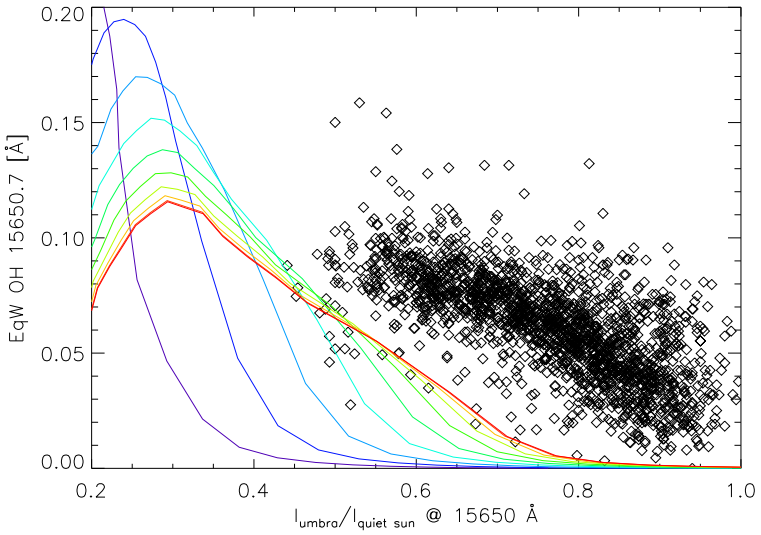

Figure 6.17 Equivalent width of the OH 15650.7 Å blend for the Livingston sunspot sample vs. continuum intensity at 15650 Å with the RH predictions for equivalent width of this line for heliocentric angles from disk center (red) to the limb (violet).

that they also underestimate the Fe I equivalent width relative to the theoretical relation indicated by the blue curve.

We are reasonably confident with the calculation of the OH populations and the lines in the synthetic spectra, however the same cannot be said about the Fe I lines. The Fe I lines have non-LTE ionization and we have calculated them assuming LTE.

As a check of the models I have compared observations from two additional sources. The infrared NSO sunspot atlas (#5) obtained with the Fourier Transform Spectrometer (FTS) on the McMath-Pierce Solar Telescope at Kitt Peak provides a high-resolution, pristine solar spectrum with exceptionally low spectral stray light (Wallace & Hinkle 2001). From the profiles (available at ftp://nsokp.nso.edu/pub/atlas/) I have measured an equivalent width of 0.0968 Å for the Fe I 15648.5 Å g=3 line which is in very good agreement with the quiet-Sun model with a temperature of 6850 K at $\tau_1.6 = 1$.

The Livingston sunspot survey observations (1995-present, e.g. Penn & Livingston (2006)) have been obtained with the McMath-Pierce main spectrograph by scanning a single germanium diode through the spectral profile at 15650 Å to produce consistent observations

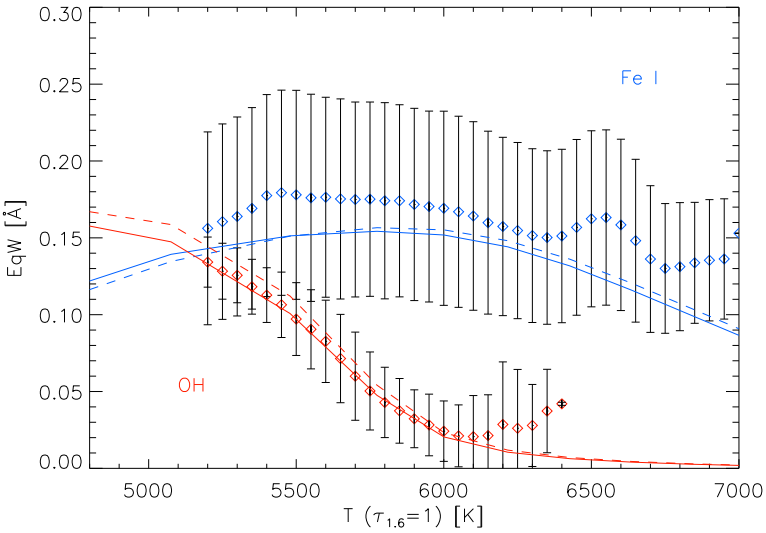

Figure 6.18 Equivalent width of the OH 15651.9 Å and Fe I 15654.8 g=3 line compiled from the sample scaled by a factor of 1.55 vs. continuum temperature at 15650 Å.

with exceptionally low noise. Line fitting has been done to obtain Gaussian full-width at half max and amplitude of the 15650.7 Å OH blend relative to the continuum and the continuum contrast ratio of the sunspot umbra. The data are plotted in Figure 6.17 for all the spots in the sample. The observations were obtained in good seeing conditions, however they are not corrected for spatial stray light so we expect this is why they lie to the right of the theoretical relations for different heliocentric angles from disk center (red) to the limb (violet). However, the Livingston observations do agree in magnitude with the theoretical values, which is consistent with a low level of spectral stray light.

From the agreement of the model results with the McMath observations, I believe we can trust the synthetic OH equivalent width for all model temperatures and the Fe I equivalent width for at least the quiet-Sun model. The measurement of Fe I equivalent width from FIRS cannot be relied on due to its uncertainty, and we must adopt the scaling from OH itself to match the data to the theory. The OH and Fe I equivalent widths, adjusted by a factor of 1.55 in comparison to the models, are shown in Figure 6.18. The agreement achieved by applying a simple scaling factor is compelling, but if we are to believe this

factor, it implies a spectral stray light fraction of 35% under the second stray light model described. This model of stray light is not particularly realistic for FIRS, where we might expect overlapping orders to fall onto neighboring slits. From the favorable comparison of FIRS 6302 Å with Hinode SOT/SP data, which shows very comparable contrast in both the continuum and in the lines, we can conclude that the possible stray fraction is at least limited to the infrared. In the future the measurement of OH may be carried out using OH lines at visible wavelengths. Additional investigation is necessary to confirm the presence of such a large spectral stray light fraction in FIRS, but for the moment we apply a universal factor of 1.55 to the OH equivalent widths and proceed with the H_2 inference under this caveat. Under this assumption the scatter in the OH equivalent width measurements of H_2 fractions of less than about 0.2% cannot be distinguished from a null detection.

Chapter 7

Survey Results

Abstract

In this section I present the relationship between magnetic field, temperature, and molecular hydrogen fraction for every sunspot observation in the survey. The results from the analysis of the observations by Milne-Eddington inversion and spectral synthesis from the radiative transfer and chemical equilibrium applied to atmospheric models are brought together. While the RT effect prevents a direct determination of the equilibrium state of the entire sunspot, the measurements of temperature and magnetic field in the umbra should describe the equilibrium state of a simplified sunspot atmosphere. The proxy measurement of OH provides an additional temperature-sensitive measure and confirms the validity of the theory that H_2 is able to form in abundance. Individual histories for each active region are discussed to provide context and determine the evolutionary state of the region for each observation.

7.1 Description of Results

7.1.1 Figures

The results for each of the sunspot observations are shown at the end of this chapter in Figures 7.3 to 7.68 listed in chronological order. The figures consist of four panels from the analysis of the IR 15650 Å observations and, if visible observations were

taken simultaneously, I include three additional panels from the analysis of the 6302 Å observations. The plots are as follows:

IR Thermal Magnetic Relation and EqW$_{OH}$ The upper left plot shows the square of the magnetic field from the inversion of the 15650 Å iron lines plotted against the continuum brightness temperature determined from the fill-factor corrected continuum intensity. Umbral (black points) and penumbral (gold points) regions have been selected based on intensity. The measured OH equivalent width, scaled by a universal factor of 1.55, is plotted in red with values given by the red scale on the right hand side of the plot. The theoretical OH equivalent determined from the synthetic spectrum is included in black.

Inferred H_2 Fraction In the upper right plot, the blue points indicate the H_2 fraction of the total particle number at the τ=1 height in the core of the Fe I 15648.5 Å g=3 line inferred from the scaled OH measurement plotted against the continuum brightness temperature. The black line shows the relation of the H_2 fraction at $\tau = 1$ in the 15648.5 Fe I line core and temperature at $\tau = 1$ in the nearby continuum determined from the model atmospheres.

IR Magnetic Field Strength Map The lower left panel in the group of four shows the map of the magnetic field strength produced by the inversion of the 15650 Å Fe I lines. The solid gold line and dotted black line show the selected penumbral and umbral boundaries respectively.

IR Intensity Map The lower right panel shows a map of the continuum intensity of the sunspot at 15650 Å. The solid gold line and dotted black line show the selected penumbral and umbral boundaries respectively.

Visible Thermal Magnetic Relation and EqW$_{OH}$ The upper left plot shows the square of the magnetic field from the inversion of the 6302 Å iron lines plotted against the continuum brightness temperature determined from the fill-factor corrected

continuum intensity. Umbral (black points) and penumbral (gold points) regions have been selected based on intensity.

Visible Magnetic Field Strength Map The lower left panel in the group of 3 shows the map of the magnetic field strength produced by the inversion for the 6302 Å Fe I lines. The solid gold line and dotted black line show the selected penumbral and umbral boundaries respectively.

Visible Intensity Map The lower right panel shows the continuum intensity map of the sunspot region at 6302 Å. The solid gold line and dotted black line show the selected penumbral and umbral boundaries respectively.

7.1.2 Tables

The properties of the sample determined from the infrared observations are summarized in Table 7.1. For each sunspot observation I list the maximum umbral field strength, minimum umbral temperature, maximum OH equivalent width scaled by a factor of 1.55, and maximum inferred H_2 fraction at $\tau = 1$ in the Fe I g=3 line core. The classification of the B^2-T relation, umbral area and total spot area are also listed. The maximum umbral magnetic field, minimum umbral temperature, umbral area, and total sunspot area are listed for the visible observations in Table 7.2.

7.1.3 Comments on Errors

For the sample as a whole, the overall agreement of the OH equivalent width resulting from the models and the measured OH equivalent width scaled by a factor of 1.55 is encouraging. This agreement is best in the larger spots. In smaller spots the correction applied by the filling fraction is larger, and it is not possible to fit the OH lines in the spectrum with good accuracy, this contributes to an underestimation of the OH equivalent width and greater uncertainty in the inferred temperature.

Table 7.1 Survey Results from Infrared Observations

No.	Date	Target	cos(μ)	B [G]	T [K]	EqW$_{OH}$ [Å]	σ_{EqW}	N$_{H2}$ [%]	σ_N	Spot Area [arcsec^2]	Umb. Area [arcsec^2]	B-T
1	2001-04-18	NOAA 9429	0.947	2931	5721	0.074	0.012	1.148	0.148	761	157	c
2	2005-03-10	NOAA 10743	0.456	2905	5663	0.118	0.021	1.278	0.246	1145	463	b
3	2005-03-10	NOAA 10741 F	0.947	2411	5987	0.014	0.003	0.349	0.085	680	54	b
4	2005-03-12	NOAA 10743	0.770	3146	5338	0.139	0.029	1.972	0.396	1642	533	d
5	2005-03-12	NOAA 10742	0.954	2823	5441	0.086	0.011	1.279	0.125	1488	277	d
6	2009-07-06	NOAA 11024	0.786	2720	5791	0.066	0.119	0.997	1.551	715	132	c
7	2009-07-07	NOAA 11024	0.728	2753	5905	0.051	0.314	0.817	···	852	112	b
8	2009-07-08	NOAA 11024	0.525	2546	5874	0.074	···	0.939	···	542	55	c
9	2009-07-09	NOAA 11024	0.407	2388	5906	0.045	0.093	0.615	···	422	23	a
10	2009-12-11	NOAA 11034	0.614	2201	6051	0.012	0.063	0.237	0.748	109	1	a
11	2009-12-12	NOAA 11034	0.818	2100	5844	0.014	0.064	0.335	1.377	77	8	a
12	2009-12-15	NOAA 11035	0.865	2722	5775	0.039	0.007	0.730	0.109	728	77	b
13	2009-12-16	NOAA 11035 F	0.830	2587	5737	0.052	0.063	0.869	···	1028	125	b
14	2009-12-17	NOAA 11035 F	0.782	2910	5396	0.115	0.012	1.490	0.162	1173	156	b
15	2009-12-17	NOAA 11034 L	0.754	2991	5539	0.123	0.021	1.612	···	1486	286	b
16	2009-12-18	NOAA 11035 F	0.695	2803	5575	0.119	0.012	1.449	0.142	705	80	c
17	2009-12-18	NOAA 11035 L	0.629	2855	5514	0.124	0.023	1.451	0.267	1482	228	b
18	2010-02-12	NOAA 11046 L	0.865	2448	5940	0.039	0.015	0.719	0.275	97	7	a
19	2010-02-12	NOAA 11045 F	0.552	2302	5924	0.041	0.002	0.637	0.025	230	7	a
20	2010-02-12	NOAA 11046 F	0.848	2676	5937	0.043	0.005	0.767	0.082	985	95	b
21	2010-02-13	NOAA 11046 L	0.834	2496	5824	0.040	0.004	0.730	0.074	187	10	a
22	2010-02-13	NOAA 11046 F	0.859	2581	5886	0.042	0.059	0.755	···	1071	88	b
23	2010-02-14	NOAA 11046 F	0.823	2604	5691	0.055	0.020	0.903	0.299	579	45	b
24	2010-02-14	NOAA 11046 L	0.765	2757	5657	0.007	···	0.162	···	251	0	a
25	2010-02-15	NOAA 11046 F	0.743	2499	5904	···	···	···	···	347	26	a

continued on the next page

Table 7.1, continued.

No.	Date	Target	cos(μ)	B [G]	T [K]	EqW$_{OH}$ [Å]	σ_{EqW}	N$_{H2}$ [%]	σ_N	Spot Area [arcsec^2]	Umb. Area [arcsec^2]	B-T
26	2010-02-15	NOAA 11046 L	0.645	2100	6800	···	···	···	···	63	0	a
27	2010-02-15	NOAA 11046 F	0.538	2140	6273	···	···	···	···	193	2	a
28	2010-02-17	NOAA 11049	0.980	2528	5835	0.018	0.003	0.455	0.074	324	12	a
29	2010-02-18	NOAA 11049 L	0.973	2158	6093	0.012	0.044	0.319	1.184	493	24	a
30	2010-02-19	NOAA 11049 L	0.904	2259	5885	0.035	0.011	0.683	0.178	297	20	a
31	2010-02-19	NOAA 11049 F	0.933	2201	5974	0.016	0.083	0.396	···	257	12	a
32	2010-02-20	NOAA 11049 L	0.833	2571	5870	0.033	0.007	0.642	0.128	304	41	c
33	2010-05-21	NOAA 11072	0.906	2691	5809	0.075	0.043	1.143	0.623	364	80	b
34	2010-05-22	NOAA 11072	0.965	2364	5807	0.031	0.005	0.652	0.104	941	109	b
35	2010-05-23	NOAA 11072	0.964	2625	5581	0.100	0.068	1.425	···	1008	209	b
36	2010-08-29	NOAA 11101	0.968	2748	5734	0.059	0.016	0.999	0.625	1140	247	b
37	2010-08-30	NOAA 11101	0.995	2733	5748	0.062	0.016	1.042	0.253	1267	280	b
38	2010-09-01	NOAA 11101	0.896	2695	5846	0.060	0.185	0.978	···	1281	212	a
39	2010-09-01	NOAA 11102	0.687	2419	5805	0.043	0.006	0.710	0.083	490	37	b
40	2010-09-01	NOAA 11103/4	0.972	2489	5613	0.012	0.002	0.312	0.046	329	6	a
41	2010-09-02	NOAA 11101	0.782	2680	5821	0.069	0.027	1.026	···	965	170	b
42	2010-09-02	NOAA 11102	0.548	2029	6063	0.033	0.014	0.549	0.250	174	7	a
43	2010-09-02	NOAA 11103/4	0.697	2362	5934	0.027	0.007	0.531	0.121	188	5	a
44	2010-09-02	NOAA 11105	0.975	2491	5797	0.040	0.011	0.773	0.189	191	20	b
45	2010-09-04	NOAA 11105	0.802	2408	5788	0.048	0.018	0.815	···	415	49	b
46	2010-09-04	NOAA 11101	0.381	2419	5911	0.048	0.645	0.628	···	437	47	b
47	2010-09-05	NOAA 11101	0.233	2288	5990	0.060	0.060	0.653	···	61	11	a
48	2010-09-12	Pore Group	0.387	2047	5877	0.027	0.007	0.414	0.112	16	0	a
49	2010-09-12	NOAA 11106	0.486	2689	5482	0.071	0.044	0.896	1.089	303	42	b
50	2010-11-28	NOAA 11130 F	0.976	2381	5700	0.029	0.010	0.630	0.203	249	40	a

continued on the next page

Table 7.1, continued.

No.	Date	Target	cos(μ)	B [G]	T [K]	EqW$_{OH}$ [Å]	σ_{EqW}	N$_{H2}$ [%]	σ_N	Spot Area [arcsec^2]	Umb. Area [arcsec^2]	B-T
51	2010-11-30	NOAA 11130 L	0.892	2662	5592	0.052	0.029	0.899	···	1226	180	b
52	2010-11-30	NOAA 11130 F	0.922	2461	5657	0.026	0.006	0.571	0.127	1403	57	b
53	2010-12-01	NOAA 11130 F	0.837	2453	5737		2.909	0.714	···	484	44	a
54	2010-12-01	NOAA 11130 L	0.776	2595	5714	0.054	0.016	0.868	0.226	1079	159	b
55	2010-12-02	NOAA 11130 L	0.624	2704	5777	0.055	0.007	0.805	0.083	610	66	c
56	2010-12-02	NOAA 11131	0.301	2700	5400	0.196	0.029	1.400	0.414	821	152	b
57	2010-12-03	Pore Group	0.788	2504	5469	0.020	0.007	0.446	0.129	190	4	a
58	2010-12-03	NOAA 11131	0.459	2950	5200	0.273	0.030	1.500	···	970	253	d
59	2010-12-03	NOAA 11130	0.426	2572	5811	0.061	0.035	0.761	···	250	34	b
60	2010-12-04	NOAA 11132	0.982	2642	5864	0.037	0.143	0.450	···	147	20	b
61	2010-12-04	NOAA 11131	0.605	3200	5300	0.224	0.032	1.500	1.936	1348	348	d
62	2010-12-05	NOAA 11132	0.974	2302	5778	0.022	0.069	0.150	···	181	5	a
63	2010-12-05	NOAA 11131	0.686	2950	5200	0.186	0.065	2.300	···	2242	524	d
64	2010-12-05	NOAA 11133	0.455	2620	5617	0.096	0.036	1.089	···	526	92	b
65	2010-12-06	NOAA 11133	0.628	2819	5450	0.112	0.022	1.341	0.433	661	150	d
66	2010-12-06	NOAA 11131	0.815	3212	5050	0.160	0.057	2.400	···	2149	652	d

Table 7.2 Survey Results from Visible Observations

No.	Date	Target	$\cos(\mu)$	B [G]	T [K]	Spot Area [arcsec^2]	Umb. Area [arcsec^2]
1	...	...	...	...	...	...	...
2	...	...	...	...	...	...	...
3	...	...	...	...	...	...	...
4	...	...	...	...	...	...	...
5	...	...	...	...	...	...	...
6	2009-07-06	NOAA 11024	0.786	2724	4610	874	102
7	2009-07-07	NOAA 11024	0.728	2550	4890	808	65
8	2009-07-08	NOAA 11024	0.525	2353	5081	503	23
9	2009-07-09	NOAA 11024	0.407	2207	5870	160	4
10	2009-12-11	NOAA 11034	0.614	2121	5917	88	0
11	...	...	...	...	...	...	...
12	2009-12-15	NOAA 11035	0.865	2602	4972	571	23
13	2009-12-16	NOAA 11035 F	0.830	2510	4918	939	50
14	2009-12-17	NOAA 11035 F	0.782	2900	3600	1137	94
15	2009-12-17	NOAA 11035 L	0.754	3100	3600	1604	176
16	2009-12-18	NOAA 11035 F	0.695	2700	3700	619	49
17	2009-12-18	NOAA 11035 L	0.629	2800	4100	1456	123
18	2010-02-12	NOAA 11046 L	0.865	2299	5712	74	0
19	2010-02-12	NOAA 11045 F	0.552	2168	5591	173	0
20	2010-02-12	NOAA 11046 F	0.848	2501	4201	785	37
21	2010-02-13	NOAA 11046 L	0.834	2353	5381	98	0
22	2010-02-13	NOAA 11046 F	0.859	2451	5224	791	17
23	2010-02-14	NOAA 11046 F	0.823	2372	4499	479	14
24	2010-02-14	NOAA 11046 L	0.765	1954	5952	122	0
25	2010-02-15	NOAA 11046 F	0.743	2330	4598	264	5

continued on the next page

Table 7.2, continued.

No.	Date	Target	cos(μ)	B [G]	T [K]	Spot Area [arcsec^2]	Umb. Area [arcsec^2]
26	2010-02-15	NOAA 11046 L	0.645	2190	6064	35	0
27	2010-02-15	NOAA 11048	0.538	1990	6018	182	0
28	2010-02-17	NOAA 11049	0.980	2265	5574	202	0
29	2010-02-18	NOAA 11049 L	0.973	1999	5866	293	0
30	2010-02-19	NOAA 11049 L	0.904	2209	4715	252	5
31	2010-02-19	NOAA 11049 F	0.933	2135	5119	202	2
32	2010-02-20	NOAA 11049 L	0.833	2295	···	193	11
33	2010-05-21	NOAA 11072	0.906	2613	4537	306	41
34	2010-05-22	NOAA 11072	0.965	2235	5306	948	31
35	2010-05-23	NOAA 11072	0.964	2622	5119	1131	115
36	2010-08-29	NOAA 11101	0.968	2774	4657	1200	184
37	2010-08-30	NOAA 11101	0.995	2733	4103	1383	216
38	2010-09-01	NOAA 11101	0.896	2584	4240	1215	147
39	2010-09-01	NOAA 11102	0.687	2321	4675	382	16
40	2010-09-01	NOAA 11103/4	0.972	1990	5856	97	0
41	2010-09-02	NOAA 11101	0.782	2637	5049	891	119
42	2010-09-02	NOAA 11102	0.548	2077	5900	105	0
43	2010-09-02	NOAA 11103/4	0.697	2082	5893	77	0
44	2010-09-02	NOAA 11105	0.975	2107	5533	151	0
45	2010-09-04	NOAA 11105	0.802	2173	5497	311	12
46	2010-09-04	NOAA 11101	0.380	2283	5579	314	25
47	2010-09-05	NOAA 11101	0.233	2139	5900	210	0
48	2010-09-12	Pore Group	0.387	1755	6191	13	0
49	2010-09-12	NOAA 11106	0.486	2457	5330	295	21
50	2010-11-28	NOAA 11130 F	0.976	2252	5543	197	10

continued on the next page

Table 7.2, continued.

No.	Date	Target	cos(μ)	B [G]	T [K]	Spot Area [arcsec^2]	Umb. Area [arcsec^2]
51	2010-11-30	NOAA 11130 L	0.892	2651	4551	1236	84
52	2010-11-30	NOAA 11130 F	0.922	2219	5600	1184	0
53	2010-12-01	NOAA 11130 F	0.837	2103	5412	376	7
54	2010-12-01	NOAA 11130 L	0.776	2490	4813	1040	78
55	...	...	...	...	...	...	...
56	2010-12-02	NOAA 11131	0.301	2400	5500	813	128
57	2010-12-03	Pore Group	0.788	2266	5600	111	0
58	2010-12-03	NOAA 11131	0.459	2400	4500	974	222
59	2010-12-03	NOAA 11130	0.426	2307	5586	298	20
60	2010-12-04	NOAA 11132	0.982	2232	5399	102	4
61	2010-12-04	NOAA 11131	0.606	3000	3900	1329	245
62	2010-12-05	NOAA 11132	0.974	2072	5852	112	0
63	2010-12-05	NOAA 11131	0.686	4000	3123	2209	423
64	2010-12-05	NOAA 11133	0.455	2526	5631	455	50
65	2010-12-06	NOAA 11133	0.628	2692	4629	593	106
66	2010-12-06	NOAA 11131	0.815	3100	3500	2151	544

For spots on the extreme limb, the magnetic field and intensity cannot always be trusted. The Milne-Eddington inversion sometimes fails to retrieve a realistic magnetic field strength. The inclination of the solar surface causes multiple velocity components to combine in a single sight line. This increases the difficulty of fitting the Milne-Eddington model can result in a severe underestimation of the magnetic field strength, especially in the visible where the Zeeman splitting is not very large compared to the line width (eg. 7.58). At the limb, the continuum contrast of sunspots is increased, this effect should be corrected by the model, but the intensity gradients due to limb darkening are larger and may not be properly removed or accounted for in the data. In addition, there are also regions of bright plage at the limb which alter the intensity profile of the sunspot and surrounding quiet-Sun. As a result, observations of sunspots at the extreme solar limb cannot be trusted.

7.2 Observed Relations

7.2.1 Relations from Infrared Data

For most of the spots the maximum field strength and the maximum OH equivalent width occur at the lowest temperature as we would expect if the strength of the OH lines is strongly linked to atmospheric temperature, and temperature is in turn controlled by magnetic suppression of convective energy transport. There are a few interesting exceptions where the strongest magnetic fields correspond to higher intensity positions on the umbral/penumbral boundary, NOAA 11035 which was flaring, and NOAA 11072 during its rapid emergence.

The sunspot thermal-magnetic relations show some interesting behavior related to the size and strength of the sunspot magnetic field. At the smallest, weakest end, the sample contains many examples of pores, both isolated and associated with larger sunspots in an active region. Observation 31 (7.33) shows a typical observation of a pore group. Almost universally, pores display a very linear relation between B^2 and T, from the boundary with the quiet-Sun all the way to the core of the pore, with a slope of about -2300 G^2/K. Their magnetic fields are too weak to maintain temperatures which are cool enough to form

significant H_2, which is in agreement with the low measured EqW_{OH} that is seen in pores. Scatter in the OH measurement at high temperatures prohibits inference of H_2 fractions below about 0.2%.

A majority of the sunspots in the sample display a three segment curve in their thermal-magnetic relation. The slope is vertical at the outer edge of the sunspot penumbra where the magnetic field is increasing rapidly, but where the magnetic-filling factor is still small, resulting in a large range of values for the magnetic field strength, whereas the intensity across the penumbra is relatively constant. Around the umbral-penumbral boundary the temperature gradient increases and the relation flattens into a slope similar to the ones seen in pores. Inside the umbral boundary the relation departs from this linear slope and the relation is often seen to increase in slope at the lowest temperatures. Very clear examples of this behavior can be seen in spots of a variety of sizes, for example Observations 6 (7.8), 14(7.16), 15(7.17), 22(7.24), and 36(7.38).

A small handful of the observations show a clear, isothermal increase in magnetic field strength at the lowest temperatures in the umbra, these cases are NOAA 9429 (7.3), NOAA 11130 (7.57), and NOAA 11049 (7.34). Somewhat more ambiguous cases for isothermal increases are shown by the following polarity spot in NOAA 11035 (7.18) and NOAA 11101 during its later observations (Fig. 7.40 and 7.43). Although they have neither the largest field strengths, nor the strongest OH lines, these cases are particularly interesting and important because they clearly violate the assumed form of the thermal-magnetic relation given in Equation 2.3. Most interestingly, all of these spots were all relatively small and in a state of slow decay.

The largest sunspots in the sample show slightly different behavior. While they still display a clear 3-segment curve, the temperature at which the umbral increase in slope occurs is much lower than for smaller spots. The linear behavior through the hot umbra has a smaller slope with an elevated magnetic field relative to the relations shown by smaller spots. NOAA 10743 (7.6) and 11131 (7.68) are the clearest cases for this behavior, although NOAA 11101 (7.43) shows a similar shallow slope through the umbral-penumbral boundary.

It is important to point out that each one of these sunspots was a long-lived α-spot with a prominent, highly ordered penumbra. While these large spots should have very significant H_2 populations, we are at a loss to explain how they maintain an altered equilibrium state and why the increase in the slope of the relation occurs so deep in the umbra.

Although the sunspot groups in the survey are often composed of pores and spots of different sizes, it is surprising that they often show a cohesive B^2-T relation given the vastly different relations shown by very large sunspots. However, there is evidence that the relation between B^2 and T may not follow the same path between different sunspots in the same observation (for example Observations 12 and 20), and even in different regions within the same umbra (Observations 41, 51). While it is possible that errors from the correction of stray light affect the appearance of the relation, it is unlikely to affect adjacent regions of the sunspot umbra very differently. To investigate this problem further it would be necessary to identify different umbral regions and produce individual B^2-T relations. While this is somewhat beyond the scope of this research, it may be interesting to consider in future analysis of the data.

7.2.2 Relations from Visible Data

The results from the visible observations are similar to those seen for the infrared, though the magnitude of the changes in the magnetic field seem smaller, and the relations are overall more linear for most of the spots observed. In fact, at 6302 Å the lines are formed 150 km higher in the atmosphere than the infrared lines, where it is cooler, and there should be more H_2. Naively one would think that the behavior would be even more accentuated, but obviously there is more to it. One caution about the visible data is that the Zeeman sensitivity of the lines is reduced, and therefore the magnetic field ambiguity covers a larger regime in the sunspot, and penumbral fields cannot be trusted. This demonstrates the importance of a good stray light determination and correction so it is possible to use the visible data quantitatively. With greater uncertainty in the temperature and no good molecular measure like OH in the infrared, it is not possible to establish the effect of H_2 from

the visible observations alone. While the results for the 6302 Å observations are provided here, they are not discussed further.

7.3 Classification

Although all of the sunspots are distinctly different, even from day to day for the same spot, based on the discussion of the results from the infrared observations above, they can be placed into 4 basic categories based on the behavior of the thermal-magnetic relation and the strength of the OH lines:

(**a**) Pores that display linear behavior in B^2 vs. T, with maximum umbral field strengths below 2500 G, minimum umbral continuum temperatures higher than 6000 K, and an inferred H_2 fraction of below 0.75%.

(**b**) Sunspots that show non-linear behavior manifesting as an increase above the B^2 vs. T relation in the umbra at a temperature of about 6000 K or higher. Maximum umbral fields are greater than 2500 K, minimum umbral temperatures are below 5800 K, and the H_2 fraction is greater than about 0.75%

(**c**) Small to mid-sized sunspots that show an isothermal increase in the magnetic field around 5800 K, with maximum umbral magnetic field strengths between 2600-2700 G and H_2 fractions just short of 1%.

(**d**) Large spots that show an increase in the slope of the B^2 vs. T relation at much lower temperature in the umbra, around 5500 K or lower, and H_2 fractions of 1.5% and higher.

The classification for each observation is summarized in the results in Table 7.1. These properties are visualized in Figures 7.1 and 7.2, where I have plotted the maximum umbral magnetic field strength vs. the minimum umbral temperature and maximum umbral H_2 fraction vs. minimum umbral temperature. In these plots the symbol for each sunspot observation is the identifying number, and the assigned classification of the B^2-T curve is indicated by the color. As we would expect, the smallest sunspots and pores (a, gold) populate the high temperature, low magnetic field strength, low H_2 end of the relation, while

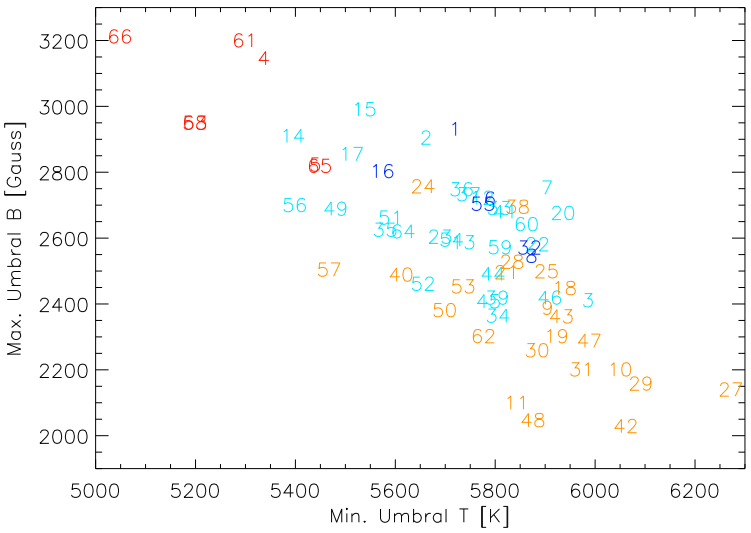

Figure 7.1 The maximum umbral field strength vs. minimum umbral temperature from the infrared data for every spot in the sample.

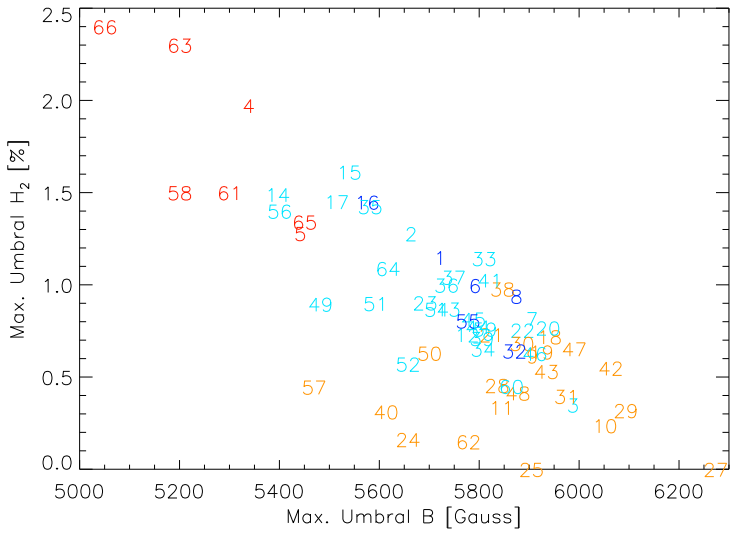

Figure 7.2 The maximum inferred umbral H_2 fraction vs. minimum umbral temperature for every spot in the sample.

the largest spots (d, red) fall at the opposite end. Mid-field and temperature objects (b, light blue) fall between them, but it is interesting to note that the spots with strong upturns (c, dark blue) fall mostly toward lower magnetic field strengths, higher temperatures, and lower H_2 fractions at the boundary between the (a) and (b) cases, indicating that they might be transitioning from one equilibrium state to another.

7.4 Individual Description of Sunspots

Here I discuss the individual sunspot histories to provide context for the observations and briefly discuss the results seen in the figures for each sunspot on an individual basis.

NOAA 9429 was an α-spot that appeared fully formed around the east limb on 13 Apr 2001 and disappeared around the west limb on 25 Apr. The Observation 1 (7.3) was taken on 18 Apr just before it passed disk center, following this it appeared to slowly decay.

This is a very early observation taken with the HSG prior to the commissioning of FIRS. The detector used for this observation appears to have a non-linear response which was not calibrated or corrected in the data. We have adjusted the intensity by hand to match the FIRS observations, and the intensity-dependent quantities, (temperature, OH equivalent with, and N(H_2)/N(Total)) should only be considered in a qualitative way. However, the umbral magnetic field strength derived from Stokes profile inversion is independent from intensity and still a reliable quantity. Despite the uncertainty in the temperature scaling, this sunspot shows a striking isothermal increase in magnetic field up to strengths of 3000 G at the lowest temperatures. This upturn can be ascribed to strong magnetic features visible in the center of the sunspot umbra seen in the magnetic field map.

NOAA 10741 starts as a β-region on the limb on 4 Mar 2005 with a large leading polarity spot. It appears to decay steadily until 8 Mar. On 9 Mar at the central meridian it developed a large following polarity pore group. The observation in Figure 7.5, taken

on the tenth is of a part of this group. Another spot of comparable size emerged next to it on 13 Mar as it traveled out of view around the west limb on 15 Mar.

NOAA 10742 was a β-spot group that appeared around the east limb on 7 Mar 2005. It increased in size as it reached the central meridian (obs. 4 of the main spot in the following group) and the following polarity decays as the leading polarity becomes concentrated in a main lead spot which disappears around the limb on 18 Mar.

NOAA 10743 was a large α-spot which appeared around the limb on 9 Mar. Observations 2 and 5 were taken shortly after on the 10 and 12 Mar. It evolved fairly quiescently as it reached the central meridian, but developed some rotationally leading pores and appeared to decay as it reached the west limb on 22 Mar.

NOAA 11024 emerged very rapidly on 3 July 2009, 15° east of the central meridian. Observations 6, 7, 8, and 9 were taken of the leading polarity spot in the β region as it passed disk center and slowly evolved as it traveled toward the west limb where it disappeared on the 12th. The following polarity spots evolves more rapidly.

NOAA 11034 was a tiny β pore group that emerged around the east limb on 10 Dec 2009. It was relatively quiescent during Obs. 10 and 11 taken on the 11th and 12th, but on the 13th disintegrated into a large pore group, decayed further on 14 Dec, and was gone the next day.

NOAA 11035 was a complex $\beta\gamma$-region that emerged rapidly on 14 Dec 2009, 10° from the central meridian. Disorganized pore groups coalesced into large leading and following polarity sunspots. The Observation 15 (7.17) of the leading spot was taken on 17 Dec when the spots were largest. After producing several C-class flares on this day the region began to decay, but the spots remained large as they traveled around the west limb on 22 Dec.

The observation of NOAA 11035 shows exceptionally high magnetic field strengths in excess of 3500 G in small features located off of the main umbra (can be clearly

identified in panel (c)), which are probably a manifestation of flare activity, while the maximum umbral magnetic field strength is more typical, 3000 G. The thermal-magnetic relation for this spot does not display a sharp upturn, the relation gradually increases in slope in the main umbra, starting at 5800 K. The OH equivalent width is in good agreement with the model and reaches high values corresponding to a H_2 fraction by number of about 2%.

NOAA 11045 was a large and rapidly evolving β-γ region that emerged 20° east of the central meridian on 6 Feb. The region was strongest on the 6 and 7 Feb, but gradually spread out and lost spots in the polarity inversion region. Obs. 19 is of the following polarity spot group before it disappeared around the limb on 14 Feb.

NOAA 11046 appeared around the east limb as a small β-region on 7 Feb 2010. It became more developed around 11 or 12 Feb with a larger following polarity group and a smaller leading polarity spot or umbra. Both the lead and following components of the region were observed as it passed disk center on the 12th (leading:18, 21, 24, 25; following:20, 22, 23, 26) and decayed over the next several days. The spots seemed to have disappeared almost entirely as they reached the limb on 18 Feb.

Although this sunspot group is composed of three umbrae with partial penumbrae, it displays a cohesive thermal-magnetic relation which shows a clear increase in slope in the umbra, at about 6100 K. A maximum umbral field strength of 2500 G is reached in the right-most umbra. The OH measurement is in very good agreement with the theoretical relation, and implies a H_2 fraction of just under 1%.

NOAA 11048 was a decaying β-γ region whose remains appeared around the limb on 14 Feb 2010. Obs. 27 was taken on its second day from the limb and the region continued to decay until it was nearly gone on 17 Feb.

NOAA 11049 was a small β-region that emerged on 16 Feb 2010 about 15° east of the central meridian. Obs. 28, 29, 30, and 32 follow the leading polarity spot during

the region's relatively quiescent evolution stage until the 20th, while obs. 31 is the only observation of the following polarity. Although the region continued to evolve, the lead and following pores remained roughly equivalent in size until 21 Feb when the lead pore became stronger and the following region diminished before the region disappeared around the west limb on 24 Feb.

This pore group shows a very linear thermal-magnetic relation and low magnetic field strengths, a 2200 G maximum is reached in the left-most pore. The OH lines are small or undetected, corresponding to 0.5% H_2 fraction at most.

NOAA 11072 was a β region that appeared on 20 May 2010 30° east of disk center, quickly developed a medium-sized leading polarity spot. Obs. 32, 33, 34 follow its early evolution on the 21st-23rd. Following 23 May umbral size decreased, and the following polarity spots disappeared. The region seemed to decay as it headed toward the limb where it disappeared from view on 28 May.

NOAA 11101 appeared around the east solar limb as a fully formed α-spot with a penumbra on 25 Aug 2010. It did not undergo significant evolution as it transited the disk and disappeared around the west limb on 5 September. Observations 36, 37, 38, 41, 46, and 47 follow its passage across the central meridian and up to the limb from Aug 29-Sep 5.

The spot shows a large scatter in the thermal-magnetic relation from fine magnetic structure which is evident near the umbra-penumbra boundary. However, it shows a higher density of points at 5800 K rising above the main B^2-T track. Agreement with the theoretical OH relation is excellent and we can infer a maximum H_2 fraction of 1%.

NOAA 11102 was a small β-group that emerged as a pore on 28 Aug 2010 5° east of the central meridian. It was largest on 31 Aug-1 Sep, and decayed and disappeared as it reached the limb on 4 Sep. Obs. 39 and 42 were taken during this decay phase. Unusually, the largest spot in the region was of following polarity.

NOAA 11103/4 were small, colocated β pore groups which emerged on 1 Sep 2010, 30° west of central meridian (obs. 40), Obs. 43 on the second day. The regions had almost entirely decayed, leaving a tiny pore and surrounding plage, as it reached the limb on 4 Sep.

NOAA 11105 emerged on 2 Sep 2010, 10° west of the central meridian (obs. 44). The β region evolved rapidly over the 3 following days (obs. 45), on the 4th day it produced a significantly larger lead spot and then disappeared from view around the limb on 7 Sep.

Unnamed Pore Group was a tiny pore group which emerged and disappeared on Sep 12, 2010 when Obs. 48 was taken.

NOAA 11106 emerged around the east limb on 11 Sep 2010, and was a β-spot of leading polarity. Obs. 49 was taken soon after on 12 Sep. The spot reached a maximum size on 14 Sep. and began to decay after the passage of disk center. The following spot disappeared on 19 Sep, and the lead spot had entirely disappeared just before it reached the west limb, leaving a plage.

NOAA 11130 was a rapidly evolving, medium sized β-region with a single lead spot and weaker following group which emerged near disk center on 27 Nov 2010. Obs. 50 was taken during this growth stage on the 28th. The region reached a maximum size on the 30th, the lead spot having developed a full penumbra (Obs. 51), and following spots with partial penumbrae (Obs. 52). It slowly decayed over the next 4 days as it reached the limb (leading Obs. 54, 55, following Obs. 53). Observation 59 was taken of the leading polarity sunspot on 2 Dec during this decay phase when the spot group with following polarity had diminished to pores.

While NOAA 11130 does not have a particularly strong magnetic field (2600 G) or strong OH lines, this spot shows a clear case of isothermal intensification of the umbral magnetic field. The sharp upturn is approximately 500 G, starting at 5800 K. The

OH line equivalent width is in very good agreement with the theoretical prediction, indicating 1% H_2.

NOAA 11131 was a large α-spot that appeared fully formed from around the east limb on 1 Dec 2010, remained roughly the same size and shape, developing a few transient pores, until it disappeared around the west limb on 13 Dec. Observations 56, 58, 61, 63, and 66 followed the beginning of its passage across the disk from the 2nd to the 6th.

NOAA 11131 is the largest spot in the sample with a maximum umbral field of 3100 G. The thermal-magnetic relation follows a significantly different path from the other spots, large magnetic field strengths are reached early in the penumbra followed by a shallow slope through a majority of the umbra, and an abrupt increase in slope at a temperature of 5300 K. The OH lines are quite strong, reaching a value corresponding to 2.5% H_2, but the equivalent width decreases in slope in cool, high magnetic strength regions. This may be a sign the reversal of OH equivalent width occurs at a lower temperature than predicted from the models.

Unnamed Pore Group emerged and disappeared on Dec 3 when Obs. 57 was taken.

NOAA 11132 was a small β-pore group, that emerged and disappeared rapidly near disk center on 4-5 Dec 2010. Obs. 60 and 62 were on these two days. The lead and following components spread out between days but did not change size appreciably.

NOAA 11133 was a smaller α-spot which emerged around the limb on 4 Dec 2010. It appeared to shrink gradually after it transited the central meridian and passed out of view around the west limb on 15 Dec. Observations 64 and 65 were taken during a rather quiescent time just after it became visible near the east limb.

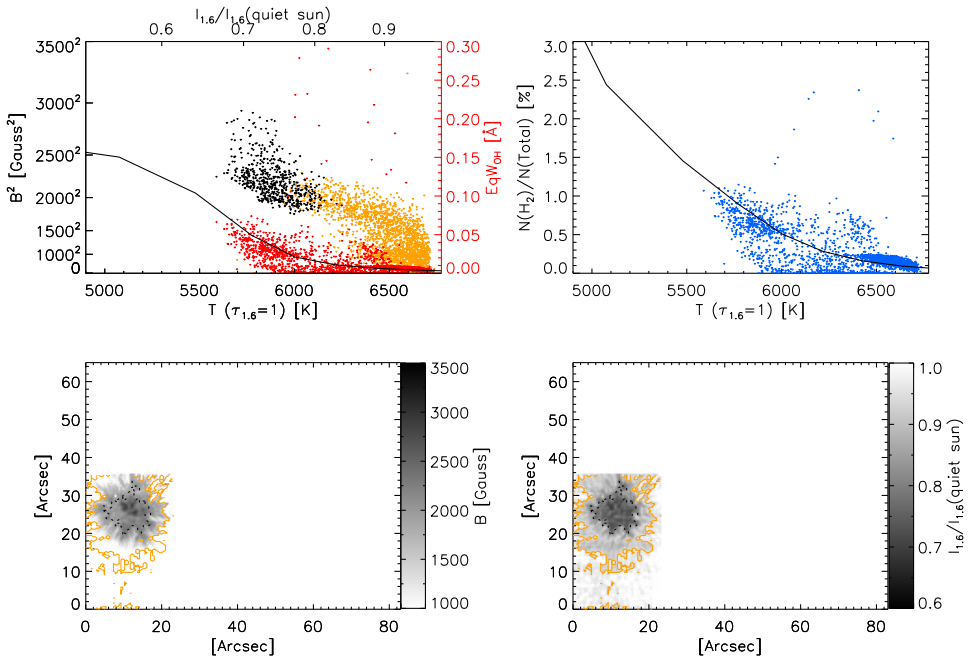

Figure 7.3 Obs. (1): NOAA 9429 observed on 2001-04-18.

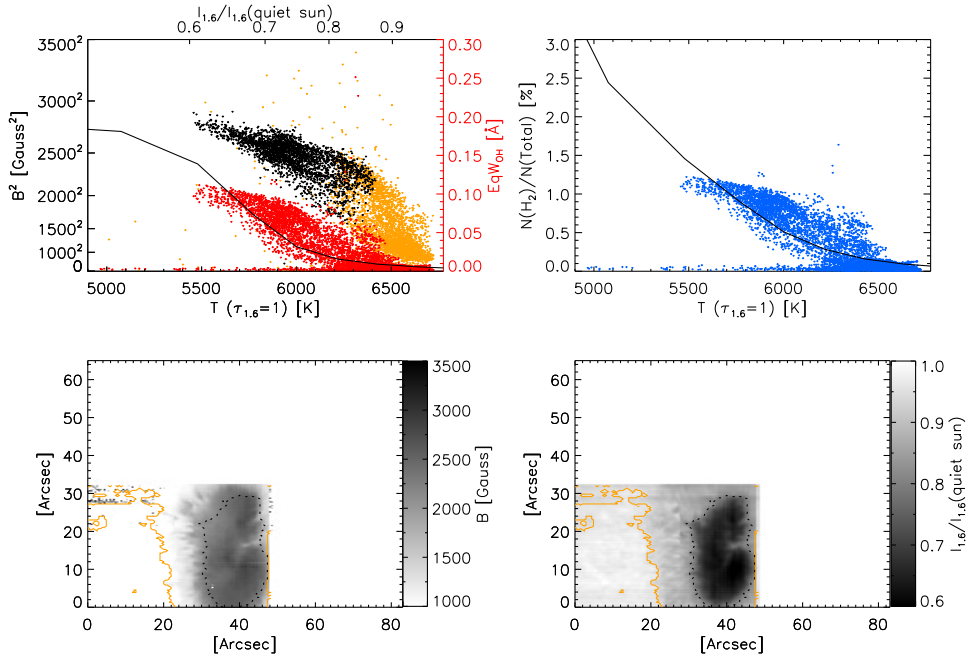

Figure 7.4 Obs. (2): NOAA 10743 observed on 2005-03-10.

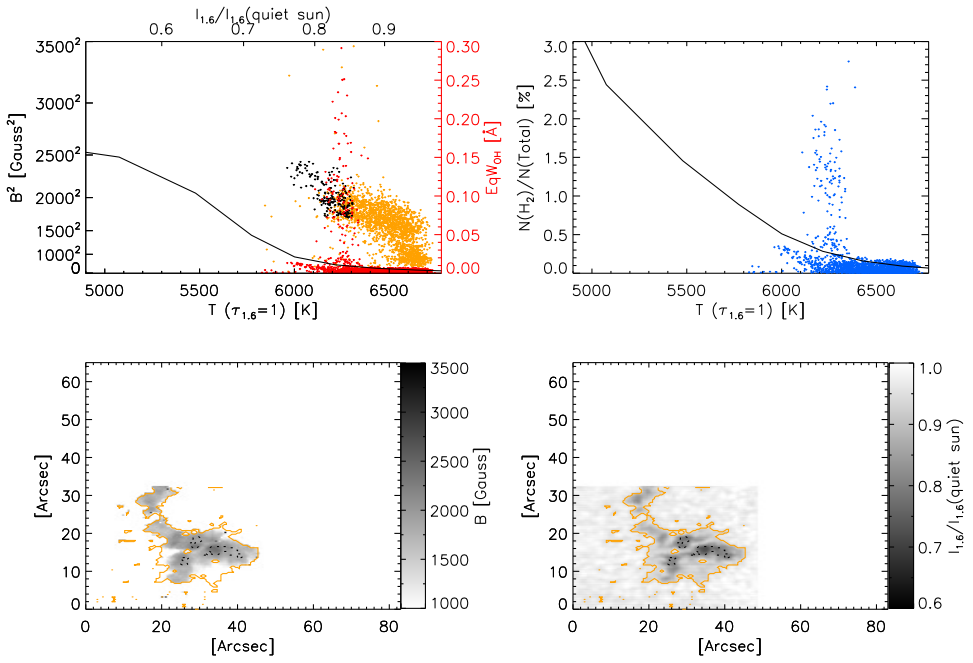

Figure 7.5 Obs. (3): Pore group in NOAA 10741 observed on 2005-03-10.

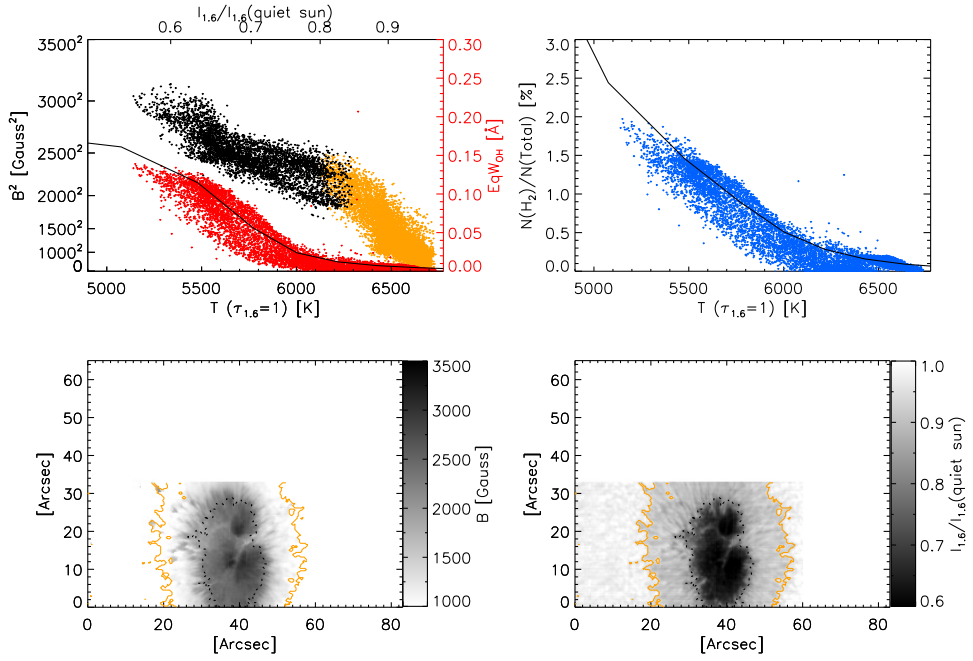

Figure 7.6 Obs. (4): NOAA 10743 observed on 2005-03-12.

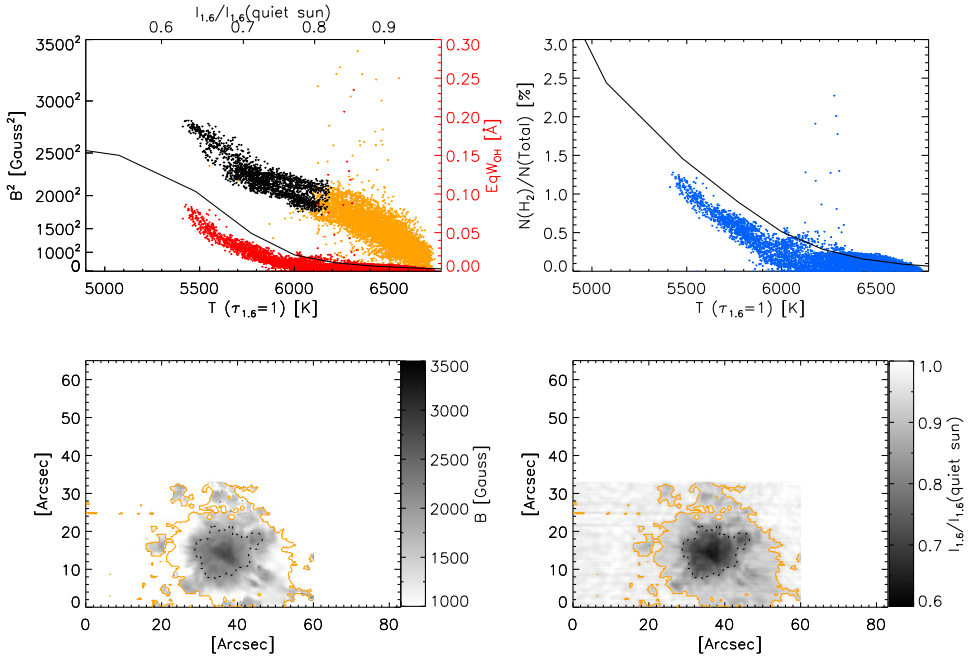

Figure 7.7 Obs. (5): NOAA 10742 observed on 2005-03-12.

