

密级 _____



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

光球活动区磁非势性及其关联耀斑

作者姓名 _____ 杨 潇

指导教师 _____ 张洪起 研究员 林钢华 研究员

_____ 中国科学院国家天文台

学位类别 _____ 理学博士

学科专业 _____ 天体物理

培养单位 _____ 中国科学院国家天文台

2013 年 11 月

Typeset by L^AT_EX 2_ε at June 19, 2014

With package C^ASthesis v0.2 of C_TE_X.ORG

Active-Region Magnetic Nonpotentiality in Solar Photosphere and Associated Flares

By
YANG Xiao

A Dissertation Submitted to
University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Doctor of Science

National Astronomical Observatories
Chinese Academy of Sciences

November, 2013

摘 要

磁场主导着太阳的演化，磁场的非势性质与诸如耀斑、日冕物质抛射等太阳爆发活动密切相关。尽管反映磁场非势性的物理量的演化并不直接触发太阳爆发，具有强非势性和复杂性的活动区由于拥有足够可供释放的自由能量，相比简单的活动区更容易爆发。从统计角度研究磁非势性的太阳周演化及其与太阳爆发的关系，对于理解太阳活动的长期演化、预测爆发活动等是非常有用的。目前，地面望远镜由于维护及操控的便捷，非常适合于长时期的观测研究，在数据积累方面相比空间望远镜仍有其优势。

本文利用怀柔太阳观测基地太阳磁场望远镜 20 余年的矢量磁场数据，从统计角度研究了太阳光球活动区磁场非势性的强度随太阳活动周的演化，一并探讨了这些非势性物理量与关联耀斑的关系。基于活动区磁非势性和关联耀斑的统计关系，我们利用一种简单的通用机器学习方法检验了这些磁非势性参量作为太阳耀斑预测因子的预测性能。所得到的主要结果有：

(1) 活动区的两个磁剪切角均值 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ ，纵向电流密度绝对值的均值 $\overline{|J_z|}$ ，平均无力因子的绝对值 $|\alpha_{av}|$ ，以及有效距离 d_E 随整体太阳活动水平变化不显著。然而，这些参量仍能反映出活动区在太阳活动峰年较低年具有更强的非势性。

(2) 电流螺度密度绝对值的均值 $\overline{|h_c|}$ 、自由磁能密度均值 $\overline{\rho_{free}}$ 、纵场加权后的有效距离 d_{Em} 与平均黑子数显现出高的正相关，且三者之间也有着相对紧密的关系。三者与年均黑子数的线性相关系数都高于 0.59。它们可以用于表征太阳活动水平。

(3) 经纵场加权后的有效距离 d_{Em} 由于加入了磁场强度信息而能更好地反映活动区的磁活动特征。

(4) 这些反映活动区磁场非势性的物理量有助于提高耀斑预测、尤其是较大级别耀斑预测的预测性能。由于它们的获取仅依赖于本地观测，且它们相比于仅从视向磁场得到的物理量具有更全面的磁场信息，将它们应用于太阳耀斑的短期预测中能更高效、更可靠的得到预测结果，从而对本地观测和其他监测使用部门提供有一定价值的参考。

并且，我们在文中强调：为避免单个指标对优化过程的误导以及对结果的误用，预测结果应该被仔细地评估，注意使用多个性能指标以公平评价预测结果。如分开训练集和测试集、 k -折交叉确认这样的步骤在提高预测模型的泛化能力方面是非常必要的。另外，不同数据集的内禀特性会使得同样的预测工具表现得非常不同，比较预报因子或预报方法需基于同样的数据环境才是有意义的。希望这些问题能得到实际操作人员和预测使用者的注意和重视。

本文的结果对认识活动区磁场的活动周演化、活动区磁非势性与耀斑的关系有一定的参考价值。同时，耀斑是灾害性空间天气的驱动源之一，从预测因子的角度提高耀斑预测性能，本文的结果也具有一定的参考意义和实用价值。

关键词： 耀斑预测，磁非势性，光球活动区，太阳耀斑，支持向量机

Abstract

Magnetic fields dominate the evolution of the Sun. The accumulation of magnetic nonpotential energy is responsible for the solar eruptions such as flares and CMEs. Though evolutions of the measures that represent the magnetic nonpotentiality do not trigger solar eruptions directly, the active regions with strong nonpotentiality and great complexity are easier to erupt than the simpler ones, due to the sufficient free energy to release. Studying the evolution of the magnetic nonpotentiality with solar cycles from a statistical point of view, is helpful to understand the long-term evolution of solar activities and to predict eruptive activities. So far ground-based telescopes, whose maintenance and manipulation for long-term observations are relatively easier than space-borne telescopes, still have advantages in data accumulation.