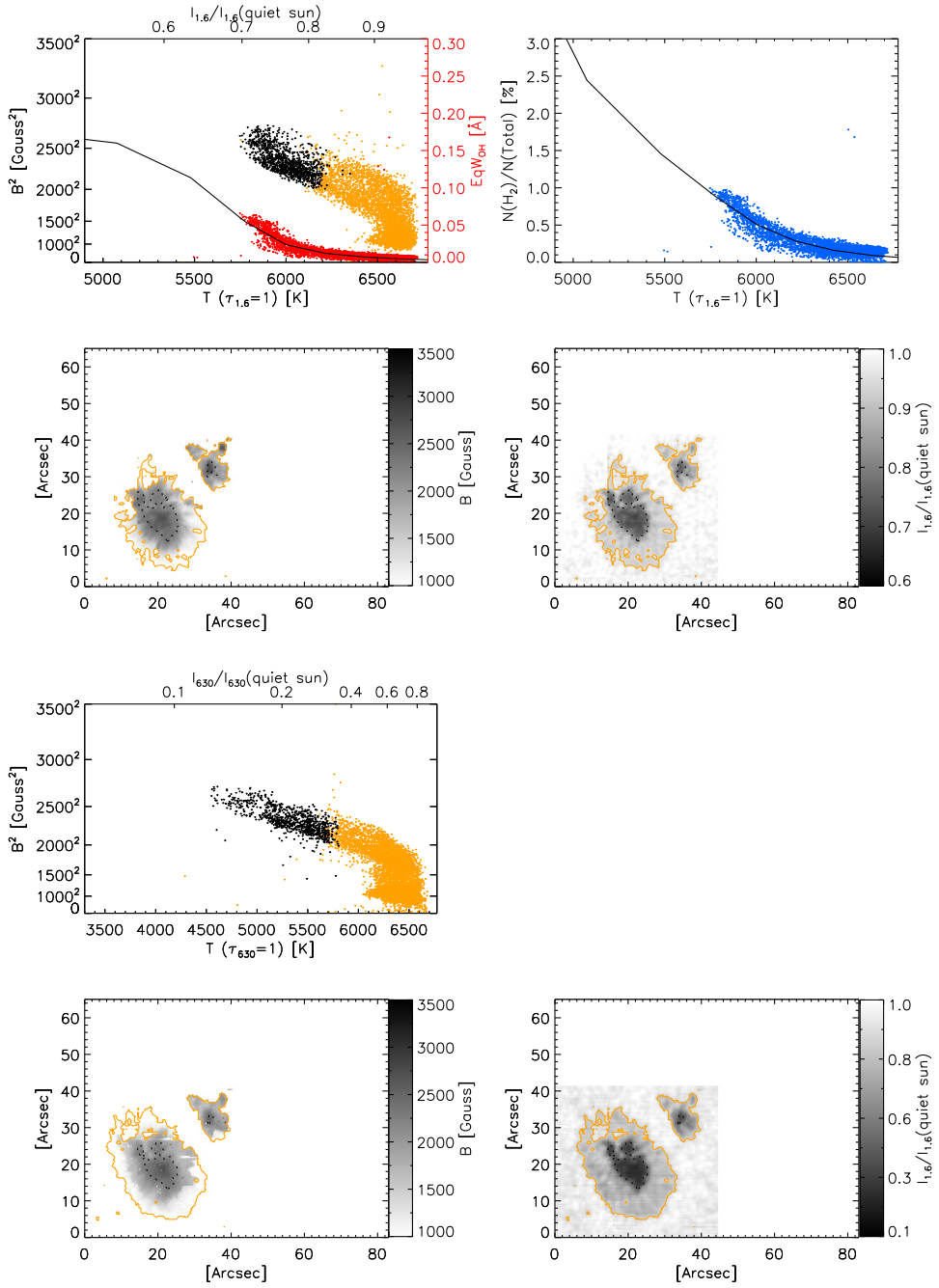

Figure 7.8 Obs. (6): NOAA 11024 observed on 2009-07-06.

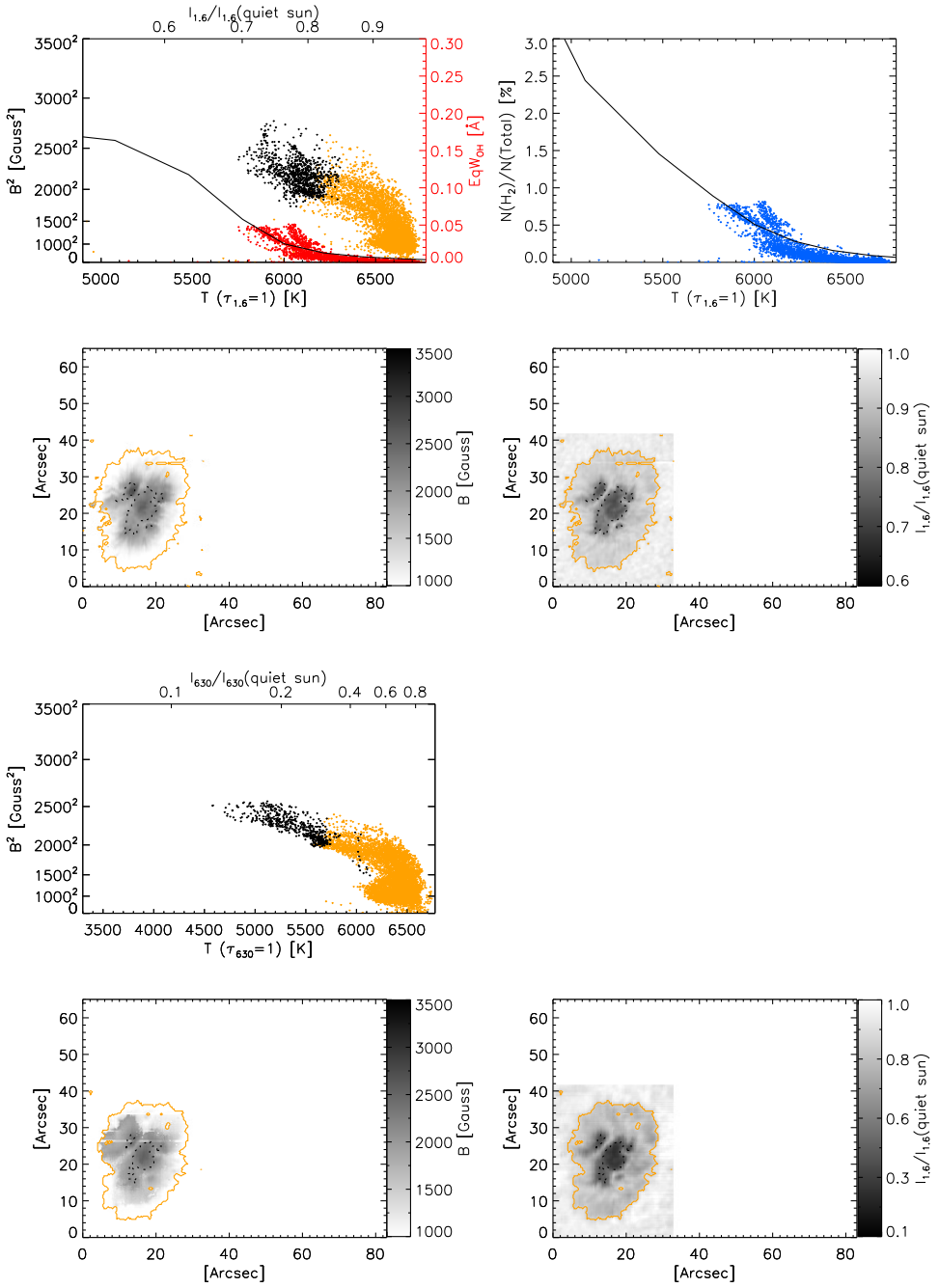

Figure 7.9 Obs. (7): NOAA 11024 observed on 2009-07-07.

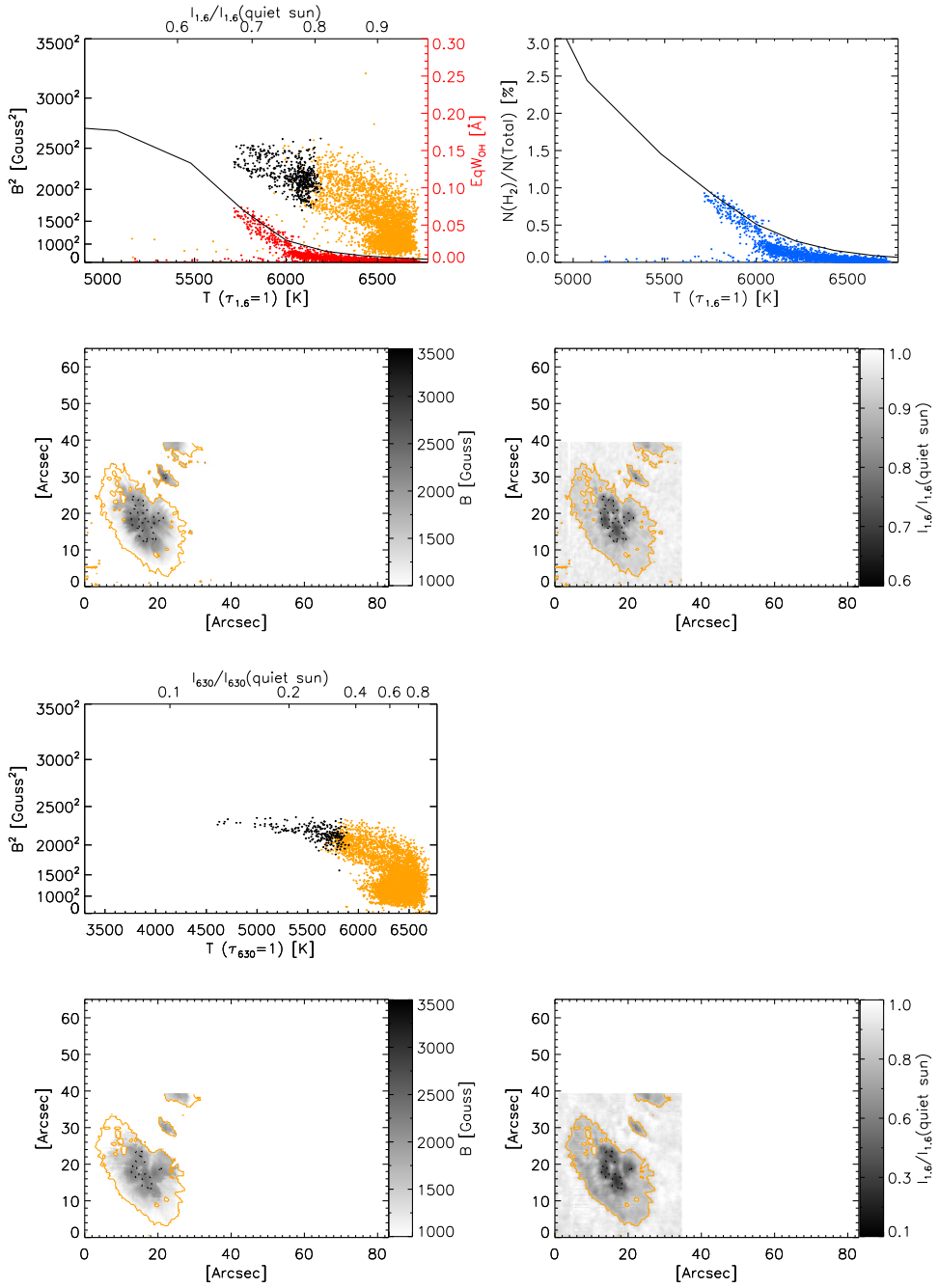

Figure 7.10 Obs. (8): NOAA 11024 observed on 2009-07-08.

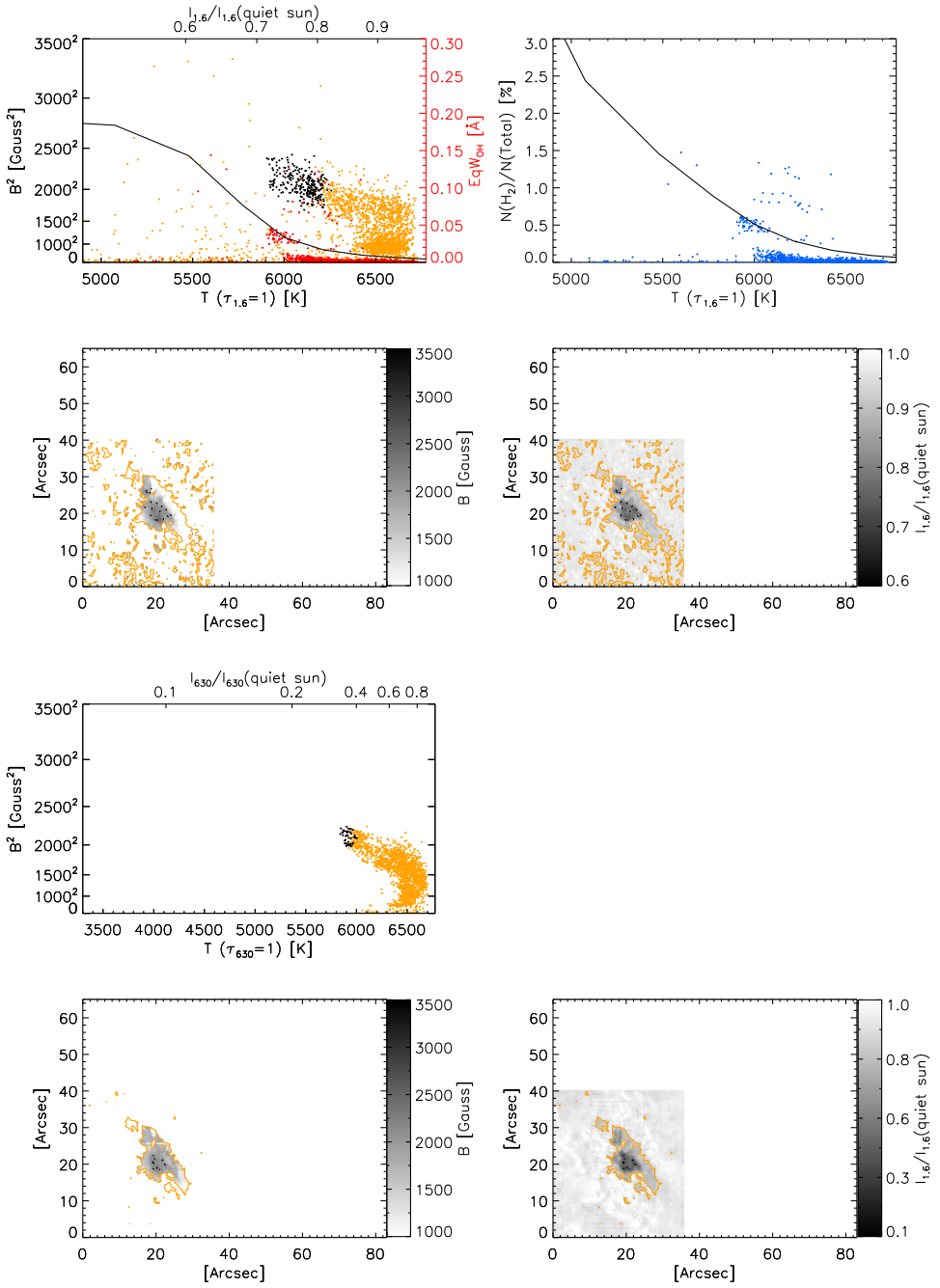

Figure 7.11 Obs. (9): NOAA 11024 observed on 2009-07-09.

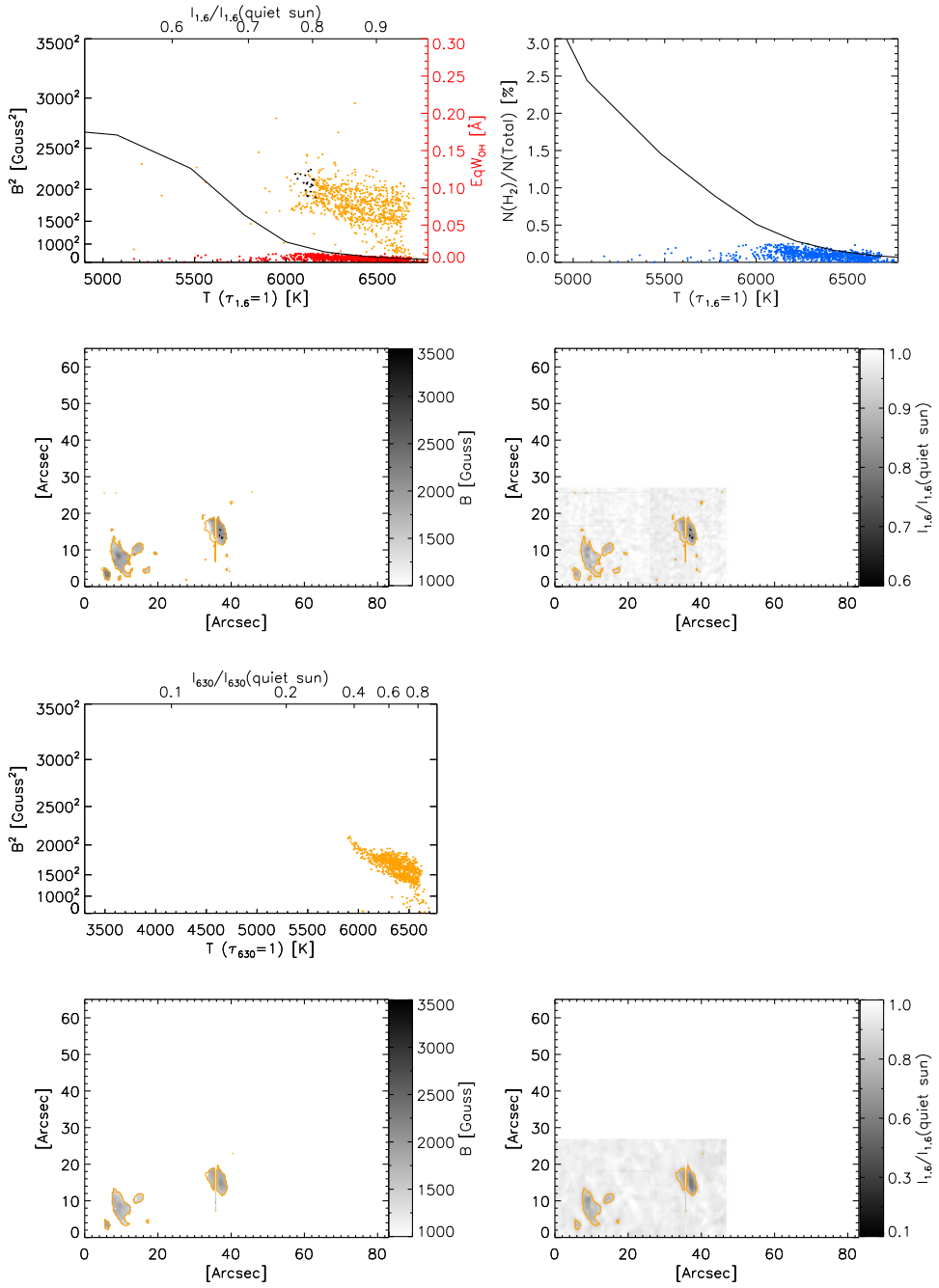

Figure 7.12 Obs. (10): NOAA 11034 observed on 2009-12-11.

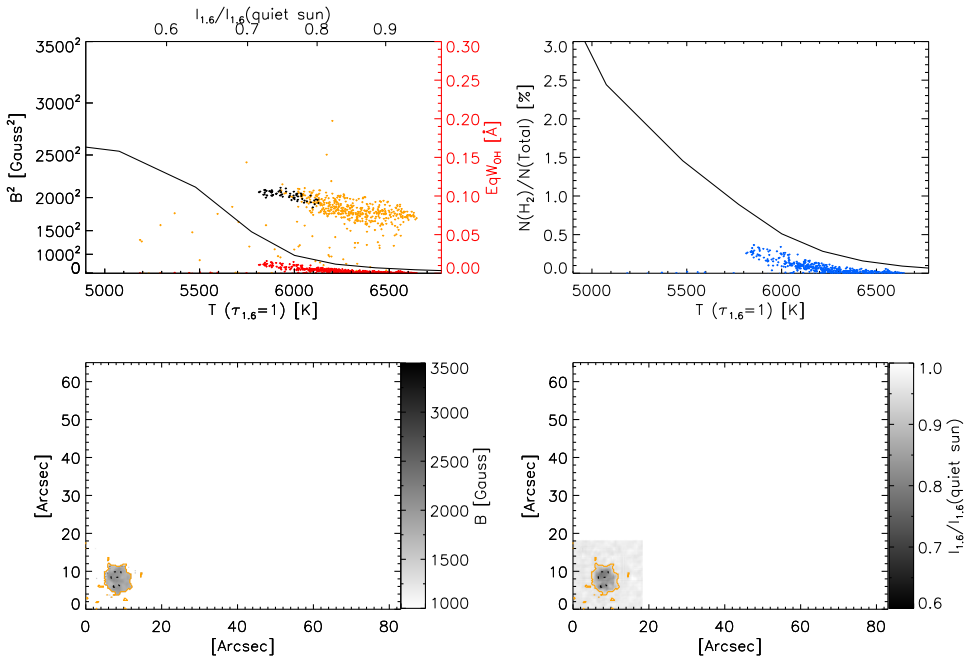

Figure 7.13 Obs. (11): NOAA 11034 observed on 2009-12-12.

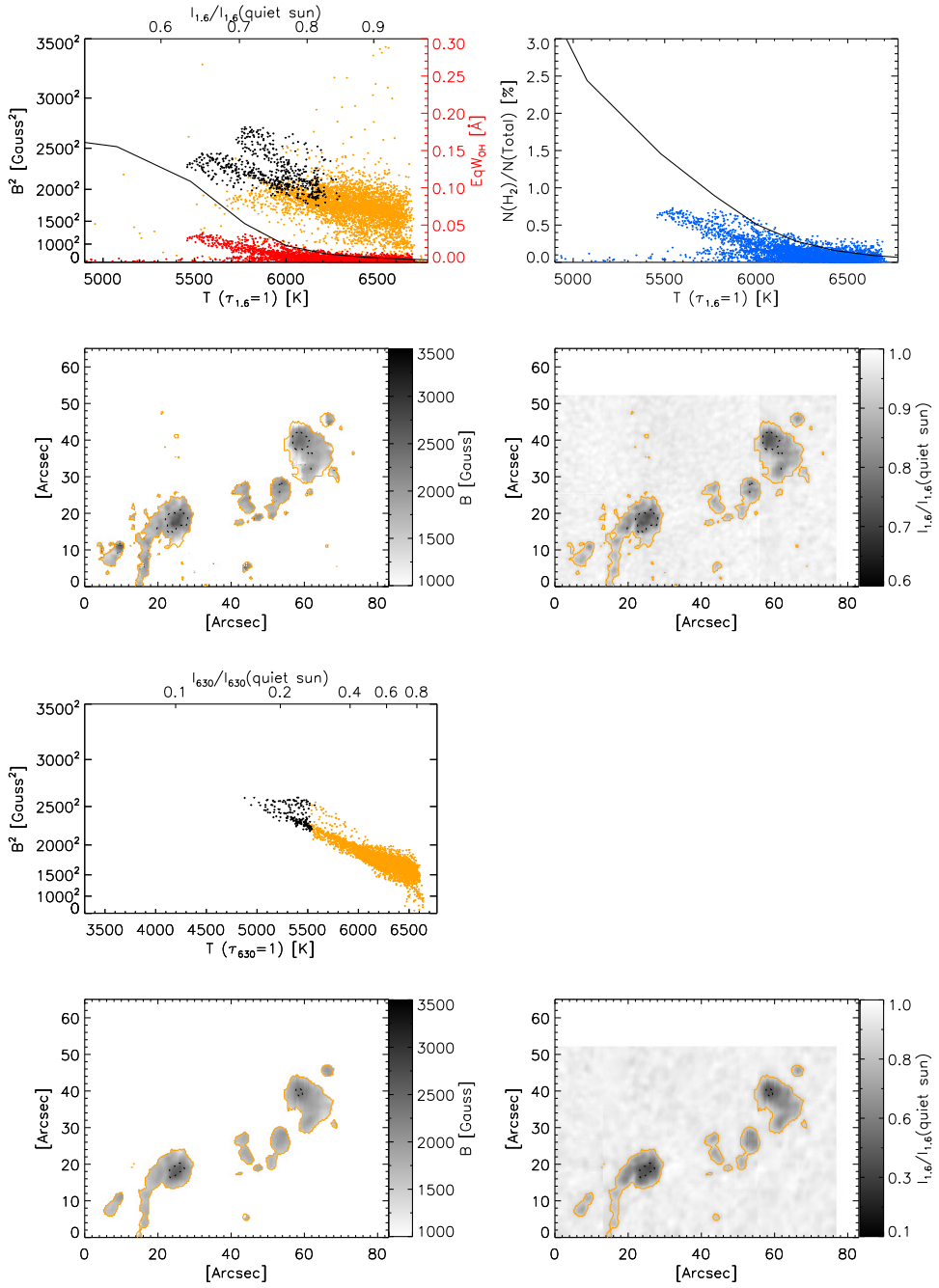

Figure 7.14 Obs. (12): NOAA 11035 observed on 2009-12-15.

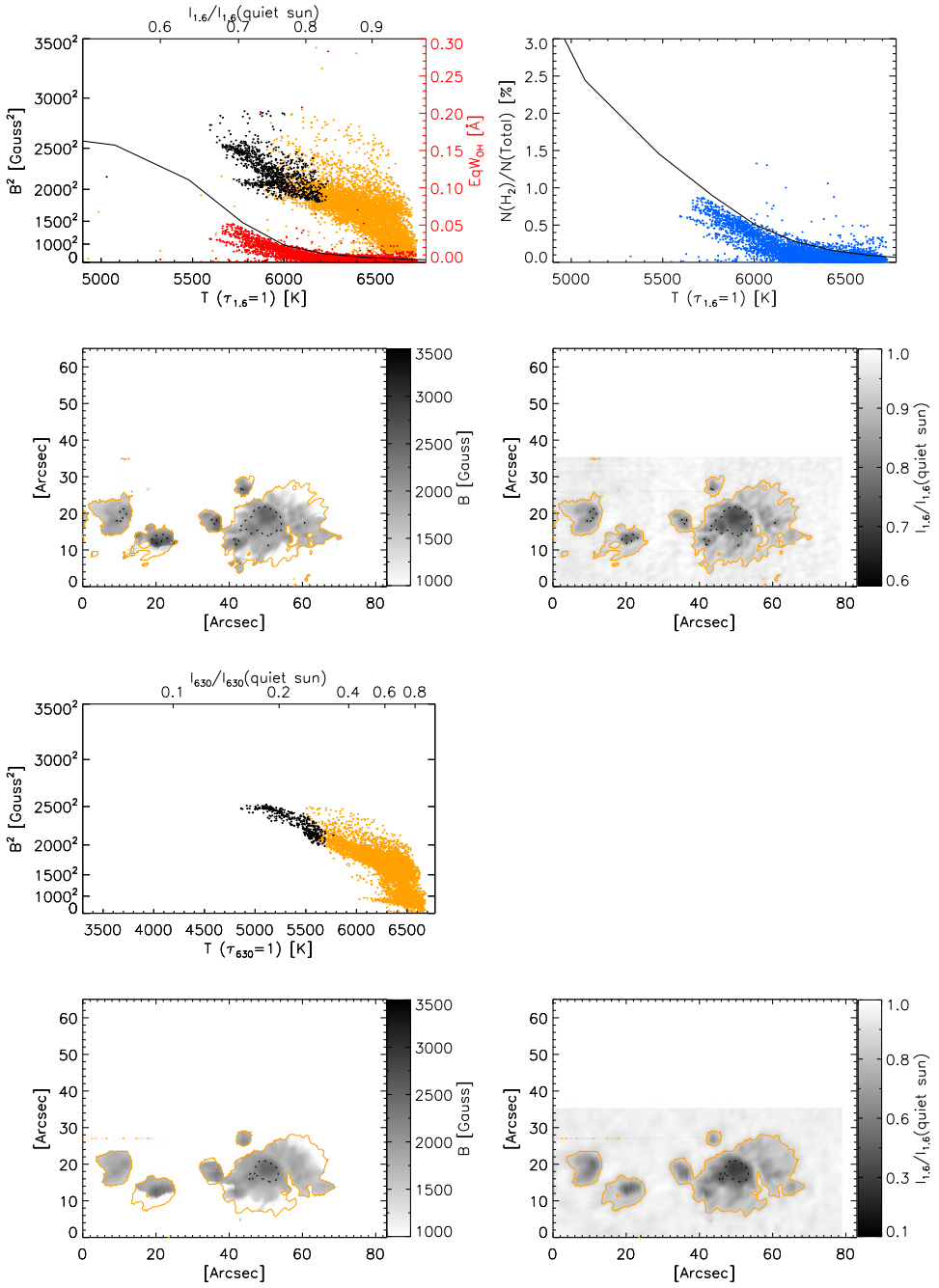

Figure 7.15 Obs. (13): The following polarity spot in NOAA 11035 observed on 2009-12-16.

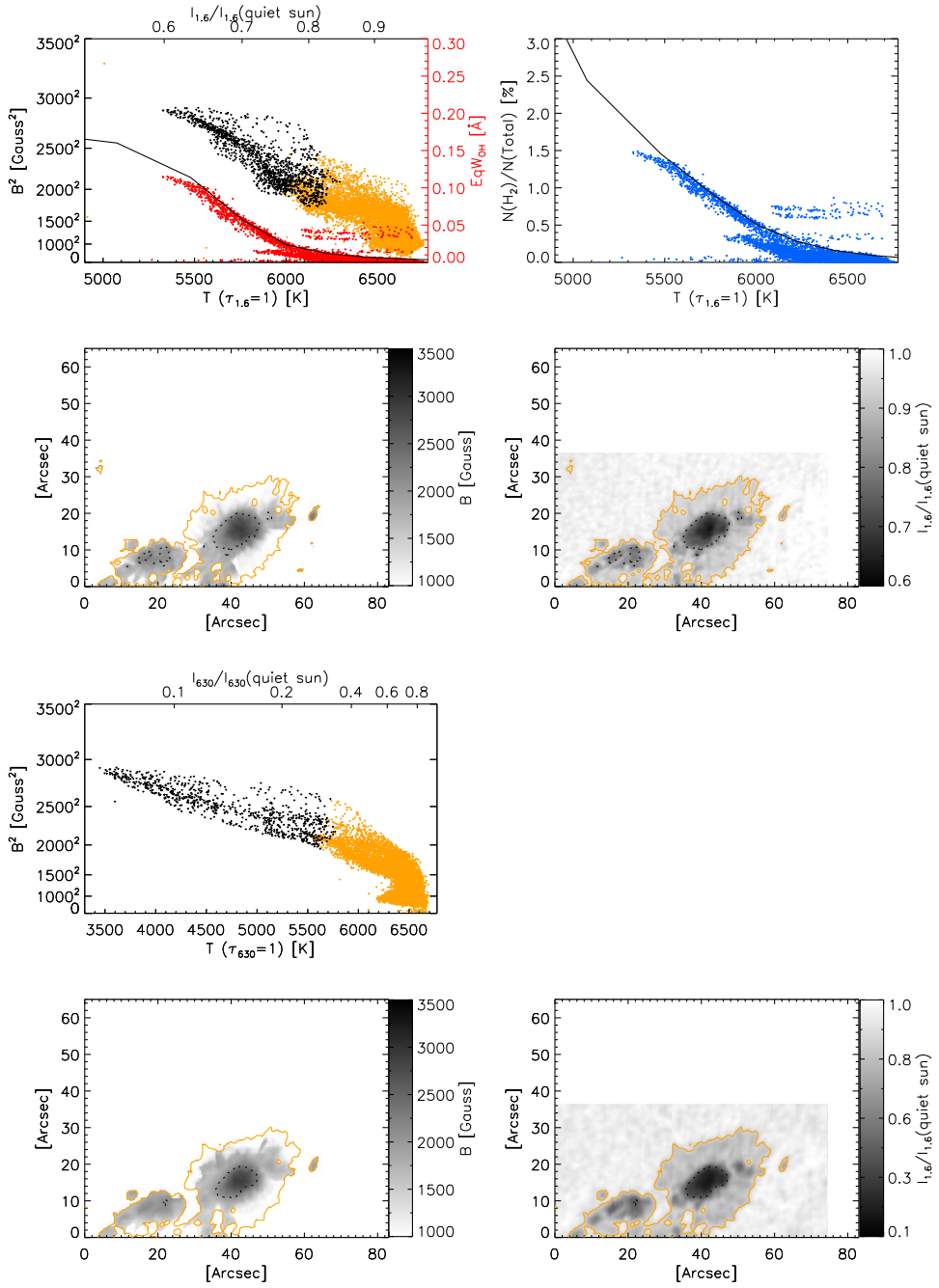

Figure 7.16 Obs. (14): The following polarity spot in NOAA 11035 observed on 2009-12-17.

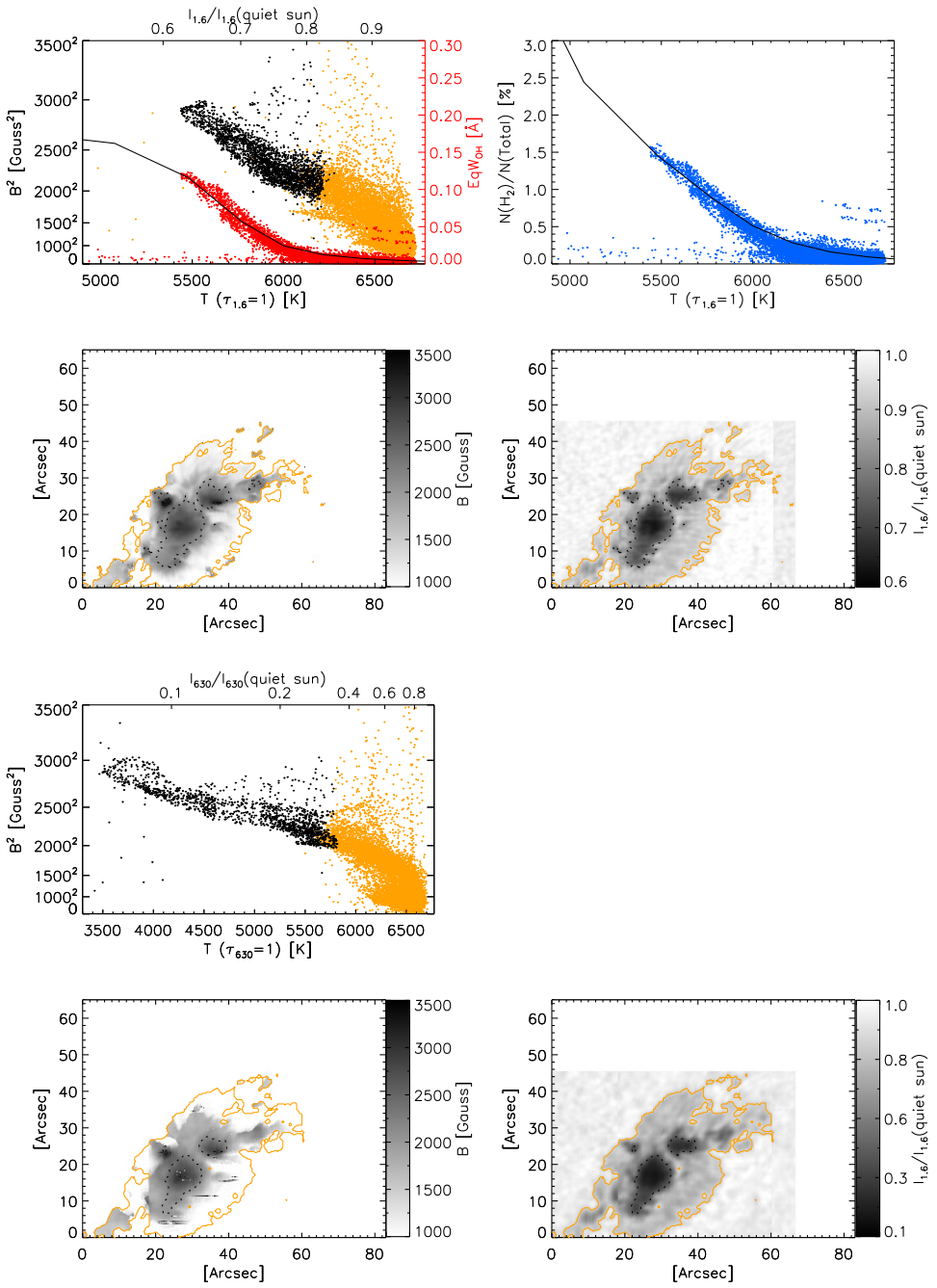

Figure 7.17 Obs. (15): The leading polarity spot in NOAA 11035 observed on 2009-12-17.

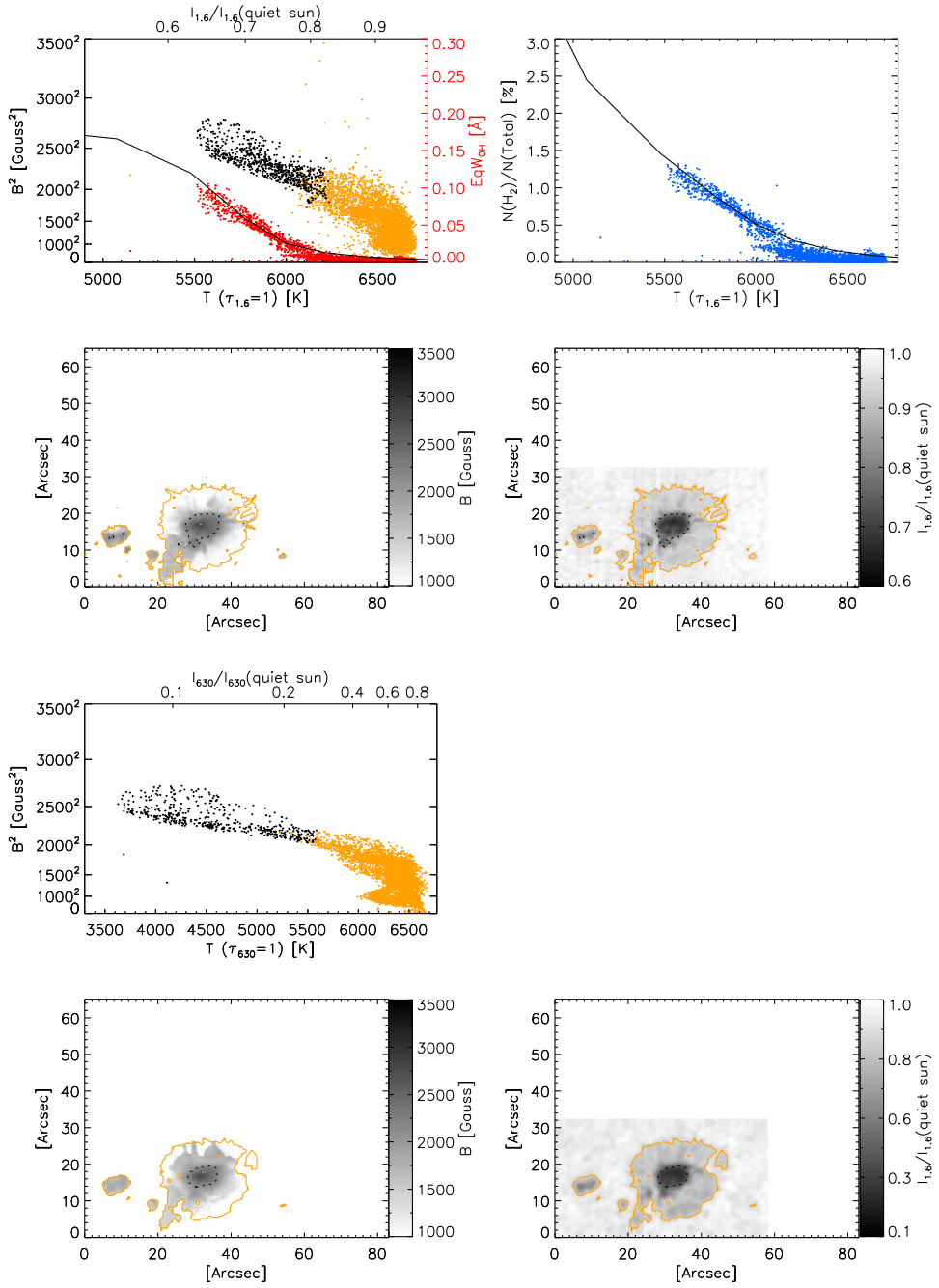

Figure 7.18 Obs. (16): The following polarity spot in NOAA 11035 observed on 2009-12-18.

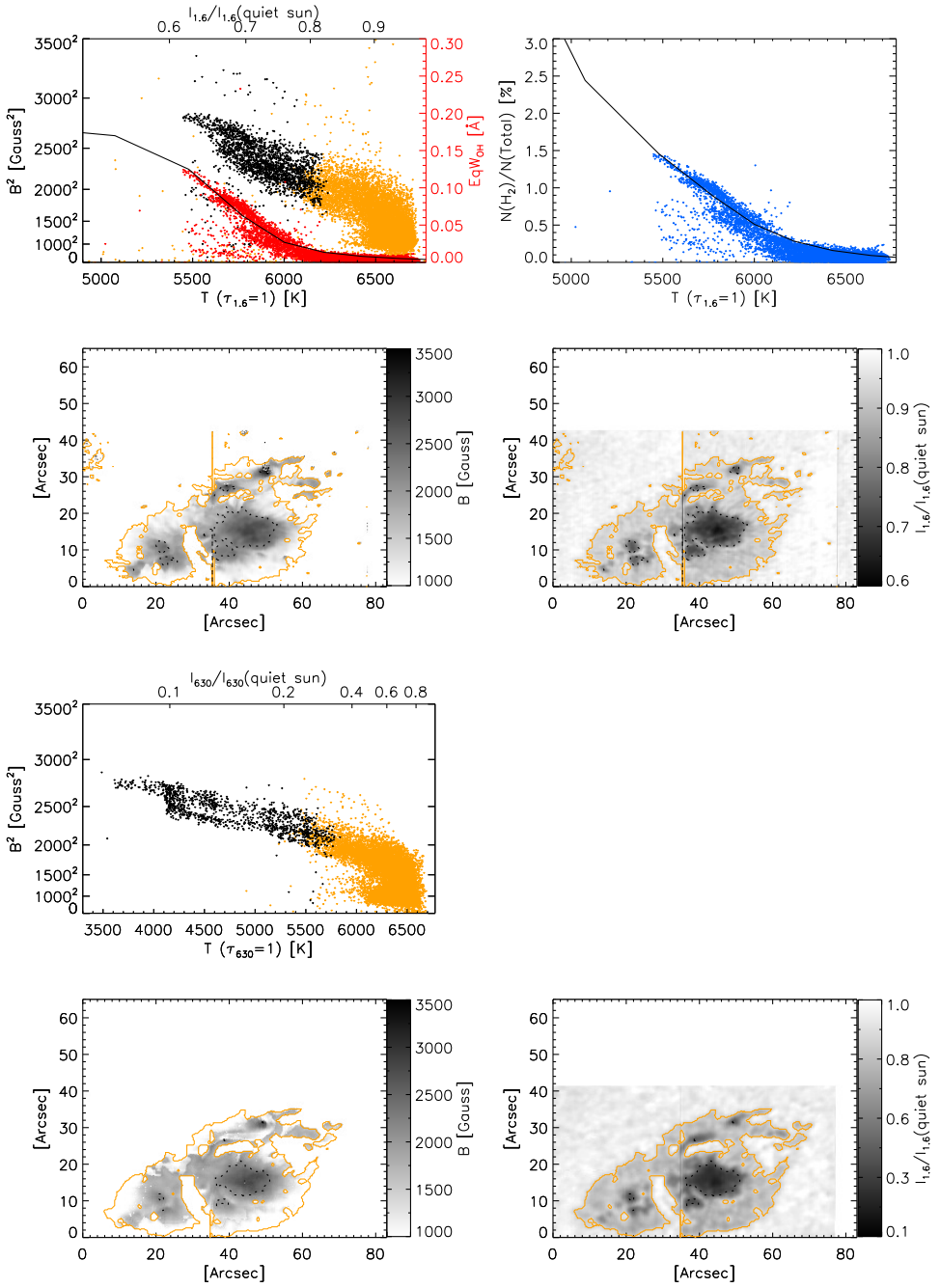

Figure 7.19 Obs. (17): The leading polarity spot in NOAA 11035 observed on 2009-12-18.

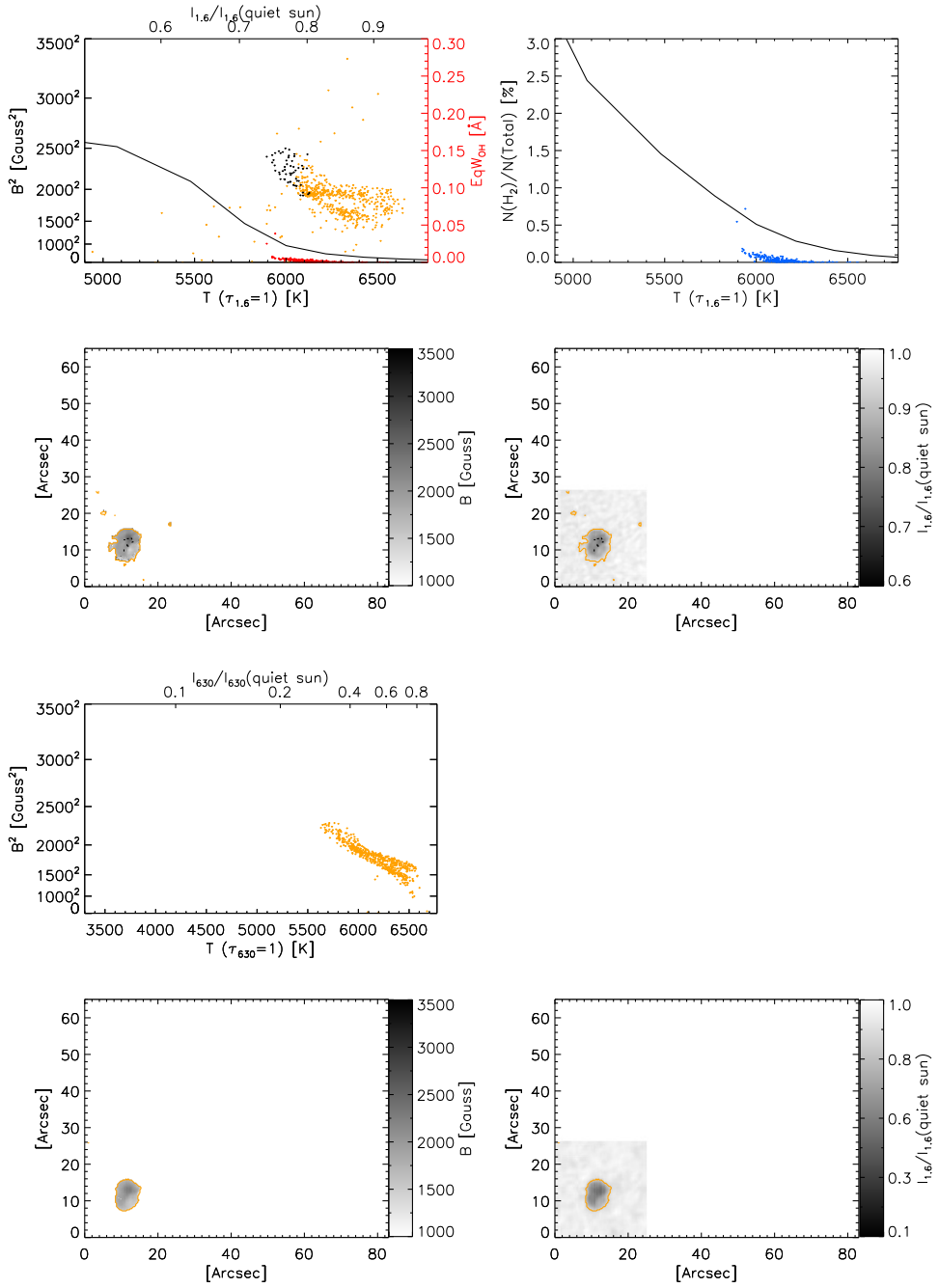

Figure 7.20 Obs. (18): The leading polarity spot in NOAA 11046 observed on 2010-02-12.

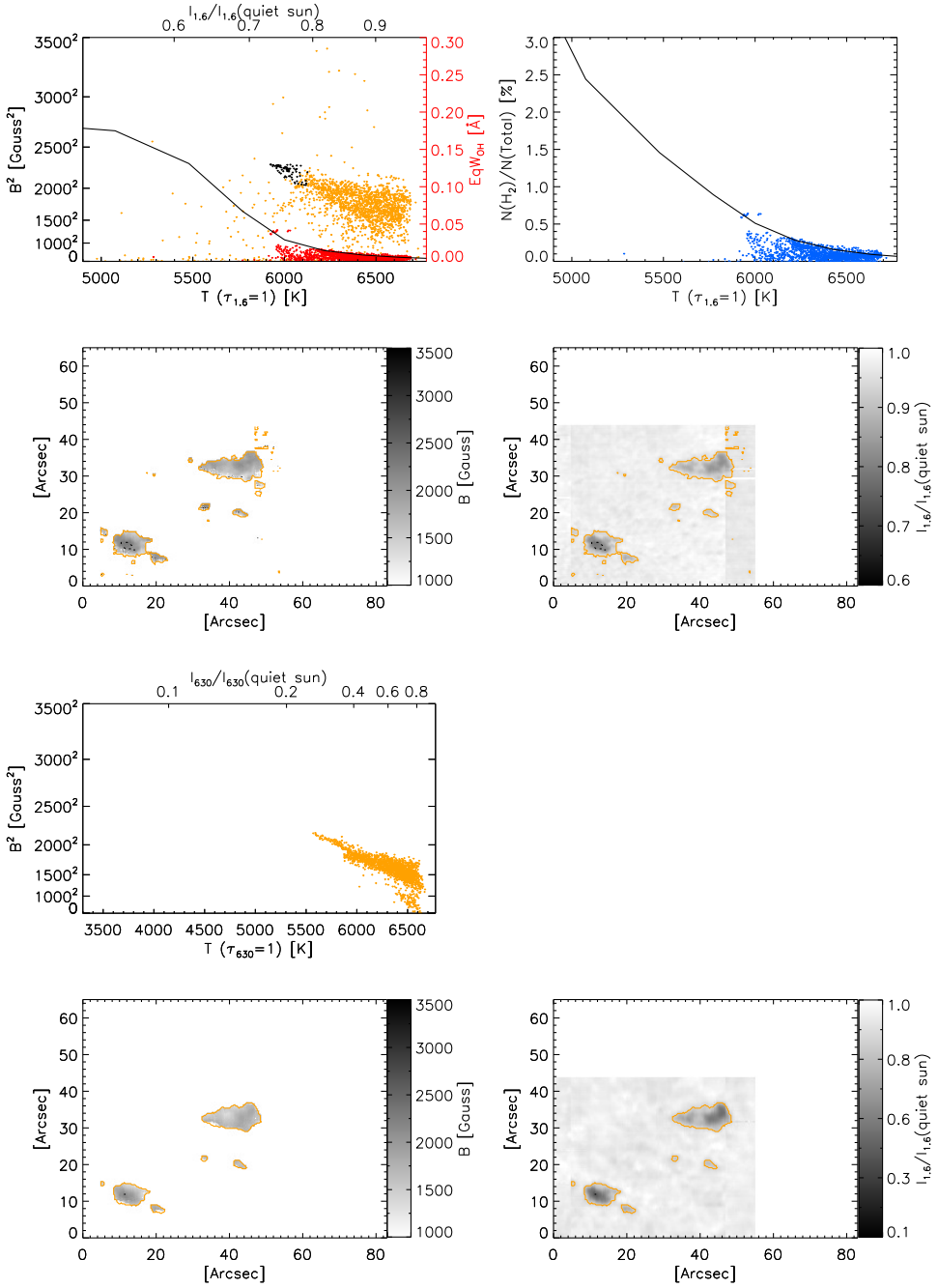

Figure 7.21 Obs. (19): The following polarity group in NOAA 11045 observed on 2010-02-12.