In this paper, based on the vector magnetograms observed by the Solar Magnetic Field Telescope at Huairou Solar Observing Station over 20 years, a statistical study is carried out on the photospheric magnetic nonpotentiality in solar active regions and its relationship with associated flares. Furthermore, in the light of the statistical relationship between magnetic nonpotentiality in active regions and associated flares, we utilize a simple kind of general learning machine to verify the flare prediction performance of these magnetic nonpotentiality parameters. The main results are:

(1) The two mean magnetic shear angles $\overline{\Delta\phi}$ and $\overline{\Delta\psi}$, the mean absolute vertical current density $\overline{|J_z|}$, the absolute twist factor $|\alpha_{av}|$, and the effective distance d_E in ARs do not change significantly with the global solar activity level. However, it is more likely that these parameters show higher values in the solar maximum than in the solar minimum.

(2) The mean absolute current helicity density $\overline{|h_c|}$, the mean free magnetic energy density $\overline{\rho_{free}}$, and the longitudinal-field weighted effective distance d_{Em} show high positive correlation with the mean sunspot number, and these parameters also have relatively close relationship with each other. The Pearson linear

correlation coefficients of the above three with the yearly mean sunspot numbers are larger than 0.59. They can be used to characterize the solar activity level.

(3) The longitudinal-field weighted effective distance d_{Em} (including the strength of the magnetic field) turns out to be much better in indicating the magnetic activities of active regions.

(4) The magnetic measures that reflect the magnetic nonpotentiality in active regions improve the flare prediction performance, especially for more powerful flares. Their acquisition only depends on the local observations, and they have more comprehensive information of magnetic fields than the measures only obtained from line-of-sight magnetic fields. Therefore, the prediction results will be more efficient and more reliable by applying them as predictors in the flare predictions, and will be very useful to provide valuable information to the local observations and the users of other monitoring departments.

In addition, we would like to emphasize in the text that: To avoid misleading the optimization work or misusing the results from a single verification measure, prediction results should be assessed carefully. *Multiple* verification measures are probably acceptable to fairly evaluate the prediction results. Steps like k -fold cross-validation and exclusively dividing training and testing sets are necessary for improving the generalization capability of the prediction models. The intrinsic properties of various data sets may make a specific tool perform rather differently, and hence, it is then significant to make comparisons on predictors or on methods in the same data environment. We hope these issues could be noticed and paid attention to by the practical staff and the users of the predictions.

The results of this paper have certain reference values to understand the solar-cycle evolutions of the magnetic fields in active regions, and the relationship between active-region magnetic nonpotentiality and flares. Meanwhile, since solar flare is a driver of disastrous space weather, our results have certain values in practical operations as references, by improving the prediction performance of solar flares from a predictor way.

Keywords: Flare Prediction, Magnetic Nonpotentiality, Photospheric Active Regions, Solar Flares, Support Vector Machine

目 录

摘要	i
Abstract	iii
目录	v
第一章 引言	1
1.1 太阳磁场与太阳活动	1
1.1.1 太阳磁场观测	1
1.1.2 太阳爆发活动	4
1.2 太阳磁场非势性研究	9
1.3 太阳耀斑预测	12
1.4 本文的研究目的与内容	15
第二章 本论文研究的工具基础	17
2.1 数据挖掘概念	17
2.2 高性能计算	19
2.3 本论文研究的数据基础	23
第三章 光球活动区磁非势性随活动周的演化及其与耀斑的关系	25
3.1 观测和数据处理	25
3.1.1 怀柔太阳观测站的观测	25
3.1.2 数据选择与预处理	26
3.2 磁非势性和磁复杂性参量	30
3.2.1 磁剪切角	30
3.2.2 纵向电流密度, 电流螺度密度, 无力因子	31
3.2.3 自由磁能密度	32

3.2.4	有效距离	32
3.3	统计分析与结果	33
3.3.1	22 – 23 太阳周非势性的强度分布	33
3.3.2	非势性与其关联耀斑	39
3.3.3	非势参量与有效距离的比较	43
3.4	结论与讨论	43
第四章	基于光球活动区磁非势性的太阳耀斑预测	47
4.1	数据与方法	47
4.1.1	数据及其预处理	47
4.1.2	作为预报因子的磁非势性参量	48
4.1.3	预测方法：支持向量分类	50
4.2	实验与结果	51
4.2.1	实验设计	51
4.2.2	性能评价	52
4.2.3	实验结果	57
4.3	结论与讨论	60
第五章	总结与展望	67
5.1	总结	67
5.2	展望	68
参考文献	69
发表文章目录	87
简历	89
致谢	91