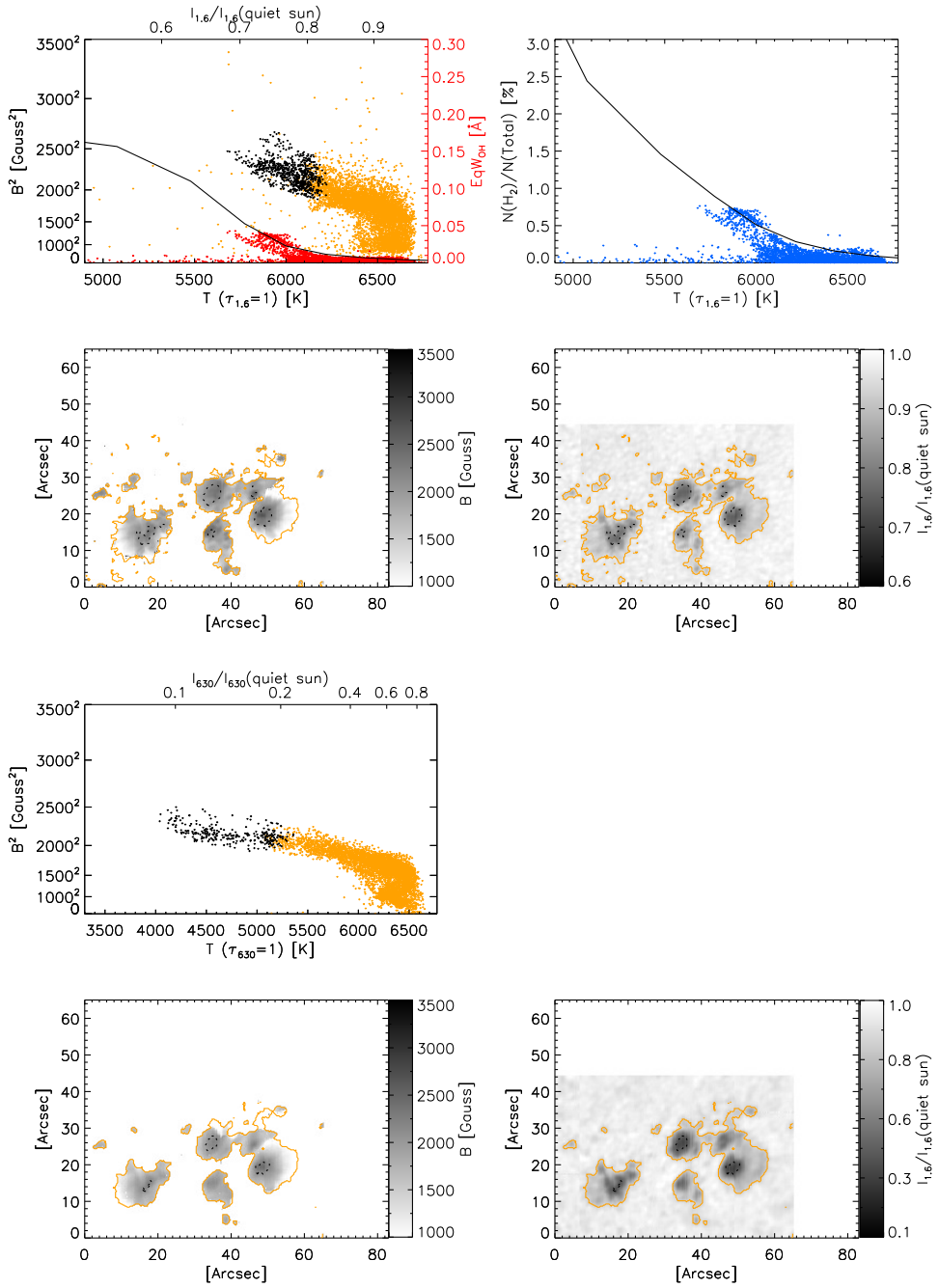

Figure 7.22 Obs. (20): The following polarity group in NOAA 10046 observed on 2010-02-12.

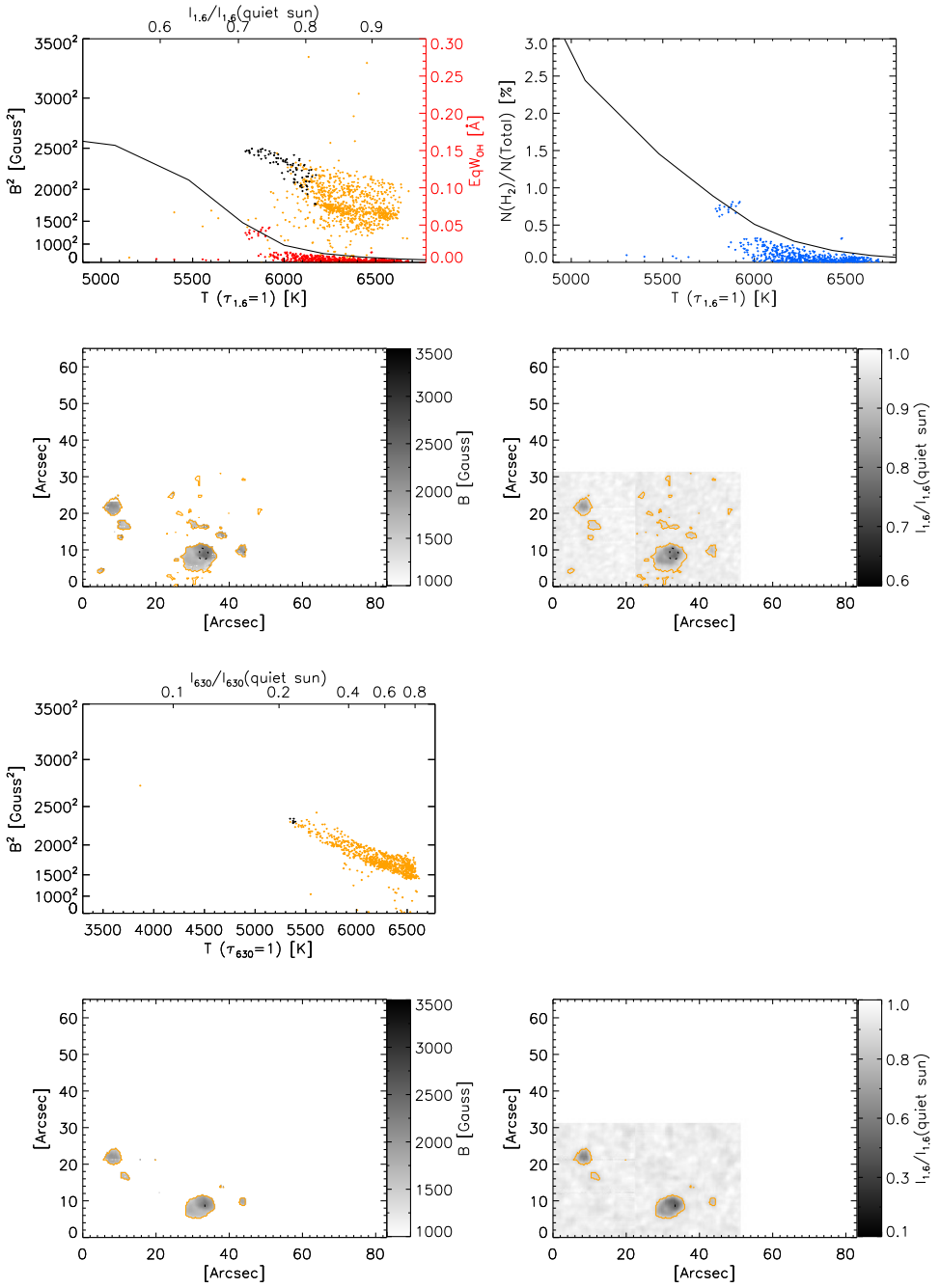

Figure 7.23 Obs. (21): The leading polarity spot in NOAA 11046 observed on 2010-02-13.

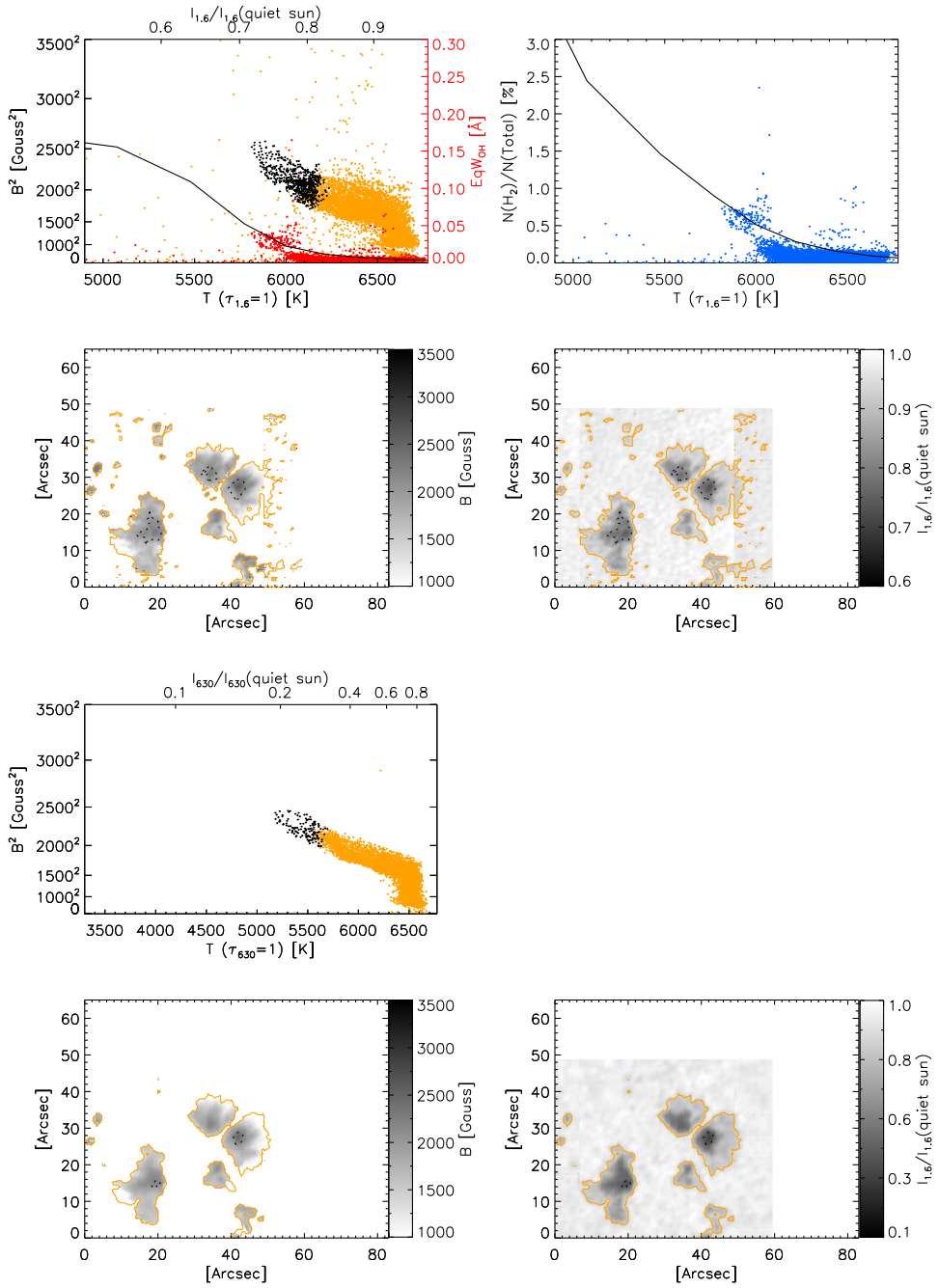

Figure 7.24 Obs. (22): The following polarity group in NOAA 11046 observed on 2010-02-13.

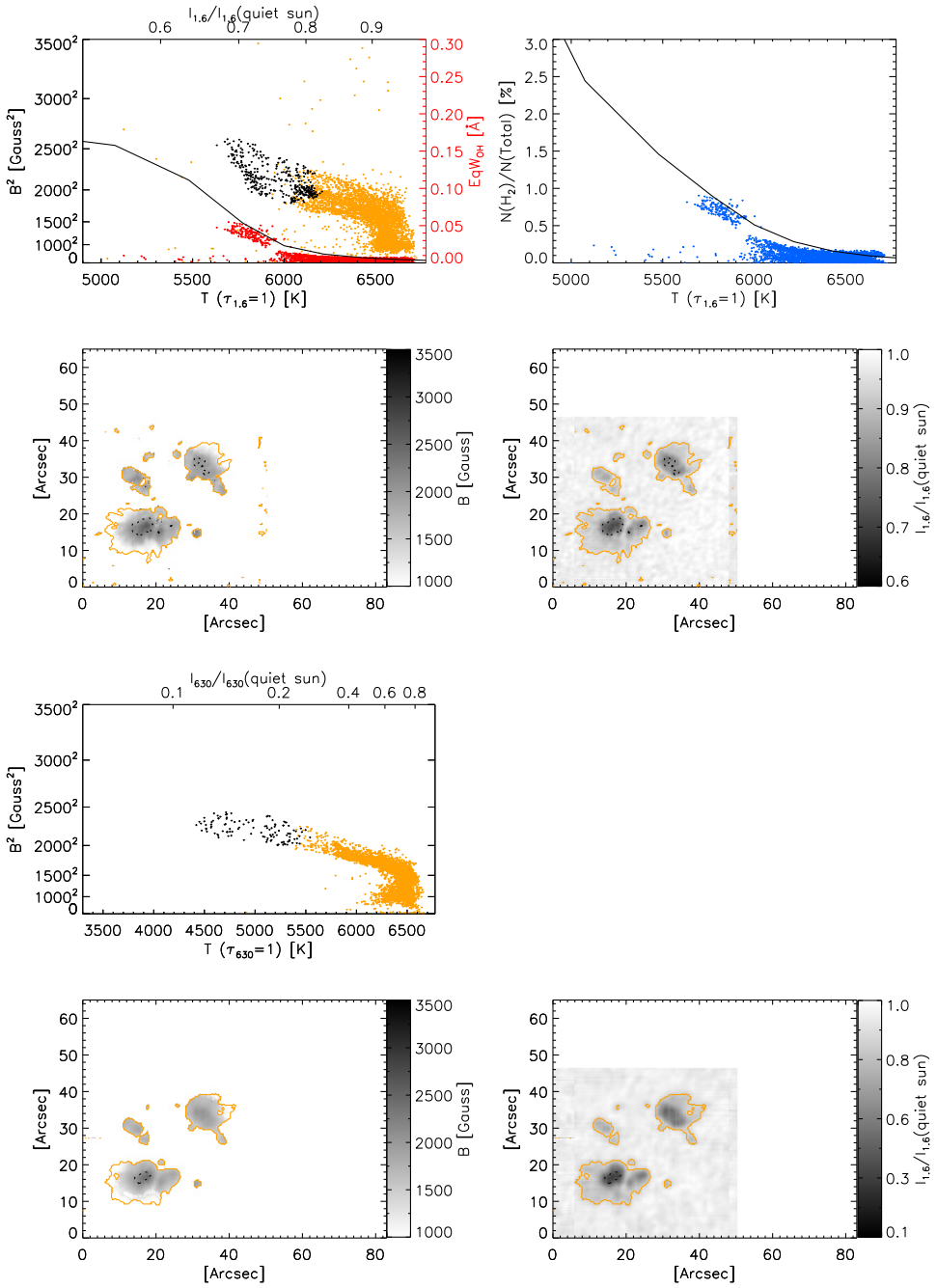

Figure 7.25 Obs. (23): The following polarity group in NOAA 11046 observed on 2010-02-14.

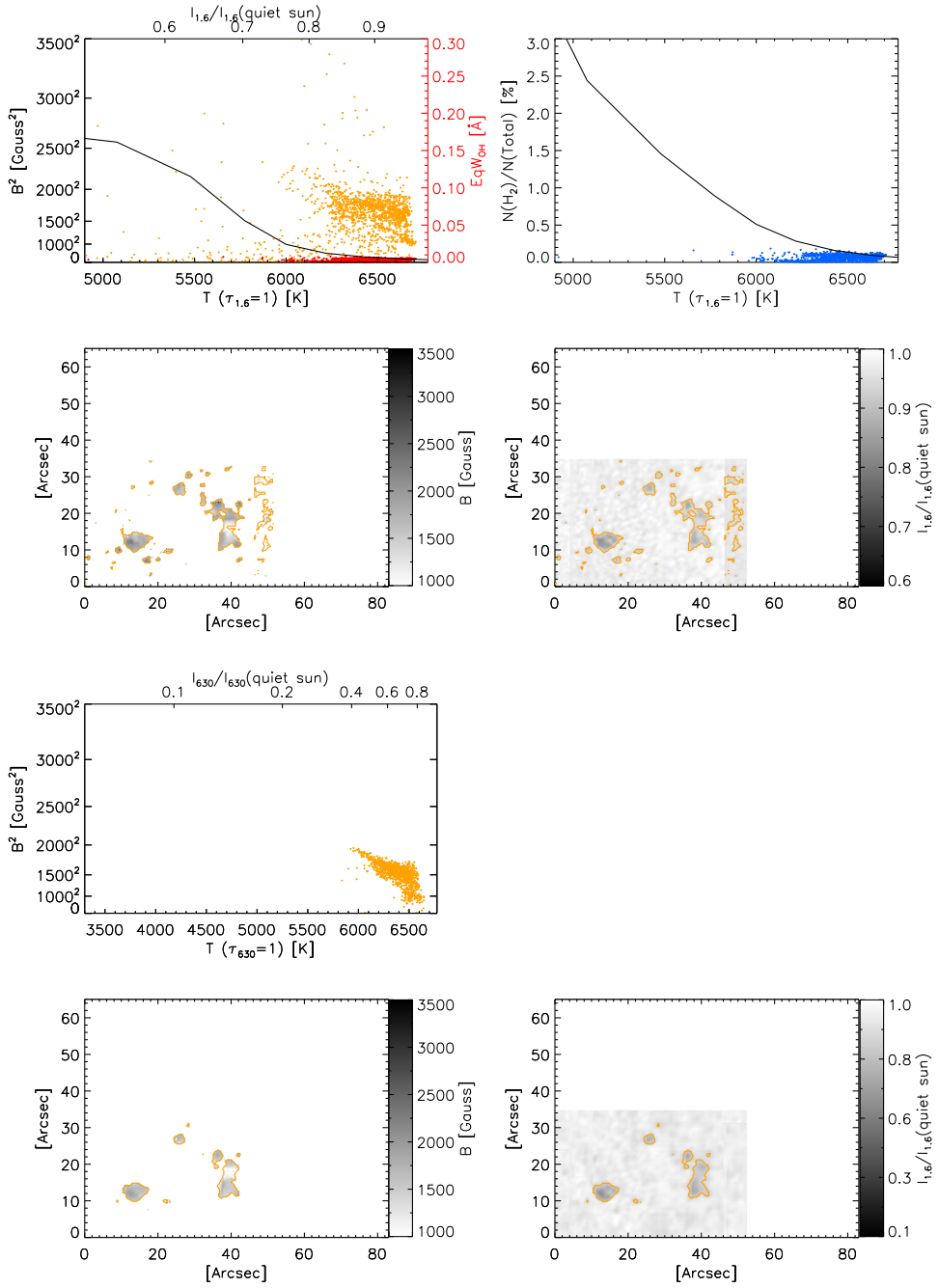

Figure 7.26 Obs. (24): The leading polarity spot in NOAA 11046 observed on 2010-02-14.

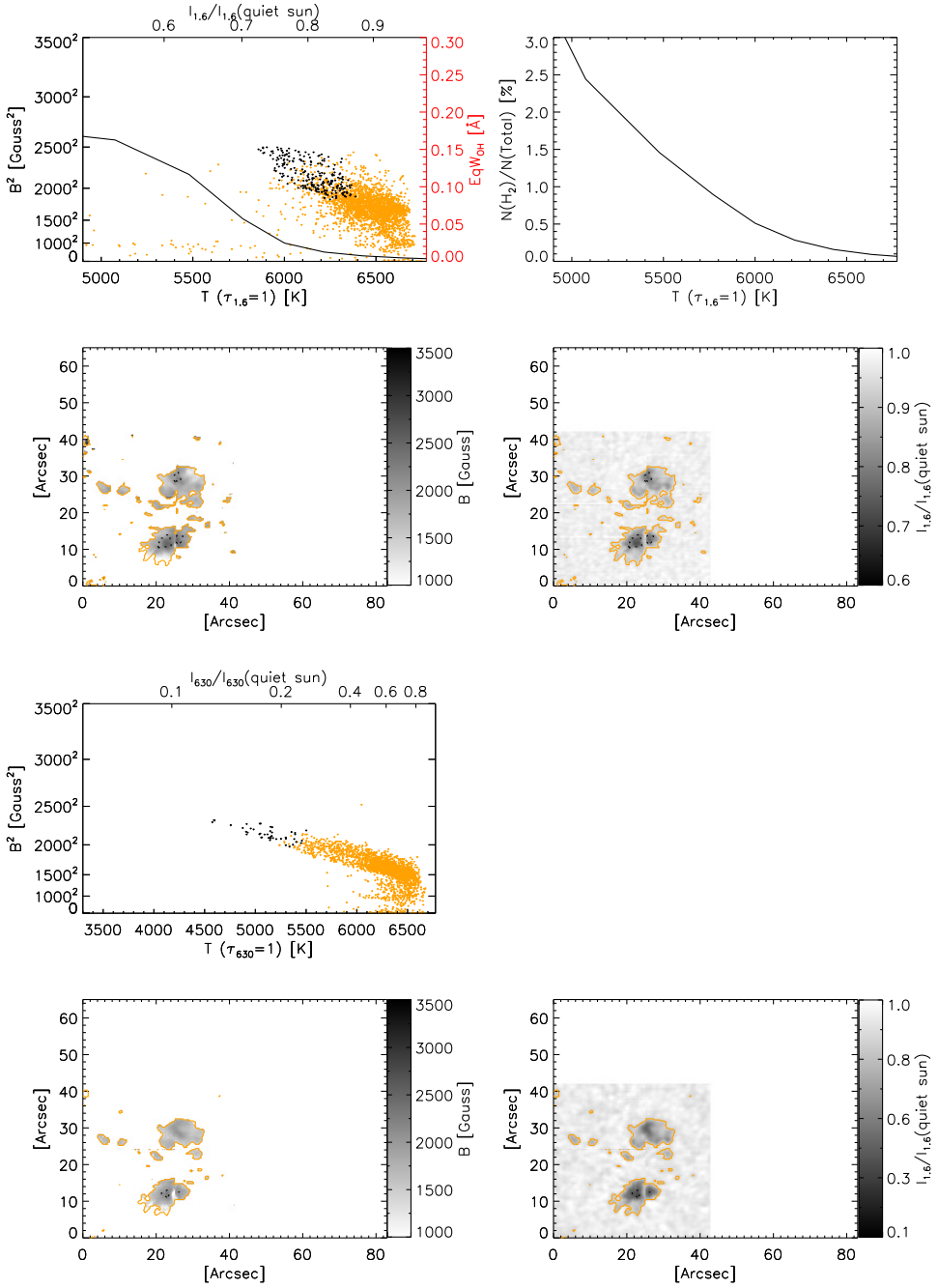

Figure 7.27 Obs. (25): The following polarity group in NOAA 11046 observed on 2010-02-15.

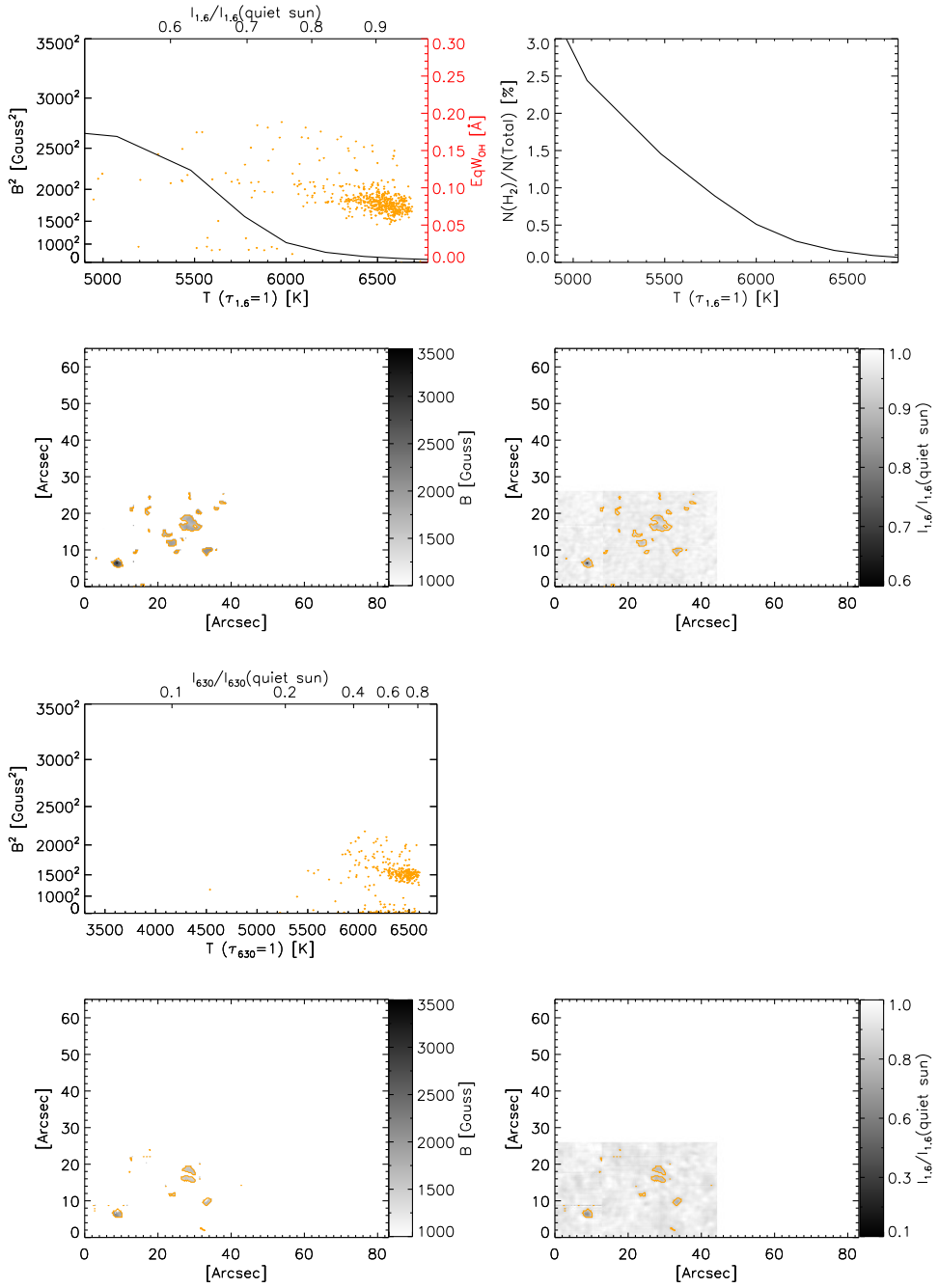

Figure 7.28 Obs. (26): The leading polarity spot in NOAA 11046 observed on 2010-02-15.

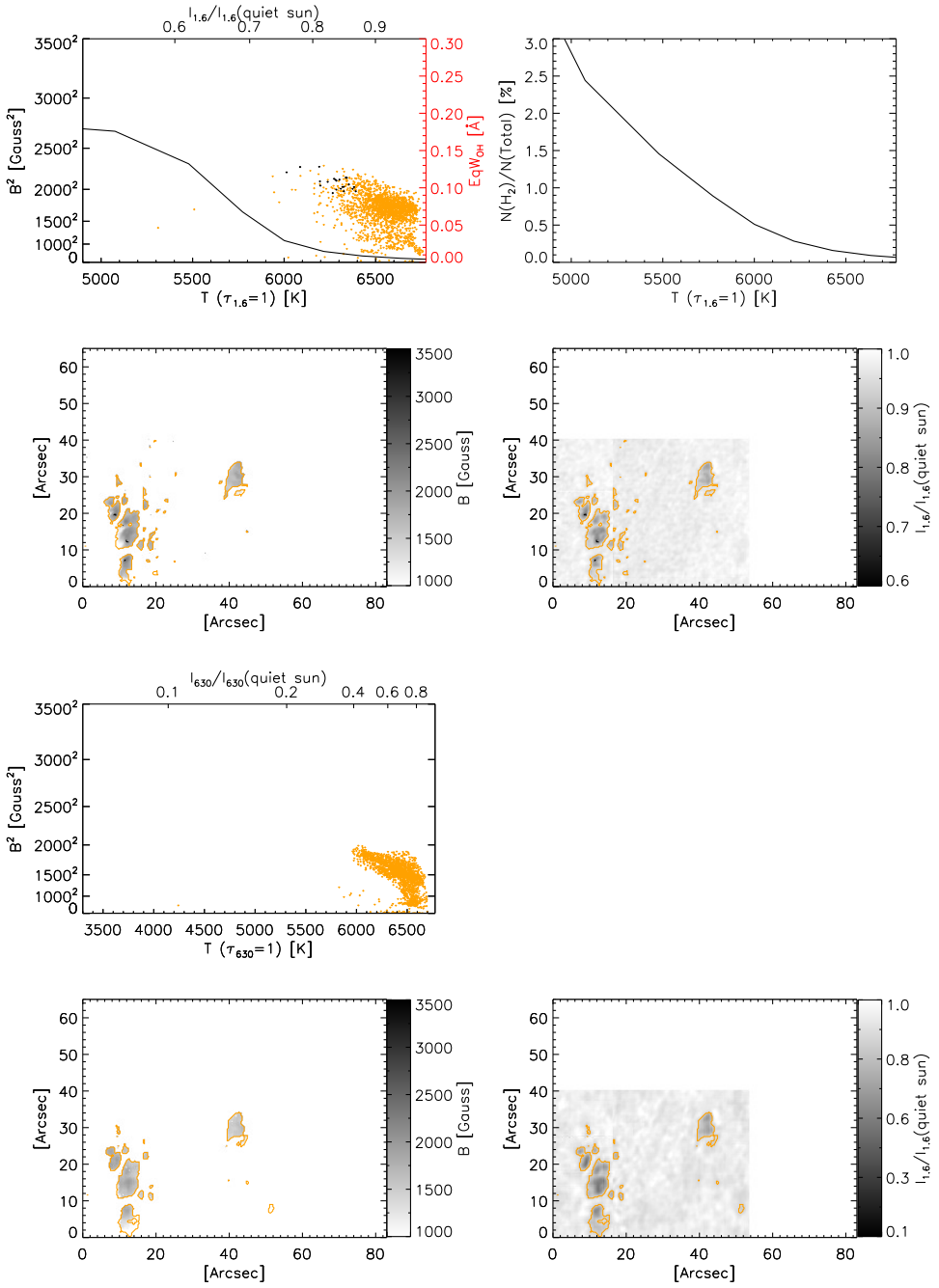

Figure 7.29 Obs. (27): NOAA 11048 observed on 2010-02-15.

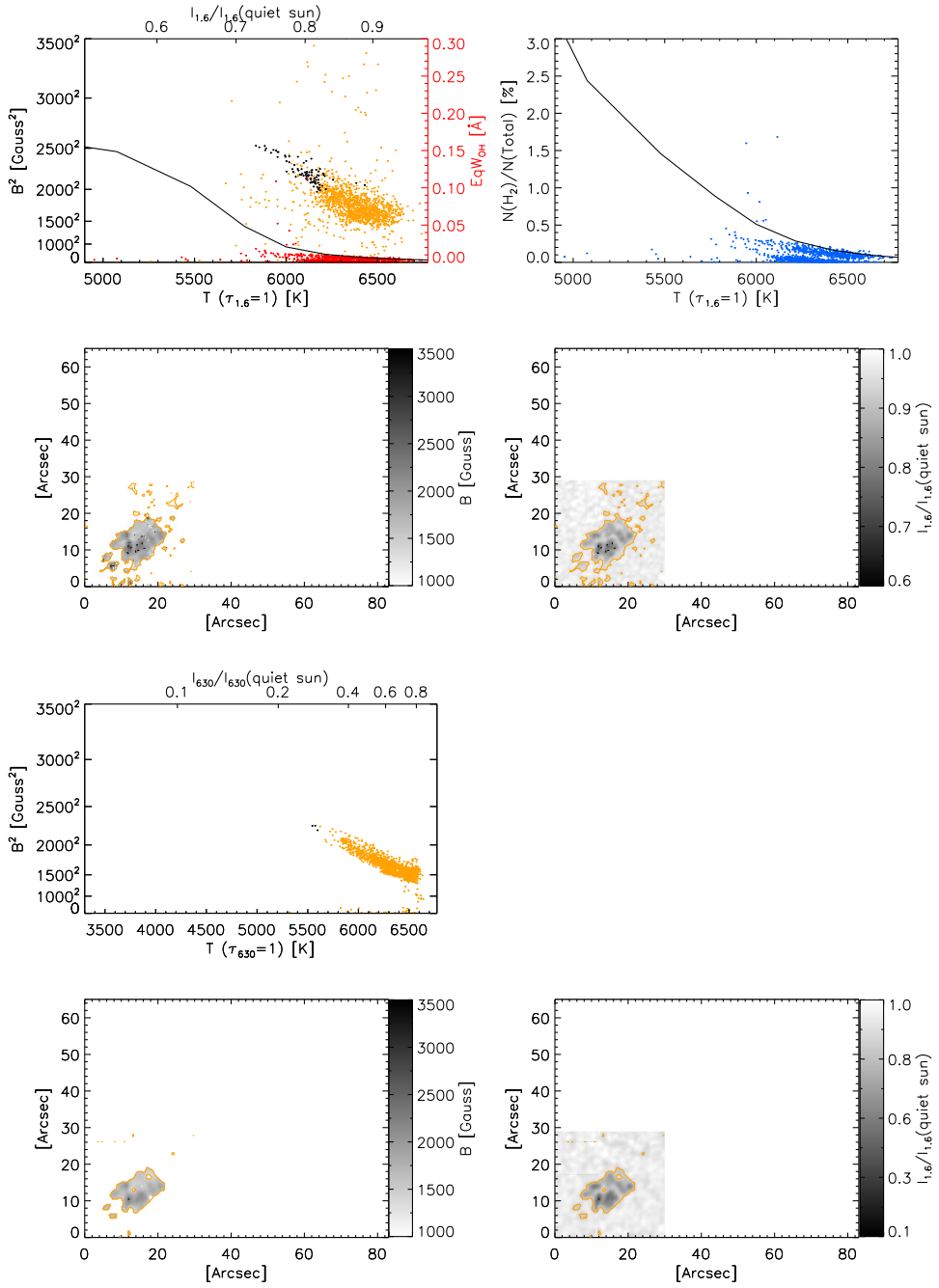

Figure 7.30 Obs. (28): NOAA 11049 observed on 2010-02-17.

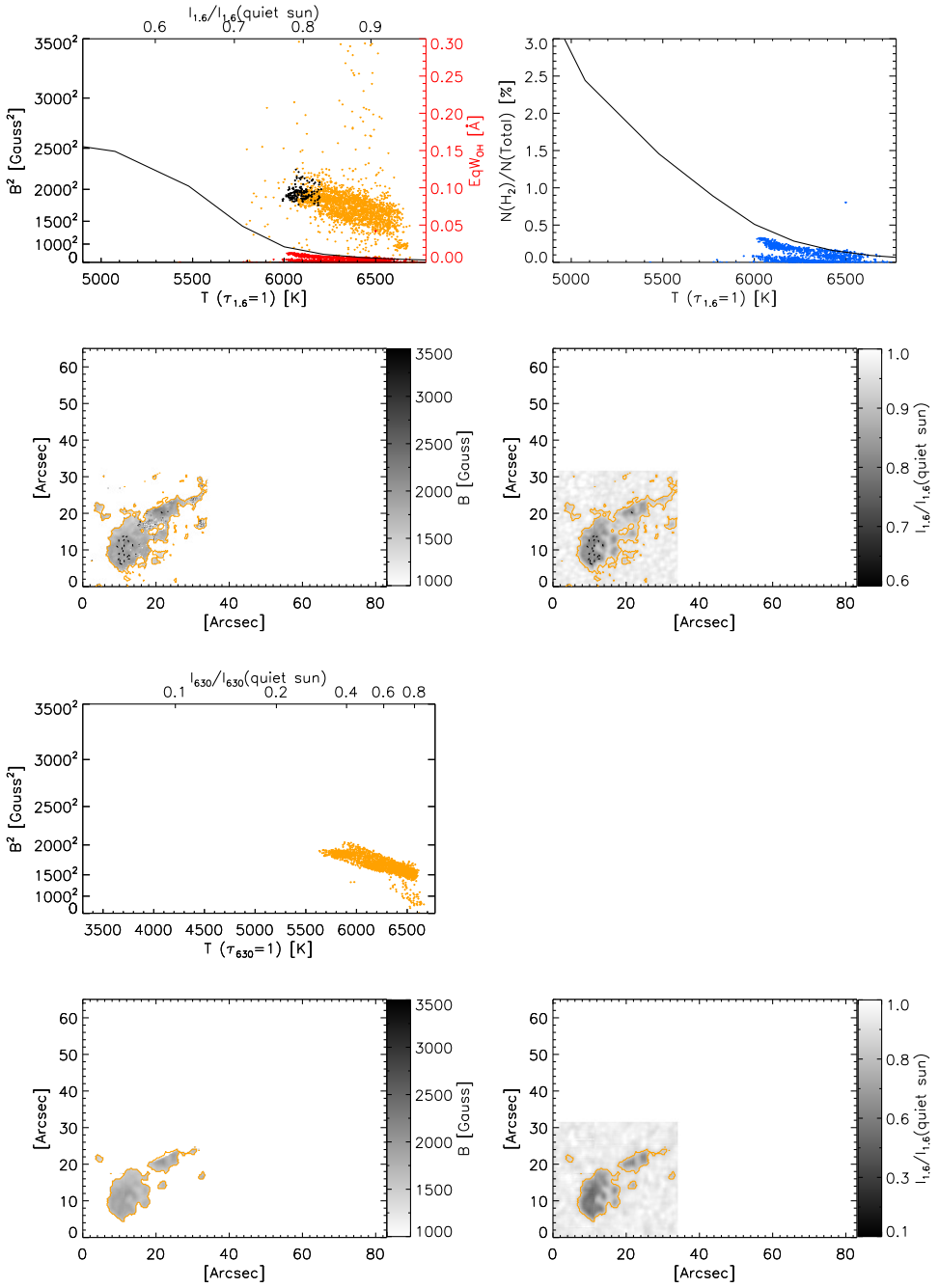

Figure 7.31 Obs. (29): NOAA 11049 observed on 2010-02-18.

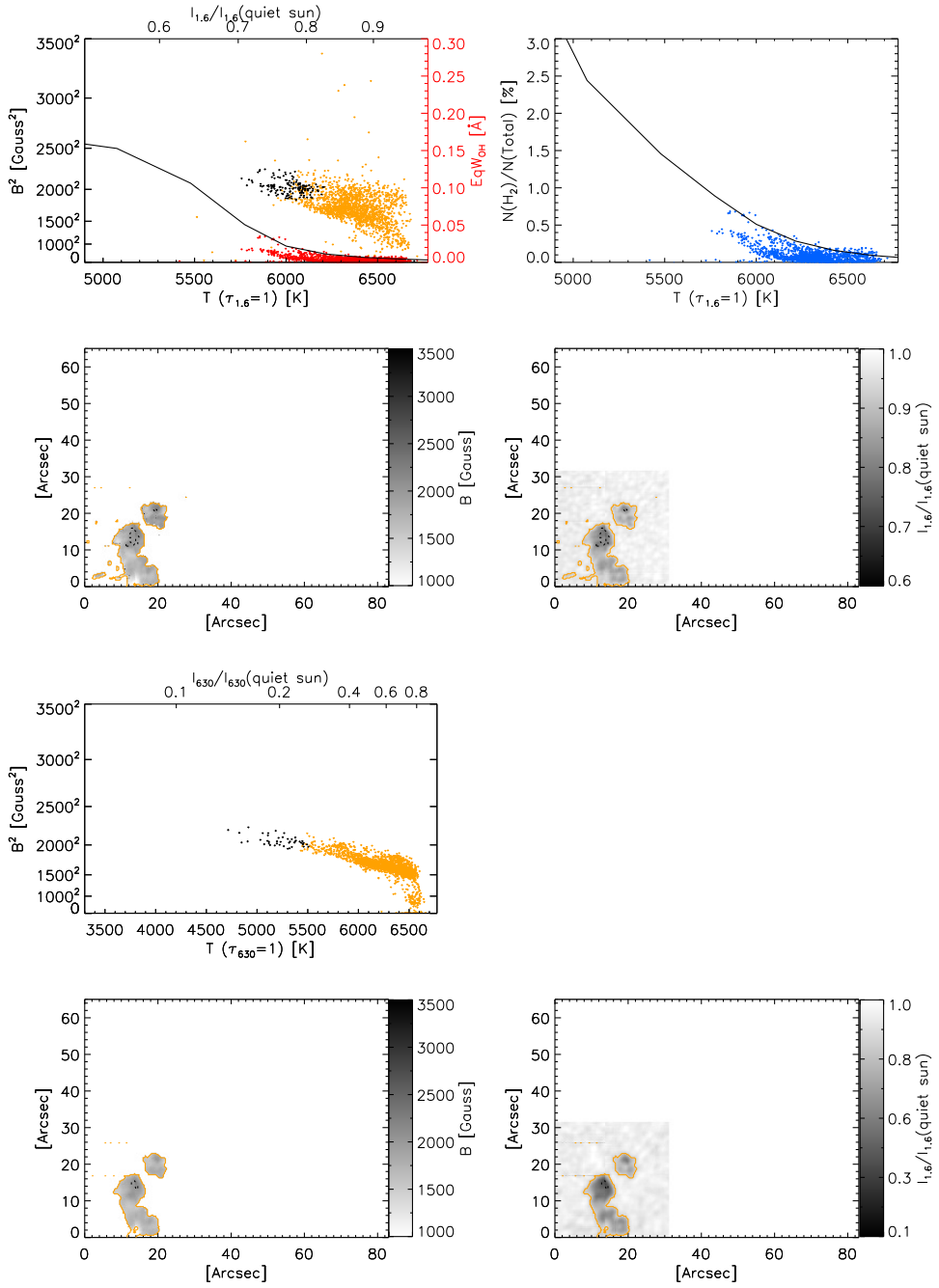

Figure 7.32 Obs. (30): The leading polarity spot in NOAA 11049 observed on 2010-02-19.

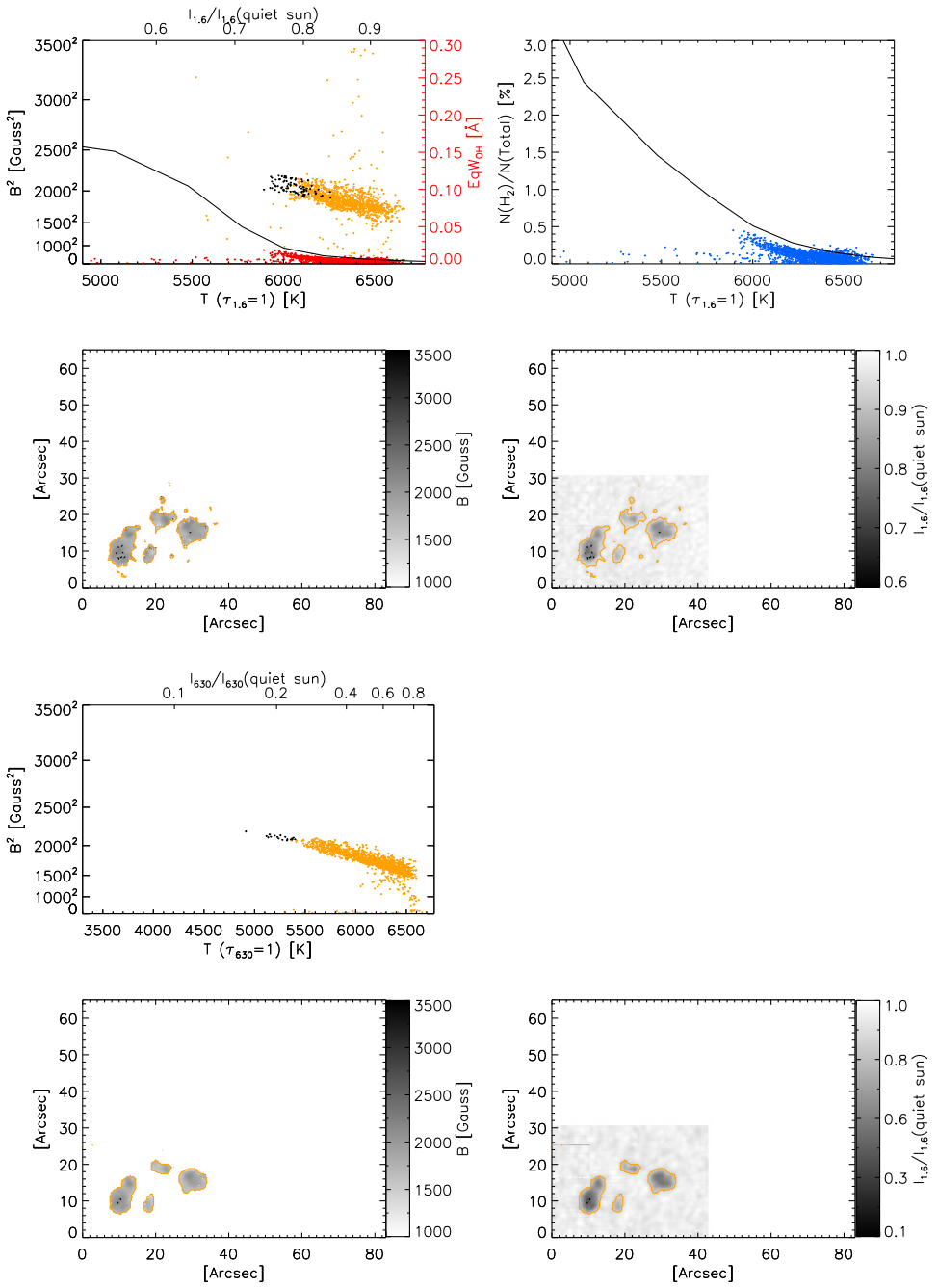

Figure 7.33 Obs. (31): The following polarity spot in NOAA 11049 observed on 2010-02-19.

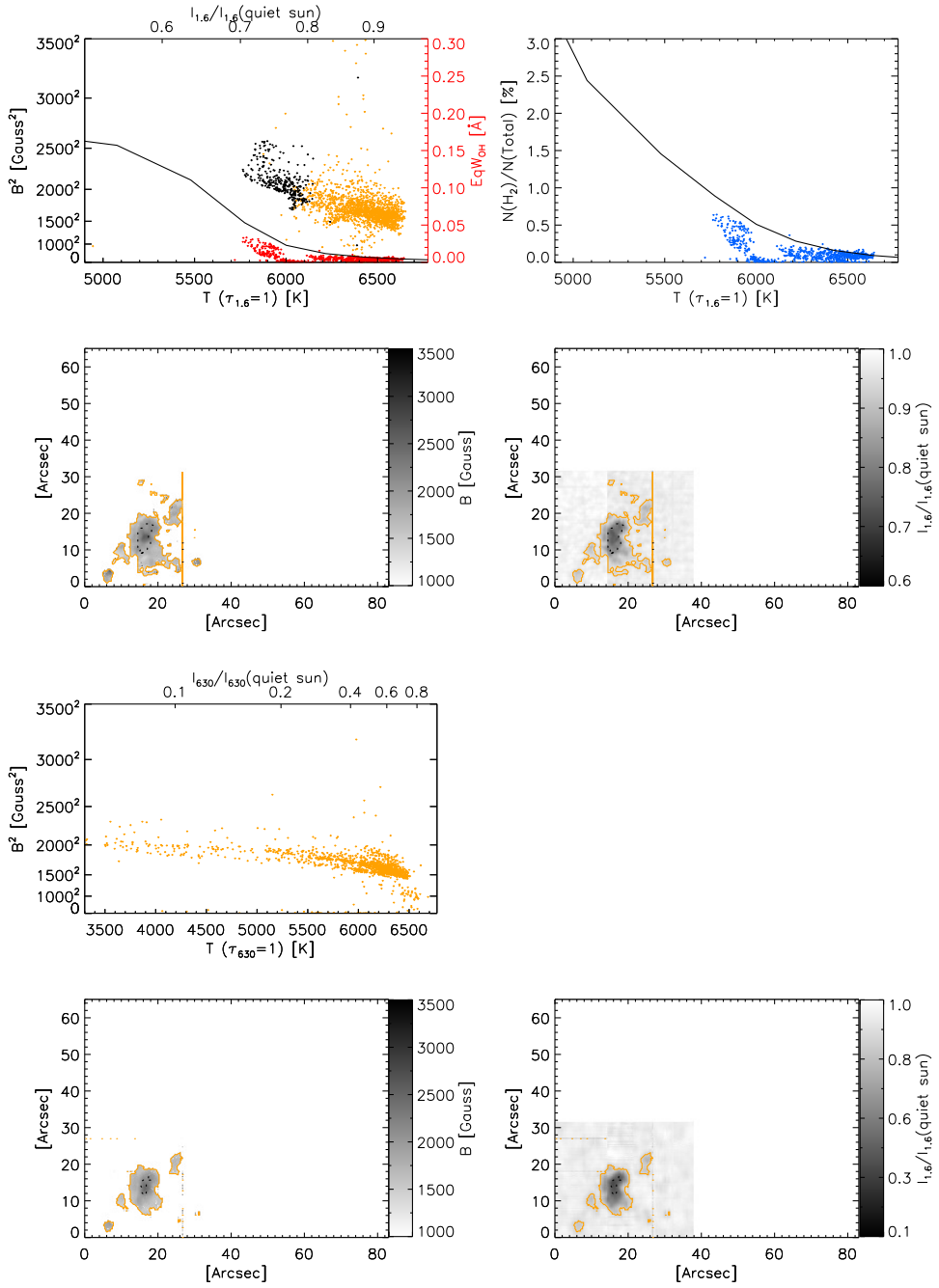

Figure 7.34 Obs. (32): The leading polarity spot in NOAA 11049 observed on 2010-02-20.

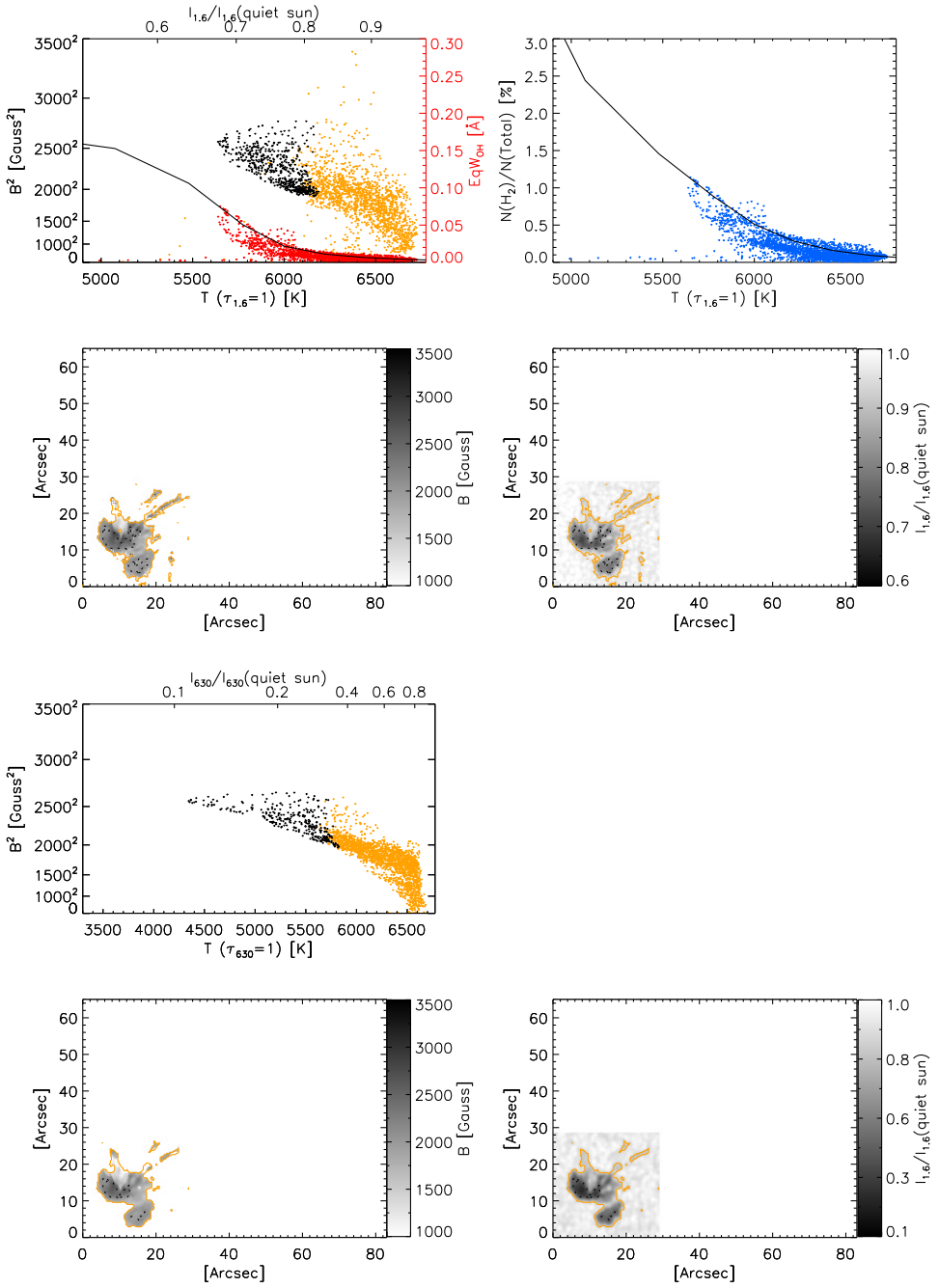

Figure 7.35 Obs. (33): NOAA 11072 observed on 2010-05-21.