表 格

1.1	H α 耀斑分类	8
1.2	软 X 射线耀斑分类	8
3.1	数据样本在 1988 年至 2008 年间的数量分布	27
3.2	八个参量年 (月) 均值分别与年 (月) 均黑子数的线性相关系数	38
3.3	22 周下降期 (1992 – 1996) 与 23 周下降期 (2002 – 2007) 各参量 年均值比较	40
3.4	八个参量间的线性相关系数	43
4.1	耀发 (f) 和无耀发 (n-f) 的样本分布	52
4.2	2 \times 2 混淆矩阵的定义	53
4.3	评价指标	55
4.4	两种情形下的两种预测结果举例	58
4.5	测试 SVC 的评价结果	59
4.6	C1.0 级以上耀斑的预测结果评价	62
4.7	C5.0 级以上耀斑的预测结果评价	63
4.8	M1.0 级以上耀斑的预测结果评价	64
4.9	M5.0 级以上耀斑的预测结果评价	65
4.10	X1.0 级以上耀斑的预测结果评价	66

插图

1.1	1610 – 2010 年平均黑子数 (引自 NASA 网站)	2
1.2	黑子蝴蝶图和日均黑子面积百分比 (1875 年至今) (引自 NASA 网站)	3
1.3	怀柔太阳观测基地全日面色球 $H\alpha$ 望远镜观测图像示例 (经消除临边昏暗改正)	5
1.4	太阳分层结构及日地空间关系 (引自 NASA 网站)	6
1.5	太阳活动对人类活动的影响 (引自 NASA 网站)	7
1.6	24 太阳周黑子数预测 (引自 NASA 网站)	12
1.7	10.7 cm (2.8 GHz) 太阳射电流量预测 (引自 NASA 网站)	13
3.1	挑选磁图示例	27
3.2	怀柔活动区号与 NOAA 活动区号的一一对应	28
3.3	对应怀柔活动区号与 NOAA 活动区号的参考资料	29
3.4	一张矢量磁图示例 — NOAA 活动区 5356, 观测于 1989 年 2 月 15 日 02:12 UT	34
3.5	1988 – 2008 年活动区样本的 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{ J_z }$ 、 $\overline{ h_c }$ 、 $ \alpha_{av} $ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的散点图	35
3.6	1988 – 2008 年活动区样本的 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{ J_z }$ 、 $\overline{ h_c }$ 、 $ \alpha_{av} $ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的年均值	36
3.7	1988 – 2008 年活动区样本中活跃样本与宁静样本分别的 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{ J_z }$ 、 $\overline{ h_c }$ 、 $ \alpha_{av} $ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的年均值	37
3.8	1988 – 2008 年月均黑子数、耀斑指数、10.7 cm 射电流量	38
3.9	$\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{ J_z }$ 、 $\overline{ h_c }$ 、 $ \alpha_{av} $ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的耀斑产率	42
3.10	$\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ 、 $\overline{ h_c }$ 和 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 、 $\overline{ h_c }$ 和 $\overline{ J_z }$ 的相关	45

第一章 引言

太阳是位于 Hertzsprung-Russell 图主序星带上的一颗典型恒星，其恒星分类为 G2V (G2 为其光谱类型，V 为其光度类型¹⁾)。得益于与地球的近距离，太阳是唯一一颗可以进行高空间分辨率观测的恒星，人们可以对它进行更细致的研究。也正因为这个近距离，太阳与夜间天文的观测无论从望远镜设计、图像采集、数据处理等多方面都存在着很大的差异。太阳物理是天体物理的一个重要分支，研究太阳的活动规律和演化机制，对认识恒星世界具有重要的普遍性意义。太阳也是检验原子物理、核物理、高温等离子体物理、磁流体力学、中微子物理、和广义相对论等的一个理想的宇宙实验室。同时，太阳还是地球及其生命的孕育者和主宰者，它是整个太阳系的核心，研究太阳活动及其对地球的影响具有重要的现实意义和实用价值。

1.1 太阳磁场与太阳活动

宇宙中 99% 以上正常物质 (除暗物质、暗能量外) 的状态为磁化等离子体。等离子体是具有整体电中性、局部存在电磁作用的介质，由自由电子和带电离子为主要成分，常被称为物质的第四态。由于宇宙中电阻较低，往往存在不同尺度的电流，故而存在不同尺度的磁场。从早期宇宙磁场、星系际磁场、星系团磁场、银河系磁场，到行星际磁场、行星磁场、恒星磁场，磁场在天体的各层次都普遍存在。太阳上动态的、瞬息万变的活动现象，来源于磁场及与磁场密切相关的等离子体的行为和演化。对太阳活动的研究，主要是研究其多种空间尺度下的磁场的发生、发展和消亡，揭示太阳各层次特征物和太阳爆发活动的产生机理和演化规律，认识和预测太阳活动对地球、对人类的影响。

1.1.1 太阳磁场观测

对太阳的观测研究始自太阳黑子。古代中国和希腊的天文学家很早就观测到了太阳黑子，后来 Galileo 最早对其在日面上的运动进行了系统地观测。四百年以来，人们一直保持追踪黑子的出现和消失，希望能知道太阳周变化过程的

¹GV 型主序恒星又叫黄矮星，质量约为 $0.8 - 1.0 M_{\odot}$ ，表面有效温度约 $5300 \text{ K} - 6000 \text{ K}$ 。

驱动机制、以及太阳活动与地球的气候变化到底有何联系。关于太阳的许多疑问至今仍然是迷，需要利用不断出现的观测手段、经过一代又一代科研人员的不断努力，去逐渐揭开答案。通过对黑子长期的观察和记录，人们了解了它在日面上出现平均约为 11 年的周期 (图 1.1²) 及它所处纬度变化的蝴蝶图 (图 1.2³)。通过对太阳表面现象的其他物理量的测量和计算，同样也可以描画出具有约 11 年周期的曲线图。

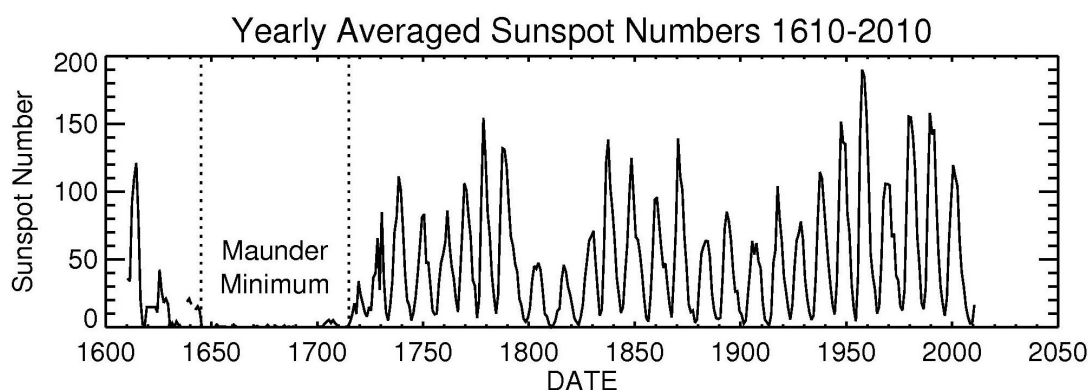


图 1.1: 1610 – 2010 年平均黑子数 (引自 NASA 网站)

太阳黑子是太阳上的局部强磁场区，活动区的概念是以黑子为核心而定义的。在太阳的不同层次上，也或强或弱地同时表现着这些强磁场区所对应的特征物，如光球上的黑子、光斑，色球上的谱斑，日冕层次的冕环等等。许多太阳爆发活动也都与以黑子为中心的活动区密切相关。研究太阳爆发的物理起源离不开对太阳活动区磁场的观测。

太阳磁场的观测要追溯到 20 世纪初，Hale (1908) [65] 利用在实验室里发现的 Zeeman 效应 (Zeeman 1897a,b [159,160]) 首次测量到太阳黑子的磁场。然而，直到 Babcock 父子系统地观测太阳磁场后 (Babcock & Babcock 1952 [12]; Babcock & Livingston 1958 [11]; Babcock 1959 [10])，太阳磁活动的面纱才被一层层揭开。太阳物理的研究便开始围绕太阳磁场观测展开。发展到现在，已形成对活动区、宁静区 (网络内、网络、极区) 等不同对象，光球、色球和日冕等不同层次，局部和全日面等不同尺度的磁场的观测和理论研究。人类目前对光球层磁场的测量比较可靠，但横场方向的 180° 不确定性、约为 100 G 的横