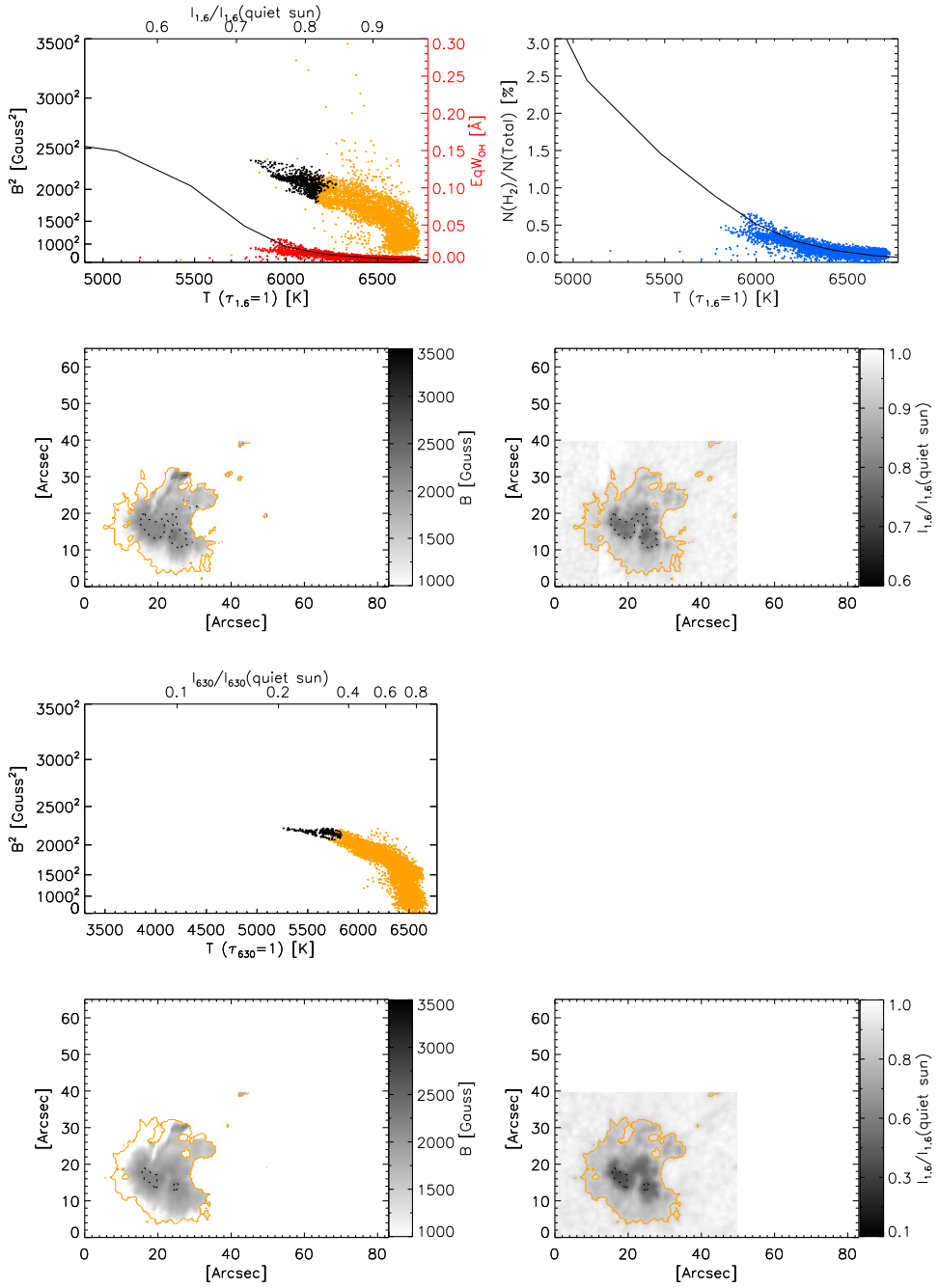

Figure 7.36 Obs. (34): NOAA 11072 observed on 2010-05-22.

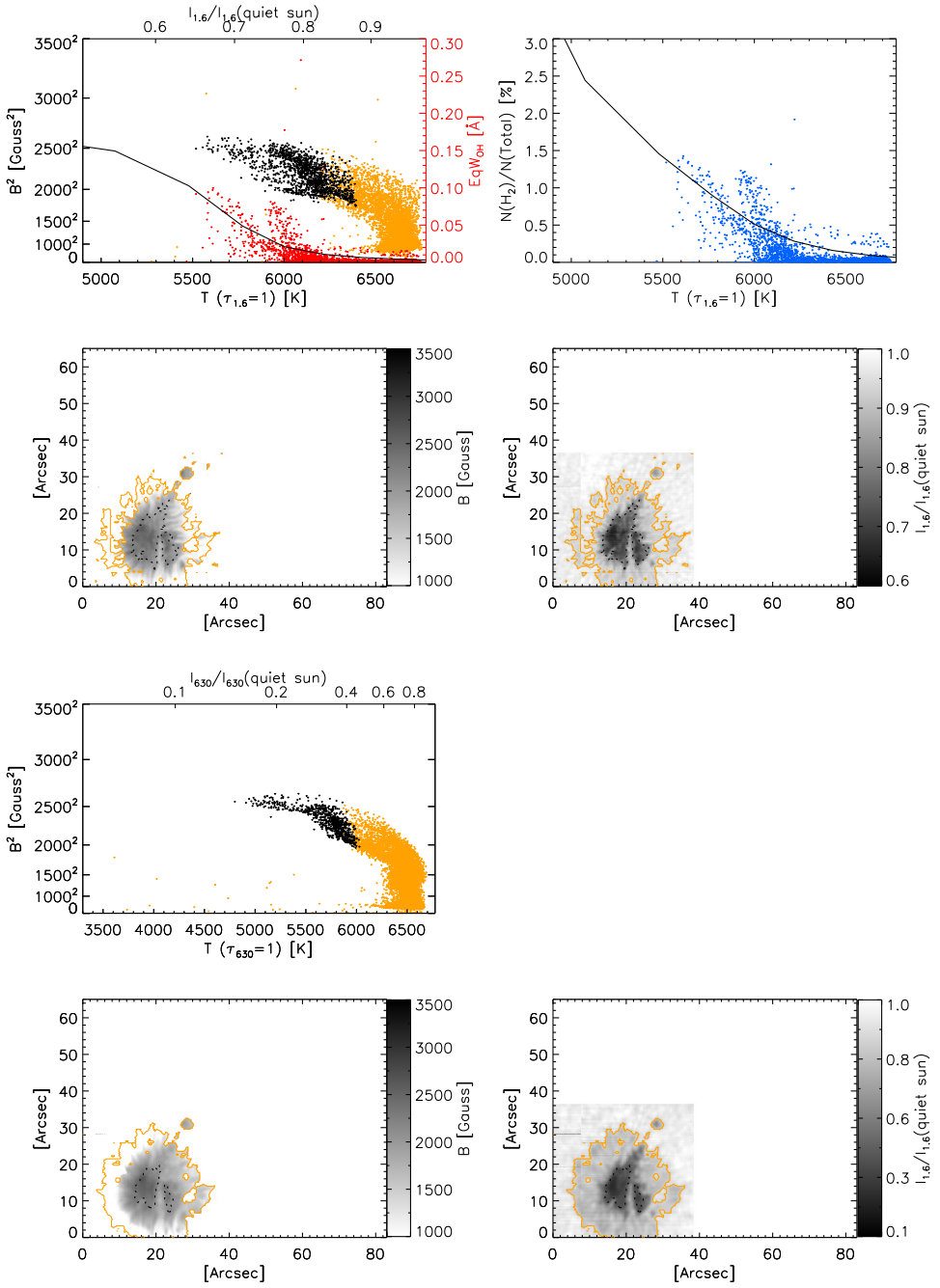

Figure 7.37 Obs. (35): NOAA 11072 observed on 2010-05-23.

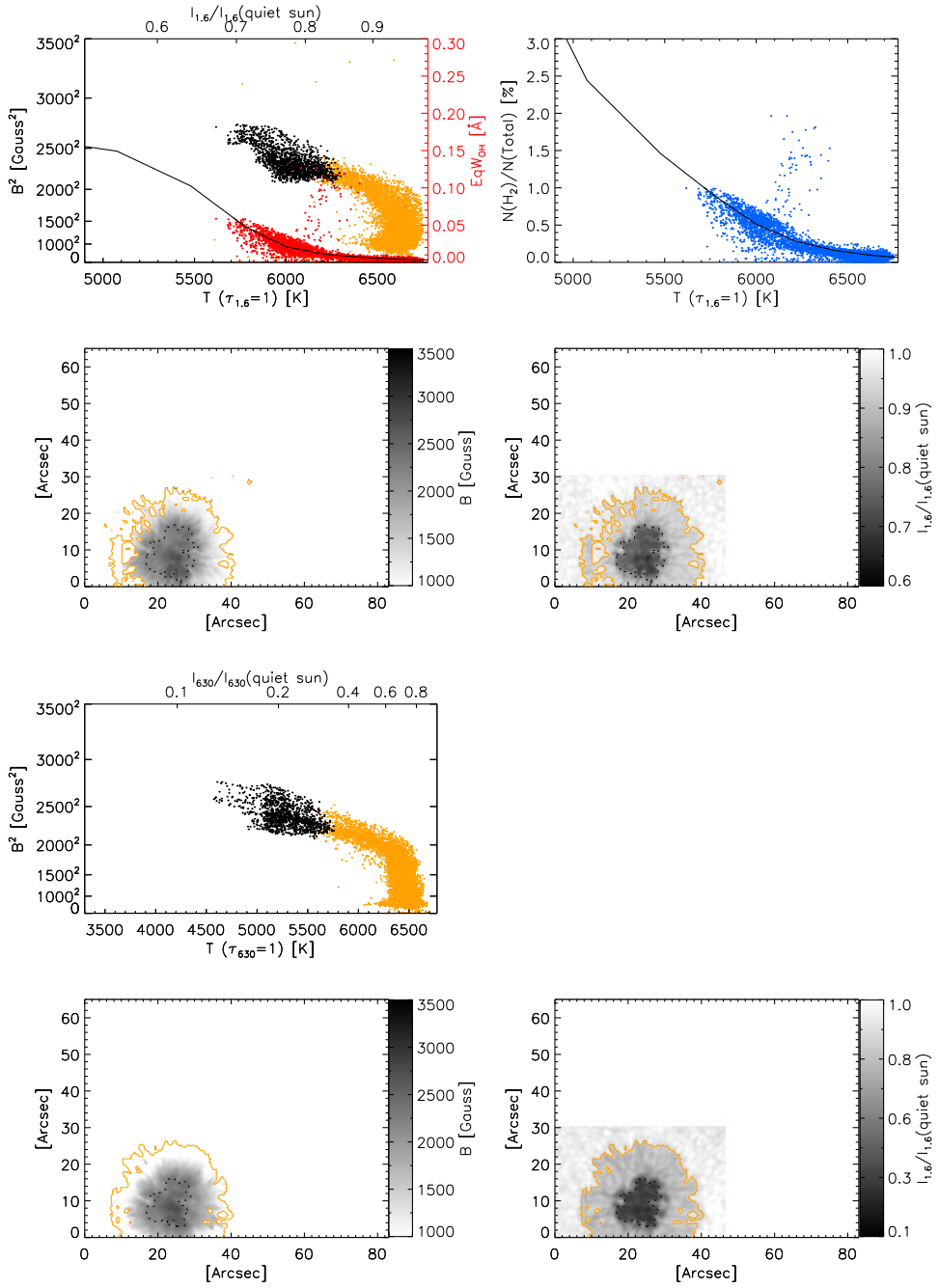

Figure 7.38 Obs. (36): NOAA 11101 observed on 2010-08-29.

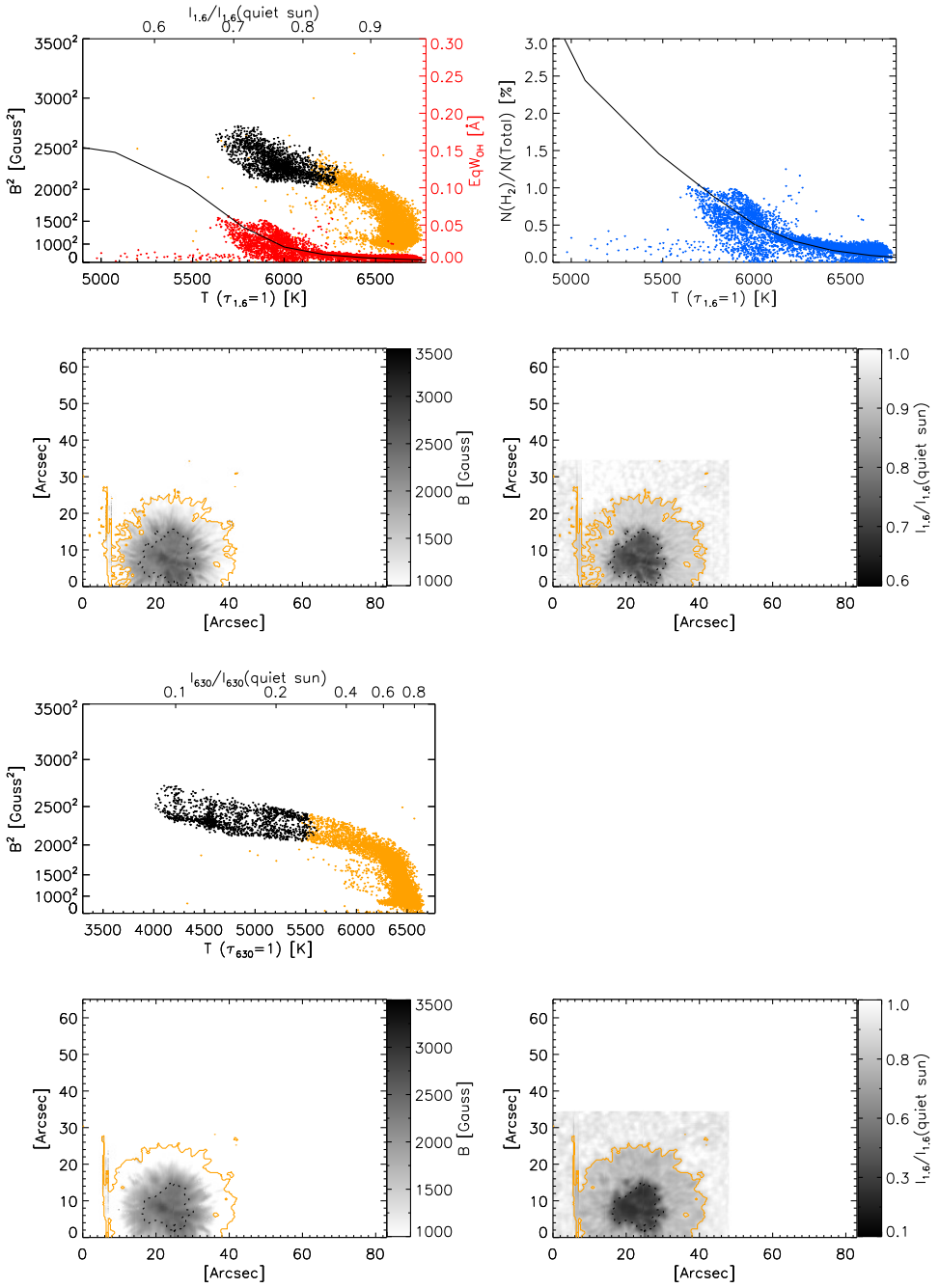

Figure 7.39 Obs. (37): NOAA 11101 observed on 2010-08-30.

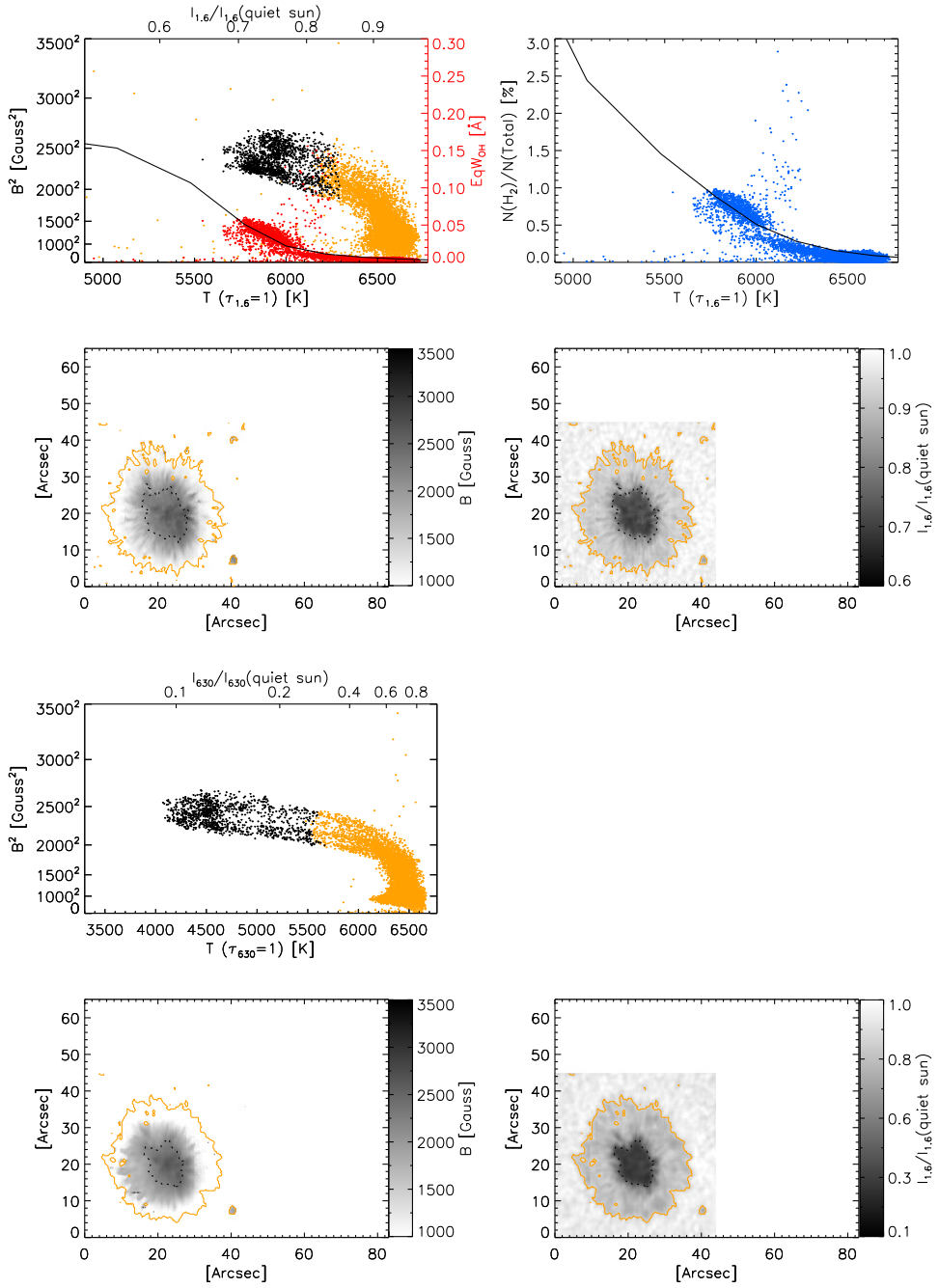

Figure 7.40 Obs. (38): NOAA 11101 observed on 2010-09-01.

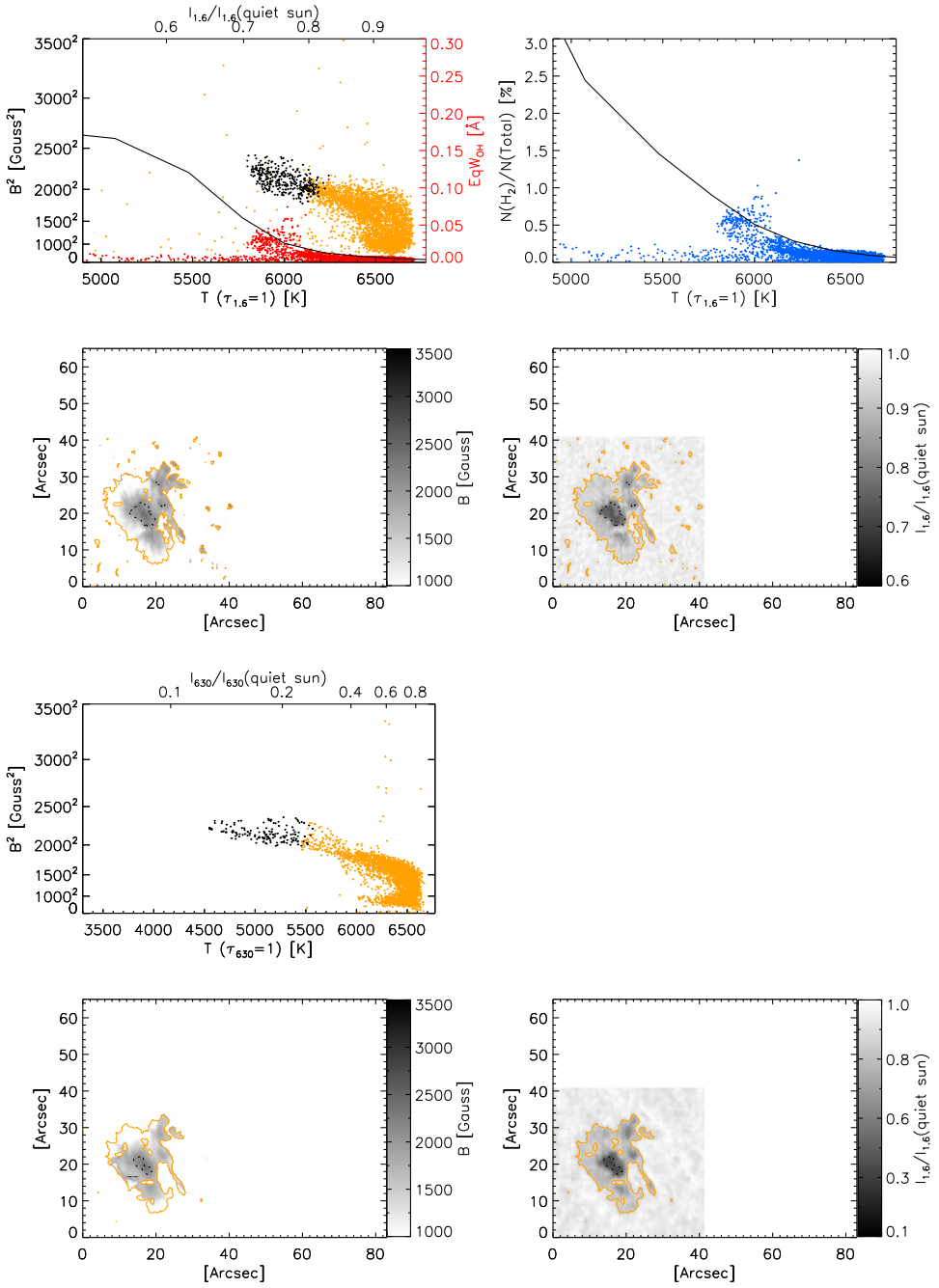

Figure 7.41 Obs. (39): NOAA 11102 observed on 2010-09-01.

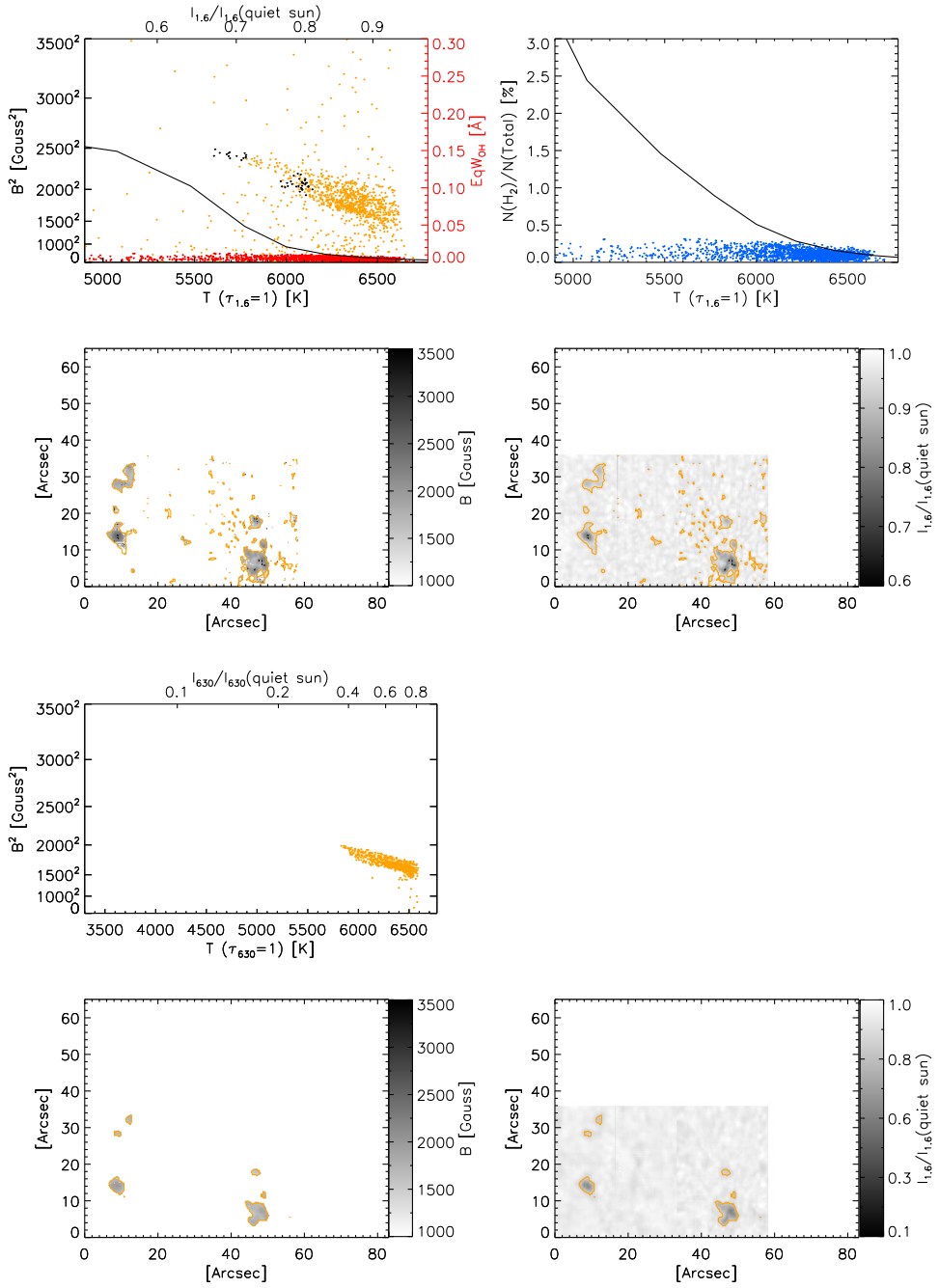

Figure 7.42 Obs. (40): NOAA 11103/4 observed on 2010-09-01.

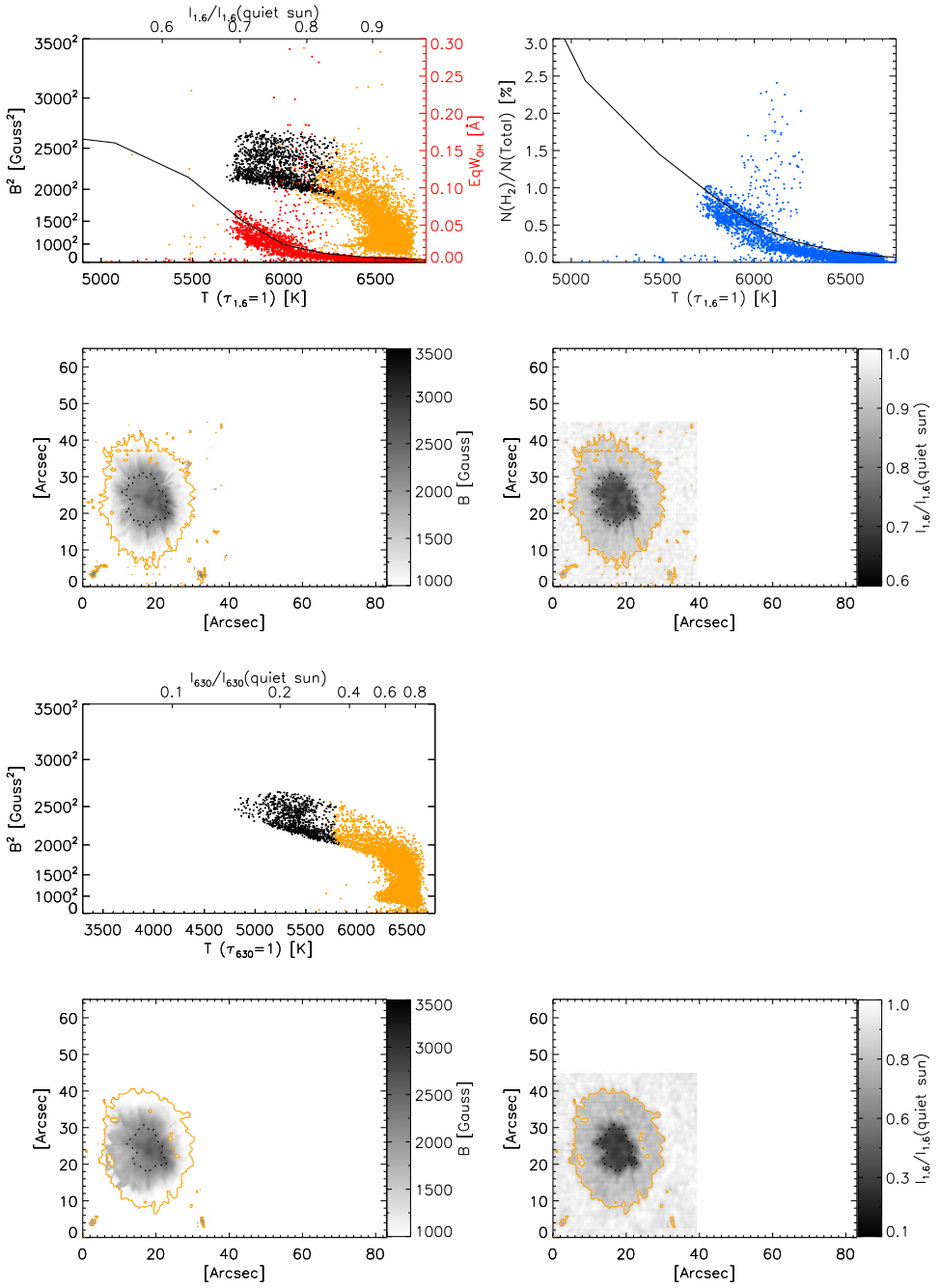

Figure 7.43 Obs. (41): NOAA 11101 observed on 2010-09-02.

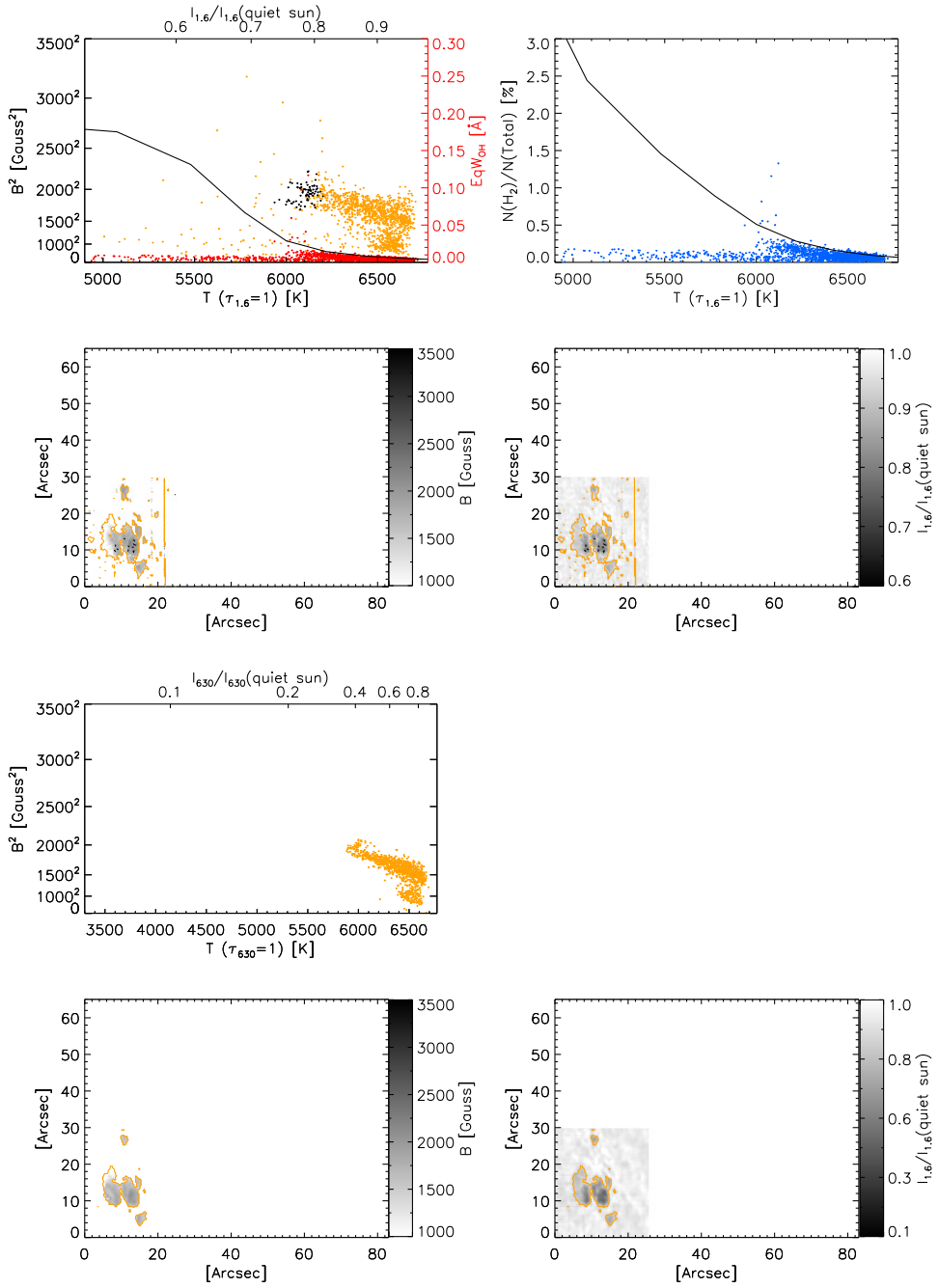

Figure 7.44 Obs. (42): NOAA 11102 observed on 2010-09-02.

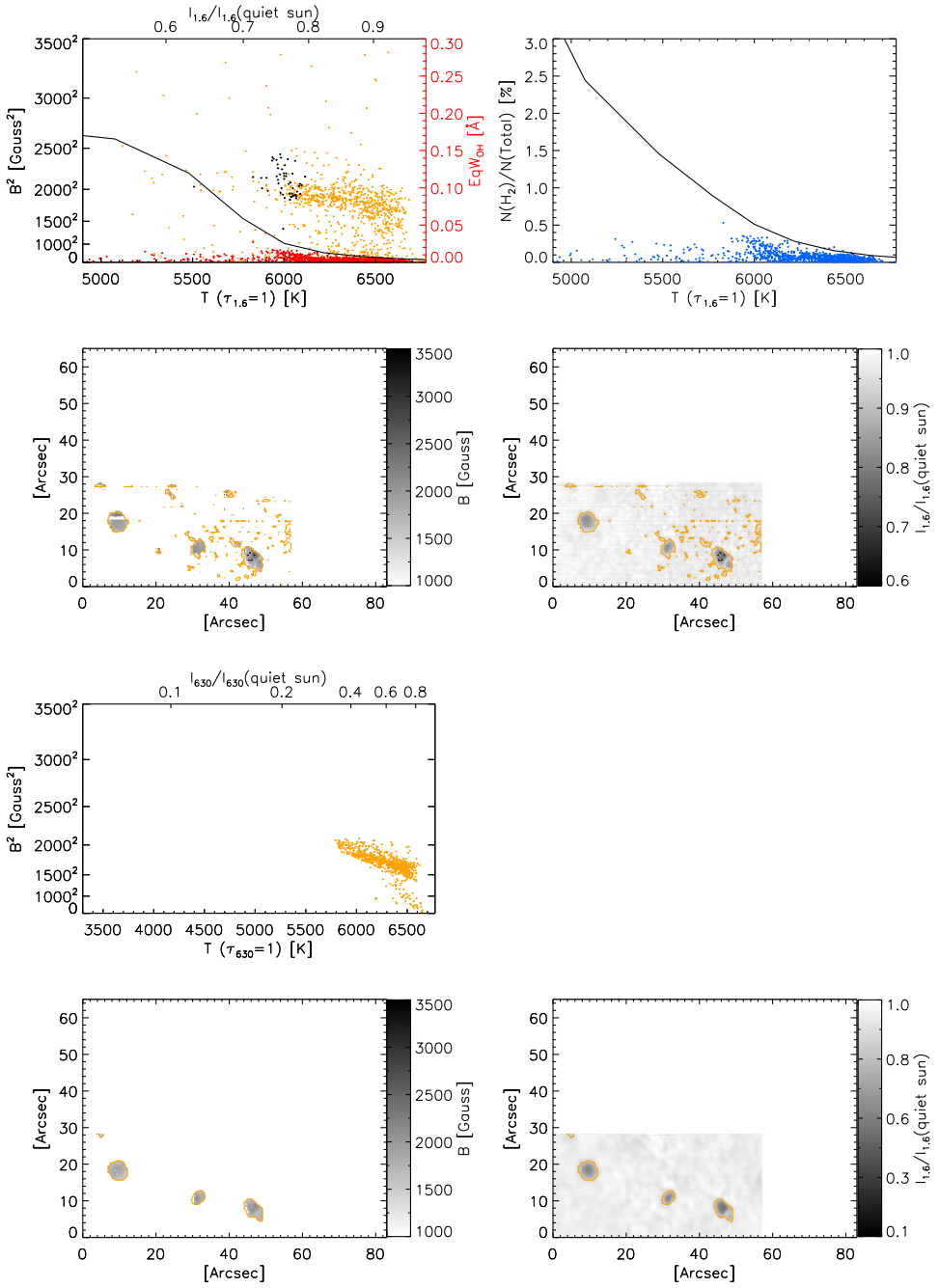

Figure 7.45 Obs. (43): NOAA 11103/4 observed on 2010-09-02.

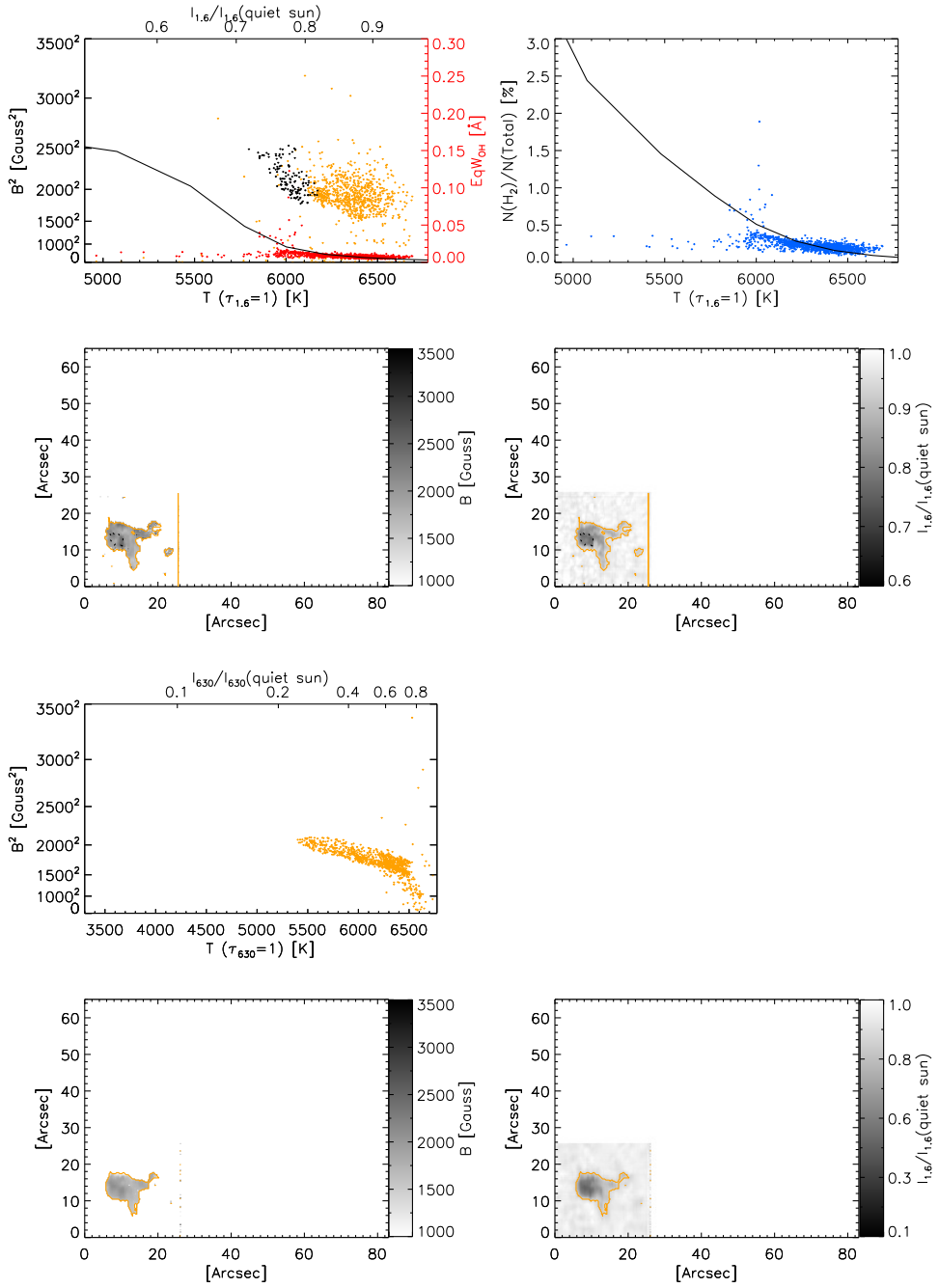

Figure 7.46 Obs. (44): NOAA 11105 observed on 2010-09-02.

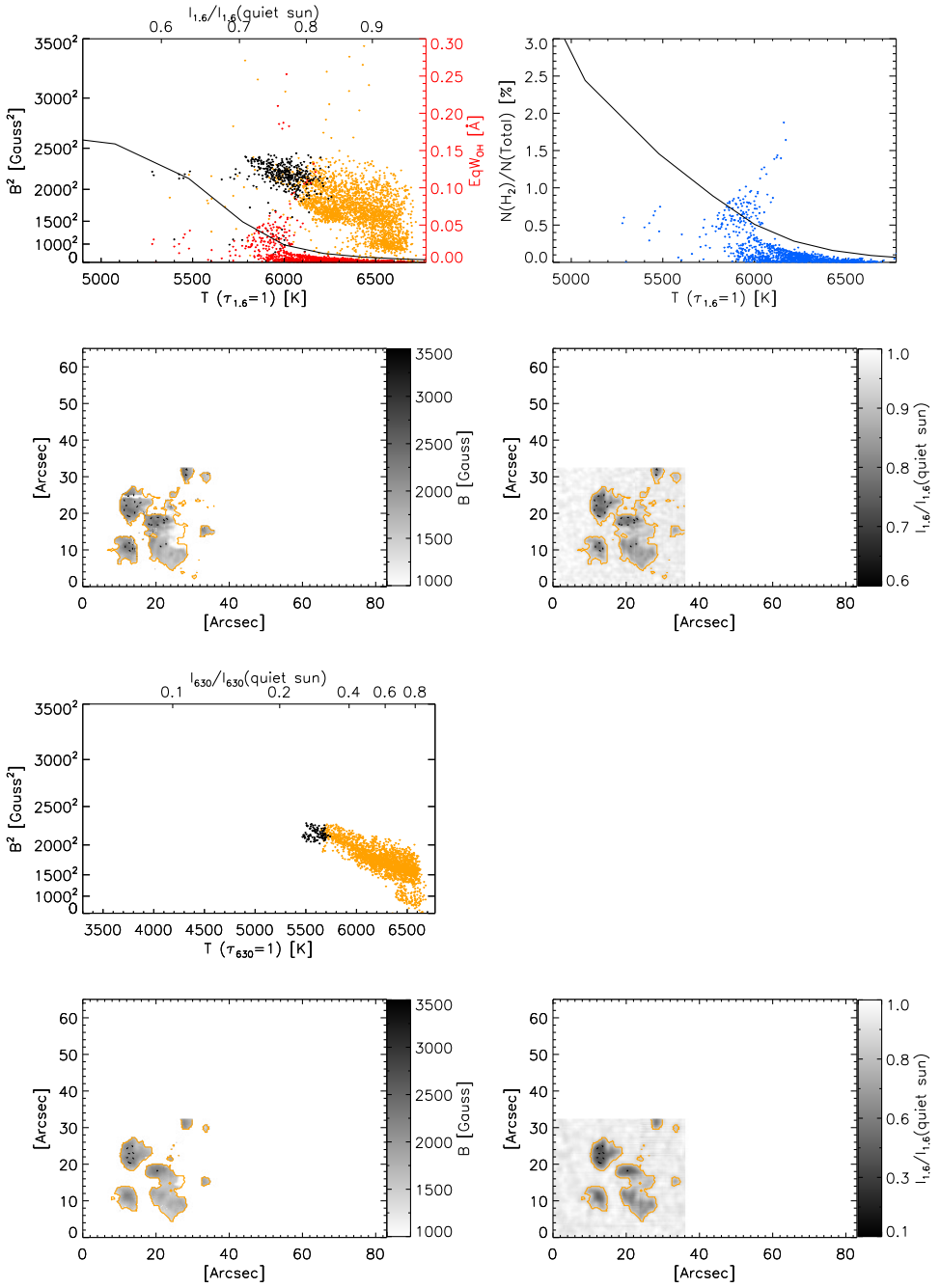

Figure 7.47 Obs. (45): NOAA 11105 observed on 2010-09-04.

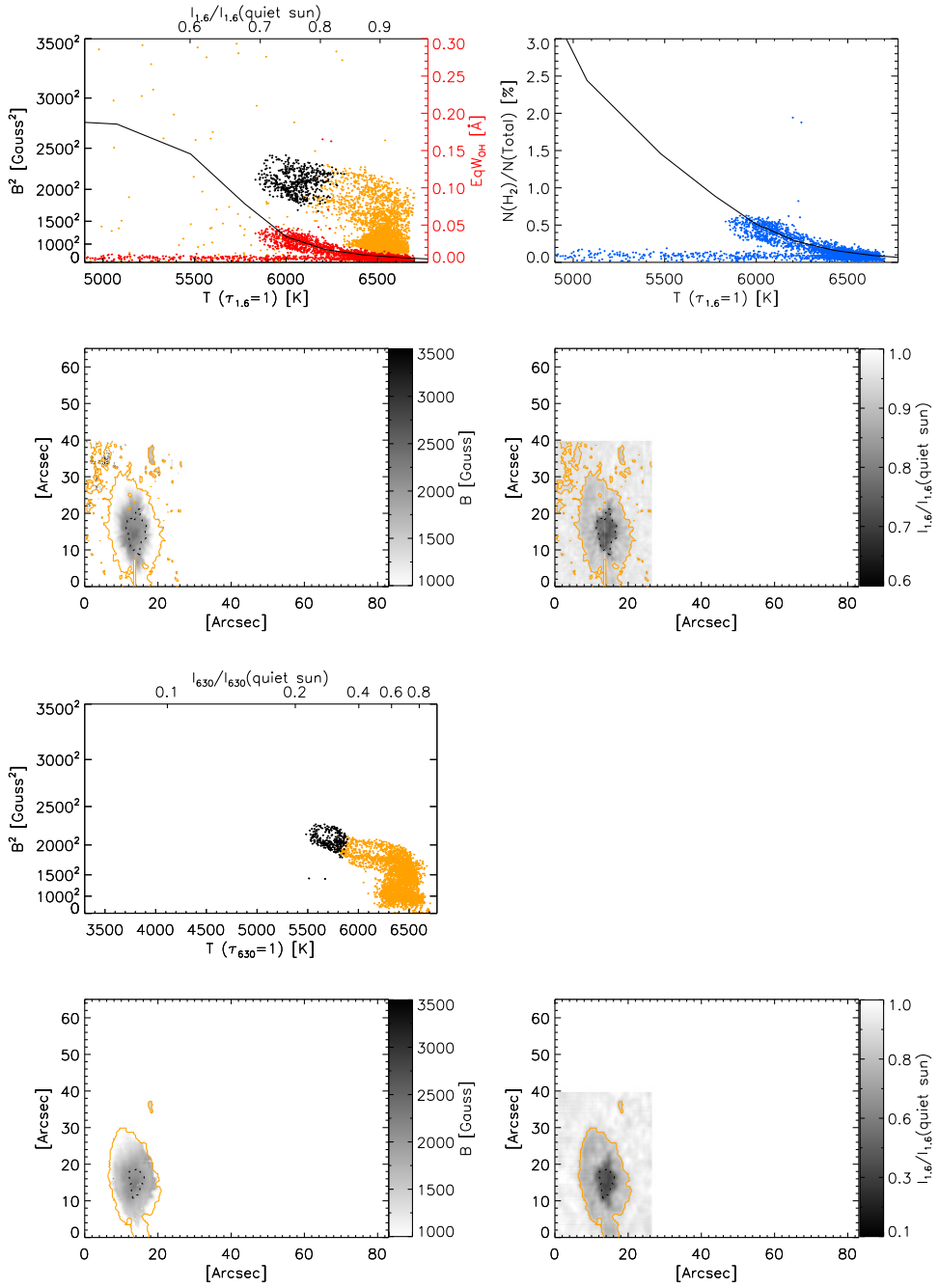

Figure 7.48 Obs. (46): NOAA 11101 observed on 2010-09-04.

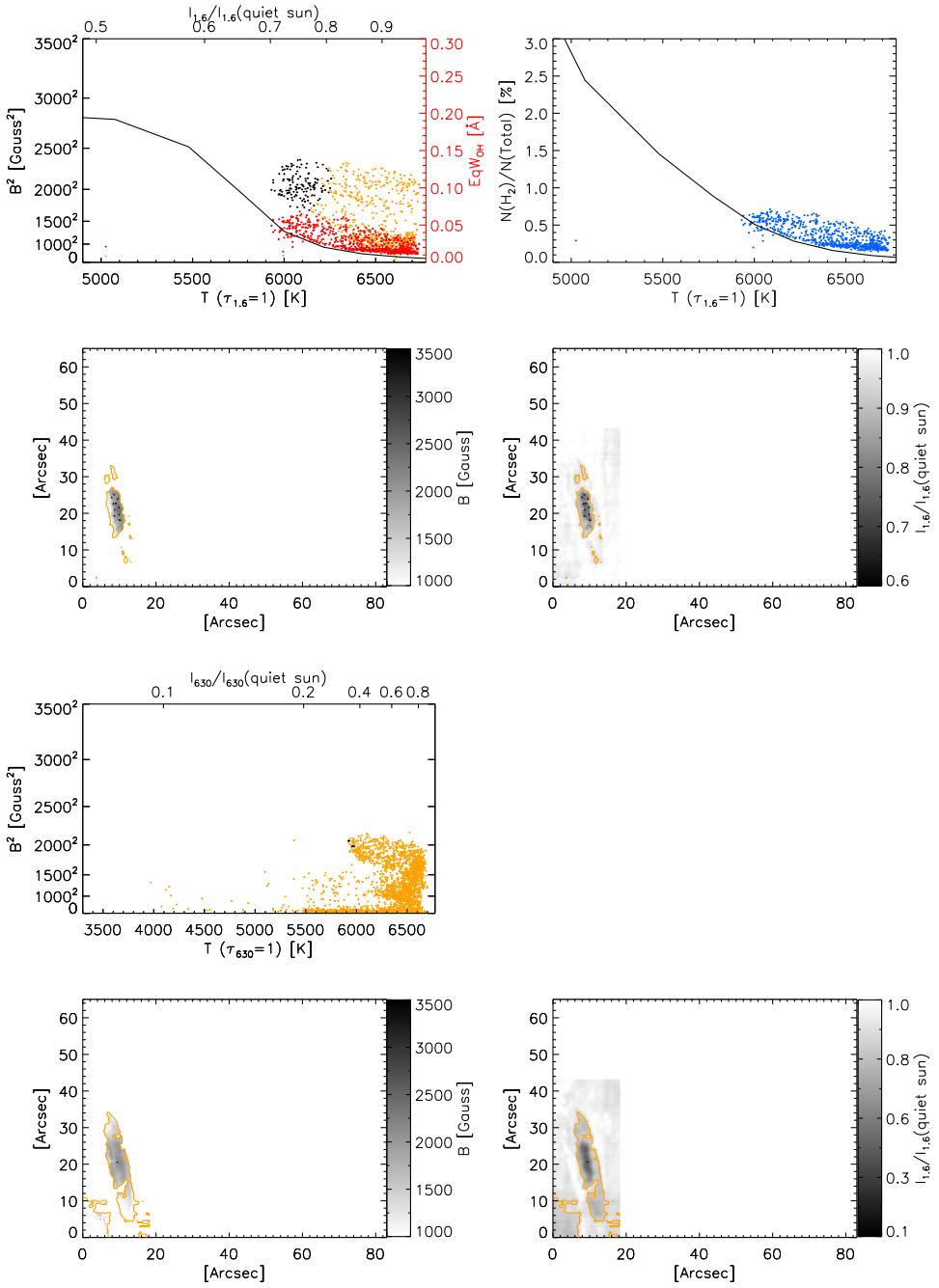

Figure 7.49 Obs. (47): NOAA 11101 observed on 2010-09-05.