²http://solarscience.msfc.nasa.gov/images/ssn_yearly.jpg

³<http://solarscience.msfc.nasa.gov/images/bfly.gif>

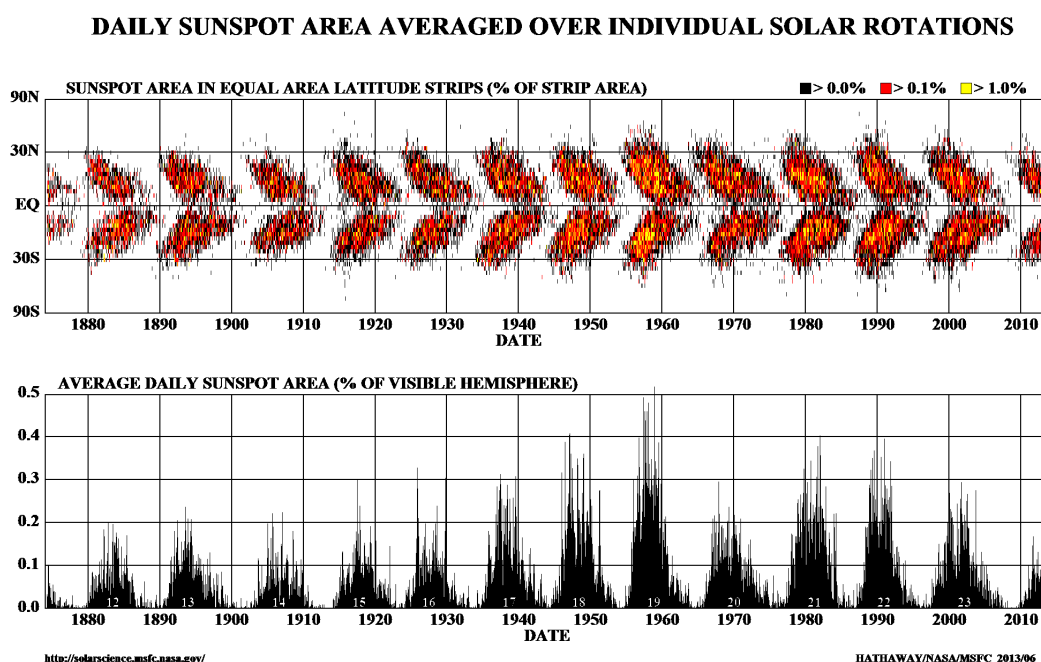


图 1.2: 黑子蝴蝶图和日均黑子面积百分比 (1875 年至今) (引自 NASA 网站)

场测量误差、磁光效应等问题依然存在，对色球层磁场的测量精度较差 (偏振信号比光球弱、谱线反演需要考虑非局部热动平衡等)，对日冕层的测量则暂时只能用射电方法做粗略的量级估计。由于许多重要的现象都是发生在日冕层次，我们了解太阳大气磁场和其中能量的产生的一个关键方法是根据理论模型从光球层磁场构造或外推日冕磁场。非线性无力场外推则需要光球矢量磁场作为初始条件。系统地进行矢量磁场观测的几个主要台站有中国科学院国家天文台的怀柔太阳观测基地 (Ai & Hu 1986 [3])，日本国立天文台的 Mitaka 观测站 (Sakurai et al. 1995 [118])，美国 NASA 的马歇尔空间飞行中心 (MSFC; Hagyard et al. 1982 [62])，美国国立太阳天文台的 Kitt 峰观测站 (Henney et al. 2006 [72])，美国夏威夷的 Mees 观测站 (Mickey et al. 1996 [97]) 等。搭载在 Hinode 空间卫星上的 SOT (Solar Optical Telescope; Tsuneta et al. 2008 [135]) 2006 年起开始得到矢量磁图资料。SDO (Solar Dynamics Observatory) 上搭载的 HMI (Helioseismic and Magnetic Imager; Scherrer et al. 2012 [121]) 也从 2010 年发射升空后开始常规观测全日面光球矢量磁场，时间分辨率为 12 分钟。

所谓 Zeeman 效应,是指原子谱线在外磁场作用下会发生分裂。Lorentz 应用经典电磁理论对该现象进行了解释,认为由于电子存在轨道磁矩,且磁矩方向在空间的取向是量子化的,因此在磁场作用下能级发生分裂,谱线分裂成间隔相等的三条谱线,这是正常 Zeeman 效应。在实验中观察到很多光谱线有时并非分裂成三条,间隔也不一定相同,这称为反常 Zeeman 效应。在天体物理中,利用 Zeeman 效应的偏振特性,可通过对线偏振光和左、右旋圆偏振光的测量,根据谱线在磁场中的裂距公式 $\Delta\lambda = 4.67 \times 10^{-5} g \lambda^2 B$ 来反推磁场的强度,其中 g 为 Landé 因子, λ 和 B 的单位分别为 cm 和 Gauss。具体的测量原理和反演过程许多教程和文章中都有介绍 (如: 苏江涛 2004 [130]), 在此不做更多的陈述。

目前太阳磁场测量通常使用的设备主要有两种,滤光器型磁像仪和光谱仪型磁像仪。滤光器型成像,时间分辨率较光谱仪型高,缺点是光谱分辨率不足,很难得到精确的磁场参数;光谱仪型成谱,可以得到更精确的磁场和热力学参数,但要想得到二维空间面的信息则需要耗费更多的时间。目前国内已有二维实时光谱仪和光纤阵列太阳望远镜在研制中,设备均可实现同时成像和成谱从而弥补两类设备的不足。怀柔太阳观测基地太阳磁场望远镜是一台滤光器型的矢量磁像仪。搭载在空间飞行器上的太阳磁场测量仪器中, Hinode 的 SP (Spectro-Polarimeter) 是光谱仪型、FG (Filtergram) 是滤光器型, SoHO (Solar and Heliospheric Observatory) 的 MDI (Michelson Doppler Imager) 和 SDO 的 HMI 是滤光器型。这些观测数据的使用者非常关心数据的可靠性,设备之间测量结果的相互比较工作非常必要 (见 Xu et al. 2007 [149] 等)。