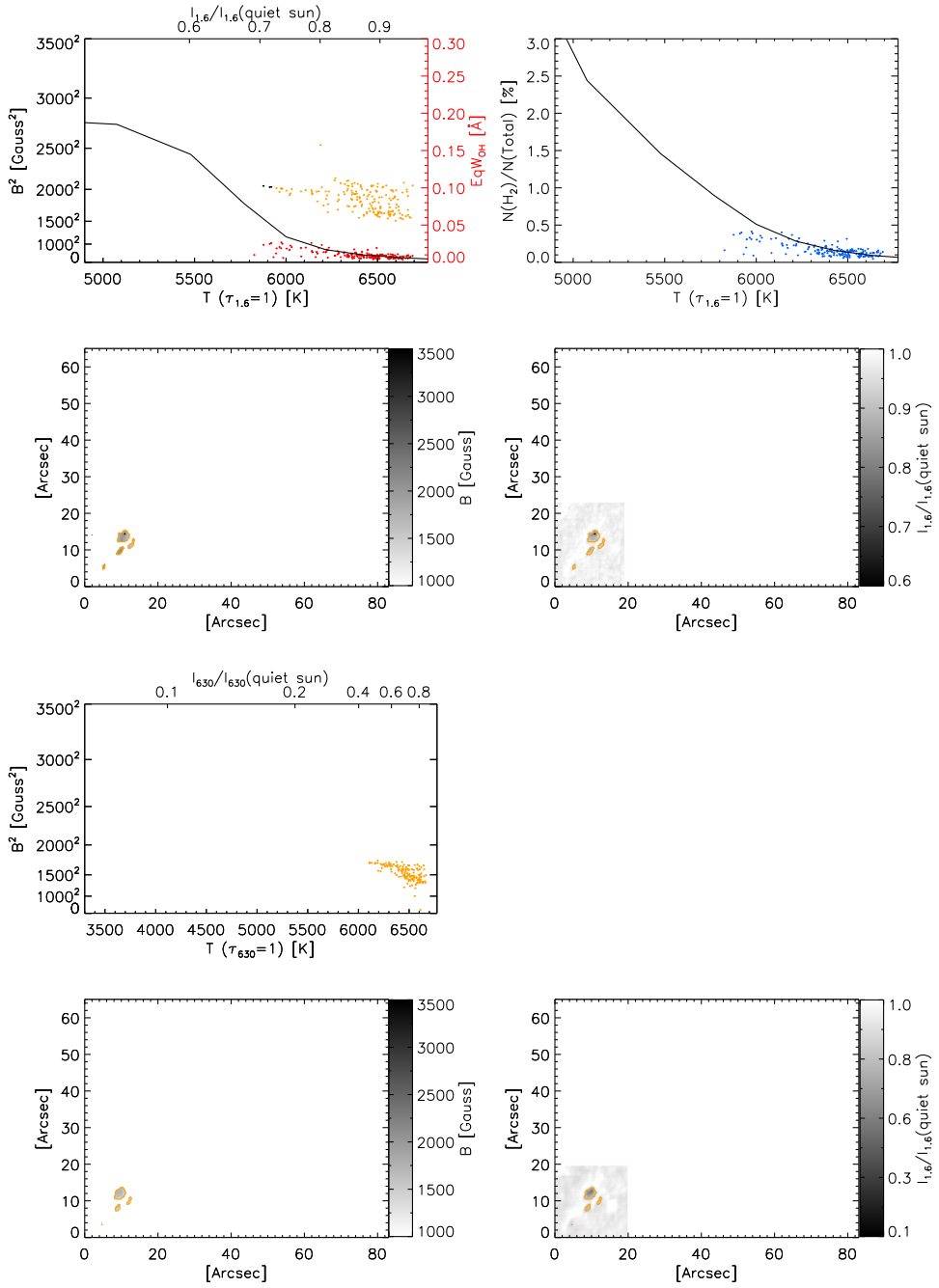

Figure 7.50 Obs. (48): An unnamed pore group observed on 2010-09-12.

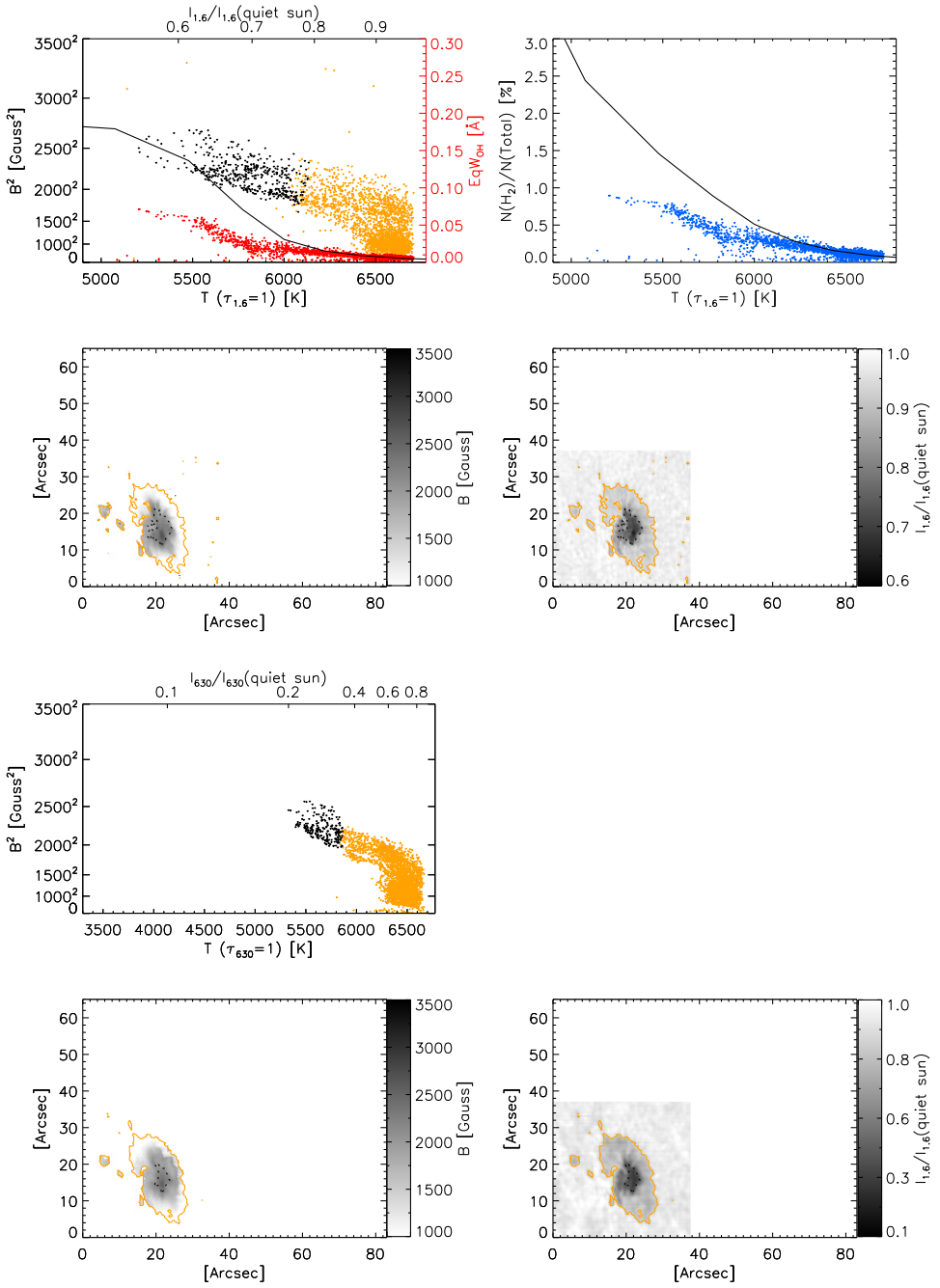

Figure 7.51 Obs. (49): NOAA 11106 observed on 2010-09-12.

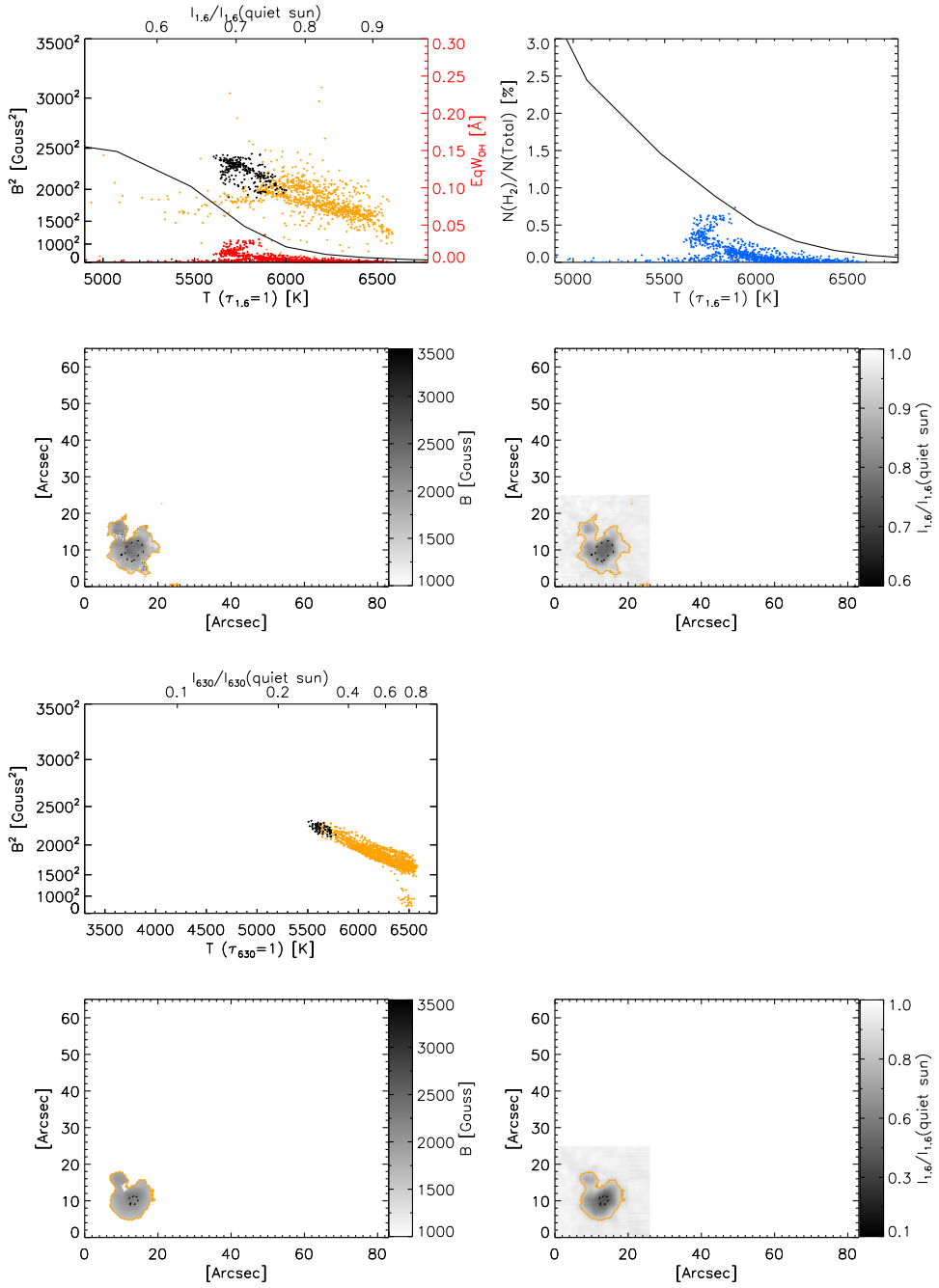

Figure 7.52 Obs. (50): NOAA 11130 observed on 2010-11-28.

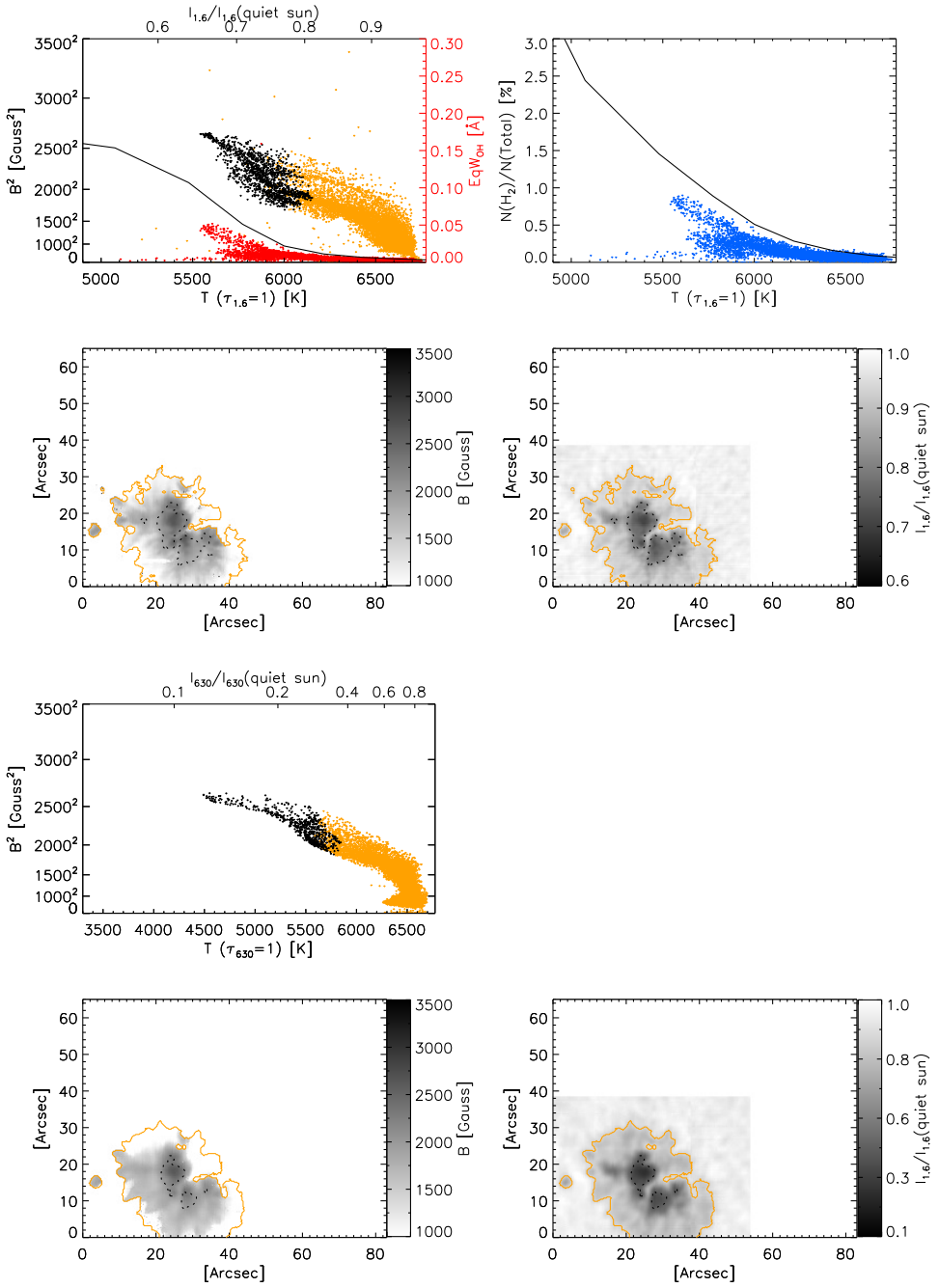

Figure 7.53 Obs. (51): The leading polarity spot in NOAA 11130 observed on 2010-11-30.

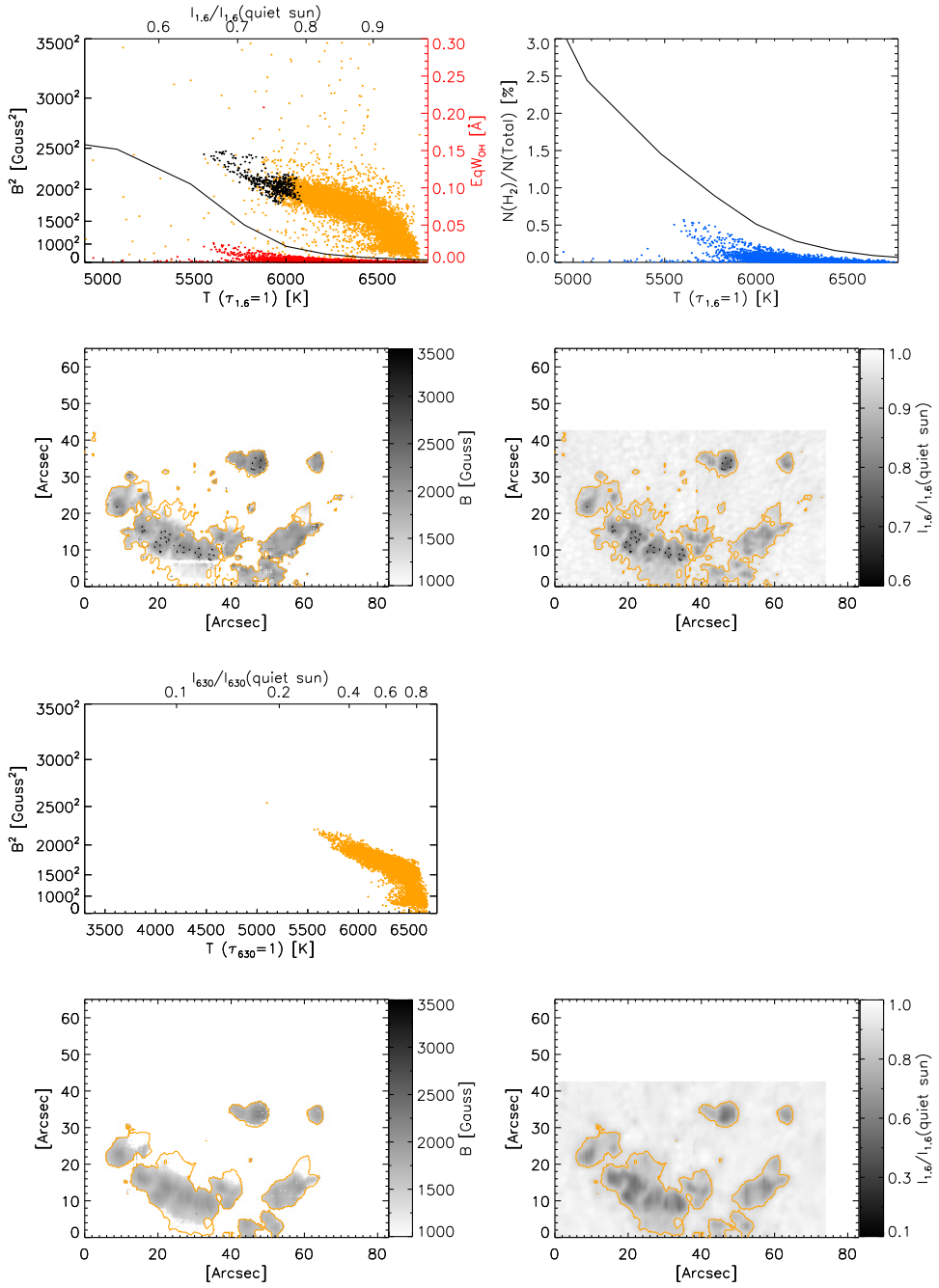

Figure 7.54 Obs. (52): The following polarity spot in NOAA 11130 observed on 2010-11-30.

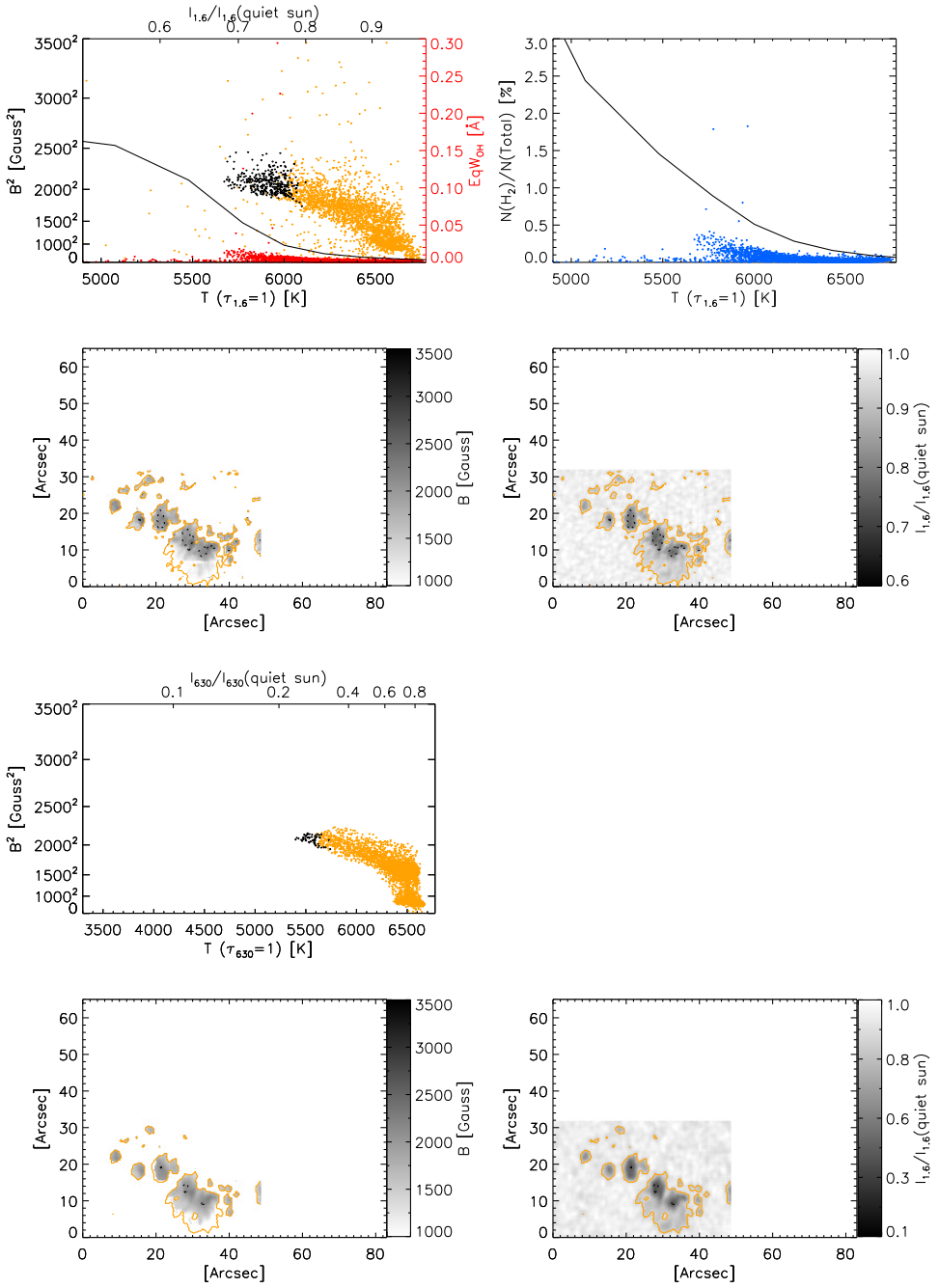

Figure 7.55 Obs. (53): The following polarity spot in NOAA 11130 observed on 2010-12-01.

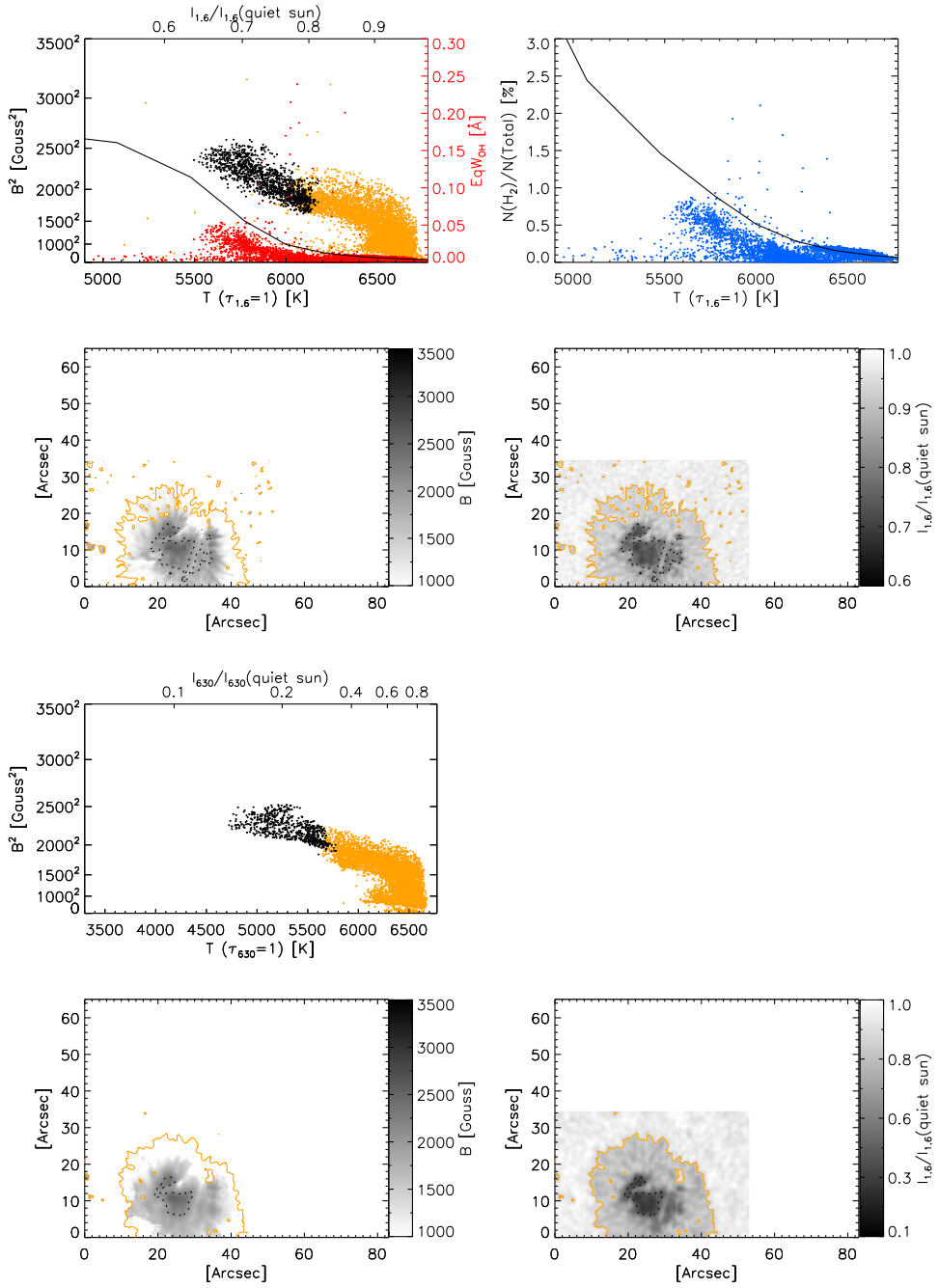

Figure 7.56 Obs. (54): The leading polarity spot in NOAA 11130 observed on 2010-12-01.

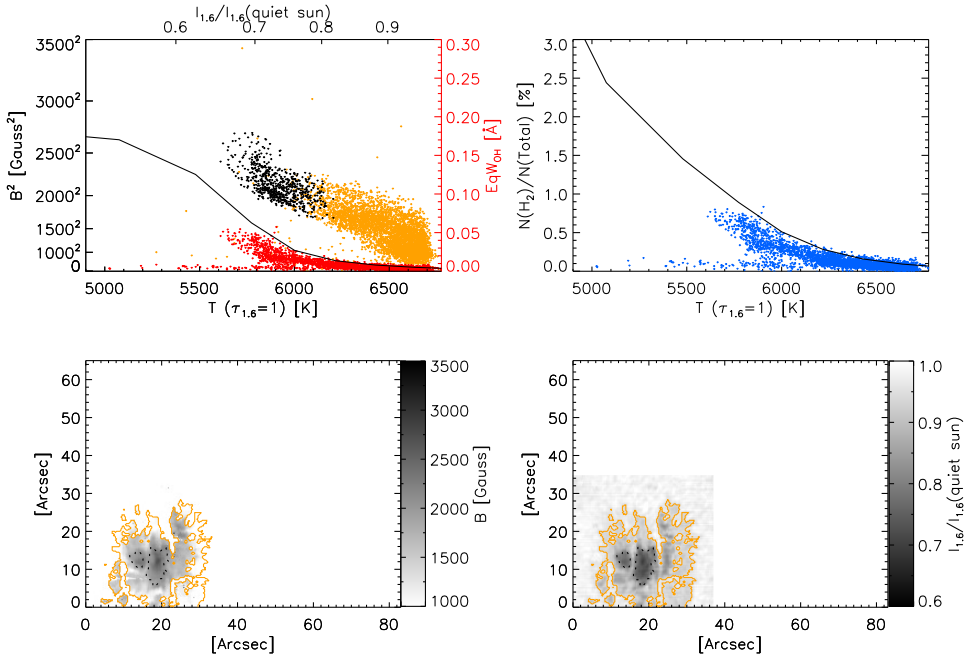

Figure 7.57 Obs. (55): NOAA 11130 observed on 2010-12-02.

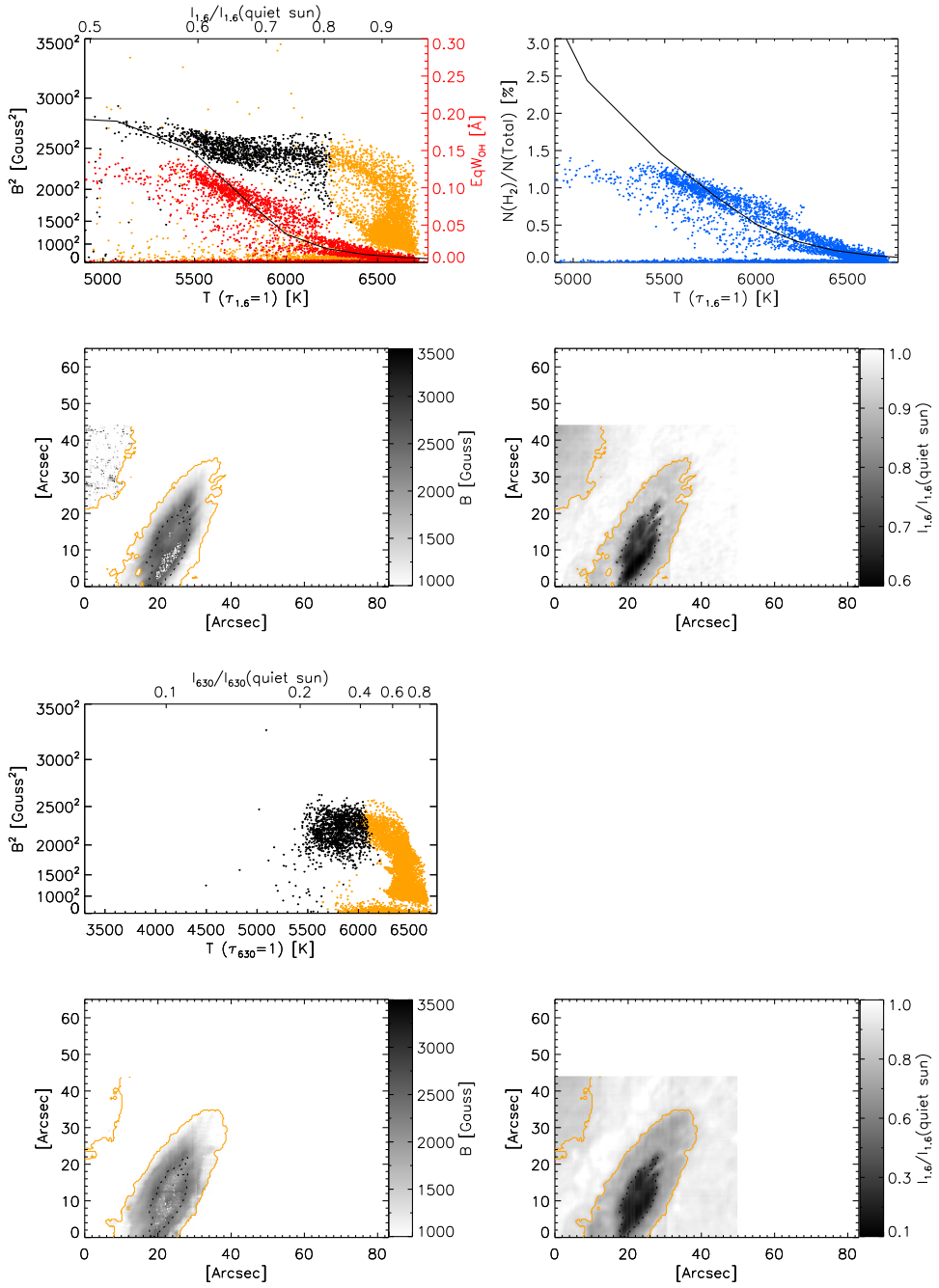

Figure 7.58 Obs. (56): NOAA 11131 observed on 2010-12-02.

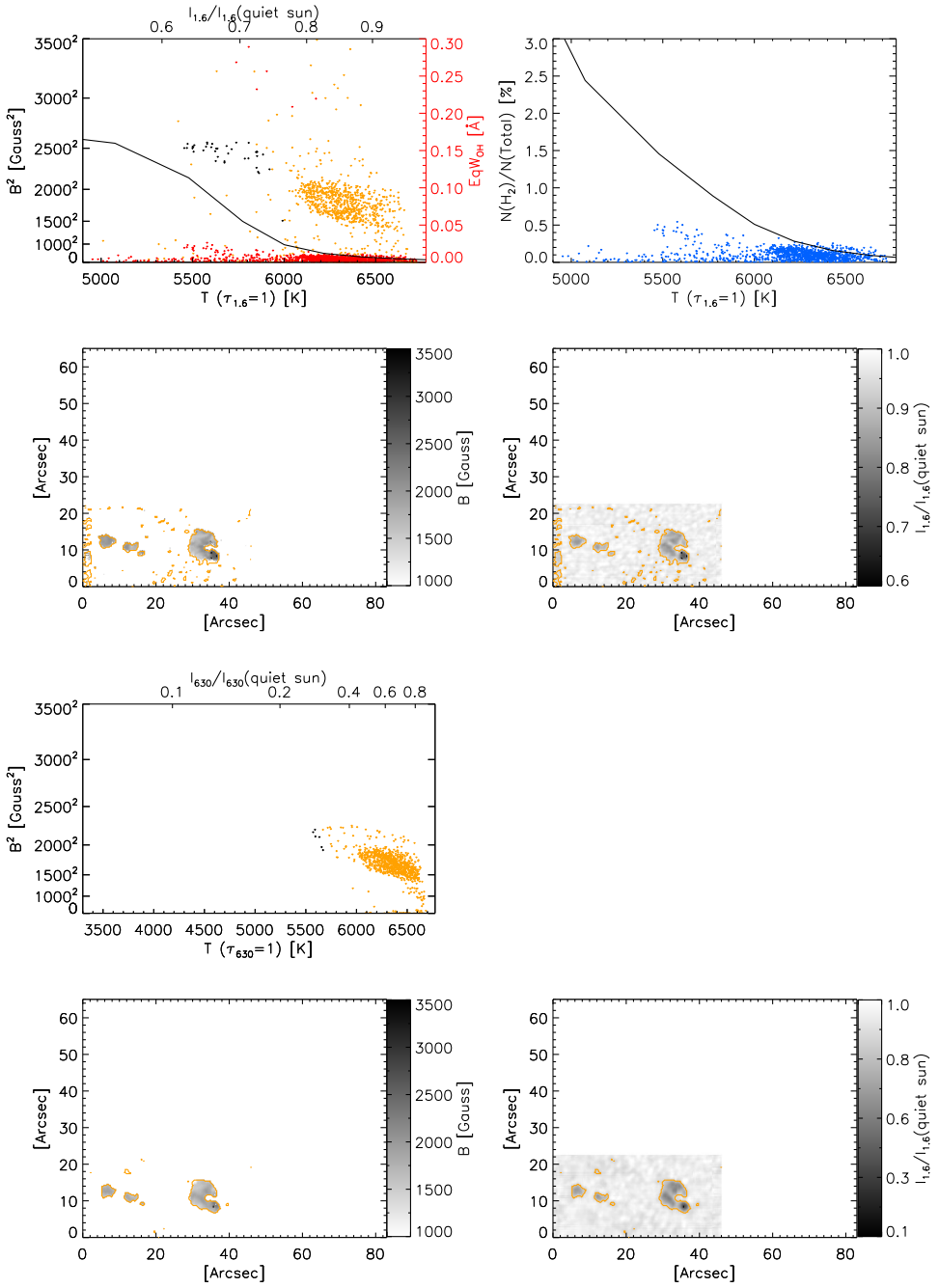

Figure 7.59 Obs. (57): An unnamed pore group observed on 2010-12-03.

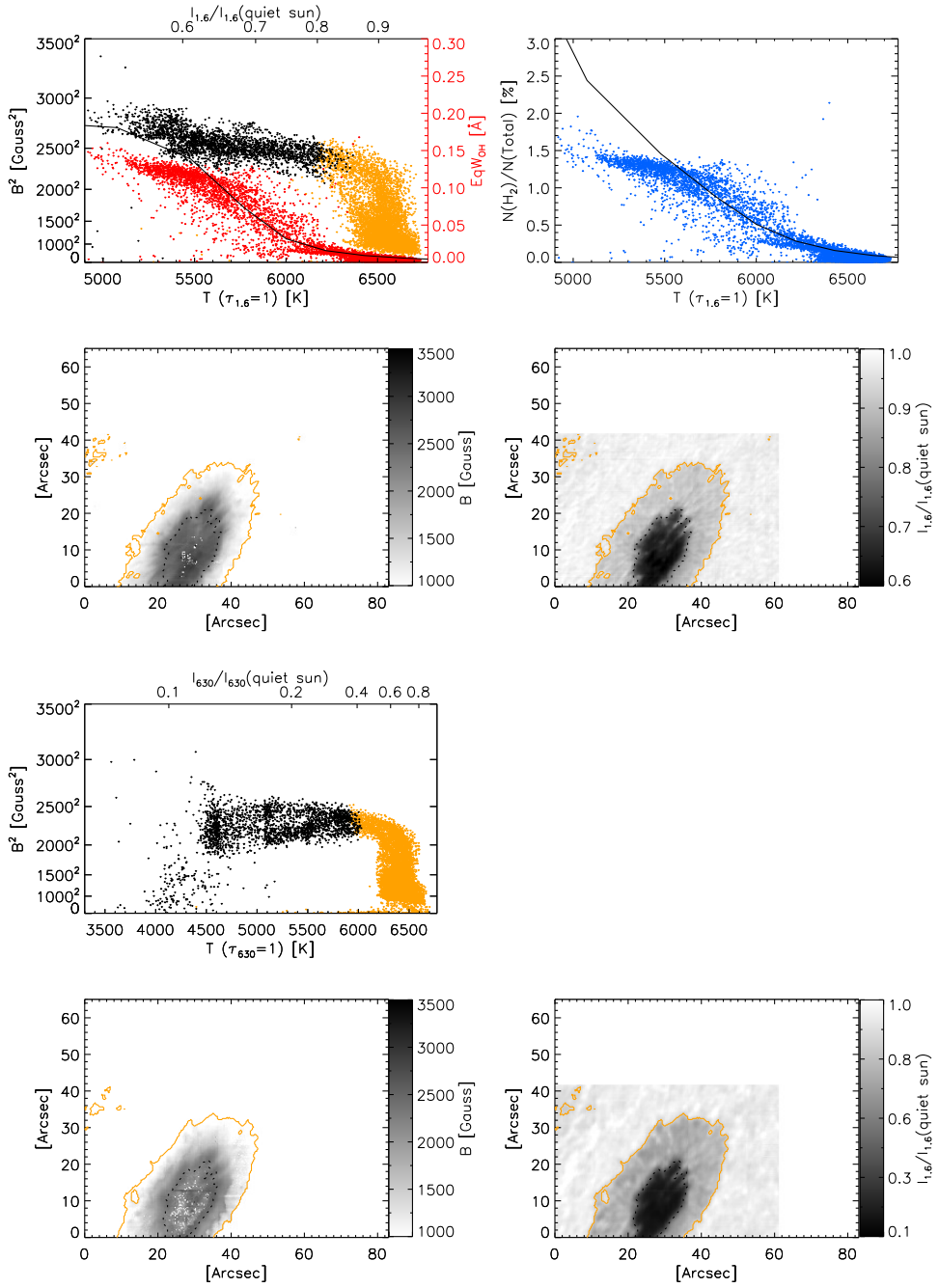

Figure 7.60 Obs. (58): NOAA 11131 observed on 2010-12-03.

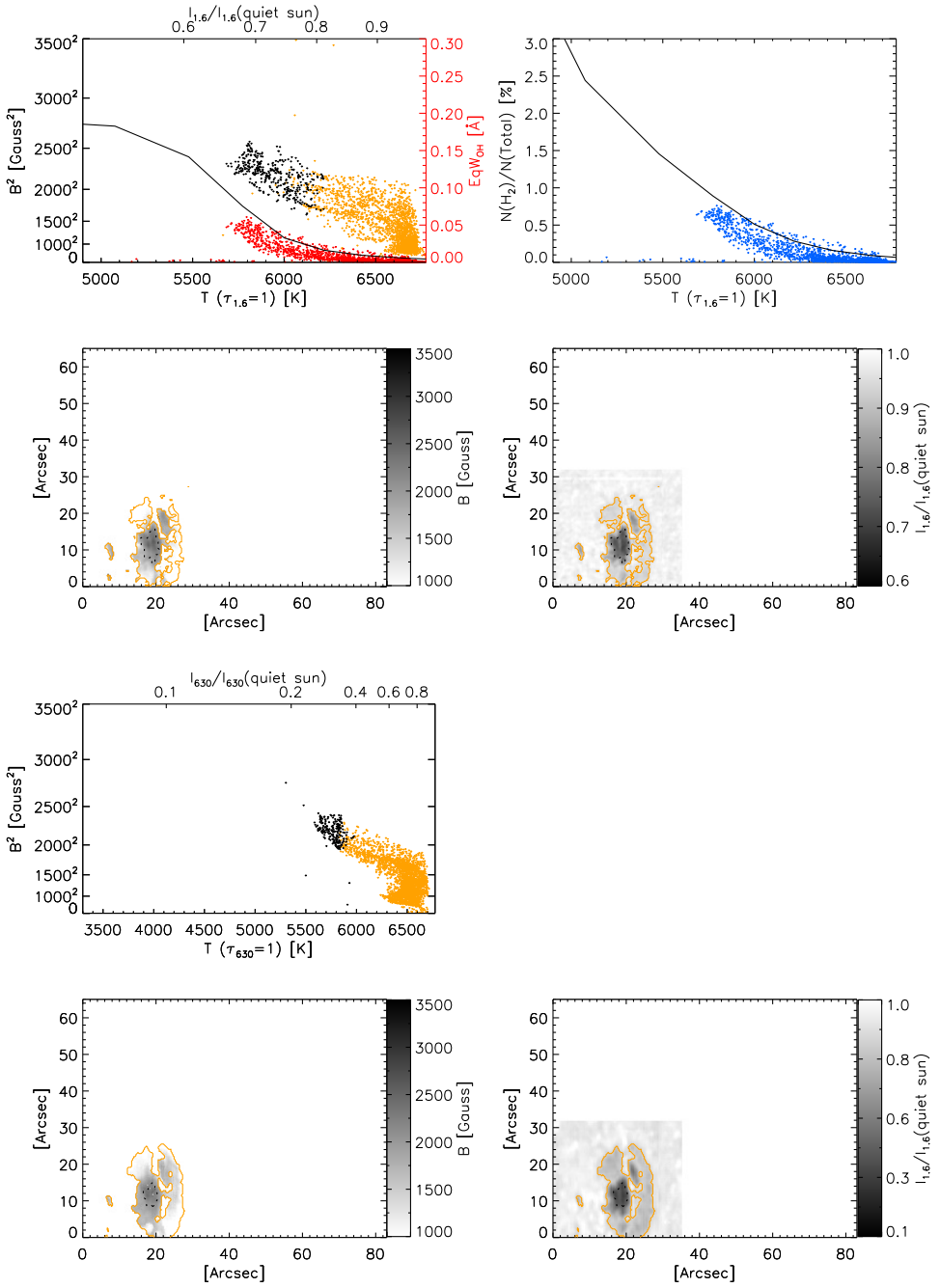

Figure 7.61 Obs. (59): NOAA 11130 observed on 2010-12-03.

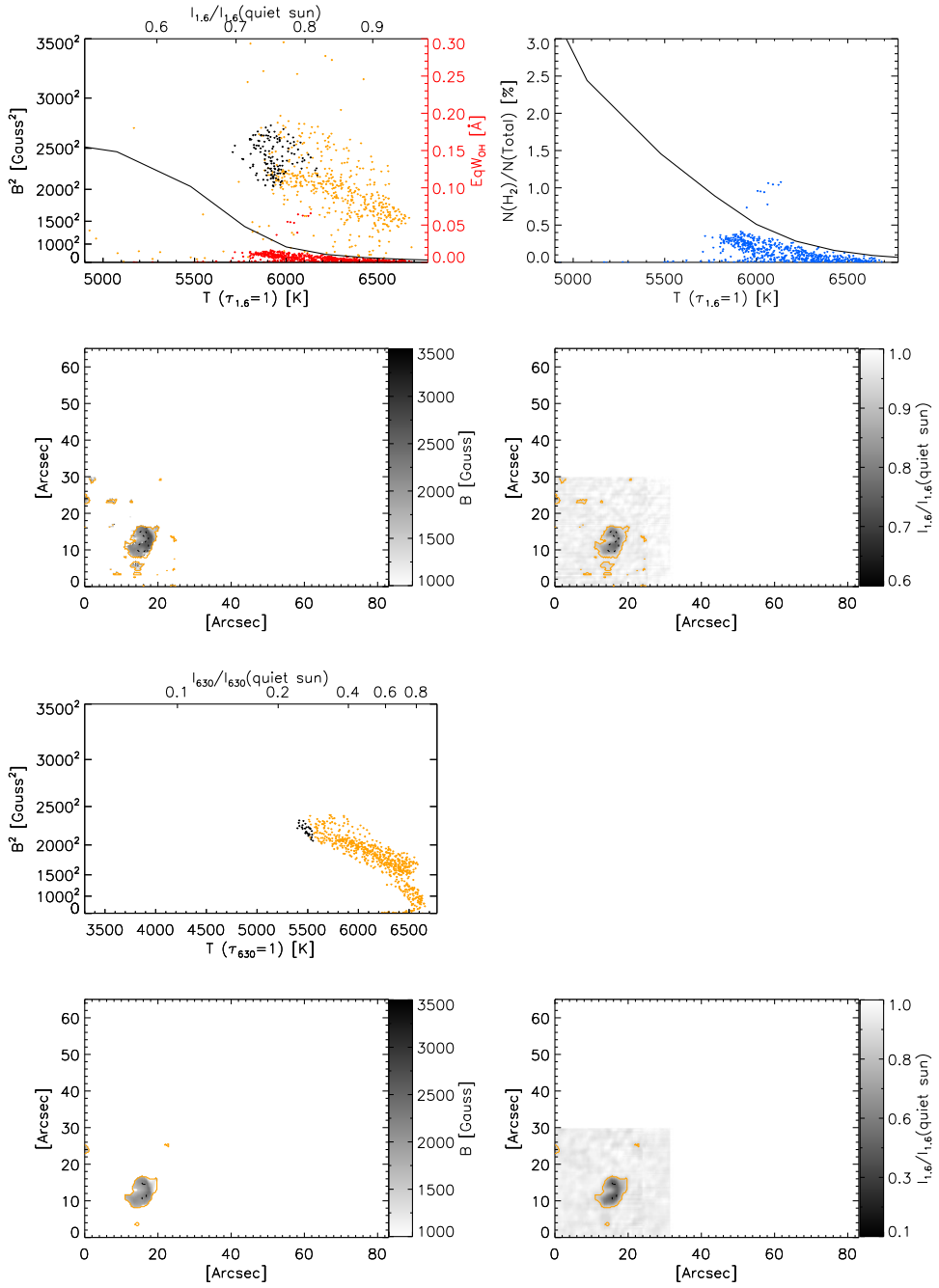

Figure 7.62 Obs. (60): NOAA 11132 observed on 2010-12-04.

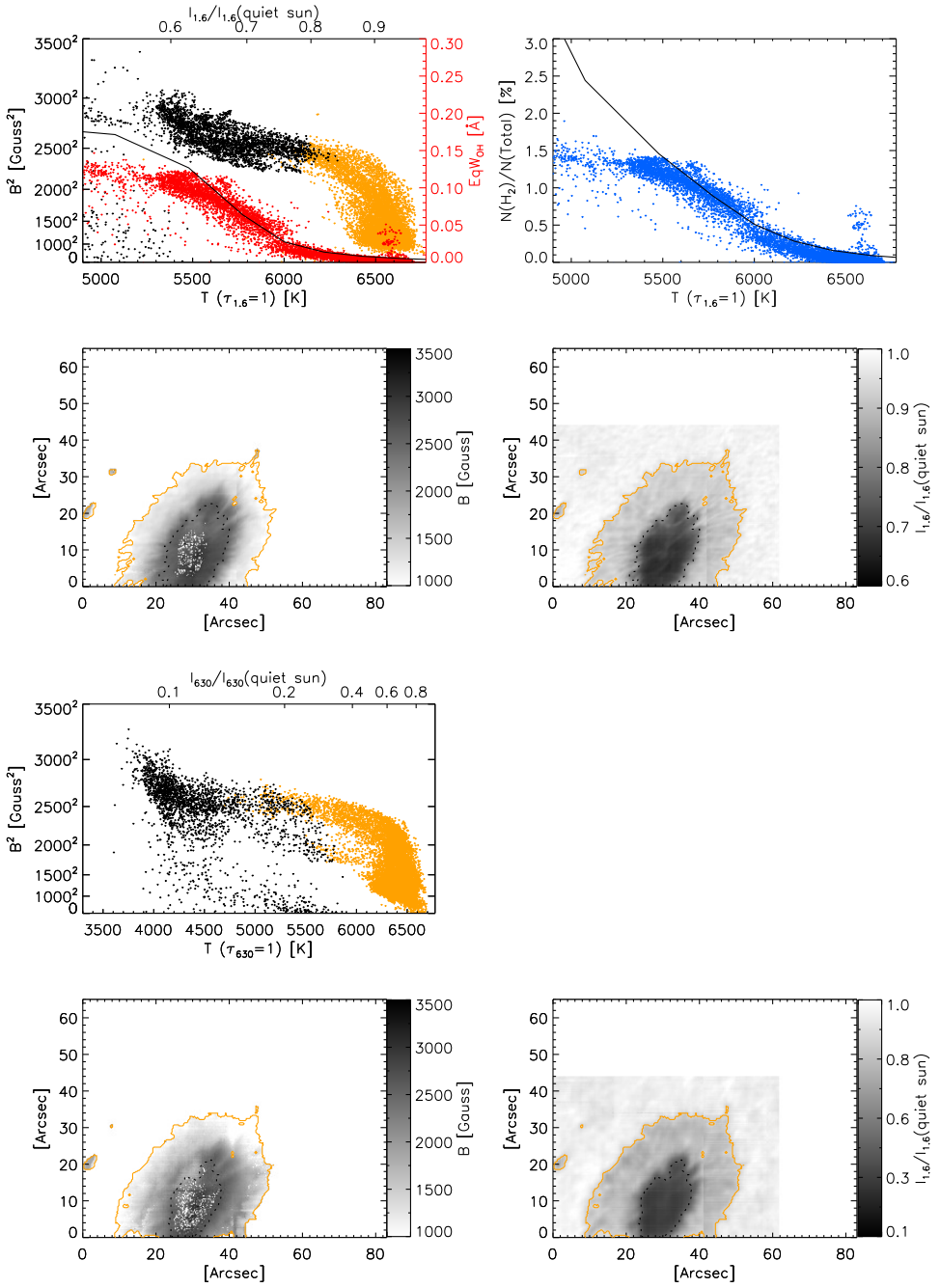

Figure 7.63 Obs. (61): NOAA 11131 observed on 2010-12-04.

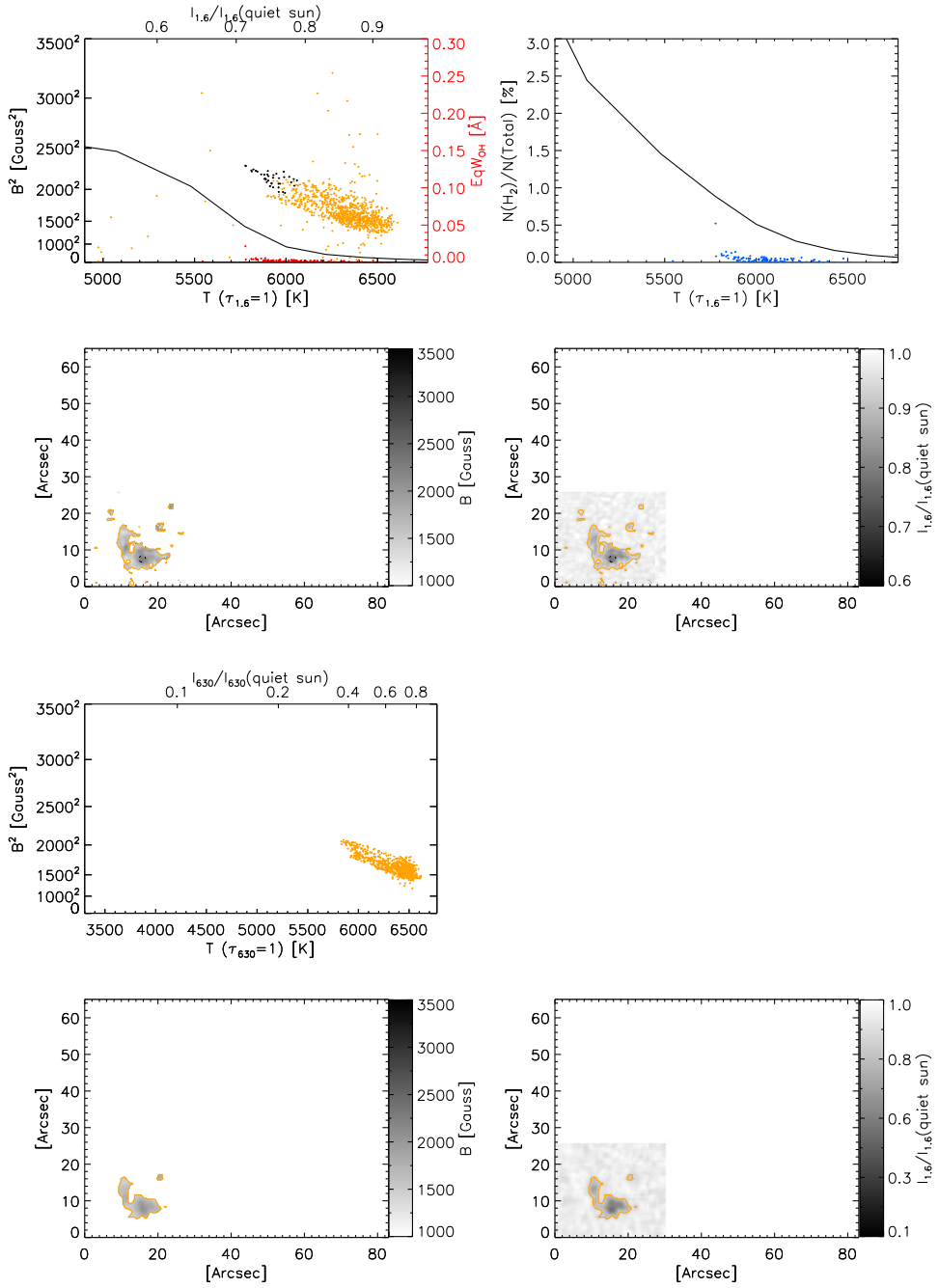

Figure 7.64 Obs. (62): NOAA 11132 observed on 2010-12-05.

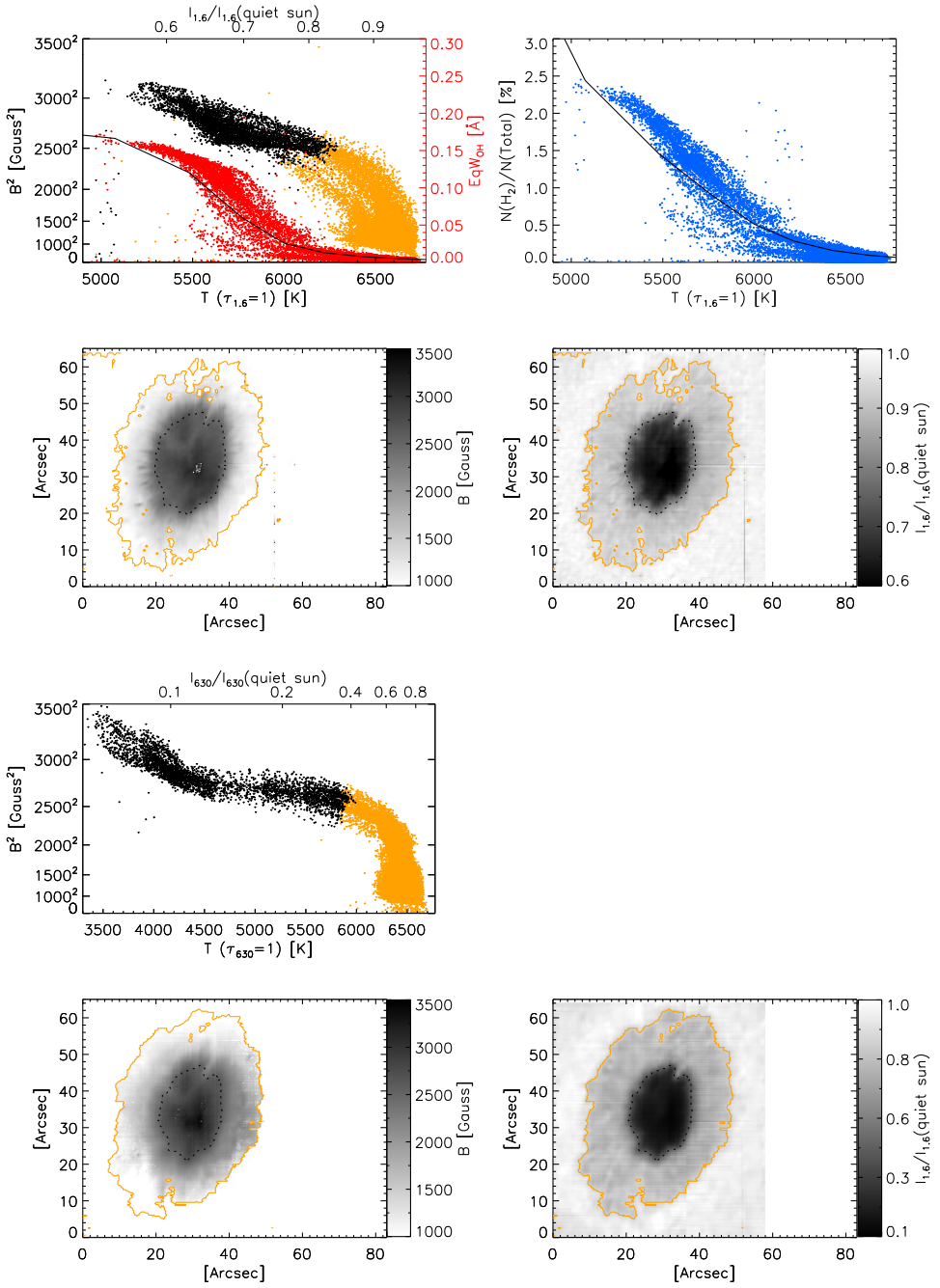

Figure 7.65 Obs. (63): NOAA 11131 observed on 2010-12-05.

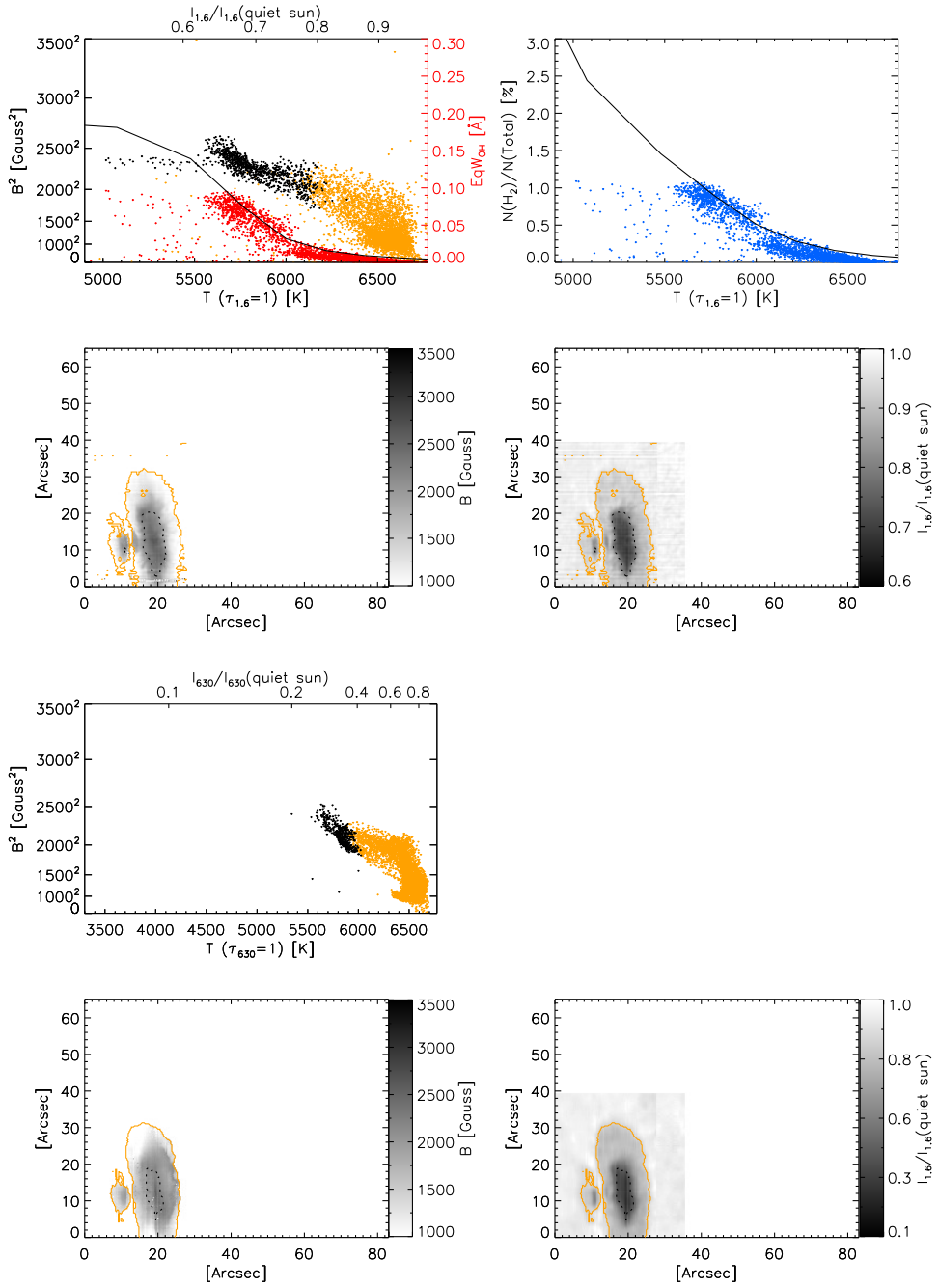

Figure 7.66 Obs. (64): NOAA 11133 observed on 2010-12-05.

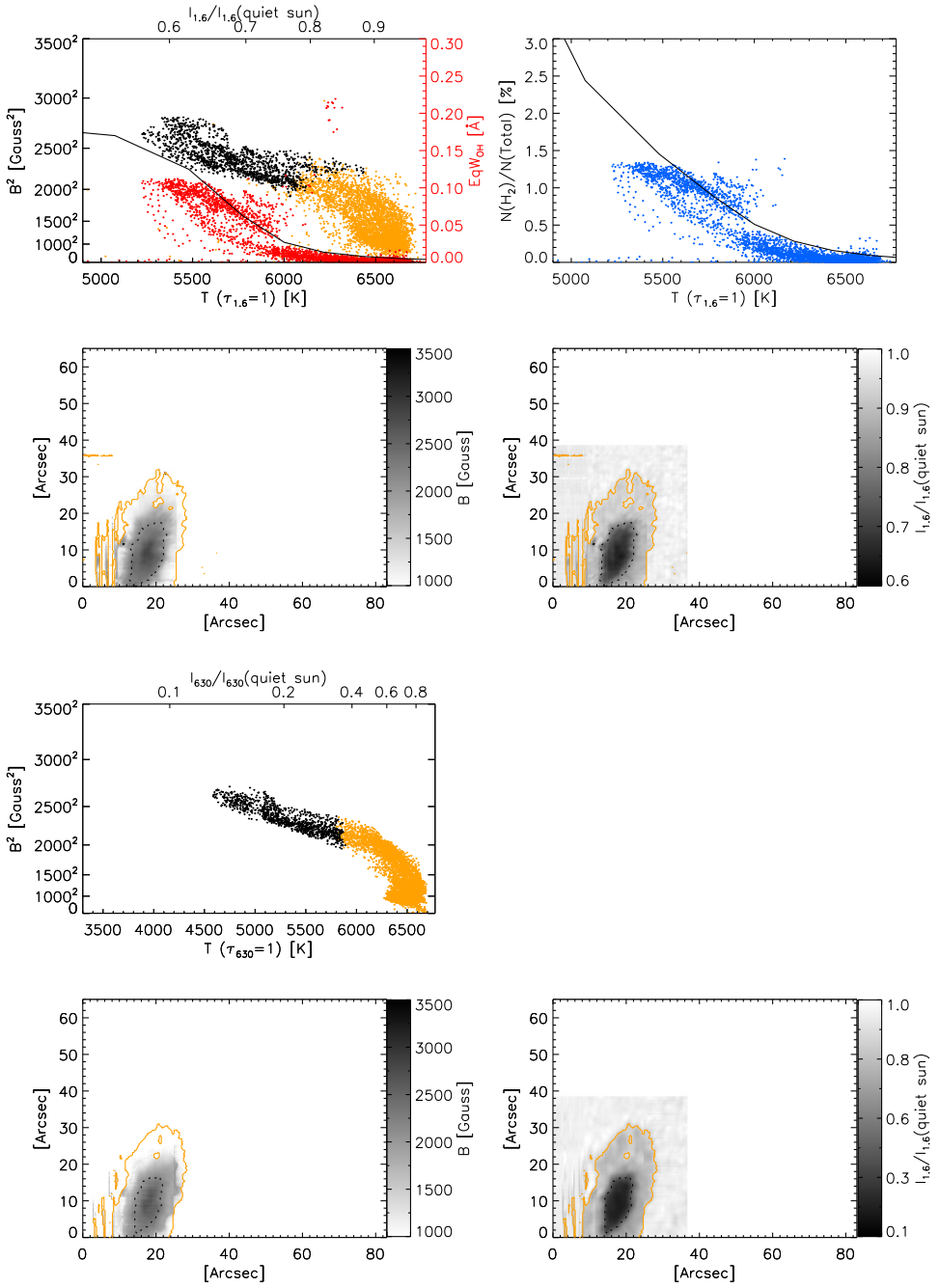

Figure 7.67 Obs. (65): NOAA 11133 observed on 2010-12-06.

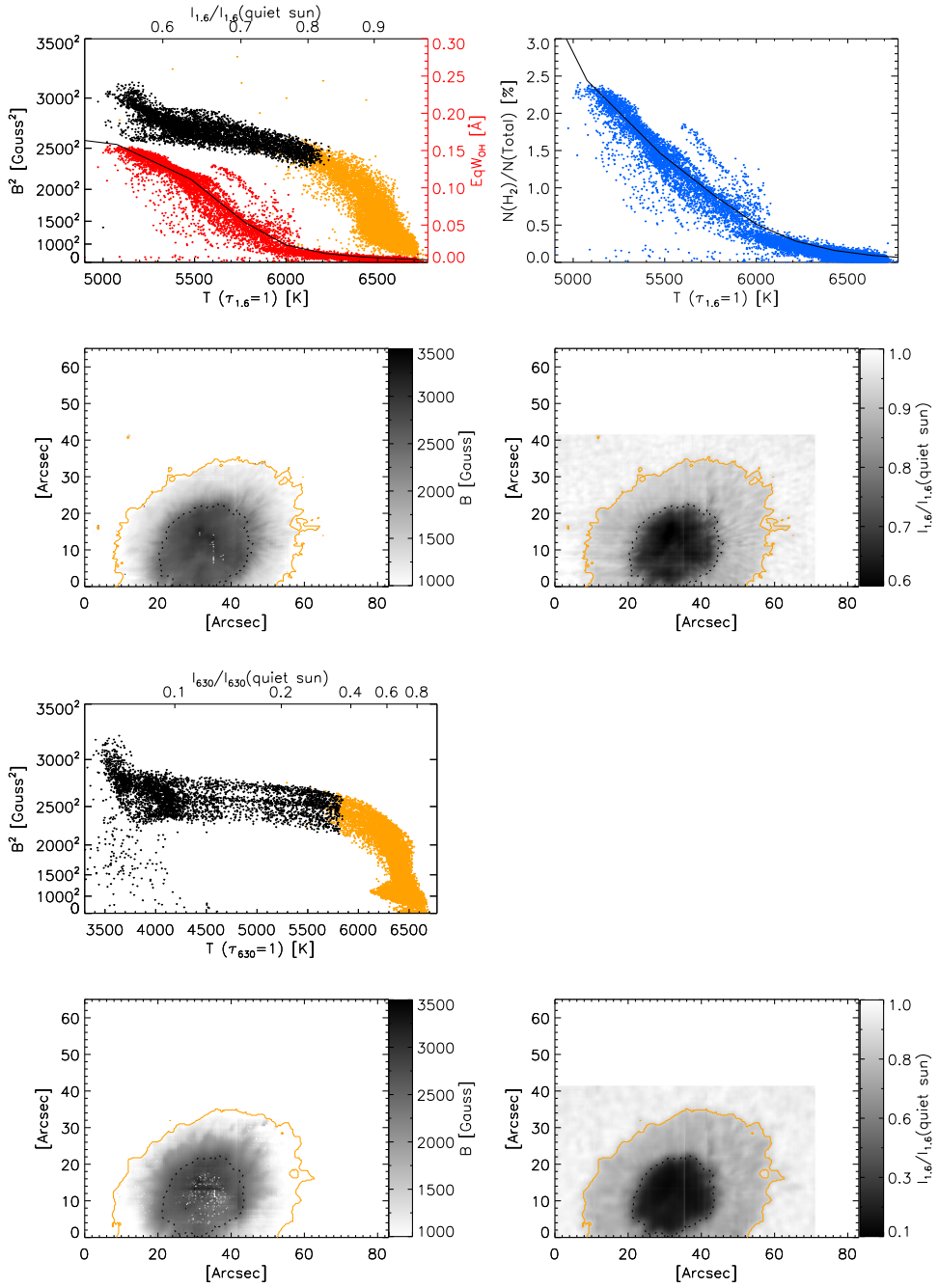

Figure 7.68 Obs. (66): NOAA 11131 observed on 2010-12-06.

Chapter 8

Discussion

Abstract

This survey and analysis provides observational evidence that significant H_2 molecule formation takes place in sunspots which are able to maintain maximum fields of greater than 2500 G. Measurements of the OH equivalent width are qualitatively consistent with predictions from radiative and chemical models and seem to imply a H_2 fraction of a few percent in the largest sunspots. The formation of this small fraction alters the equilibrium of gas pressure and magnetic pressure in the sunspot isothermally, resulting in an increase of the slope of the thermal-magnetic relation at temperatures lower than 6000 K (in the 15650 Å continuum) where the H_2 fraction begins to rapidly increase with temperature. In this chapter I suggest that the formation of H_2 in the sunspot umbra causes a rapid intensification of the magnetic field without a significant decrease in temperature which would explain the increase in slope of the observed thermal-magnetic relation.

8.1 The Role of H_2 in Sunspot Formation

During the initial stage of sunspot emergence and cooling, the formation of H_2 may trigger a temporary "runaway" magnetic field intensification process. As magnetic flux emerges and strengthens, the horizontal overturning motion of the convection zone is suppressed by the vertical magnetic field, and the atmosphere cools to form a sunspot. The gas temperature

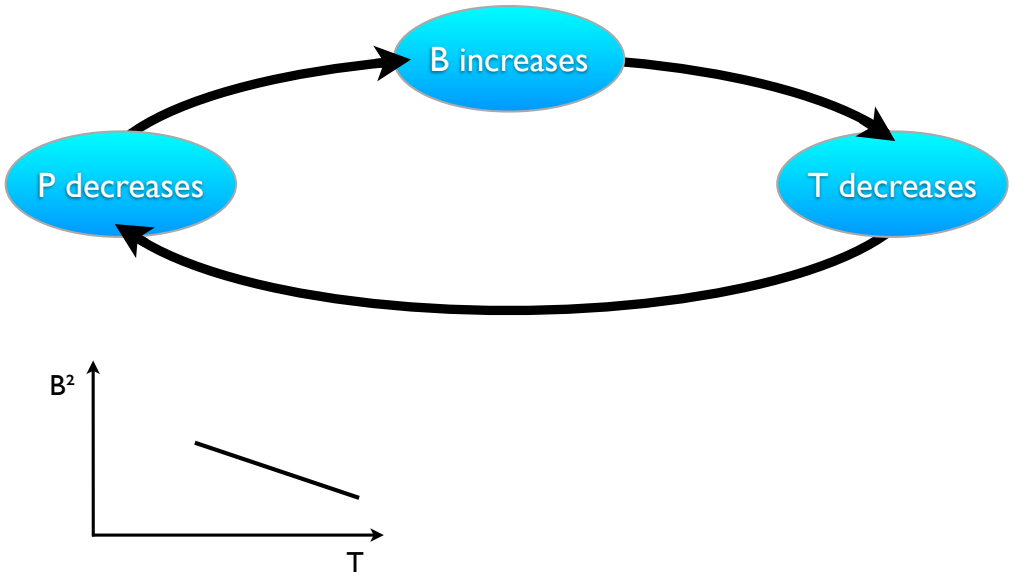

Figure 8.1 A diagram showing the interplay of quantities relevant to the magnetohydrostatic equilibrium in sunspots and an example of the thermal-magnetic relation in an emerging sunspot before the onset of significant H_2 formation.

linearly decreases in response to increases in the magnetic pressure in the sunspot. This initial phase is depicted in the diagram in Figure 8.1.

When sufficiently low temperatures are reached in the sunspot, H_2 begins to form in substantial numbers in the coolest parts of the umbra. The dissociation energy released into the atmosphere is rapidly dissipated by radiative cooling due to the low opacity of the photosphere. The total particle number of the gas is reduced as free hydrogen atoms combine to form H_2. The decrease in the particle number causes a reduction in the total kinetic pressure ($P_g = nkT$) inside the sunspot without a corresponding reduction in temperature, while the gas pressure external to the sunspot remains constant. The decrease in the gas pressure inside the sunspot causes the region where molecule formation is greatest to shrink in size, and due to the high electrical conductivity of the atmosphere the magnetic fields are dragged along with the plasma (the "frozen-in field" effect) as the forces reach a new equilibrium, resulting in an enhanced magnetic field density, as depicted in Figure 8.2.

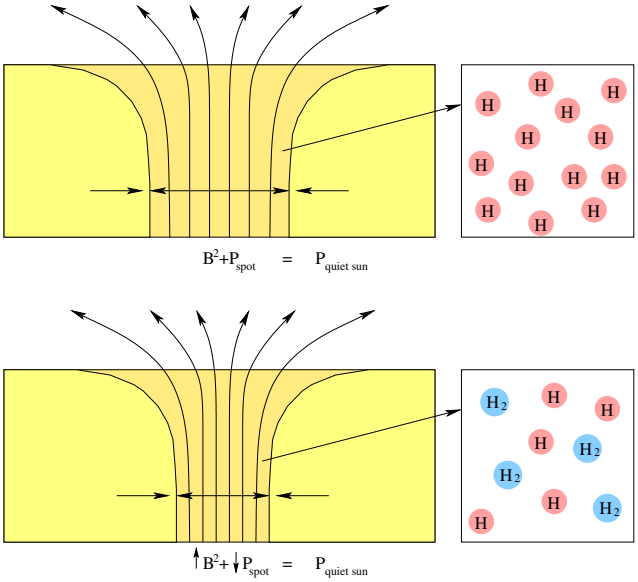

Figure 8.2 A diagram demonstrating the results of molecule formation in a magnetic flux tube.

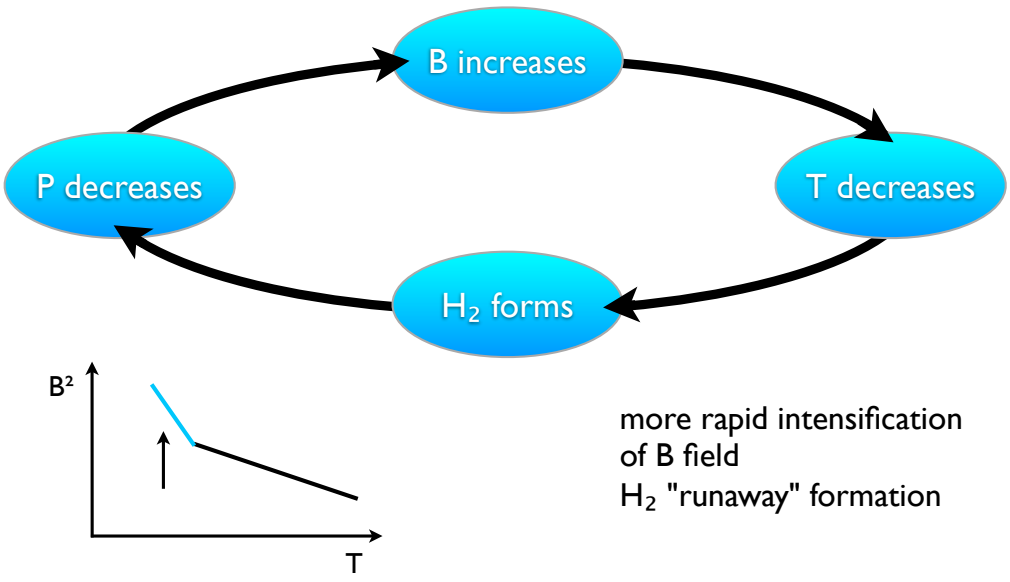

Figure 8.3 A diagram showing the interplay of quantities relevant to the magnetohydrostatic equilibrium in sunspots and an example of the thermal-magnetic relation in a growing sunspot after the onset of significant H_2 formation.

The increase in magnetic field density further inhibits the convective heating of the sunspot atmosphere, which leads to further cooling. The feedback loop caused by molecule formation is depicted in Figure 8.3. This "runaway" magnetic field intensification process is most likely a temporary phenomenon which is arrested by the transfer of radiative energy from surrounding hotter regions, as we have no observational evidence of sunspots which are cool enough, or have large enough magnetic fields to be completely depleted of atomic hydrogen. It is possible that the process will also stop when convection is completely quenched by the magnetic field.

8.2 H_2 and Sunspot Decay

While the formation of H_2 may initiate a more rapid intensification of the sunspot magnetic field during sunspot emergence, I speculate that during the decay phase in a sunspot which has already formed a substantial H_2 population, the highest concentrations of molecular gas would tend to maintain the magnetic field against decay and extend the lifetime of the sunspot. As the magnetic field in a sunspot weakens, regions of the umbra once cool become warm and H_2 dissociates back into atomic hydrogen. The more rapid increase in pressure in warmer regions of the umbra would compress the remaining cool regions, concentrating the magnetic field and maintaining the cool interior of the sunspot against convective heating. A diagram of this process is shown in Figure 8.4.

8.3 H_2 in Equilibrium

For sunspots which can be considered to be neither forming or decaying, the presence of molecules still alters the basic characteristic of the gas in the regions where molecules form. If the lower number density of molecular regions is to be maintained, higher magnetic field strengths are required. Therefore we may expect to see a discontinuity in the thermal-magnetic relation around the temperature where H_2 begins to be a significant component in the sunspot atmosphere.

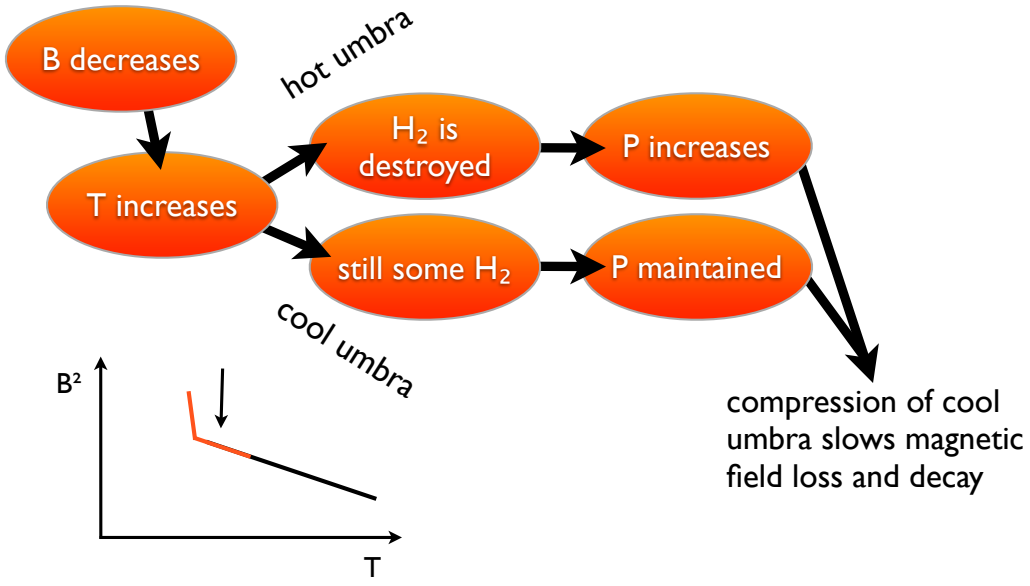

Figure 8.4 A diagram showing the interplay of quantities relevant to the magnetohydrostatic equilibrium in sunspots and an example of the thermal-magnetic relation in a sunspot during the decay phase of the magnetic field and the destruction of H_2.

8.4 Contributions to MHSE From Other Effects

It is possible that two other effects contribute to the non-linearity of the thermal-magnetic relation. In three previous studies (Martínez Pillet & Vázquez 1993; Solanki et al. 1993; Mathew et al. 2004) the non-linearity of the thermal-magnetic relation was interpreted purely as a consequence of the radiative transfer effect, which was described in Chapter 2. In a sunspot, a cooler atmosphere with the same density is more optically thin. In addition the magnetic field provides pressure support which allows the internal pressure of a sunspot to be much less than the external pressure of the quiet-Sun. The magnetic field and continuum measurements therefore originate from a much greater geometrical depth than in the quiet-Sun. In seeing deeper into the solar atmosphere we are able to see relatively hotter regions, and stronger magnetic fields (the magnetic field increases with depth about 0.5 G/km (Kneer 1972)). Therefore the observed temperature would seem to decrease less

rapidly, or equivalently the magnetic field strength would rise more quickly, than in a single geometrical layer relative to (for example) the radial distance. This effect would therefore tend to increase the slope of the thermal-magnetic relation, however it should be largest where the temperature is changing rapidly, at the umbral-penumbral boundary, not in the center of the umbra.

Martínez Pillet & Vázquez (1993), Solanki et al. (1993), and Mathew et al. (2004) assumed that the magnetic field and temperature measurements fully described the equilibrium condition of the sunspots they observed. Neglecting the curvature force and changes in the gas density, they determined the height at which the gas pressure in a quiet-Sun atmospheric model (i.e. Gingerich et al. (1971) and Spruit (1974)) was equivalent to the magnetic and gas pressure internal to the sunspot. The maximum depth determined was 400 km, about 200-400 km less than expected based on limb measurements of the Wilson depression in sunspots. If H_2 is causing an isothermal decrease in the gas pressure in a sunspot, and corresponding increase in magnetic field strength, this would cause the internal pressure of the sunspot to be over-estimated. When comparing to the quiet-Sun model, the authors would infer a greater depth from the over-estimated pressure. Including H_2 formation in the gas model of sunspot would increase the discrepancy with Wilson depression observations.

Through all of this work I have also neglected the curvature force. Increased contributions to the horizontal support from the curvature force in outlying regions may cause the magnetic pressure in the sunspot core to seem boosted, contributing to the non-linearity of the thermal-magnetic relation. The authors mentioned above assumed that the discrepancy between the Wilson depression measured in sunspot at the limb (800 km) and the inferred depth of the sunspot atmosphere (400 km) meant the curvature force in a sunspot must be much greater than previously thought, contributing up to 1/2 the magnetic support. However, it is difficult to imagine that the curvature force in the sunspot umbra should contribute to the magnetic support of the sunspot to the same order of magnitude as the magnetic pressure, and it is difficult to verify the behavior of the curvature force

observationally. Detailed 3D MHD sunspot simulations are necessary to determine the contribution of the radiative transfer effect and the curvature force to the non-linearity of the thermal-magnetic relation.

8.5 Explaining the Observed Sunspots

The behavior of weaker sunspots and pores is easily explained by the fact that they have no H_2. Their equilibrium is controlled simply by the magnetic field and (atomic) gas pressure. Pores are not much affected by the curvature force because their fields are almost entirely vertical, which can be seen from the maps of their vector magnetic fields. Their atmospheres are only 500-1000 K cooler than the quiet-Sun, and the radial temperature gradients so small, that the RT effect does not modify the observed B and T. Therefore their behavior is true to the simplified model of sunspot equilibrium and is entirely linear throughout their evolution as long at they remain above the temperature of significant H_2 formation.

The few examples of sunspots that show an isothermal increase in their umbral magnetic pressure are unambiguous cases of H_2 formation. The rise in B^2 occurs at a constant temperature, therefore it must be occurring at a single geometrical height and cannot be affected by the RT effect. In addition these are small, round spots with simple magnetic field geometries, and can be more appropriately approximated by a magnetic flux tube, so the contribution to equilibrium from the curvature force in their umbrae must also be small. The increases in B in these sunspots occur at a specific temperature where the models predict a corresponding rise in the H_2 fraction. In addition the sunspots are *all* in a decay phase. In a number of these sunspots the sharp increase in the magnetic field is attributable to discrete, high magnetic field cores which are clearly identifiable in the sunspot umbrae. The formation of small cores agrees well with the proposed scenario of destruction of H_2 in sunspots, where small regions of H_2 can remain cool due to strong, compressed magnetic fields. This may indicate that H_2 concentration occurs in sub-arcsecond discrete cores at

or below the resolution limit of the observations, and in these features the magnetic fields are stronger, the temperatures are lower, and the H_2 fractions are larger that what can be observed. These spots are at a cusp. They are small enough, warm enough, and their magnetic fields are weak enough that the gas pressure in the spot is much larger than the magnetic field, while still being cool enough to have H_2 still existing in small cores. Even the 1% alteration in the sunspot gas pressure caused by the molecules requires a large change in the magnetic field to maintain equilibrium, which is why they are so prominent in the survey.

Sunspots in a growth phase might show the same behavior at a certain size as they quickly pass through the onset of molecule formation. The emergence of flux occurs relatively quickly and due to observational bias, by the time an emerging sunspot is large enough to be noticed and observed it has already passed this stage. However the decay of sunspot magnetic field occurs slowly in spots which have already been identified and are under observation.

In sunspots which are already large and cool, far from the initial onset of H_2 formation or its complete destruction, in states of growth, decay, or quiescence, the formation and destruction of H_2 necessitates a change in temperature in addition to an increase in the magnetic field strength. The result is an increase in the slope of the thermal-magnetic relation where molecules have formed relative to the regions where they have not. In *all* of the larger sunspots in the sample and increase in slope is apparent, usually occurring around continuum temperatures 5800-6000 K coincident with the temperature of major H_2 formation discussed in Chapter 6.

The behavior of the largest sunspots is more difficult to understand. Although they should have large H_2 populations which strongly affect the observed sunspot equilibrium condition between B^2 and T, increases in the slope of their thermal-magnetic relations around 5800-6000 K are not apparent. Instead they show increases in the magnetic pressure at temperatures of 5500 K or higher. In addition the slopes of the thermal-magnetic relation in the linear regime are much shallower, and occur at larger magnetic pressure. The effects

of radiative transfer are more of a concern in large sunspots that show steeper temperature gradients through the penumbra and umbra. Due to their size and complex geometry the effects of the curvature force may be greater. However, in the very coolest regions of the umbra in these sunspots the assumption of a small RT effect and small curvature force should still be valid. The H_2 scenario presented here does not adequately explain them. Because we have so few observations of very large sunspots it is necessary to carry out additional investigation into the equilibrium conditions for these extreme cases in order to fully understand the sunspot equilibrium problem.

8.6 Other Effects of H_2 Formation

8.6.1 Alteration of Umbral Heat Capacity

The formation of H_2 appears to be the most likely cause of the sharp increase in slope of the thermal-magnetic relation in sunspots. In addition to this magnetic intensification process, the formation of molecules increases the heat capacity of the sunspot atmosphere. For an ideal atomic gas at constant volume, the heat capacity is $C_v \propto 3/2N$ where N is the particle number. Due to the additional degrees of freedom of the H_2 molecule the factor of 3/2 increases to 7/2 (for both rotational and vibrational modes). Therefore, the formation of a H_2 fraction of 10% would ideally raise the heat capacity of the gas by 13% over an equivalent number density of atomic gas particles. This non-thermal reservoir for energy may have an important effect on the local radiative output of the Sun. Consequently, modeling of the MHS equilibrium condition of sunspots in the form of Equation 2.1 must include a multiple-component atmospheric model with the proper equation of state to account for the altered thermodynamics of the sunspot atmosphere due to the formation of H_2.

8.6.2 Neutral Winds in Sunspots

It has been hypothesized that there may exist a neutral wind in sunspots flowing from the umbra to the penumbra and quiet-Sun (Kuhn & Morgan 2006). Neutral species are

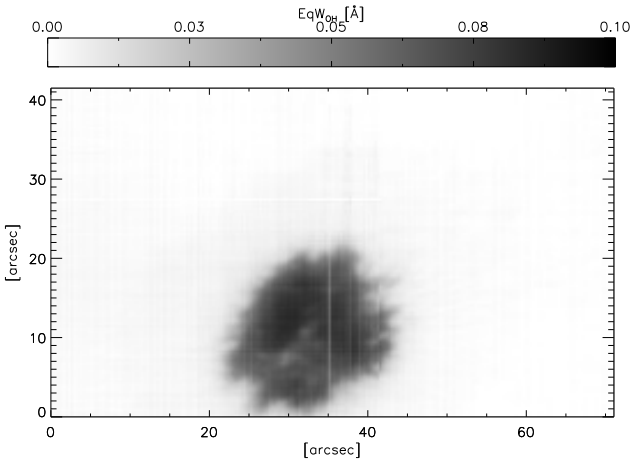

Figure 8.5 A map of the OH equivalent width in the large sunspot NOAA 11131, Obs. 66.

not bound to follow the magnetic field lines, which permits motion perpendicular to the direction of the magnetic field. Kuhn & Morgan (2006) hypothesize that an over-abundance of neutral species in the umbra leads to diffusion of the neutral population into non-neutral regions where they become ionized or dissociated which produces a continuous outflow. It's believed that this effect should be visible as an over-density of neutral species at the umbra-penumbra boundary in a sunspot.

The survey of sunspots that I have obtained seems an ideal opportunity to answer this question. As neutral species, both H_2 and OH should participate in ambipolar diffusion. A map of the OH equivalent width in a large sunspot is shown in Figure 8.5. No obvious strengthening of the OH equivalent width is present at the umbral-penumbra boundary. However, the observations neither confirm or deny the effect of ambipolar diffusion in the sunspot penumbra. An over-density of H_2 would be difficult to observe at the umbra-penumbra boundary precisely because the RT effect there is strongest, and the equivalent width of the OH lines is small compared to the uncertainty. The H_2 fraction by number in the photosphere is also extremely small, and the signature of the neutral wind may not be detectable. This problem may also be answered definitively by realistic MHD sunspot models which account for a molecular gas component.

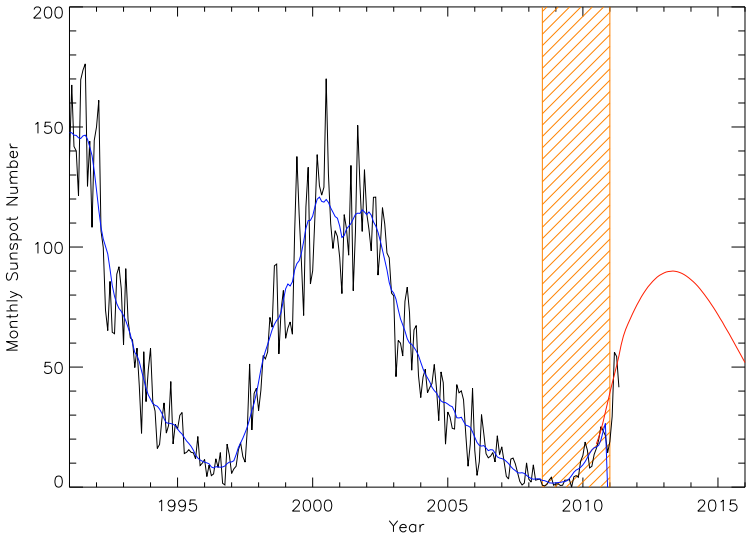

Figure 8.6 Monthly (black) and smoothed monthly (blue) sunspot numbers from 1991 to 2011, and the revised prediction for cycle 24 (red). The orange shaded region indicates the period of time over which the FIRS sunspot survey was conducted. Sunspot data comes from the NOAA Space Weather Prediction Center.

8.7 The Current Solar Cycle

The observations taken during this dissertation were obtained sporadically through the minimum phase of solar cycle 23 and the beginning of solar cycle 24. The sample does not contain very large sunspots with very high magnetic field strengths or many complex configurations. Indeed, the current solar cycle may be the lowest level of activity seen since the early 20th century, and is certainly the deepest minimum in modern memory. The activity levels for the current solar cycle have continued to fall below predicted levels. Figure 8.6 shows the monthly sunspot number from 1991 to 2011, from the peak of solar cycle 23 to the onset of solar cycle 25, the black line shows the data, the smoothed data is shown by the blue line, and the current prediction for cycle 23 is shown in red.

Due to the unexpected decrease in the activity of the current cycle, the long term activity cycle of the Sun has come under close scrutiny. Indicators from a wide variety of techniques seem to indicate that solar activity during the next solar cycle will be very low

or nonexistent. Measurements of the Sun's global vibrational modes indicate that certain modes associated with the strength of next cycle are weak or missing (Hill et al. 2011). High precision 15650 Å intensity spectra of sunspots taken over a long baseline from 1998 to 2005 show that during the past activity cycle sunspot magnetic fields have gotten weaker, and sunspot umbrae are becoming brighter and hotter (Penn & Livingston 2006; Livingston et al. 2011) independent from the normal solar activity cycle. Penn & Livingston (2006) measure a 52 G/year decrease in sunspot magnetic fields, an increase in temperature from 5100 to 5700 K, and a corresponding decrease of about 1/2 in the OH line depth, which is in agreement with the modeling conducted for this dissertation. Extrapolating this trend to the next solar cycle, they predict that solar magnetic fields at the surface may be too weak to form sunspots at all.

This prediction of decreasing sunspot magnetic field strength makes the work carried out in this dissertation of particular importance. As sunspots become weaker, and warmer, the fraction of H_2 formed in the umbra will be correspondingly less and it will no longer be possible to study the effect of H_2 on sunspot evolution. A continued observational effort as solar cycle 24 enters its maximum phase is of particular importance in order to understand the horizontal force balance in the largest sunspots with darker, cooler umbrae.

Chapter 9

Conclusions

Abstract

During the course of this dissertation research I have brought together the results from many high quality spectropolarimetric observations of magnetic and molecular diagnostics in sunspots obtained with FIRS, a new instrument. I have analyzed them with an original Milne-Eddington inversion code and carried out analysis with atmospheric models in order to definitively confirm the presence of molecular hydrogen in sunspots. I have accomplished what no other work on this subject has yet attempted, to describe the equilibrium conditions in a comprehensive sample of sunspots and pores, to prove the importance of the role that molecular hydrogen plays in sunspot evolution, and fully explain the equilibrium state of sunspots. In this chapter I summarize the results of this dissertation and discuss future avenues of research that now warrant investigation.

9.1 H_2 and Sunspot Magnetohydrostatic Equilibrium

The comparison of observations and model atmospheres allow us to infer a H_2 fraction of up to at least 3% in the largest sunspots in the sample.

The decrease in particle density due to the formation of H_2 molecules forces the magnetic field to intensify more rapidly during sunspot formation. During the decay phase in a

sunspot H_2 prevents magnetic field loss. As sunspot's magnetic field weakens and the atmosphere warms, H_2 dissociates, the particle number increases, and the pressure increases rapidly in the hottest outer edges of the umbra. The cool interior of the umbra which still contains H_2 is compressed by this surrounding higher pressure, maintaining a high magnetic flux density and cooler temperatures. **The intensification of the magnetic field due to the formation of molecular hydrogen contributes to the longevity of sunspots magnetic fields.**

The intensification of the magnetic field due to H_2 is clearest in small sunspots which display unambiguous increases in their magnetic fields at constant temperature. Several of these observations show the strong magnetic field is concentrated in discrete structures. This may imply cooler temperature, stronger magnetic fields and larger H_2 fractions are present in sub-arcsecond fine structure which is beyond the resolution limit of these observations.

All sunspots with umbral magnetic field strengths greater than 2600 G display an increase in slope in the relation between their magnetic pressure and temperature, corresponding to an increase of several hundred Gauss, which usually occurs in the coolest part of the umbra around a continuum temperature of 5800-6000 K at 1.6 μm. **The position of the change in slope in the thermal-magnetic relation coincides with the temperature where the concentration of H_2 begins to steeply increase.**

The scenario presented here fails to explain the behavior seen in the magnetic pressure-temperature relations for the largest spots in the sample, which show increases in magnetic pressure at much lower temperatures than we would expect if they were caused by H_2.

The formation of molecular hydrogen in sunspots may have other exciting consequences for sunspot physics. In particular an increase in the heat capacity of the gas due to the increased degrees of freedom of the molecular gas relative to the atomic gas is of great interest. **The storage of energy in the rotational and vibrational modes of the molecule which do not contribute to the thermal signature of the gas may be**

able to explain where at least some of the energy goes when a sunspot forms, contributing to the sunspot irradiance problem.

9.2 Prospects for Future Work

There are many questions still left unanswered by the research presented in this dissertation, and several clear ways in which it should continue.

9.2.1 Other Science Results from the Sunspot Survey

The sunspot survey obtained with FIRS, fully reduced and analyzed, is rich with possibilities for additional science. There are several interesting candidates which make it possible to investigate sunspot formation and evolution on an individual basis. In particular, NOAA 11035 is an interesting candidate for further study of the photospheric magnetic fields during flares.

Comparison of Multi-Height Observations with MHD Sunspot Models

Full-MHD atmospheric models of sunspots provide another way to investigate the influence of H_2 formation on sunspot MHS equilibrium, to determine the importance of the RT effect in the determination of the apparent atmospheric temperature, and the contribution of the curvature force in the umbra. I have already begun to conduct a preliminary comparison with a MHD model sunspot from Matthias Rempel like those described in Rempel (2010).

Atmospheric models generated in the MHD sunspots simulation were fed to RH to determine the intensity and $\tau = 1$ height of the 15650 Å continuum. The resulting intensity image of the MHD spot is shown in Figure 9.1. By eye, I've selected the umbral and penumbral boundaries from this intensity image. Azimuthal averages of continuum intensity and temperature are shown in Figure 9.2, the vertical and horizontal components of the magnetic field are shown in Figure 9.3 and magnetic field gradients with height are shown

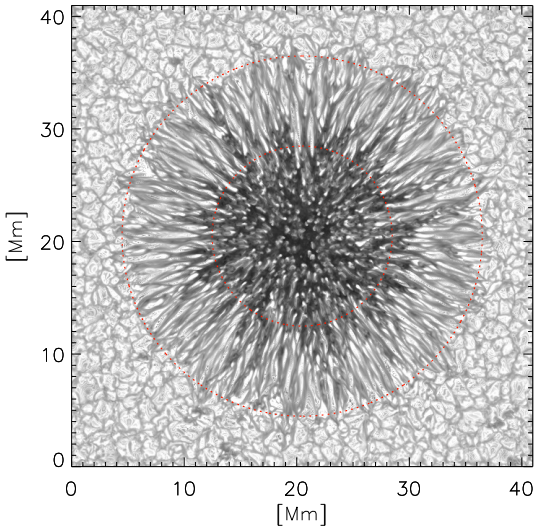

Figure 9.1 Intensity at the $\tau = 1$ in the 15650 Å continuum, normalized to the quiet-Sun. The selected umbral and penumbral boundaries are indicated by the red dotted lines.

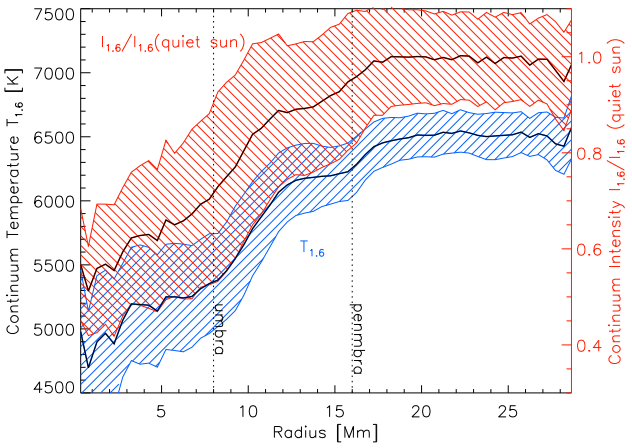

Figure 9.2 Azimuthally averaged continuum temperature and intensity (blue and red heavy lines respectively) at 15650 Å as a function of radius in the Rempel spot. The 1-σ scatter for each quantity is indicated by the hatched area.

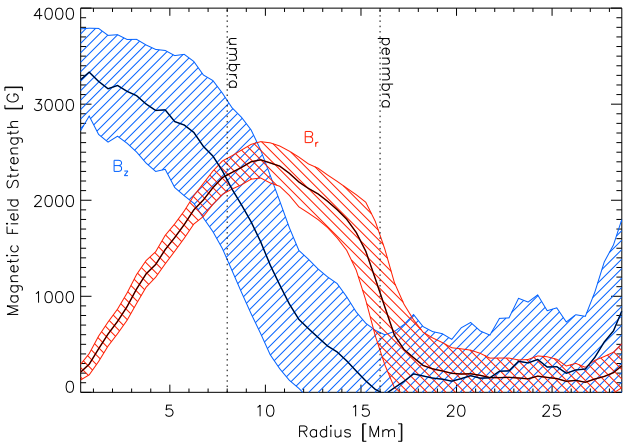

Figure 9.3 Azimuthally averaged vertical (blue) and horizontal (red) components of the magnetic field at 15650 Å as a function of radius in the Rempel spot. The 1-σ scatter for each quantity is indicated by the hatched area.

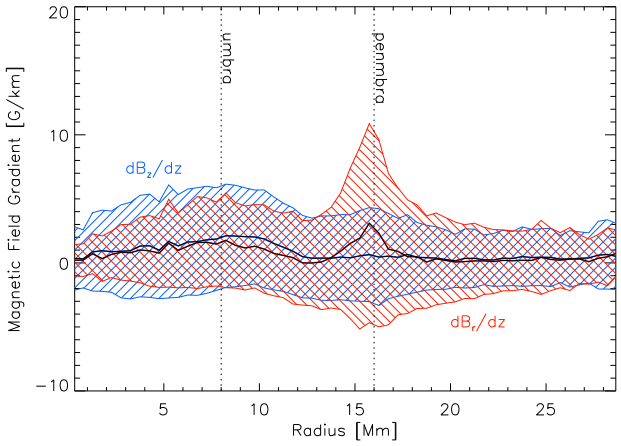

Figure 9.4 Azimuthally averaged vertical (blue) and horizontal (red) magnetic field gradients with height at 15650 Å as a function of radius in the Rempel spot. The 1-σ scatter for each quantity is indicated by the hatched area.

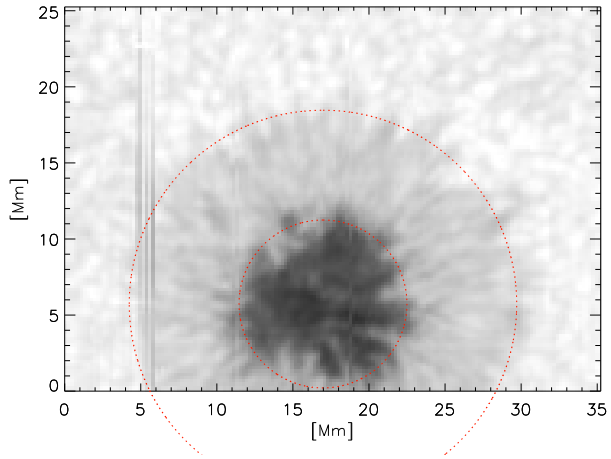

Figure 9.5 Map of the continuum intensity of NOAA 11101 taken on 2010-08-30. The selected umbral and penumbral boundaries are indicated by the red dotted lines.

in Figure 9.4. In these plots the umbral and penumbral boundaries are indicated by the dotted lines, and an elegant method of determining these boundaries can be seen. The umbral boundary coincides with where the vertical and horizontal field are roughly equal, and the penumbral boundary where the average vertical field drops to zero.