1.1.2 太阳爆发活动

太阳大气的不同层次上存在着一些互相关联的特征现象,这些现象一般都可归属到太阳活动中,包括渐变型活动和爆发型活动。渐变型如黑子、日珥 (暗条)、冕环、光斑、谱斑等。这里我们主要讨论爆发型活动。太阳大气的局部区域出现的能量辐射增强而形成的瞬变现象,或者太阳大气中进行的激烈扰动的物理过程,为太阳爆发型活动。爆发型活动一般包括耀斑爆发、日珥 (暗条) 爆发、日冕物质抛射 (Coronal Mass Ejection, CME) 等。

1.1.2.1 太阳耀斑

太阳耀斑是太阳表面大气 (大致在过渡区或低日冕层) 中发生的一种急遽

的不稳定过程，在短时间内 (约 $10^2 - 10^3$ 秒) 释放出量级约为 $10^{29} - 10^{33}$ 尔格的能量，它引起局部区域的瞬时加热，并引起各种电磁辐射及粒子辐射的突然增强。人类首次有记录的耀斑观测来自于 1859 年 9 月 1 日两个独立的观测者 Carrington R. C. 和 Hodgson R.。这个耀斑为一次较为罕见的白光耀斑 (即耀斑强到在可见光波段的很大范围内可以观测到很强的连续辐射，可以用白光望远镜观测到)。20 世纪 60 年代后，空间设备上天，人类对耀斑的观测逐渐扩展到全电磁波段。而在这之前，人们对太阳耀斑的主要观测集中于几条光学谱线的成像观测，尤其是氢线 $H\alpha$ $\lambda 6563 \text{ \AA}$ 。到现在， $H\alpha$ $\lambda 6563 \text{ \AA}$ 仍然是用于太阳色球活动监测非常重要的一条谱线，包括怀柔太阳观测基地全日面色球望远镜在内的的全球 $H\alpha$ 网络 (Global H-alpha Network, GHN)⁴ 每日更新着最新的太阳色球 $H\alpha$ 的观测。怀柔太阳观测基地的网站上也时刻在发布最新的色球 $H\alpha$ 观测资料⁵ (图 1.3)。

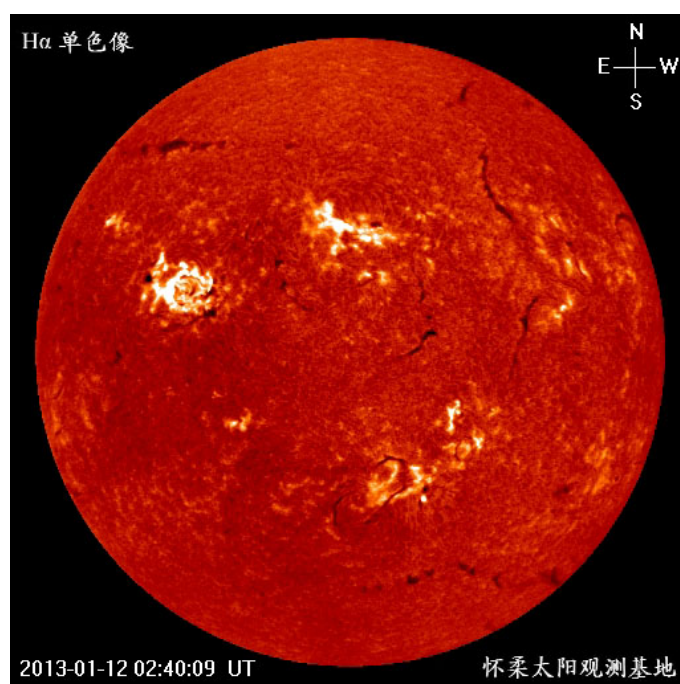


图 1.3: 怀柔太阳观测基地全日面色球 $H\alpha$ 望远镜观测图像示例 (经消除临边昏暗改正)

太阳耀斑涉及到许多非常复杂的物理过程，包括提供给耀斑爆发的可释放

⁴http://swrl.njit.edu/ghn_web/

⁵<http://sun.bao.ac.cn/observation/latesthalph.php>

能量的积累、等离子体不稳定性和各种波对耀斑爆发的触发、耀斑过程所引起的高能粒子的加速和传播等，它们激发产生的从 γ 射线、X 射线、紫外、可见光、红外一直到射电波段的全波段的电磁辐射的增强、耀斑区域太阳大气动力学变化、等离子体运动、物质抛射等，它们的产生机制、发展演化等，都是研究人员所关心的问题。另外，耀斑所引起的 X 射线辐射增强将引起地球电离层的变化，高能粒子辐射也会造成地球轨道附近高能粒子污染并干扰地球的磁层，其扰动还会继续向下传播导致包括平流层和对流层在内的地球低层大气的物理状态的变化。由于人类发展到今天，包括航空、航天、无线电通讯在内的很多技术都受到大气变化的影响，太阳活动通过这种方式影响人类活动，预测太阳剧烈活动的发生，对可能产生的地球物理效应进行预测和监测，能尽可能地减少损失，故而是十分必要的。(图 1.4 和图 1.5)

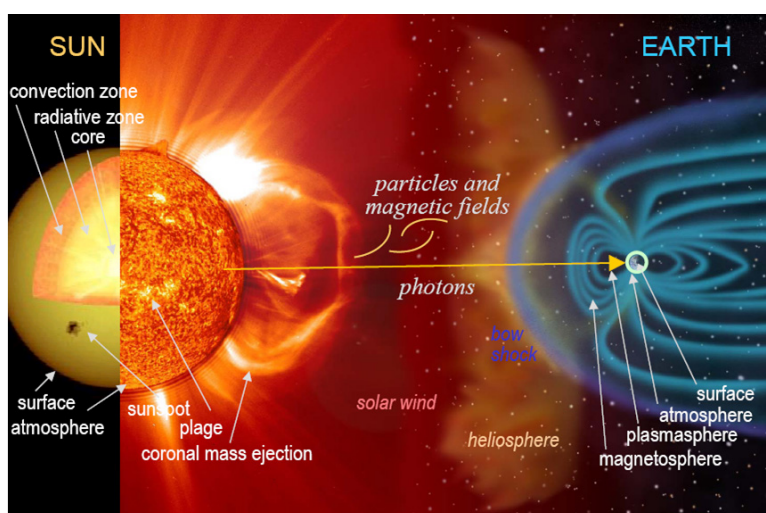


图 1.4: 太阳分层结构及日地空间关系 (引自 NASA 网站)

对于太阳耀斑级别划分，普遍使用的主要有两种：软 X 射线耀斑分类、和 $H\alpha$ 耀斑分类。

由于大气辐射窗口的允许，在地面进行光学波段观测非常便利， $H\alpha$ 分类使用较早。在 $H\alpha$ 太阳单色成像观测中，依据耀斑在亮度极大时的面积从小到大依次划分为 S、1、2、3、4 五个等级，依据耀斑亮度的差别将其分为 F (faint, 较暗)、N (normal, 中等)、B (brilliant, 很亮) 三个等级，两种划分并排放置或将后者作为前者的下标，这样就构成了一个耀斑的 $H\alpha$ 类别。面积的确定需要先对耀斑处做投影改正，即需换算到耀斑处在日面中心的面积，再根

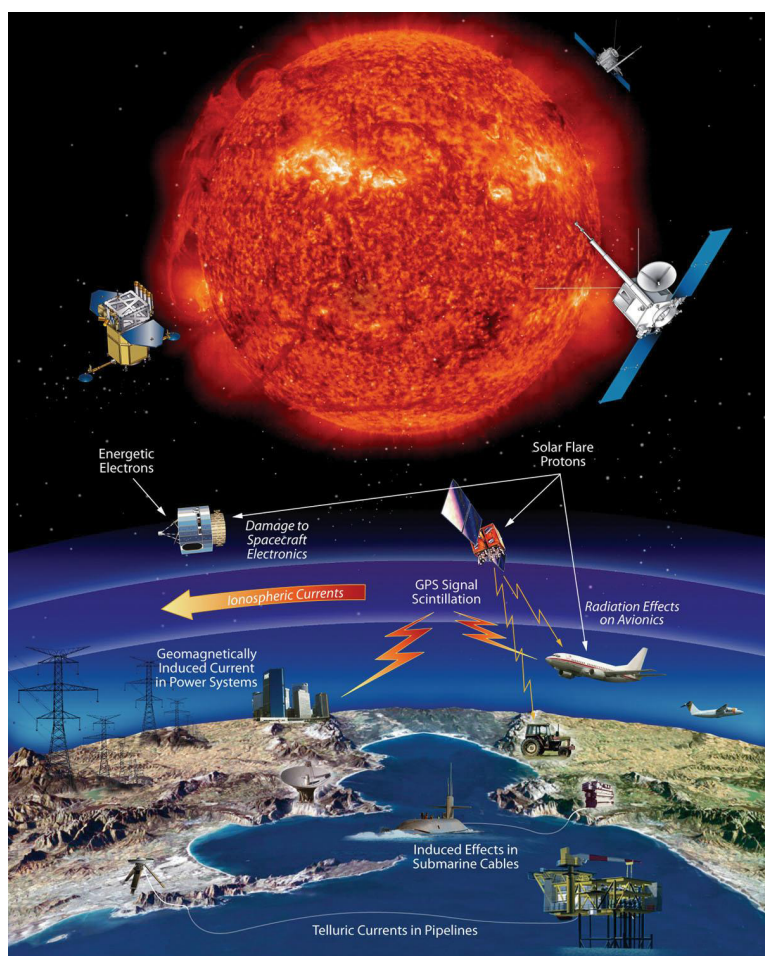


图 1.5: 太阳活动对人类活动的影响 (引自 NASA 网站)