These relations can be qualitatively compared with an observation of a very round sunspot at disk center, for example NOAA 11101 on 2010-08-30 whose intensity map is shown in Figure 9.5. The magnetic field vector of this observation has been disambiguated to obtain the vertical and horizontal components of the magnetic field in local solar coordinates. Then the azimuthal averages of the same quantities can be compared, and with the exception of the magnetic field gradients they seem to agree. The 6302 Å data was used to obtain the magnetic field gradient, while for the MHD model it was calculated for the $\tau = 1$ height in the 15650 Å continuum. For the real sunspot, the magnitude of the vertical and horizontal gradients are found to be twice as big as they are for the larger model spot, and dB_r/dz has the opposite sign from dB_z/dz.

These mysteries can be sorted out later. RH could potentially produce the full-Stokes spectrum with more intensive calculations. Generation of the full-Stokes spectrum,

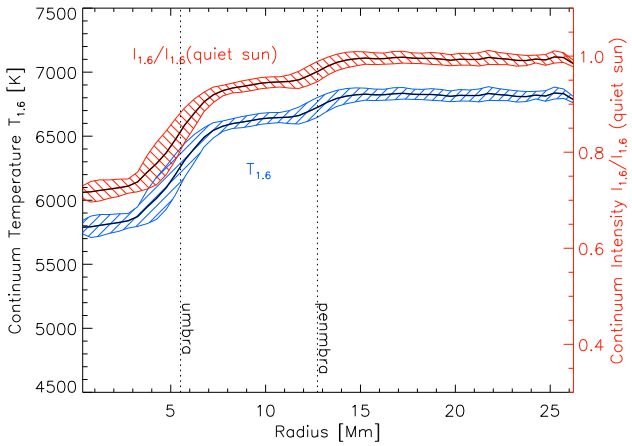

Figure 9.6 Azimuthally averaged continuum temperature (blue) and intensity (red) at 15650 Å as a function of radius in NOAA 11101. The 1-σ scatter for each quantity is indicated by the hatched area.

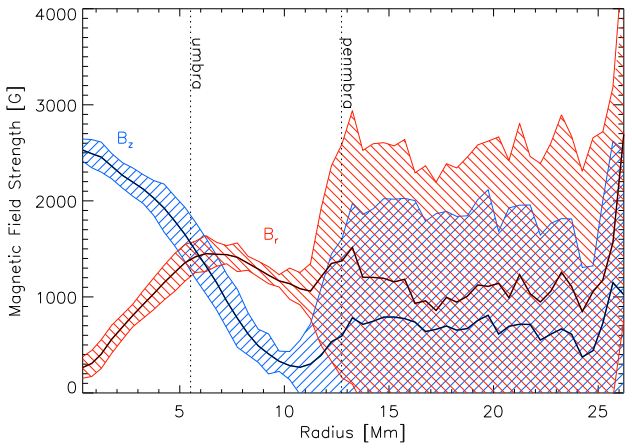

Figure 9.7 Azimuthally averaged vertical (blue) and horizontal (red) components of the magnetic field at 15650 Å as a function of radius in NOAA 11101. The 1-σ scatter for each quantity is indicated by the hatched area.

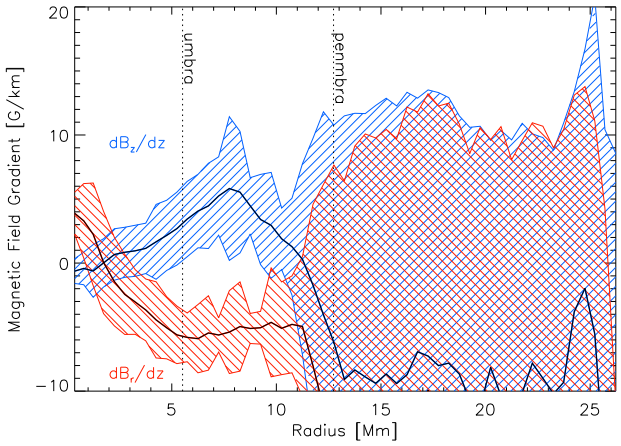

Figure 9.8 Azimuthally averaged vertical (blue) and horizontal (red) magnetic field gradients determined from the 15650 and 6302 Å observations as a function of radius in NOAA 11101. The 1-σ scatter for each quantity is indicated by the hatched area.

degradation with an instrumental profile, and the application of inversion techniques is the obvious next step to test both the inversion technique and the realism of the MHD models.

9.2.2 Extending the Sunspot Survey

Extended survey observations conducted with FIRS are particularly important if the activity of the next solar cycle is as low as forecast. Additional observations with FIRS should focus on large sunspots in order to fully sample the overturn in the OH relation, which can be seen to occur around a continuum temperature of 4000-5000 K (Figure 6.7), and understand the strange behavior of the thermal-magnetic relation in very large spots. If the effects of stray light in FIRS can be understood, it might be possible to make greater use of the 6302 Å observations. It may help to use data from a thoroughly calibrated and widely used source such as the SOT/SP on Hinode to investigate the force equilibrium in sunspot umbrae further. If the uncertainty from stray light is low, then the temperature-dependent H_2 fraction could be determined directly from the observed intensity without a molecular observation.

9.2.3 OH Measurements in the Visible and near-IR

With the possibility of a large stray light fraction in the FIRS infrared channel, it is difficult to definitively confirm the validity of the RH code, and clearly establish the relationship with H_2. OH absorption lines are scattered throughout the visible and infrared, and it should be a fairly simple task to observe these with FIRS or another instrument, and generate the necessary diagnostics with RH. The OH lines at 3.98 μm in the near-IR provide an interesting prospect. These are the strongest OH lines in the sun and have depths of 30% of the continuum intensity in sunspot umbrae while abundant nearby Ni I and Si I lines provide good prospects for magnetic field measurements. Due to their wavelength, they cannot be observed with the Dunn, whose quartz vacuum window cuts off the infrared around 2.2 μm. However, they could be observed with the NSO Infrared Camera (NIRC) on the McMath-Pierce Solar Telescope main spectrograph, whose open design allows for observations at long wavelengths to be carried out. The poor daytime seeing at Kitt Peak is no obstacle to correct with a simple tip-tilt tracker at these long wavelengths, therefore good spatial resolution in sunspots can still be achieved. Coordinated magnetic field observations could be conducted with FIRS, or one of the many Zeeman-split Si I or Ni I lines near 4 μm could be used to obtain a simultaneous magnetic field measurement.

9.2.4 Extension to Chromospheric Heights: FIRS and IBIS

A large majority of the FIRS observations taken since 2009 have complementary observations taken with the Ca II 8542 Å channel on IBIS. The analysis of these observations was beyond the scope of this dissertation, but provide complementary magnetic field measurements at chromospheric heights. Multi-height data are a powerful tool for investigating the 3-D nature of magnetic fields and understanding the connection between magnetic fields which pass through the photosphere and chromosphere and into the corona.

9.2.5 Multi-Wavelength Inversions

With FIRS and IBIS together it is possible to look at magnetic field diagnostics in significantly different wavelength and atmospheric regimes: 6302, 8542,10830, 15650 Å. Although the possibility of taking multi-height data is real (and more such instruments are in development), the tools to analyze and understand these observations are few and underdeveloped. Inversion codes such as NICOLE have the capability to invert multi-wavelength data, but the results need to be tested, proven, and understood. The Rybiki-Hummer radiative transfer and chemical equilibrium code is able to describe the formation of a wide variety of spectral lines in different physical situations. It would make a very powerful forward model for an inversion technique.

9.2.6 Coordinated Magnetic Field Observations with IRIS

An interesting prospect for future coordinated observations with FIRS and IBIS is the Interface Region Imaging Spectrograph (IRIS), a new scanning UV spectrograph being built at Montana State University that will provide unprecedented 1" resolution of ionized atomic emission lines. FIRS and IBIS could provide corroborating magnetic field measurements in the chromosphere and photosphere, while IRIS can accurately obtain information on the temperature and velocity of structures in the chromosphere. In addition, IRIS will cover the lost wavelength ranges from 1332-1358 and 1390-1406 Å which include several of the fluorescent H_2 lines reported in Bartoe et al. (1979) and the emission lines which excite them. Although it may not be possible to establish the photospheric H_2 abundance in sunspots from these observations, interesting science with H_2 awaits us in the chromosphere.

9.2.7 The Growth of Starspots

In stellar atmospheres slightly cooler than the Sun, the temperature changes resulting from starspot magnetic fields would result in much larger H_2 fractions, and it's possible that the evolution of spots in such atmospheres is driven by the atmospheric transition to the molecular phase. While it's interesting to speculate about what magnetic phenomena

might be like elsewhere, it may be possible to investigate the growth rate of starspots by identifying highly active magnetic candidates using the high quality light curves resulting from planetary transit surveys such as the Kepler mission, and following up with high cadence photometry and spectropolarimetry to determine growth rates and magnetic field strengths.

Thank you dear reader for sticking with it from the beginning to this very sweet end!

Appendix A

Derivation of the Horizontal Magnetohyrdostatic Force Balance Equation

Based on observations of sunspots, coronal loops, and other solar phenomena we know that the solar plasma is tied to magnetic fields. Gas and magnetic fields move together and it has been described as magnetic field being "frozen" into the surrounding gas. This has interesting implications, the foremost being the application of magnetic pressure on the gas. To better understand the underlying physics operating in sunspots the magnetohydrodynamic force balance equation and the physical justification for "frozen" fields are obtained from first principles based on the arguments presented in Jackson (1962) and Cowling (1976). The specific case of the pressure balance in a cylindrically symmetric sunspot with a vertical magnetic field is explored.

For a fluid of arbitrary density ρ, arbitrary and unknown initial velocity vector field $\vec{v}$, mass must be conserved:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \tag{A.1}$$

and the force balance equation is simply:

$$\rho \frac{d\vec{v}}{dt} = -\nabla P + \frac{1}{c}(\vec{J} \times \vec{B}) + \vec{F_\nu} \tag{A.2}$$

where $\vec{J}$ is the current density of the fluid, $\vec{B}$ is the magnetic field vector, and $\vec{F_\nu}$ is a drag force due to the viscosity of the fluid. We should also remember our basic E & M equations:

$$\nabla \times \vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0 \tag{A.3}$$

$$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{J} + \frac{1}{c}\frac{\partial \vec{D}}{\partial t} \tag{A.4}$$

$$\nabla \cdot \vec{B} = 0 \tag{A.5}$$

$\vec{E}$ is the electric field vector and $\frac{\partial \vec{D}}{\partial t}$ is the displacement current. The assumption customary to the magnetohydrodynamic approximation is that changes in the magnetic field are small, therefore the displacement current can be ignored. Now to see how the field can be "frozen" into the material we want to consider the effect of a magnetic field which has been imposed on a charged fluid, specifically, what happens when the magnetic field changes or moves.

We'd like to rewrite Equation A.3 in terms that we know or can measure by eliminating the electric field. For a fluid in motion the current density is given by Ohm's Law for charges moving relative the the magnetic field:

$$\vec{J} = \sigma(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}) \tag{A.6}$$

where σ is the conductivity of the fluid. Ohm's Law can be rewritten to eliminate the current density:

$$\vec{E} + \frac{\vec{v}}{c}\vec{B} = 0 \tag{A.7}$$

Substituting $\vec{E}$ with a rearranged version of Equation A.6, A.3 can be rewritten as:

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \times (\frac{c}{\sigma}\vec{J} - \vec{v} \times \vec{B}) \tag{A.8}$$

and substituting with $\vec{J}$ from A.4 we now get:

$$\frac{\partial \vec{B}}{\partial t} \quad = \quad \nabla \times (\vec{v} \times \vec{B}) - \nabla \times \frac{c^2}{4\pi\sigma}(\nabla \times \vec{B})$$

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) - \frac{c^2}{4\pi\sigma}(\nabla(\nabla \cdot \vec{B}) - \nabla^2 \vec{B}) \tag{A.9}$$

$\nabla \cdot \vec{B}$ is of course equal to zero, so we end up with:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) + \frac{c^2}{4\pi\sigma}\nabla^2 \vec{B} \tag{A.10}$$

For very large fluids, such as those in the Sun or other astronomical settings, it is difficult for the fluid to hinder the flow of charges and the conductivity is very large ($\sigma \to \infty$). The term on the right in Equation A.10 becomes vanishingly small:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) \tag{A.11}$$

$$\frac{\partial \vec{B}}{\partial t} = \vec{B}(\nabla \cdot \vec{v}) \tag{A.12}$$

This equation now shows that the magnetic flux through any given moving element must be constant in time, thus, that the fields are dragged along with the material, and the material is dragged along with the fields, and by definition this is only for motion perpendicular to the magnetic field.

To find the form of the pressure balance equation we must rewrite the $\frac{1}{c}(\vec{J} \times \vec{B})$ term from Equation A.2.

$$\begin{aligned} \frac{1}{c}(\vec{J} \times \vec{B}) &= -\frac{1}{4\pi}\vec{B} \times (\nabla \times \vec{B}) \\ \frac{1}{c}(\vec{J} \times \vec{B}) &= \frac{1}{8\pi}(2(\vec{B} \cdot \nabla)\vec{B} - \nabla(B^2)) \end{aligned} \tag{A.13}$$

Where the identity $\frac{1}{2}\nabla(\vec{B} \cdot \vec{B}) = (\vec{B} \cdot \nabla) + \vec{B} \times (\nabla \times \vec{B})$ has been used to simplify the expression. This now makes the force balance equation:

$$\rho\frac{d\vec{v}}{dt} = -\nabla P + \frac{1}{8\pi}(2(\vec{B} \cdot \nabla)\vec{B} - \nabla(B^2)) + \vec{F_\nu} \tag{A.14}$$

B^2 is called the magnetic pressure and $(\vec{B}\cdot\nabla)\vec{B}$ is called the magnetic tension force. It is easy to see that for a given field that obeys equation B5 the magnetic pressure force acts in a direction perpendicular to the magnetic field vector and the magnetic tension force acts parallel to it. For a static fluid viscosity vanishes and the force balance equation becomes:

$$\nabla P = \frac{1}{8\pi}(2(\vec{B}\cdot\nabla)\vec{B} - \nabla(B^2)) \tag{A.15}$$

To obtain the horizontal component of the pressure balance equation for the static state, the integration of the magnetohydrostatic force balance equation is the next necessary step. However in order to do this we must assume a geometry for the magnetic field. Let us consider an ideal sunspot which is small and round and has no penumbra. The functional form of the magnetic field is $\vec{B} = B_r(r,z)\hat{r} + B_z(r,z)\hat{z}$. However inside the sunspot the magnetic field has a large vertical component that is constant with height, but has a radial dependency and drops to zero at the edge. The radial component of the magnetic field changes with height in the spot, but remains fairly constant over the radial distance. Therefore the gradients of the vertical magnetic field with height and radial magnetic field with radius will be small. The radial component of the magnetic pressure becomes:

$$\begin{aligned} \nabla B^2 &= \frac{\partial B_z^2(r,z)}{\partial r}\hat{r} \\ \int \nabla B^2 dr &= \int_r^a \frac{\partial B_z(r,z)^2}{\partial r'} dr' = B_z^2(a,z) - B_z^2(r,z) \end{aligned} \tag{A.16}$$

The magnetic field of couse vanishes at the sunspot boundary ($B_z(a) = 0$). The radial component of the magnetic tension becomes:

$$\begin{aligned} 2(\vec{B}\cdot\nabla)\vec{B} &= 2B_z(r,z)\frac{\partial B_r(r,z)}{\partial z}\hat{r} + 2B_r(r,z)\frac{\partial B_z(r,z)}{\partial r}\hat{z} \\ \int \quad 2(\vec{B}\cdot\nabla)\vec{B} \quad dr &= 2\int_r^a B_z(r',z)\frac{\partial B_r(r',z)}{\partial z}dr' = F_c(r,z) \end{aligned} \tag{A.17}$$

where a is the maximum radial extent of the sunspot. The magnetic tension is non-zero because there is a small component of the magnetic field in the radial direction. This

additional pressure is called the curvature force F_c due to the fact that it is caused by magnetic field lines curving from the vertical to radial direction. The pressure provided by the thermal motions of the gas is given by the ideal gas equation of state:

$$P(r,z) = \int \nabla P dr = \int_r^a \frac{\partial P}{\partial r} dr = P(a,z) - P(r,z) \tag{A.18}$$

where $P(a,z) = P_p(z) = nkT_p(z)$ is the pressure of the photosphere and is a constant independent of distance from the sunspot center, and $P(r,z) = P_s(r,z) = nkT_s(r,z)$ is the pressure inside the spot. The final result is:

$$\frac{1}{8\pi}(B_z^2(r,z) + F_c(r,z)) = nk(T_p(z) - T_s(r,z)) \tag{A.19}$$

There are two important points which become apparent in the course of this derivation. The first is that the curvature force is a poorly determined quantity because it requires knowledge of the vertical structure of the magnetic field. This is information which we gain from using two sets of diagnostic lines in the visible and the infrared formed approximately 60 km apart in the sunspot atmosphere (as we see from the separation of the $\tau = 1$ heights in Figure 6.10). The second point is that all of the gas responds to the magnetic pressure and we know this is not true if 10% of the gas is neutral molecular hydrogen. The non-magnetic portion of the gas (along with its radial dependence) needs to be included if it is in fact significant.

Appendix B

A Method for the Calibration of Instrumental Polarization

B.1 Introduction

The Kuhn technique (Kuhn et al. 1994) used with the FIRS data is one of many ways to determine and remove the instrumental polarization. Because it relies on line symmetry it cannot tell us everything about the optical system, and it is not reliable in situations where the magnetic field is weak, such as quiet sun observations. Complicated methods exist to fully characterize the polarization properties of the telescope and instrument system (Bernasconi 1997; Elmore et al. 2010; Socas-Navarro et al. 2011). We wish to determine the instrumental polarization in the most simple, direct, and accurate way possible with the fewest number of initial assumptions. A linear polarizer and quarter-wave plate can be tuned to input pure Q, U, and V states into the optical path to be characterized. The rotation of these optical elements results in a Stokes vector which is a predictable combination of sine-like terms. Measuring the Stokes vector resulting from the instrument, we can retrieve the Müller matrix using a non-linear least-squares fit based on the known functional form of the matrix of the calibration optics assuming a static instrument Müller matrix.

The calibration optics themselves are imperfect and need to have well defined parameters such as unknown efficiencies, position errors. Wavelength dependent effects may contribute to a poor fit of the instrument Müller matrix. Independent testing of the calibration optics would remove much of the guesswork but this is not always possible to do this easily and quickly with a complicated instrument (although it probably only needs to be done once). Not all of the parameters of the calibration optics are known a priori or can be left as free parameters of the fitting procedure. A poor selection of free parameters may lead to non-unique solutions and a poor reproduction of the Müller matrix. I will motivate a calibration scheme where which considers the effect of the following three parameters on the accuracy of retrieval of the Müller matrix:

1. Error in the position of the linear polarizer (az0)
2. Error in the position of the wave plate (az1)
3. Retardance of the wave plate (phi)

where only the orientation of the linear polarizer must be accurately determined to 1° or better independently of the calibration.

B.2 Non-Linear Least-Squares Fit to the Instrument Müller Matrix and Additional Parameters

The non-linear least-squares fit to the instrument Müller matrix and the additional free parameters were written in IDL and implemented the *curvefit* function which uses a gradient-expansion algorithm. The fit functions are given in the appendix. We regret that *curvefit* does not allow for calculation of the covariance matrix as some other non-linear fitting functions do, but it does provide the best fit of any comparable function tested.

B.3 Calibration with a Linear Polarizer

First we will discuss the simple scenario of determining just the linear polarization crosstalk of a system. A single rotation of a linear polarizer allows us to determine 12 of the 16

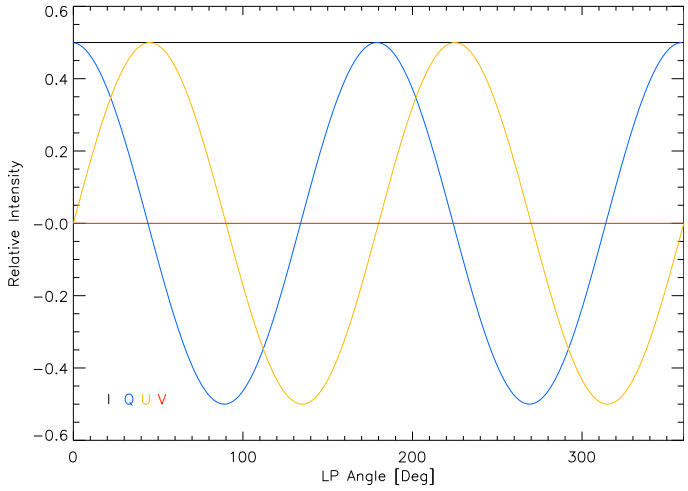

Figure B.1 The variation of the components of the Stokes vector resulting from a perfect linear polarizer as it is rotated through 360° applied to totally unpolarized light.

components of the instrument Müller matrix (I→I, I↔Q, I↔U, I→V, Q→Q, Q↔U, Q→V, U→U, U→V). If there is an error in the position of the linear polarizer, can we determine what this error is from the measurements and correct for it? In fact the answer is no. We can see that the Stokes vector that results from a perfect linear polarizer is a linear combination of sine and cosine terms, as shown in figure B.1.

The observed Stokes vector is the product of the unknown Müller matrix of the instrument and the linear polarizer acting on what is assumed to be unpolarized light:

$$M_{inst} M_{lp}(\theta_{lp}) S_0 = S_{obs} \tag{B.1}$$

where $S_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$,

$$M_{inst} = \begin{pmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{pmatrix} \tag{B.2}$$

and

$$M_{lp}(\theta) = \begin{pmatrix} 1 & c_2\theta_{lp} & s_2\theta_{lp} & 0 \\ c_2\theta_{lp} & c_2^2\theta_{lp} & s_2\theta_{lp}\ c_2\theta_{lp} & 0 \\ s_2\theta_{lp} & s_2\theta_{lp}\ c_2\theta_{lp} & s_2^2\theta_{lp} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{B.3}$$

disregarding any constant terms. The sine and cosine terms are abbreviated $cos(2\theta_{lp}) = c_2\theta_{lp}$, $cos^2(2\theta_{lp}) = c_2^2\theta_{lp}$, $sin(2\theta_{lp}) = s_2\theta_{lp}$, and $sin^2(2\theta_{lp}) = s_2^2\theta_{lp}$. The resulting equations for each component of the Stokes vector have the form:

$$\begin{aligned} S_{obs,i} = \quad & m_{i0} \\ + \quad & m_{i1}\ c_2\theta_{lp} \\ + \quad & m_{i2}\ s_2\theta_{lp} \end{aligned} \tag{B.4}$$

If we assume that there is an unknown shift in the position angle of the linear polarizer, we can see that there is no way to tell the difference between instrumental crosstalk from the Q and U states and the shift in the polarizer position, because a linear combination of sine and cosine curves of the same frequency can also be written as a phase shift of a single sine curve with the right amplitude. Stated in more physical terms, we cannot distinguish the difference between a rotation caused by the instrument or a rotation due to the calibration optics. In order to characterize the linear polarization accurately the position of the linear polarizer must be known beforehand.

B.4 Calibration with a Linear Polarizer and Quarter-Wave Plate

In the previous example we only considered the linear polarization. A linear polarizer and quarter-wave plate provide a good way to characterize the linear and circular polarization

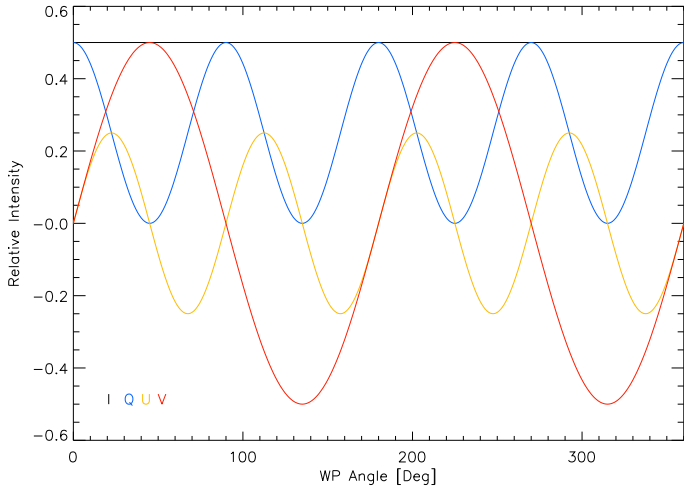

Figure B.2 The variation of the Stokes vector resulting from a perfect linear polarizer and quarter-wave plate as the quarter-wave plate is rotated through 360°.

in an optical system. By leaving the linear polarizer at a fixed position (0°) and rotating the wave plate through the full range of angles we input the Stokes vectors shown in Figure B.2

and can obtain the full Müller matrix based on a fit of the output. Again disregarding any constant terms we find the equation for the observed stokes vector is a combination of constants and sine and cosine terms:

$$M_{inst}M_{vr}(\theta_{vr},\phi)M_{lp}(\theta_{lp})S_0 = S_{obs} \tag{B.5}$$

where $S_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$,

$$M_{vr}(\theta,\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_2^2\theta_{vr} + s_2^2\theta_{vr}\ c\phi & s_2\theta_{vr}\ c_2\theta_{vr}\ (1-c\phi) & -s_2\theta_{vr}\ s\phi \\ 0 & s_2\theta_{vr}\ c_2\theta_{vr}\ (1-c\phi) & s_2^2\theta_{vr} + c_2^2\theta_{vr}\ c\phi & c_2\theta_{vr}\ s\phi \\ 0 & s_2\theta_{vr}\ s\phi & -c_2\theta_{vr}\ s\phi & c\phi \end{pmatrix} \tag{B.6}$$

$$M_{inst} = \begin{pmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{pmatrix} \tag{B.7}$$

The solution for every term in the observed Stokes vector has the form:

$$\begin{aligned} S_{obs,i} = \quad & m_{i0} \\ + \quad & m_{i1}((c_2^2\theta_{vr} + s_2^2\theta_{vr}\ c\phi)\ c_2\theta_{lp} + s_2\theta_{vr}\ c_2\theta_{vr}\ (1 - c\phi)\ s_2\theta_{lp}) \\ + \quad & m_{i2}(s_2\theta_{vr}\ c_2\theta_{vr}\ (1 - c\phi)\ c_2\theta_{lp} + (c_2^2\theta_{vr} + s_2^2\theta_{vr}\ c\phi)\ s_2\theta_{lp}) \\ + \quad & m_{i3}(s_2\theta_{vr}\ s\phi\ c_2\theta_{lp} - c_2\theta_{vr}\ s\phi\ s_2\theta_{lp}) \end{aligned} \tag{B.8}$$

However there are now three additional terms that need to be considered: the position of the linear polarizer, the position of the quarter-wave plate, and the wavelength dependent retardance of the wave plate. Are these terms degenerate with each other or any of the Müller matrix terms? The problem with the linear polarizer is the same as before. $sin^2\theta$ and $sin\theta cos\theta$ are separated by 1/4 in phase, and the constants in linear combinations of these terms cannot be distinguished from the phase shift terms. However the contribution from Stokes V is a sine curve of half the frequency is introduced, the degeneracy is broken for the wave plate, and we can easily fit its angle off-set and retardance in addition to the experimental instrument Müller matrix. From the functional form it is difficult to see the interplay of the variables so we will apply our fitting routine to simulated data as an exploratory example.

B.5 Obtaining the Müller Matrix of Simulated Data

We have synthesized the polarization calibration data given the Müller matrix of a theoretical instrument:

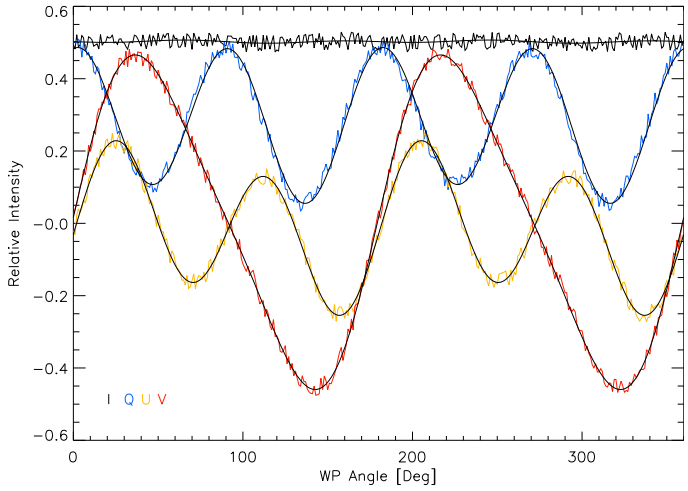

Figure B.3 The simulated Stokes vector resulting from the instrument matrix given above and a linear polarizer(az0 = 0°) and wave plate(az1 = 0°, phi = 80°) as the wave plate is rotated through 360°. Simulated data in color, fit in black, noise=5%.

$$M_{inst} = \begin{pmatrix} 1.000 & 0.003 & -0.021 & 0.001 \\ -0.006 & 0.973 & 0.113 & 0.052 \\ 0.015 & -0.074 & 0.932 & 0.135 \\ 0.005 & 0.024 & 0.351 & 0.890 \end{pmatrix} \tag{B.9}$$

and the parameters of the calibration optics. Figure B.3 shows and example of the fit to the simulated Stokes vector. Figures B.4 - B.12 show the parameters resulting from the non-linear least squares fit of the matrix equation containing the 16 components of the Müller matrix and one or two additional free parameters (WP retardance, WP azimuth, or LP azimuth).

Each plot shows the true value of the parameter from the synthesized data in black and the fitted value is over-plotted in red. All of the Müller matrix terms in the solution are assumed to be free parameters. One of the parameters of the calibration optics is allowed to be a free parameter and is also fitted, the fit assumes the other two are a constant default value (WP retardance 90°, WP azimuth off-set 0°, LP azimuth off-set 0°). The calibration optics are assigned a range of values which are varied in turn.

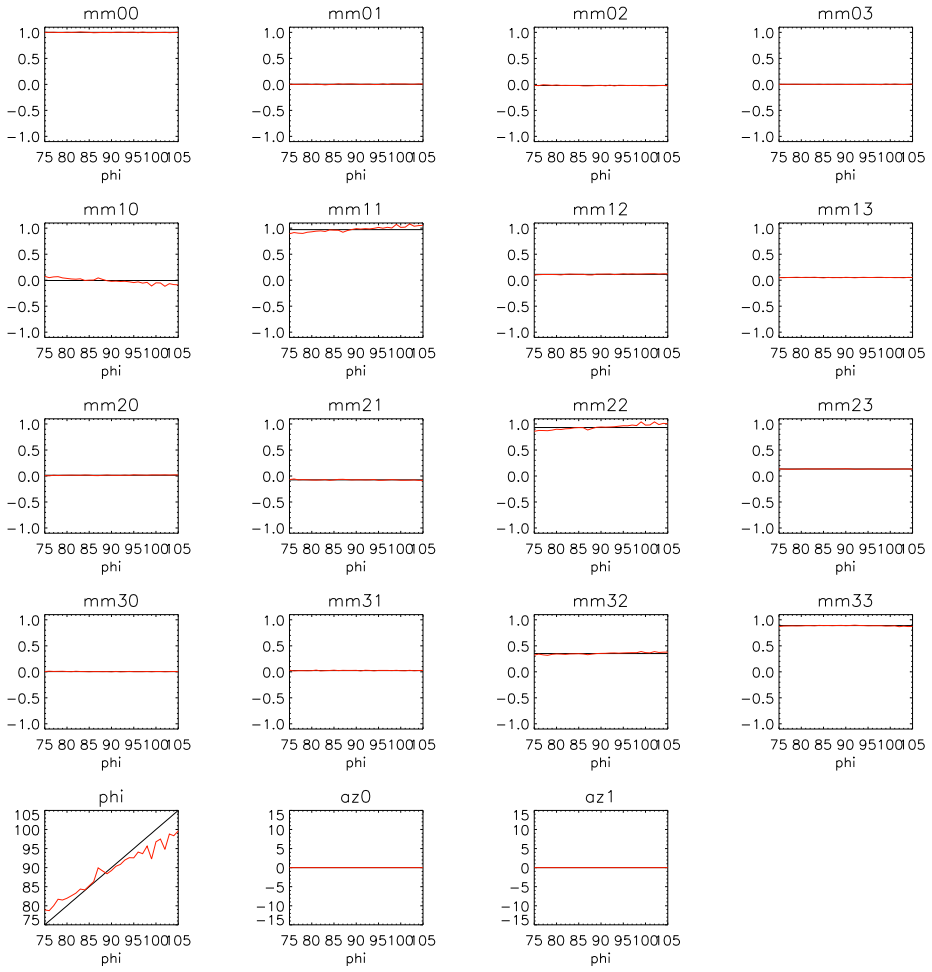

Figure B.4 The wave plate retardance is a free parameter. The wave plate retardance input into the synthesized instrument Müller matrix is varied from 75° to 105°. Fitted quantities are shown in red, true quantities are shown in black.

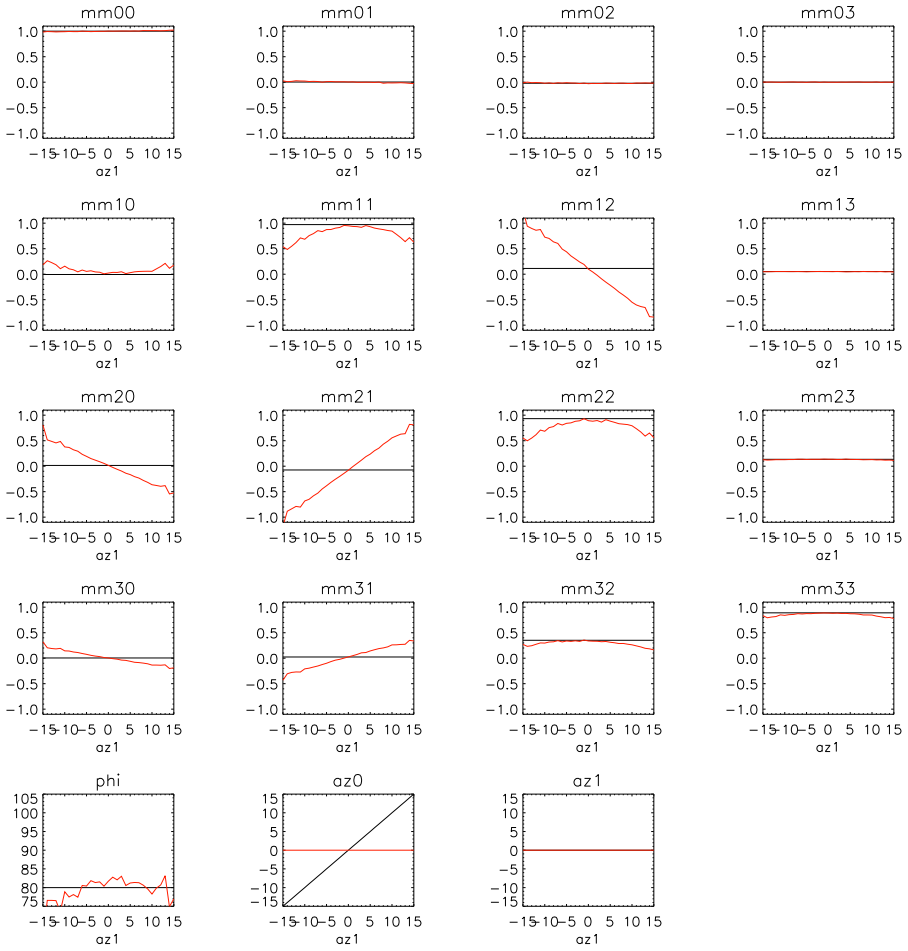

Figure B.5 The wave plate retardance is a free parameter. The wave plate azimuth off-set input into the synthesized instrument is varied from $-15°$ to $15°$. Fitted quantities are shown in red, true quantities are shown in black.

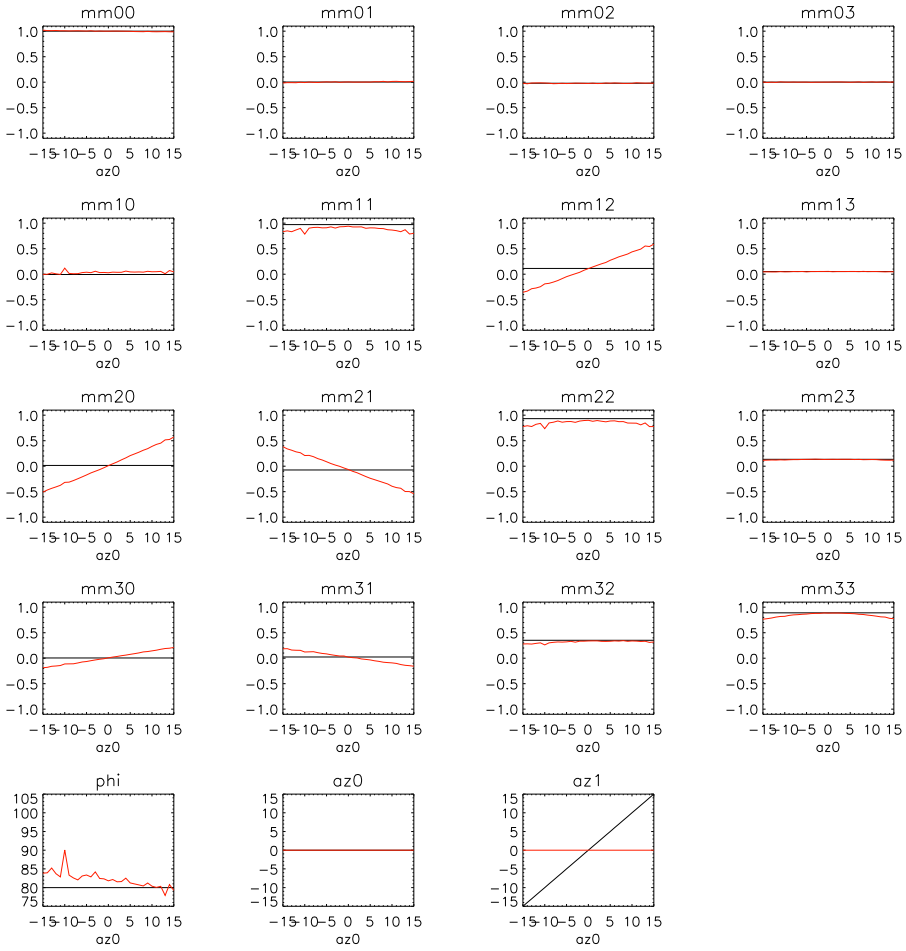

Figure B.6 The wave plate retardance is a free parameter. The linear polarizer azimuth off-set input into the synthesized instrument is varied from $-45°$ to $45°$. Fitted quantities are shown in red, true quantities are shown in black.

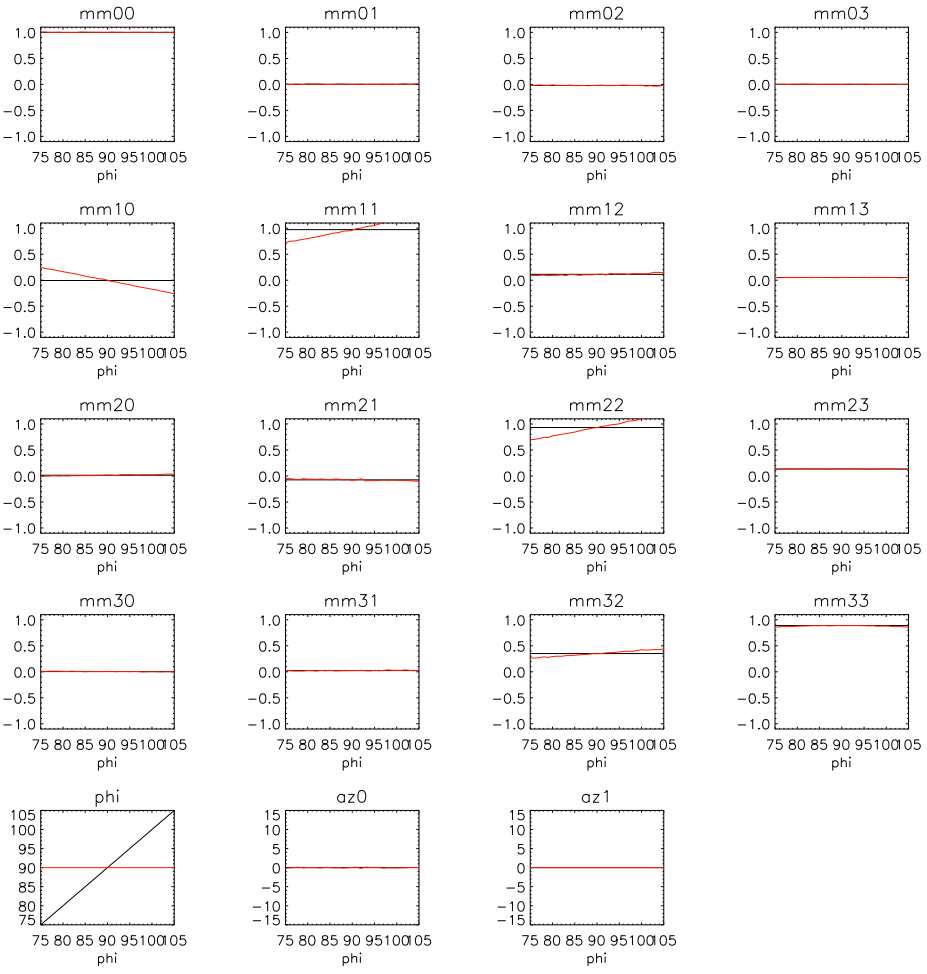

Figure B.7 The wave plate azimuth off-set is a free parameter. The wave plate retardance input into the synthesized instrument is varied from 75° to 105°. Fitted quantities are shown in red, true quantities are shown in black.

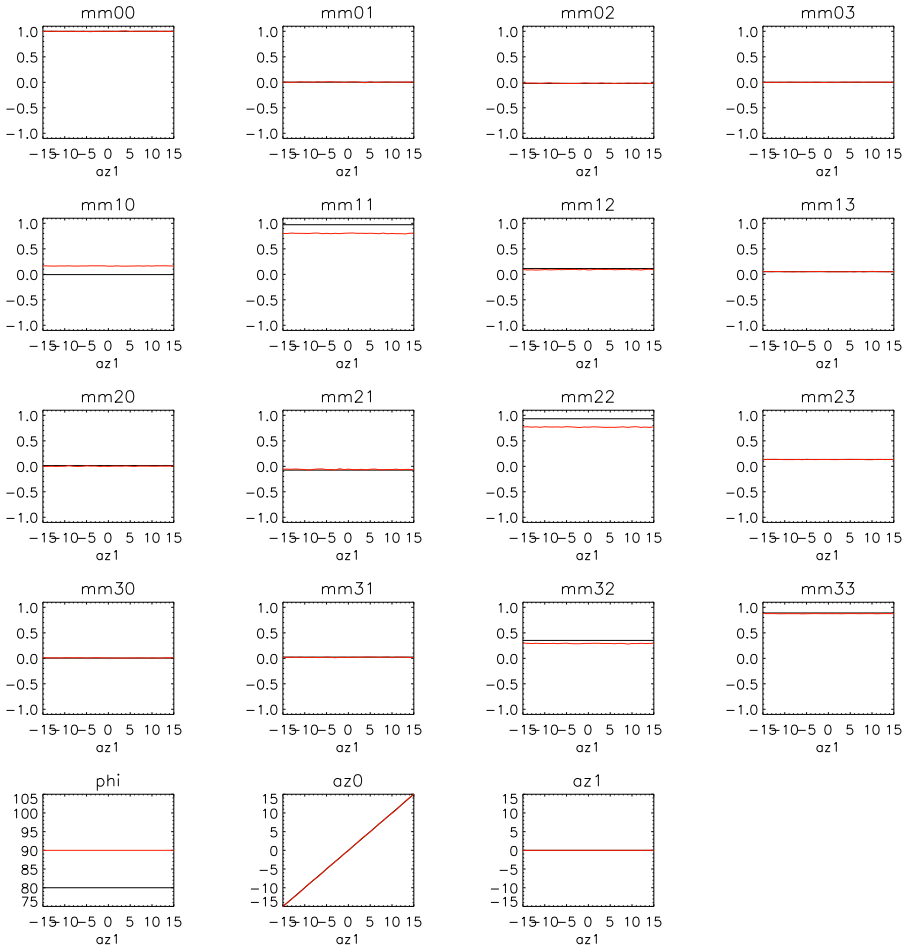

Figure B.8 The wave plate azimuth off-set is a free parameter. The wave plate azimuth off-set input into the synthesized instrument is varied from $-15°$ to $15°$. Fitted quantities are shown in red, true quantities are shown in black. The blue and yellow lines show the fit to the Müller matrix for WPR +/- 10.

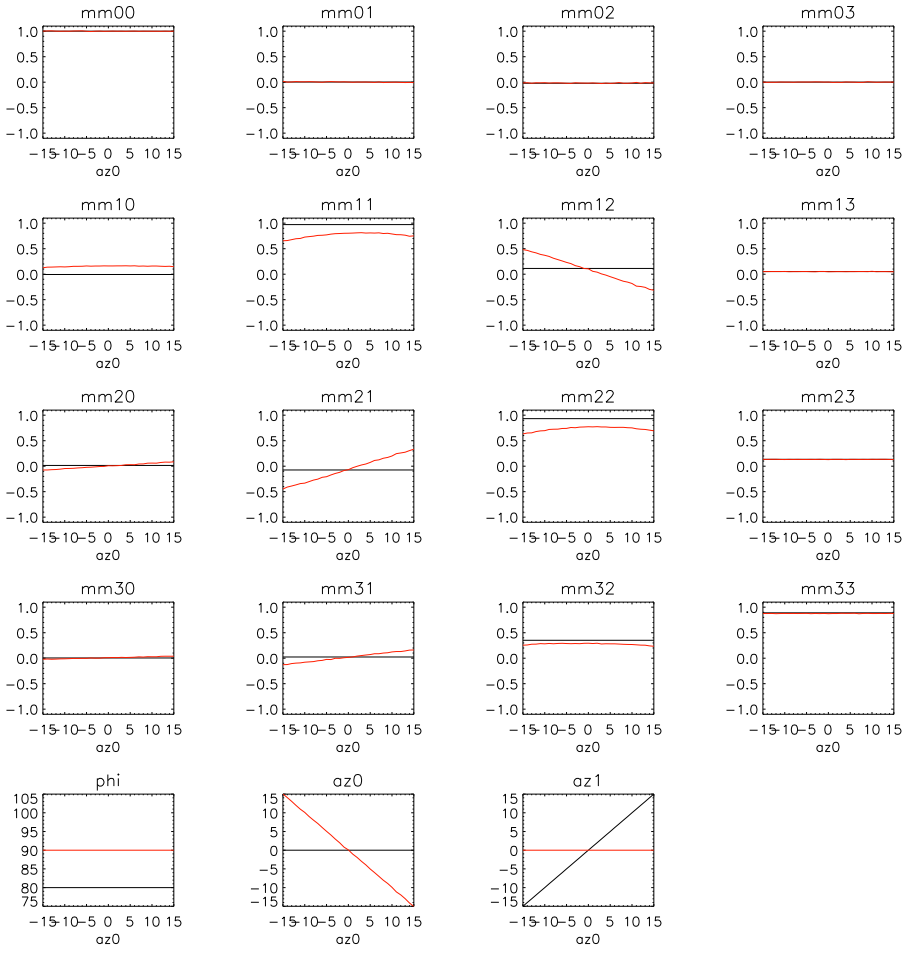

Figure B.9 The wave plate azimuth off-set is a free parameter. The linear polarizer azimuth off-set input into the synthesized instrument is varied from $-15°$ to $15°$. Fitted quantities are shown in red, true quantities are shown in black.

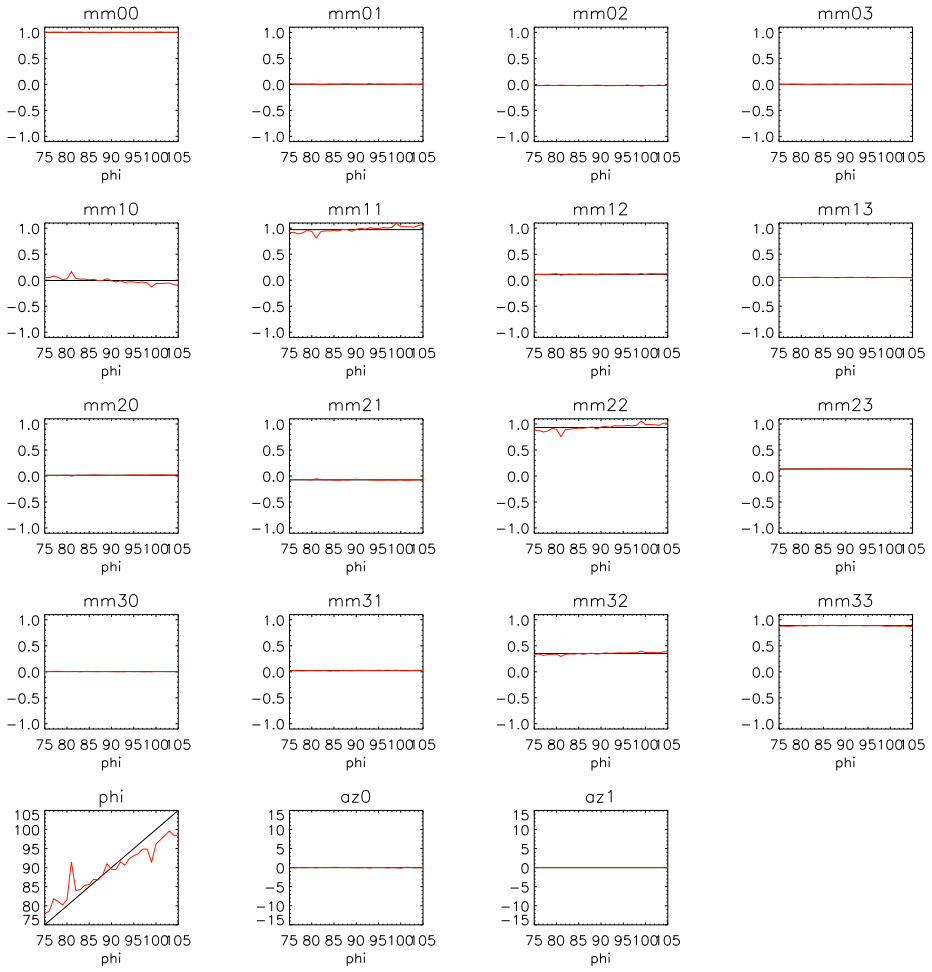

Figure B.10 The wave plate retardance and azimuth off-set are free parameters. The wave plate retardance input into the synthesized instrument is varied from 75° to 105°. Fitted quantities are shown in red, true quantities are shown in black.

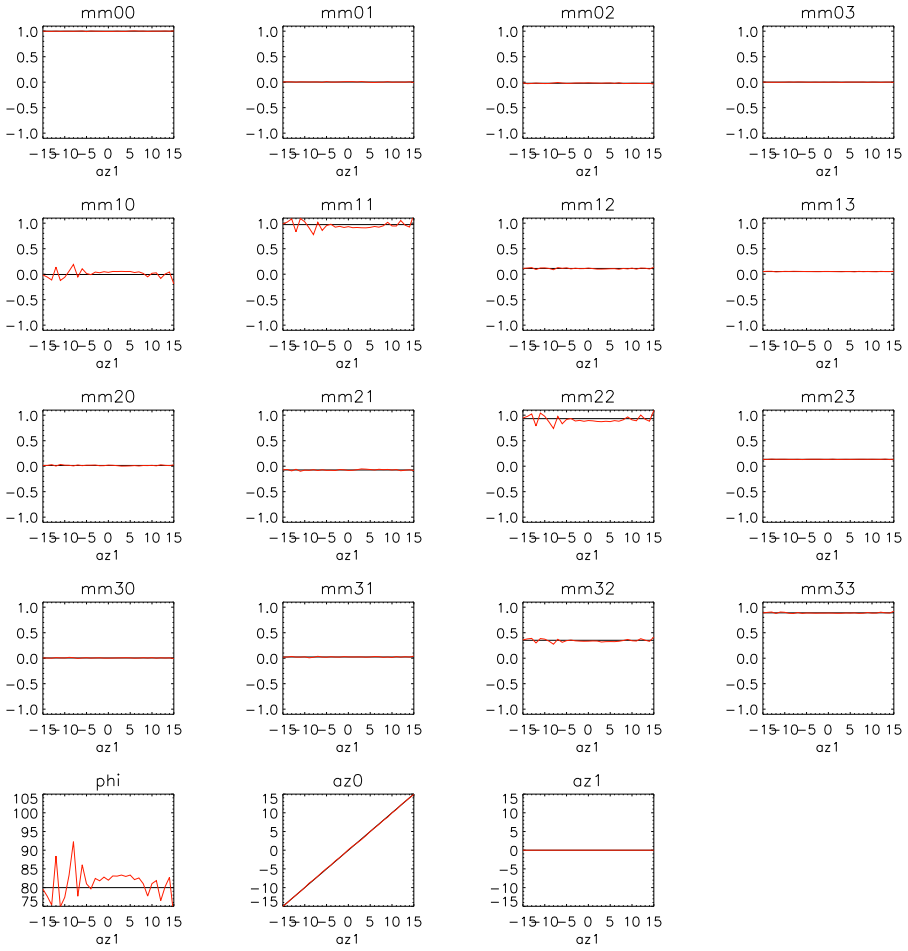

Figure B.11 The wave plate retardance and azimuth off-set are free parameters. The wave plate azimuth off-set input into the synthesized instrument is varied from $-15°$ to $15°$. Fitted quantities are shown in red, true quantities are shown in black.

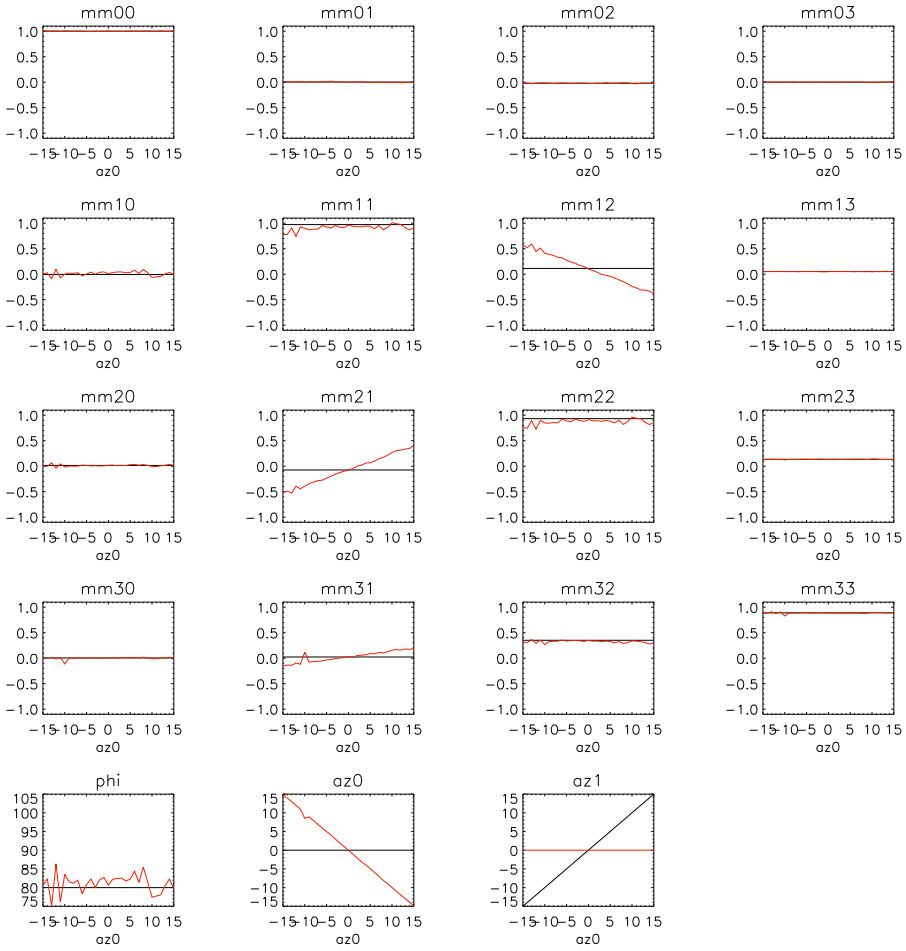

Figure B.12 The wave plate retardance and azimuth off-set are free parameters. The linear polarizer azimuth off-set input into the synthesized instrument is varied from $-15°$ to $15°$. Fitted quantities are shown in red, true quantities are shown in black.

To get an idea of the errors associated with the fit a 100 iteration Monte Carlo simulation was performed with phi and az_1 treated as free parameters of the fit (true values: phi=80, az_1=5.0, az_0=0.0). The resulting instrument Müller matrix and parameter values with their associated errors in parentheses are as follows:

$$M_{inst} = \begin{pmatrix} 1.000(0.003) & 0.003(0.005) & -0.020(0.005) & 0.001(0.002) \\ 0.02(0.03) & 0.94(0.03) & 0.110(0.008) & 0.052(0.002) \\ 0.013(0.004) & -0.072(0.006) & 0.90(0.03) & 0.134(0.002) \\ 0.005(0.003) & 0.024(0.005) & 0.34(0.01) & 0.886(0.004) \end{pmatrix} \quad \text{(B.10)}$$

$$phi = 81(1) \quad \text{(B.11)}$$

$$az_1 = 5.00(0.07) \quad \text{(B.12)}$$

B.6 Proposed Calibration Method

We have developed this procedure to determine the three parameters and obtain an accurate fit to the instrument Müller matrix:

0. Accurately set the azimuth of the LP so that the polarization direction is aligned with the +Q direction of the polarimeter. We can rely on the marking on the polarizer to get roughly $\pm 1^\circ$ alignment.

Or, if we want to be sure, then we should have the LP on a motorized rotation stage. Then if we remove the modulator of the polarimeter (and keeping the exit polarizer), we can accurately determine the axis of the LP by rotating the input LP through 360 deg.

1. With accurate alignment of the LP, we can determine the azimuth offset of the WP, even if we are not sure about the retardance of the WP.

2. With accurate alignment of the LP and the WP, we can determine the retardance of the WP.

3. With accurate knowledge of the LP and WP, we can determine the instrument Müller matrix.

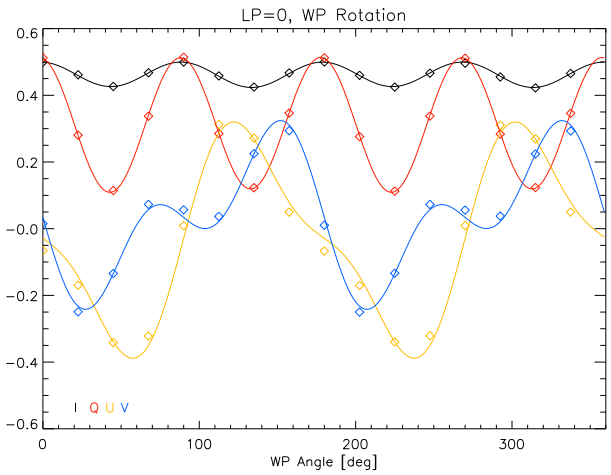

Figure B.13 The Stokes vector obtained by the FIRS visible arm using the prime focus polarization optics (diamonds). The fit is indicated by the solid line.

B.7 Polarization Calibration of the DST

The Dunn Solar Telescope induces a high level of polarization crosstalk between linear and circularly polarized states due to several oblique reflections in the primary optical elements. Following the telescope prime focus complicated secondary optics in the adaptive optics bench and instrument (of your choice) bench further confuse the polarization signal. A polarization calibration scheme was devised by Skumanich et al. (1997) for the Advanced Stokes Polarimeter and has been adopted for polarization calibration of other instruments on the DST. We propose to use this pointing dependent polarization model for the telescope before prime focus. Using the linear polarizer and retarder at prime focus we can calibrate the polarization due to the stationary optics downstream of prime focus. This calibration has been performed for the visible arm of FIRS ($\lambda = 630.2$ nm). The fit Stokes vector is shown in figure B.13 and yields the following Müller matrix and wave plate retardance:

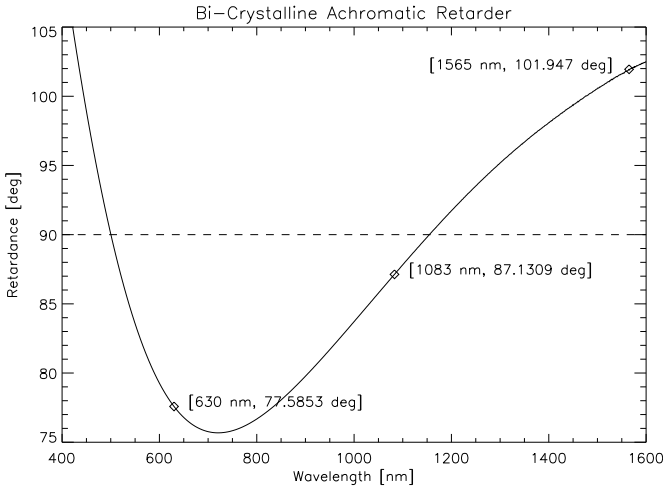

Figure B.14 Retardance of the bi-crystalline achromatic retarder at DST prime focus reproduced from measurements taken by David Elmore.

$$M_{inst} = \begin{pmatrix} 0.815057 & 0.183540 & -0.0196540 & 0.00275667 \\ 0.0464163 & 0.974786 & -0.150017 & -0.0103388 \\ -0.0738421 & 0.0158909 & 0.511176 & -0.618462 \\ 0.0943793 & -0.0266977 & -0.715407 & -0.365516 \end{pmatrix} \quad \text{(B.13)}$$

$$phi = 78.9681 \quad \text{(B.14)}$$

The experimentally determined retardance agrees within the expected error to that measure by David Elmore, which is reproduced in figure B.14.

B.8 Scheme for Polarization Calibration of ATST

Because the retardance and orientation of the wave plate do not need to be known quantities, this method is applicable to polarization calibrations involving experimental optics. It is easy enough to construct a well-aligned sheet polarizer for a large instrument aperture (4 m), however it is more difficult to do the same thing with wave plates. The retardant properties of plastic is well known and we believe that a single sheet, or layered sheets, of

plastic could be used with a large sheet polarizer to provide a uniform circular polarization input for a rough calibration of ATST from the very first optical surface to the analyzing instrument.

Appendix C

Polarized Radiative Transfer and Approximation to the Faraday-Voigt Profiles

Here I summarize the basic changes made to the simple inversion code by H. Lin, who uses linear combinations of Gaussian functions to approximate polarized line shapes. For the visible iron lines at 630.2 nm, this description is not adequate, as the lines suffer from significant line damping and the magneto-optical Faraday effect. This effect causes a "reversal" to appear in the Stokes V line core, among other effects. A detailed treatment of the equations of polarized radiative transfer is given in Jefferies et al. (1989). In their formulation, there are two integrals that must be solved numerically:

$$H(a,v) = \frac{a}{\pi}\int_{-\infty}^{+\infty} \frac{e^{-y^2}dy}{(v-y)^2+a^2} \tag{C.1}$$

$$2F(a,v) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{(v-y)e^{-y^2}dy}{(v-y)^2+a^2} \tag{C.2}$$

where v is the dimensionless spectral parameter and the damping constant $a = \gamma/4\pi\Delta v_D$ where v_D is the Doppler width of the line in question. $v = (v_o - v)/\delta v_D$ for the undisplaced component of the Zeeman split triplet, and $v = (v_o \mp v_L - v)/\Delta v_D$ for the blue(negative) and red(positive) shifted components.

Numerical approximations to these integrals are given in Matta & Reichel (1971). I have coded this as an IDL procedure, for a given value of a and array of v values the procedure returns the corresponding array for $H(a, v)$ and $F(a, v)$.

Appendix D

Derivation of the η_0 Scaling

Based on the assumption that the lines are formed at the same height and under the same conditions some of the fitted parameters for the lines are assumed to be identical (velocity, Doppler width, damping parameter), and others are assumed to scale between the lines (magnitude of magnetic splitting, line to continuum absorption ratio). By using two or more lines we are able to achieve more accurate fitting than is possible with just one line (Solanki et al. 1992). The scaling factor of the parameter η_0 has not been made clear by other authors, so we include a full derivation of it here. η_0 is defined by Jefferies et al. (1989) as the ratio of the line to continuum absorption coefficients ($\kappa_{line}/\kappa_{cont.}$). Starting from our favorite radiative transfer text (Rybiki & Lightman 1979), the line absorption coefficient is:

$$\kappa_{line} = \frac{h\nu}{4\pi}(n_l B_{lu} - n_u B_{ul}) \tag{D.1}$$

where h is Plank's constant and ν is the frequency of the line, n_l and n_u are the population of lower and upper levels in the atom. B_{lu} and B_{ul} are Einstein's coefficients for absorption and emission, which can be related in terms of the oscillator strength of the transition in the following way:

$$B_{lu} = \frac{4\pi^2 e^2}{m_e h\nu c} f_{ul} \ , \quad B_{ul} = B_{lu}\frac{g_l}{g_u} \tag{D.2}$$

where e is the electron charge and m_e is the mass, c is the speed of light, f_{ul} is the oscillator strength, and g_l and g_u are the degeneracy of the upper and lower levels. Substituting these relations into equation D.1 we get:

$$\begin{aligned} \kappa_{line} &= \frac{h\nu}{4\pi} n_l B_{lu}(1 - \frac{n_u}{n_l}\frac{g_u}{g_l}) \\ &= \frac{\pi e^2}{m_e c} f_{ul}(n_l - n_u \frac{g_u}{g_l}) \end{aligned} \tag{D.3}$$

For lines formed in LTE, we can simply use Boltzmann's equation to determine the population fraction in a given level i:

$$\frac{N_i}{N_{tot}} = \frac{g_i \ e^{-E_i/kT}}{Z(T)} \tag{D.4}$$

where N_{tot} is the total particle number and $Z(T)$ is the partition function (which is the same for all levels). Substituting this into equation D.3 and rewriting the energy difference between the levels as a function of frequency then yields:

$$\kappa_{line} = \frac{\pi e^2}{m_e c} f_{ul} \frac{g_l \ e^{-E_l/kT}}{Z(T)} (1 - e^{-h\nu/kT}) N_{tot} \tag{D.5}$$

Now the ratio of this value for two lines which are close in wavelength gives us the scale factor of η_0 between them. The continuum absorption can be assumed to be constant over small increments in wavelength, so this term can be neglected. The frequency dependent terms cancel for negligibly small differences and we get:

$$\eta_{0_1}/\eta_{0_2} = g_{l_1} f_{ul_1}/g_{l_2} f_{ul_2} \ e^{-(E_{l_1} - E_{l_2})/kT} \tag{D.6}$$

The value of $log(g_l f_{ul})$ for the magnetic lines are tabulated in many places (e.g. Borrero et al. (2003)), as are the energy of the lower level for the transition. Because we do not wish to add additional parameters to the already complex function, we assume an average value of T (5000 K).

Appendix E

Inversion Precision Tests

In this appendix I show additional tests of the effects of systematics on the precision of the 2CMO inversion technique. The plots in Figures E.1 to E.5 show the results for the tests applied to the 6302 Å data, while Figures E.6 to E.11 show the results for tests applied to the 15650 Å data. In each plot the standard deviation of magnetic field parameters (magnetic field strength, filling factor, azimuth, inclination) produced by the 2CMO inversion code is plotted against the mean value determined for every spatial position. The black points indicate umbral positions, gold points indicate penumbral positions, and blue points indicate quiet sun positions. Table E.1 lists the details of each test. Intensity information was not available for Test 1 for both the visible and infrared observations, so the quiet sun, umbral, and penumbral positions could not be identified so the points from all spatial positions appear blue.

Table E.1 Summary of Monte Carlo Tests with 2CMO

Test	Description
1	A set of 16 consecutive observations was taken with FIRS at a single slit position.
2	Spectra were resampled by random fractions of a pixel in wavelength.
3	Poisson noise corresponding to S/N=500 was introduced.
4	Poisson noise corresponding to S/N=1000 was introduced.
5	Random Q↔V and U↔V crosstalk up to 10% was introduced into the spectra.
6	Fringes with random amplitude (up to 10% intensity) and phase were introduced to Stokes I.

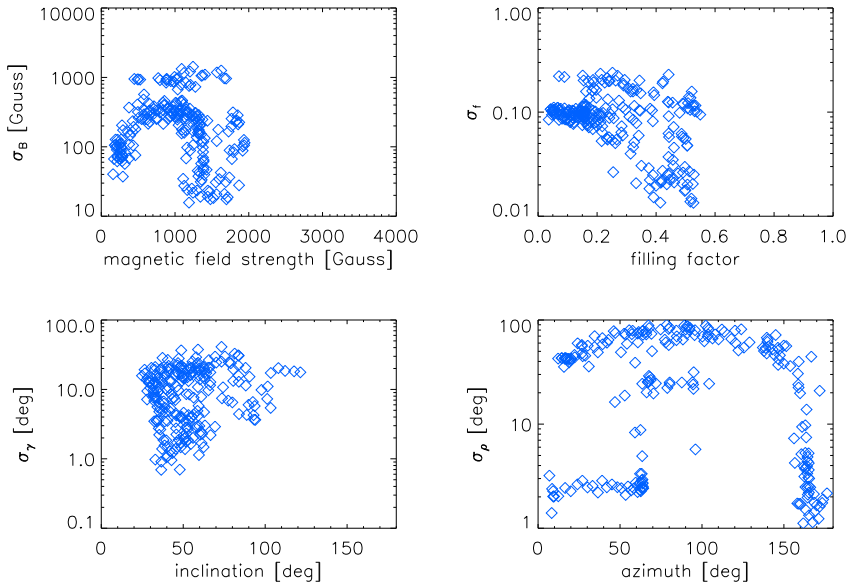

Figure E.1 The standard deviation vs. mean of the inverted magnetic field parameters for Test 1 of the 6302 Å inversion precision of 2CMO.

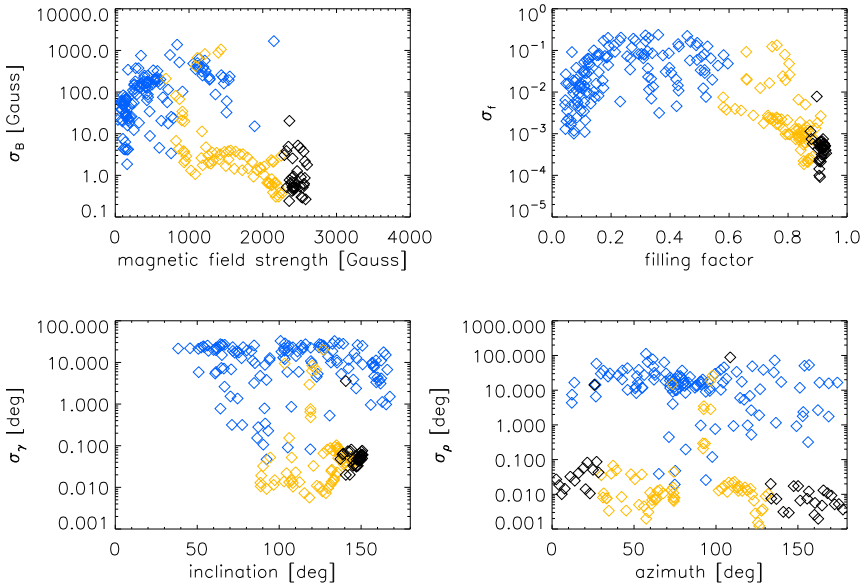

Figure E.2 The standard deviation vs. mean of the inverted magnetic field parameters for Test 2 of the 6302 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

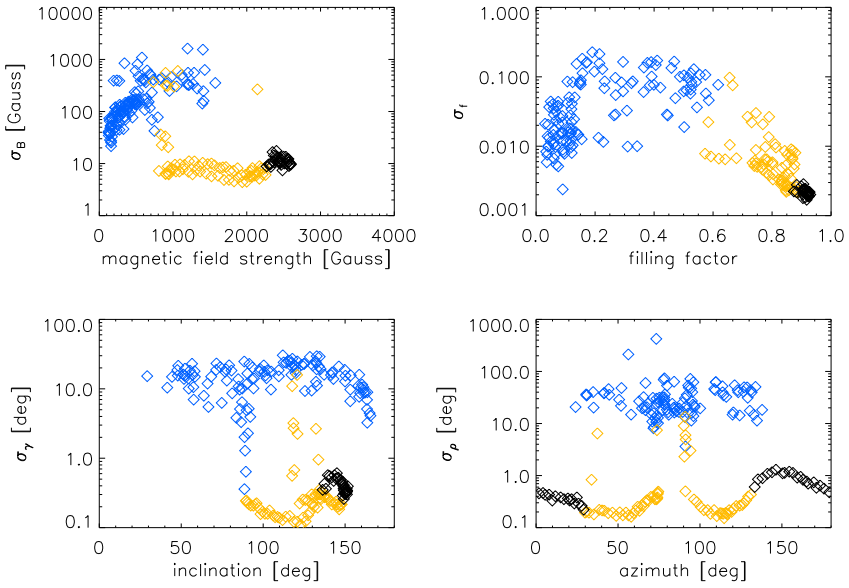

Figure E.3 The standard deviation vs. mean of the inverted magnetic field parameters for Test 3 of the 6302 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

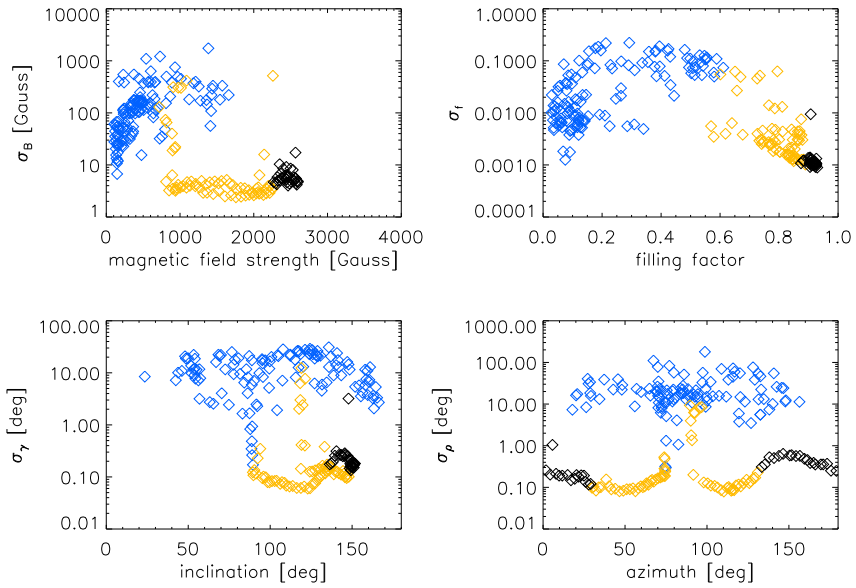

Figure E.4 The standard deviation vs. mean of the inverted magnetic field parameters for Test 4 of the 6302 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

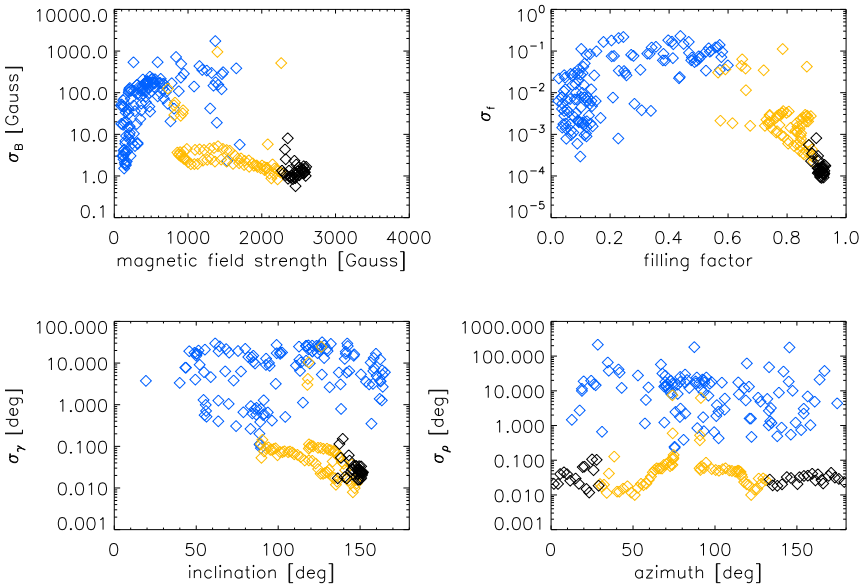

Figure E.5 The standard deviation vs. mean of the inverted magnetic field parameters for Test 5 of the Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

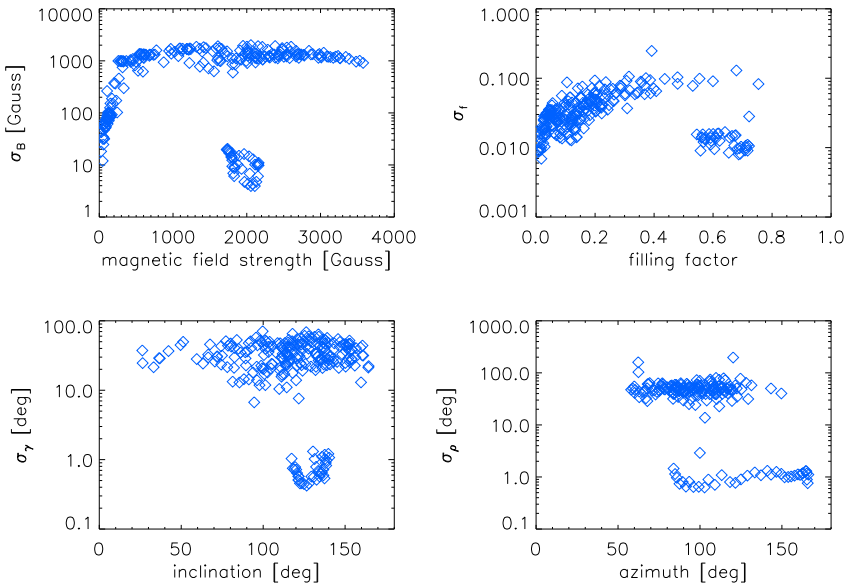

Figure E.6 The standard deviation vs. mean of the inverted magnetic field parameters for Test 1 of the 15650 Å inversion precision of 2CMO.

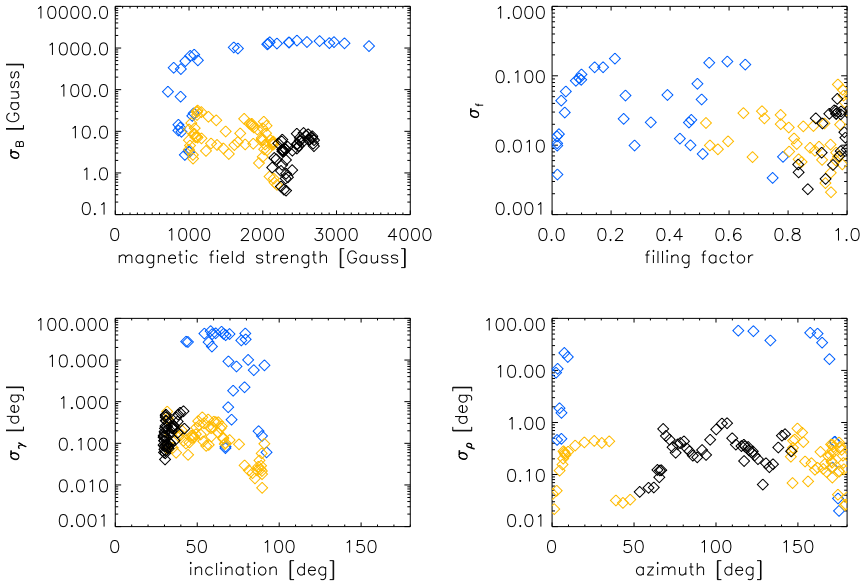

Figure E.7 The standard deviation vs. mean of the inverted magnetic field parameters for Test 2 of the 15650 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

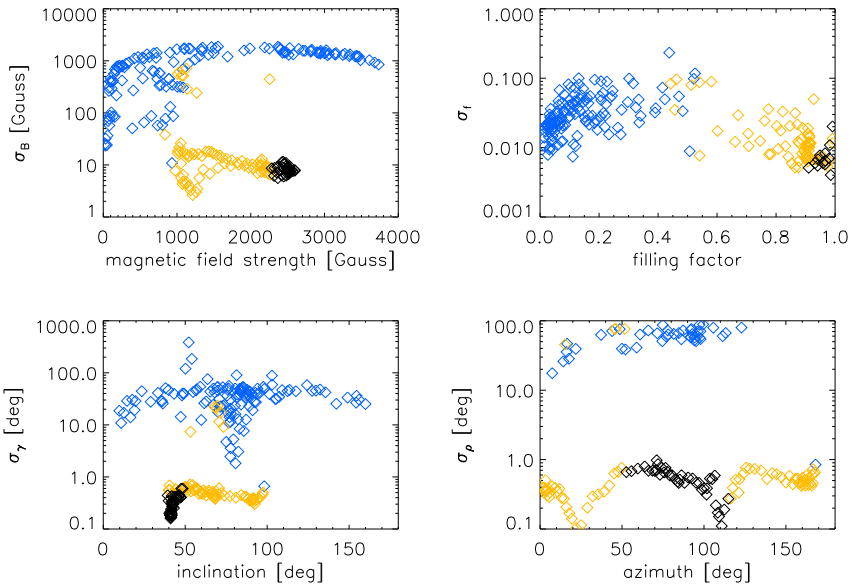

Figure E.8 The standard deviation vs. mean of the inverted magnetic field parameters for Test 3 of the 15650 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

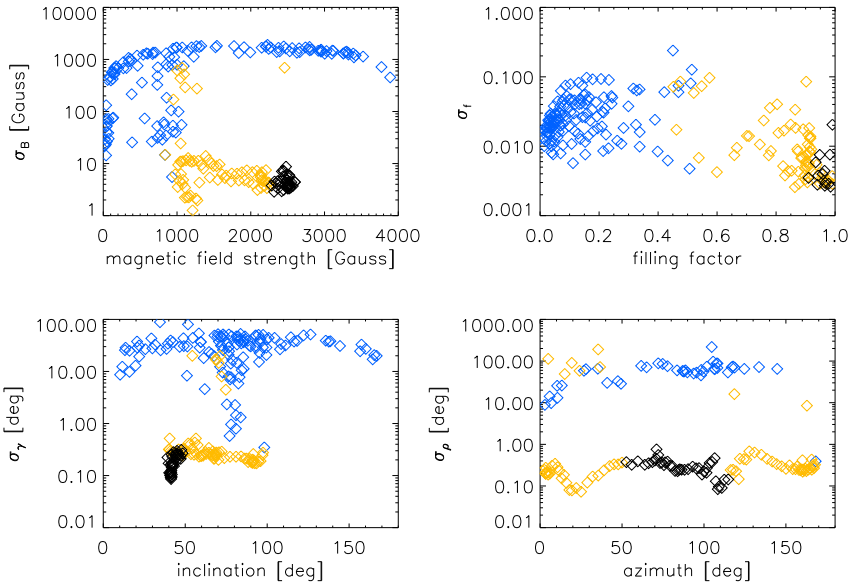

Figure E.9 The standard deviation vs. mean of the inverted magnetic field parameters for Test 4 of the 15650 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

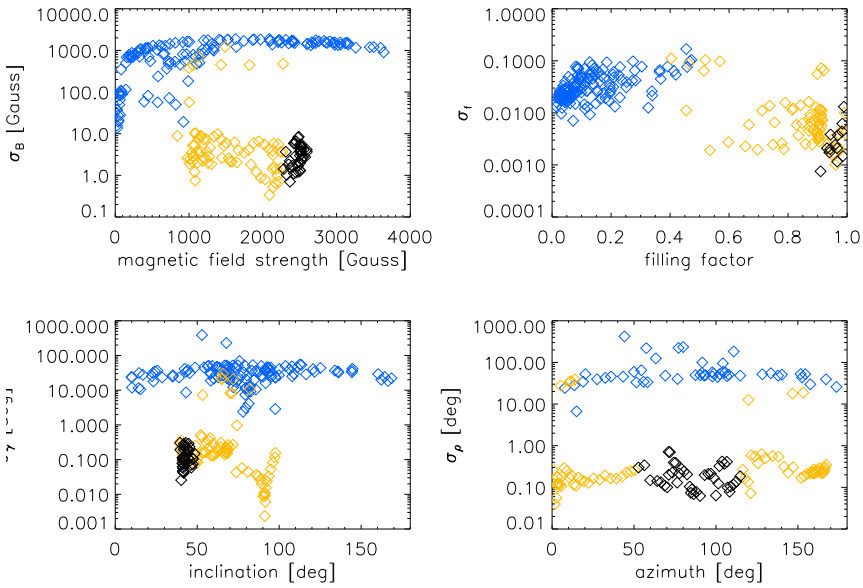

Figure E.10 The standard deviation vs. mean of the inverted magnetic field parameters for Test 5 of the 15650 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

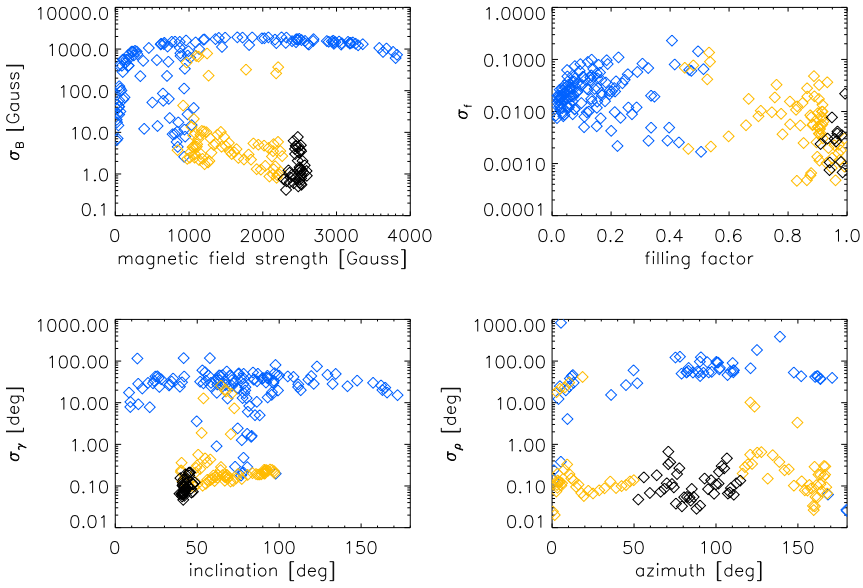

Figure E.11 The standard deviation vs. mean of the inverted magnetic field parameters for Test 6 of the 15650 Å inversion precision of 2CMO. Quiet-Sun positions are shown in blue, penumbral positions are shown in gold, and umbral positions are shown in black.

Appendix F

Notes on the Measurement of Equivalent Width

Stray light is a problem in sunspot observations (Martínez Pillet 1992) which affects the continuum contrast in the sunspot as well as the equivalent width of spectral lines. The equivalent width of a spectral line is a measure of line strength and is defined as the width of a rectangle reaching from zero intensity to the continuum level, and having the same area contained between the line and the continuum. In more mathematical terms, the equivalent width is continuum-normalized area of the line-shape, i.e. the integral between the line and the continuum level. For example, assuming a constant continuum the integral would be:

$$w = \int 1 \; - \; I(\lambda)/I_c \;\; d\lambda \tag{F.1}$$

where $I(\lambda)$ is the wavelength dependent spectral intensity, I_c is a constant representing the value of the continuum, and λ is the dependent variable, wavelength. Notice that the integral is written so the equivalent width will be positive for an absorption line. The integral of a Gaussian with unit width and unit amplitude has an area of $\sqrt{\pi}$. It follows that the equivalent width of a Gaussian line shape of the form $ae^{-(\lambda-\lambda_0)^2/2\sigma^2}$should be:

$$a\sigma\sqrt{2\pi} \tag{F.2}$$

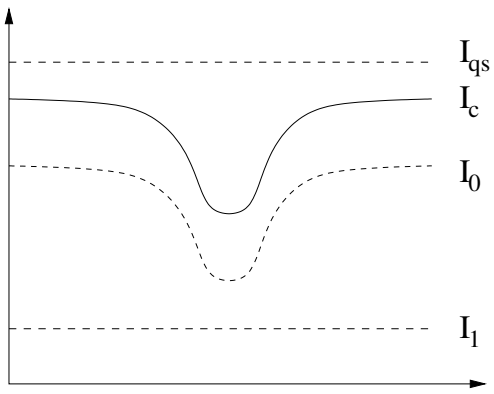

Figure F.1 A diagram of the different intensity levels involved in the calculation of equivalent width.

where a is the amplitude of the Gaussian normalized to the quiet-Sun and σ is the width.

Errors can enter into the measurement of equivalent width in a few ways: improper estimation of the continuum level for the line; poor selection of the relevant wavelengths; contamination from neighboring lines; and stray light. The equivalent width is very sensitive to small errors in the determination of continuum intensity. For example, if we are trying to measure a line with a width of 0.1 Å integrating over a 1 Å window around the line, and we overestimate the continuum by 1%, then the determined equivalent width increases by 10%. Stray light contamination from brighter regions will boost the continuum level relative to the lines in dimmer regions, reducing the equivalent width. Changes in spatial and spectral resolution should only redistribute the line flux, so if there exists an uncontaminated measurement of the equivalent width of a line (e.g. from a model or solar atlas), then it can be used to quantify the stray light level, relative to the intensity, and remove it.

Let us consider the correction of equivalent width for the case of the scattered light. Figure F.1 demonstrates the different intensity levels. The resulting continuum intensity observed is the sum of contributions from the pristine spectral intensity I_0 and a stray light component I_1 ($I_c = I_0 + I_1$). The equivalent width of the line in the uncontaminated

spectrum would of course be:

$$w_0 = \int 1 - I_0(\lambda)/I_0 \; d\lambda \tag{F.3}$$

With stray light contamination the equivalent width is altered to become:

$$w_1 = \int 1 - \frac{I_0(\lambda) + I_1}{I_0 + I_1} \; d\lambda \tag{F.4}$$

This can be rewritten in terms of the pristine equivalent width if we divide the fraction through by I_0 and move the constant factor of $1/(1 + I_1/I_0)$ outside of the integral.

$$\begin{aligned} w_1 &= \int 1 - \frac{I_0(\lambda)/I_0 + I_1/I_0}{1 + I_1/I_0} \; d\lambda \\ &= \frac{1}{1 + I_1/I_0} \int 1 + I_1/I_0 - (I(\lambda)/I_0 + I_1/I_0) \; d\lambda \\ &= \frac{1}{1 + I_1/I_0} \int 1 - I(\lambda)/I_0 \; d\lambda \\ w_1 &= \frac{w_0}{1 + I_1/I_0} \end{aligned} \tag{F.5}$$

From this relation it naturally follows that $I_1/I_0 = w_0/w_1 - 1$. Then for a measurement of equivalent width, and a known ideal equivalent width the stray contribution can be determined.

Now how we proceed depends on the stray light model that we pick. If we assume that the contamination comes from overlapping orders in the spectrograph with basically the same intensity structure as the original spectrum, the continuum contrast is preserved and the equivalent width at any given position corrects by a constant factor of the stray light fraction. However if the stray light level is constant then the continuum contrast of solar features will need to be corrected. The observed continuum contrast with stray light is really:

$$c_1 = (I_0 + I_1)/(I_0(qs) + I_1) \tag{F.6}$$

Then the true contrast can be obtained:

$$c_0 \equiv I_0/I_1(qs) = c_1(1 + I_1/I_0(qs)) - I_1/I_0(qs) \tag{F.7}$$

If we call the quantity $\alpha = I_1/I_0 = w_0/w_1 - 1$

$$c_0 = c_1(1 + \alpha_{qs}) - \alpha_{qs} \tag{F.8}$$

Then the equivalent width corrected elsewhere in the spot would be:

$$\begin{aligned} w_0 &= w_1(1 + I_1/I_0) \\ &= w_1\left(1 + \frac{I_1}{I_0(qs)}\frac{I_0(qs)}{I_0}\right) \\ &= w_1\left(1 + \frac{\alpha_{qs}}{c_1(1 + \alpha_{qs}) - \alpha_{qs}}\right) \end{aligned} \tag{F.9}$$

To put some numbers in, if the equivalent width observed is 90% that of the known value in the quiet-Sun, them the stray fraction is 10% of the quiet-Sun continuum. The contrast of a solar feature which was observed at a continuum contrast of 0.6 is actually 0.56, and the equivalent width in the feature is larger by a factor of 1.18.

References

Ahmad, Q. R., Allen, R. C., Andersen, T. C., Anglin, J. D., & 175 more authors. 2002, Phys. Rev. Lett., 89, 1301

Alfvén, H. 1943, Arkiv f. Math., Astron. o. Fys., 29, 1

Asensio Ramos, A., Martínez González, M. J., & Rubiño-Martín, J. A. 2007, A & A, 476, 959

Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D. 2004, A & A, 417, 751

Athay, R. G., Gurman, J. B., Henze, W., & Shine, R. A. 1982, ApJ, 261, 864

Auer, L. H., Heasley, J. N., & House, L. L. 1977, Sol. Phys., 55, 47

Babcock, H. W. 1961, ApJ, 133, 572

Balthasar, H. & Schmidt, W. 1993, A & A, 279, 243

Bartoe, J.-D. F., Brueckner, G. E., & Jordan, C. 1979, MNRAS, 187, 463

Bellot Rubio, L. R. 2006, in Solar Polarization 4, ed. R. Casini & B. W. Lites, ASPC

Berdyugina, S. V., Solanki, S. K., & Frutiger, C. 2003, A & A, 412, 513

Bernasconi, P. N. 1997, PhD thesis, Swiss Federal Institute of Technology

Bethe, H. A. 1939, Phys. Rev., 55, 434

Biermann, L. 1941, Astron. Ges., 76, 248

Black, J. H. & van Dishoeck, E. F. 1987, ApJ, 322, 412

Borrero, J. M., Bellot Rubio, L. R., Barklem, P. S., & del Toro Iniesta, J. C. 2003, A & A, 404, 749

Cabrera Solana, D., Bellot Rubio, L. R., Beck, C., & del Toro Iniesta, J. C. 2006, in ASPC, Vol. 358, Solar Polarization 4, ed. R. Casini & B. W. Lites, 25

Carrington, R. C. 1859, MNRAS, 20, 13

Collados, M. 2003, in Polarimetry in Astronomy, ed. S. Fineschi, Vol. 4843, 55

Collados, M., Trujillo Bueno, J., & Asensio Ramos, A. 2003, in ASPC, Vol. 307, Solar Polarization 3, ed. J. Trujillo Bueno & J. Sanchez Almeida, 468

Cowling, T. G. 1976, Magnetohydrodynamics, Monographs on Astronomical Subjects (Hilger)

Cox, A. N., ed. 2000, Allen's Astrophysical Quantities, 4th edn. (Springer-Verlag New York, Inc.)

Curdt, W., Brekke, P., Feldman, U., Wilhelm, K., Dwivedi, B. N., Schühle, U., & Lemaire, P. 2004, A & A, 375, 591

Davis, R. & Cox, A. N. 1991, Solar neutrino experiments, ed. A. N. Cox, W. C. Livingston, & M. S. Matthews (University of Arizona Press, Tucson, AZ), 51

Deinzer, W. 1965, ApJ, 141, 548

del Toro Iniesta, J. C. 2003, AN, 324, 383

del Toro Iniesta, J. C. & Ruiz Cobo, B. 1996, Sol. Phys., 164, 169

del Toro Iniesta, J. C., Ruiz Cobo, B., Bellot Rubio, L. R., & Collados, M. 1995, A & A, 294, 855

Eddington, A. S. 1924, MNRAS, 84, 308

Elmore, D. F., Lin, H., Socas-Navarro, H., & Jaeggli, S. A. 2010, in Ground-based and Airborne Instrumentation for Astronomy III, ed. I. S. McLean, S. K. Ramsay, & H. Takami, Vol. 7735, SPIE

Elmore, D. F., Lites, B. W., Tomczyk, S., Skumanich, A. P., Dunn, R. B., Schuenke, J. A., Streander, K. V., Leach, T. W., Chambellan, C. W., & Hull, H. K. 1992, in SPIE, Vol. 1746, Polarization analysis and measurement, ed. D. H. Goldstein & R. A. Chipman, 22

Evershed, J. 1909, MNRAS, 69

Foukal, P. 1990, Solar Astrophysics (John Wiley & Sons, Inc.)

Freedman, R. A. & Kauffman, W. J. 2008, Universe: Stars & Galaxies (W. H. Freedman & Company)

Gautier III, T. N., Fink, U., Treffers, R. R., & Larson, H. P. 1976, ApJL, 207, 129

Georgoulis, M. K. 2005, ApJL, 629, 69

Gingerich, O., Noyes, R. W., & Kalkofen, W. 1971, Sol. Phys., 18, 347

Grevesse, N. & Sauval, A. J. 1994, in Lecture Notes in Physics, Vol. 428, Molecules in the Stellar Environment, ed. U. G. Jorgensen (Springer-Verlag), 196

Gurman, J. B. & House, L. L. 1981, Sol. Phys., 71, 5

Hale, G. E. 1908, ApJ, 28, 315

Hauschildt, P. H., Baron, E., & Allard, F. 1997, ApJ, 483, 390

Herzberg, G. 1950, Spectra of Diatomic Molecules, 2nd edn., Molecular Spectra and Molecular Structure (Van Nostrand Reinhold Company, New York)

Hill, F., Howe, R., Komm, R., Christensen-Dalsgaard, J., Larson, T. P., Schou, J., & Thompson, M. J. 2011, SPD

Ichimoto, K., Lites, B., Elmore, D., Suematsu, Y., Tsuneta, S., Katsukawa, Y., Shimizu, T., Shine, R., Tarbell, T., Title, A., Kiyohara, J., Shinoda, K., Card, G., Lecinski, A., Streander, K., Nakagiri, M., Miyashita, M., Noguchi, M., Hoffmann, C., & Cruz, T. 2008, Sol. Phys., 248, 233

Innes, D. E. 2008, A & A L, 481, 41

Jackson, J. D. 1962, Classical Electrodynamics (John Wiley & Sons, Inc.)

Jaeggli, S. A. 2010, FIRS Data Manual, Institute for Astronomy, University of Hawai'i

Jaeggli, S. A., Lin, H., Mickey, D. L., Kuhn, J. R., Hegwer, S. L., Rimmele, T. R., & Penn, M. J. 2010, in MmSAI, Vol. 81, Chromospheric Structure and Dynamics: From Old Wisdom to New Insights, 763

Jefferies, J., Lites, B. W., & Skumanich, A. 1989, ApJ, 343, 920

Jordan, C., Brueckner, G. E., Bartoe, J.-D. F., Sandlin, G. D., & VanHoosier, M. E. 1978, ApJ, 226, 687

Kneer, F. 1972, A & A, 18, 47

Kopp, G. & Rabin, D. 1993, ApJ, 69, 69

Kosugi, T., Matsuzaki, K., Sakao, T., Shimizu, T., Sone, Y., Tachikawa, S., Hashimoto, T., Minesugi, K., Ohnishi, A., Yamada, T., Tsuneta, S., Hara, H., Ichimoto, K., Suematsu, Y., Shimojo, M., Watanabe, T., Shimada, S., Davis, J. M., Hill, L. D., Owens, J. K., Title, A. M., Culhane, J. L., Harra, L. K., Doschek, G. A., & Golub, L. 2007, Sol. Phys., 423, 3

Kuhn, J. R., Balasubramaniam, K. S., Kopp, G., Penn, M. J., Dombard, A. J., & Lin, H. 1994, Sol. Phys., 153, 143

Kuhn, J. R., MacQueen, R. M., Streete, J., Tansey, G., Mann, I., Hillebrand, P., Lin, H., Edmunds, D., & Judge, P. 1999, ApJ, 521, 478

Kuhn, J. R. & Morgan, H. 2006, in ASPC, Vol. 354, Solar MHD Theory and Observations: A High Spatial Resolution Perspective, ed. J. Leibacher, R. F. Stein, & H. Uitenbroek, 230

Kupke, R., Labonte, B. J., & Mickey, D. L. 2000, Sol. Phys., 191, 97

Landi degl'Innocenti, E. & Landi degl'Innocenti, M. 1985, Sol. Phys., 97, 239

Landolfi, M. & Landi degl'Innocenti, E. 1982, Sol. Phys., 78, 355

Lin, H. 1995, ApJ, 446, 421

Lin, H., Penn, M. J., & Kuhn, J. R. 1998, ApJ, 493, 978

Lites, B. W., Casini, R., Garcia, J., & Socas-Navarro, H. 2007, in MmSAI, Vol. 78, Solar Magnetism and Dynamics and Themis User Meeting, 148

Lites, B. W., Elmore, D. F., Seagraves, P., & Skumanich, A. P. 1993, ApJ, 418, 928

Livingston, W. 2002, Sol. Phys., 207, 41

Livingston, W. C., Penn, M. J., & Svalgard, L. 2011, SPD

Maltby, P., Avrett, E. H., Carlsson, M., Kjeldseth-Moe, O., Kurucz, R. L., & Loeser, R. 1986, ApJ, 306, 284

Martínez González, M. J., Collados, M., Ruiz Cobo, B., & Beck, C. 2008, A & A, 477, 953

Martínez Pillet, V. 1992, Sol. Phys., 140, 207

Martínez Pillet, V. & Vázquez, M. 1993, A & A, 270, 494

Mathew, S. K., Solanki, S. K., Lagg, A., Collados, M., Borrero, J. M., & Berdyugina, S. 2004, A & A, 422, 693

Mäthger, L. M. & Hanlon, R. T. 2006, Biol. Lett., 2, 494

Matta, F. & Reichel, A. 1971, Math. Comp., 25, 339

Meyer, F., Schmidt, H. U., & Weiss, N. O. 1977, MNRAS, 179, 741

Moore, R. & Rabin, D. 1985, ARAA, 23

Muñoz-Jaramillo, A. 2011, The Solar Magnetic Cycle: Past, Present, and Future

Orozco Suárez, D., Bellot Rubio, L. R., & del Toro Iniesta, J. C. 2010, A & A, 518, 3

Penn, M. J., Cao, W. D., Walton, S. R., Chapman, G. A., & Livingston, W. 2003a, Sol. Phys., 215, 87

Penn, M. J., Ceja, J. A., Bell, E., Frye, G., & Linck, R. 2003b, Sol. Phys., 213, 55

Penn, M. J. & Livingston, W. 2006, ApJL, 649, 45

Penn, M. J., Walton, S., Chapman, G., Ceja, J., & Plick, W. 2002, Sol. Phys., 205, 53

Rempel, M. 2010, in Physics of Sun and Star Spots, IAU Symposium No. 273 (arXiv:1011.0981v1)

Rempel, M., Schüssler, M., & Knölker, M. 2009, ApJ, 691, 640

Rimmele, T., Richards, C., Hegwer, S., Fletcher, S., Gregory, S., Moretto, G., Didkovsky, L. V., Denker, C. J., Dolgushin, A., Goode, P. R., Langlois, M., Marino, J., & Marquette, W. 2004, SPIE, 5171, 179

Ruiz Cobo, B. 2007, in Modern Solar Facilities – Advanced Solar Science, ed. F. Kneer, K. G. Puschmann, & A. D. Wittmann, 287

Ruiz Cobo, B. & del Toro Iniesta, J. C. 1992, ApJ, 398, 375

Rybiki, G. B. & Lightman, A. P. 1979, Radiative Processed in Astrophysics (John Wiley & Sons, Inc.)

Sabine, E. 1852, PTRSL, 142, 103

Sanchez Almeida, J., Ruiz Cobo, B., & del Toro Iniesta, J. C. 1996, A & A, 314, 295

Schühle, U., Brown, C. M., Curdt, W., & Feldman, U. 1999, in ESASP, Vol. 446, 8th SOHO Workshop: Plasma Dynamics and Diagnostics in the Solar Transition Region and Corona, 617

Schwabe, H. 1843, AN, 20, 283

Scoville, N., Kleinmann, S. G., Hall, D. N. B., & Ridgway, S. T. 1983, ApJ, 275, 201

Skumanich, A. & Lites, B. W. 1987, ApJ, 322, 473

Skumanich, A., Lites, B. W., Martínez Pillet, V., & Seagraves, P. 1997, ApJS, 110, 357

Socas-Navarro, H. 2001, in ASPC, Vol. 236, Advanced Solar Polarimetry – Theory, Observation, and Instrumentation, ed. M. Sigwarth, 487

Socas-Navarro, H. 2003, in ASPC, Vol. 307, Solar Polarization, ed. J. Trujillo Bueno & J. Sanchez Almeida, 330

Socas-Navarro, H., Elmore, D., & Asensio Ramos, A. 2011, A & A *(submitted)*, 531, 2

Solanki, S. K. 2003, A & A Rev., 11

Solanki, S. K., Ruedi, I., & Livingston, W. 1992, A & A, 263, 312

Solanki, S. K., Walther, U., & Livingston, W. 1993, A & A, 277, 639

Sonnett, S. 2007, "Observing Molecular Hydrogen in the Solar Photosphere," Institute for Astronomy, 699 paper

Spinrad, H. 1966, ApJ, 145, 195

Spruit, H. C. 1974, Sol. Phys., 34, 277

Stanchfield, D. C., Thomas, J. H., & Lites, B. W. 1997, ApJ, 477, 485

Tatum, J. B. 1966, PDAO, 13, 1

Thomas, J. H. & Weiss, N. O. 2004, ARAA, 42, 517

Tsuneta, S., Suematsu, Y., Ichimoto, K., Shimizu, T., Otsubo, M., Nagata, S., Katsukawa, Y., Title, A., Tarbell, T., Shine, R., Rosenberg, B., Hoffmann, C., Jurcevich, B., Levay, M., Lites, B., Elmore, D., Matsushita, T., Kawaguchi, N., Mikami, I., Shimada, S., Hill, L., & Owens, J. 2008, Sol. Phys., 249, 167

Uitenbroek, H. 2000a, ApJ, 531, 571

—. 2000b, ApJ, 536, 481

Unno, W. 1956, PASJ, 8, 108

Wallace, L. & Hinkle, K. 2001, Sunspot Umbral Spectra in the Region 4000 to 8640 cm-1 (1.16 to 2.50 microns), Tech. Rep. 01-001, National Solar Observatory

Watson, F., Fletcher, L., Dalla, S., & Marshall, S. 2009, Sol. Phys., 260, 5

Westendorp Plaza, C., Del Toro Iniesta, J. C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B. W., & Skumanich, A. 2001, ApJ, 547, 1130

Wolf, R. 1852, AN, 35, 369

Wood, K. & Fox, G. K. 1995, Inv. Prob., 11, 795