据表 1.1 定级；而亮度的确定则没有确定的标准。

X 射线耀斑分类则产生于空间卫星上天之后。地球大气对 X 光波段并非透明，而地球电离层对太阳 $1 - 8 \text{ \AA}$ 软 X 射线辐射强度的变化反应比较敏感，因此太阳耀斑的软 X 射线分类可以大致反映出耀斑产生地球物理效应的能力。根据耀斑发生过程的软 X 射线峰值流量的量级而将其主要分为 A、B、C、M、X 五个级别 (表 1.2)。上世纪六七十年代，太阳监控卫星 Orbiting Geophysical Observatory (OGO)、Orbiting Solar Observatory (OSO)、Solar Radiation (SOLRAD) 卫星、Synchronous Meteorological Satellite (SMS) 等都曾连续地观测过太阳软 X 射线流量 (Kahler & Kreplin 1991 [77]; Garcia 1994 [51])。美国 NOAA (National Oceanic and Atmospheric Administration) 的

表 1.1: H α 耀斑分类

面积 (10^{-6} 太阳半球面)	面积 (平方度)	等级 (1966 年 1 月 1 日前)	等级 (1966 年 1 月 1 日后)
< 100	< 2.06	1 ⁺	S
100 – 250	2.06 – 5.15	1	1
250 – 600	5.15 – 12.4	2	2
600 – 1200	12.4 – 24.7	3	3
> 1200	> 24.7	3 ⁺	4

GOES (Geostationary Operational Environmental Satellites) 系列空间卫星则从 1975 年至今系统地记录了全日面的软 X 射线积分流量情况，其软 X 射线耀斑记录被业内研究者普遍使用和参考。

表 1.2: 软 X 射线耀斑分类

级别	峰值强度 I (瓦·米 ⁻²)	峰值强度 I (尔格·厘米 ⁻² ·秒 ⁻¹)
A	$I < 10^{-7}$	$I < 10^{-4}$
B	$10^{-7} \leq I < 10^{-6}$	$10^{-4} \leq I < 10^{-3}$
C	$10^{-6} \leq I < 10^{-5}$	$10^{-3} \leq I < 10^{-2}$
M	$10^{-5} \leq I < 10^{-4}$	$10^{-2} \leq I < 10^{-1}$
X	$I \geq 10^{-4}$	$I \geq 10^{-1}$

1.1.2.2 CME 等其他爆发活动

人们对日冕的常规研究起始于日冕仪的发明，对 CME 的系统研究开始于上世纪七八十年代。CME 是太阳大气中剧烈的、尺度最大的瞬变活动现象，在短时间内将日冕物质抛射到行星际空间。与耀斑相比，CME 似乎更直接地影响着日地空间环境，朝向地球的晕状 (halo) CME 更是全方位地引起地球大气和人类空间飞行器的响应。有些日珥 (暗条)，宁静区日珥或活动区日珥，在缓慢演化过程中会突然发生爆发性膨胀或向外抛射，即为日珥 (暗条) 爆发。

在爆发日珥抛射物质的过程中，有一部分物质可达到逃逸速度从而进入行星际空间，另也有一部分物质会沿磁力线回落至日面。CME 除一部分与耀斑相联系以外，也有很大一部分与日珥 (暗条) 爆发相联系，它们都是太阳表面的活动在太阳更外层大气的反映与延伸。

1.2 太阳磁场非势性研究

太阳磁场的非势性，是指太阳的实际磁场对势场位型的偏离。不少物理量都能从某个方面反映出这种偏离，比如，磁剪切角、电流、螺度等。对这些物理量的研究都有着数十年的历史，有着大量的文献，包括理论分析、个例研究和统计研究等。

通常所说的在太阳爆发中可供释放的自由磁能，定义为实际的总磁能 E 与其对应的势场能 E_p 之差，即：

$$E_{\text{free}} = E - E_p,$$

其中，

$$E = \frac{1}{8\pi} \int_V \mathbf{B}^2 dV,$$

$$E_p = \frac{1}{8\pi} \int_V \mathbf{B}_p^2 dV.$$

\mathbf{B} 为实际磁场， \mathbf{B}_p 为纵向磁场对应的势场。实际磁场对势场的偏离越大，即磁场的非势性越强，则可供释放的自由磁能储备越多，爆发的可能性也越大。

活动区磁场的非势性产生于其磁场的不断演化。光球和光球以下流体的运动会带动位于光球表面的磁力线足点的运动，表现出来就是黑子本影之间的相对运动，以及大黑子自身的旋转，这种剪切和扭绞运动导致磁场位型偏离势场。光球活动区下方新磁流的浮现，也会使原有磁场受到干扰，这不仅有可能导致原有磁场的非势性增强，甚至也可能是耀斑爆发的诱因之一。有些新浮现的黑子，其磁场在浮现上来之前就可能已经是非势的。对于非势性产生的观测实例和理论数值模拟在许多研究中都有探讨 (如：Low & Nakagawa 1975 [89]; Krall et al. 1982 [78]; Tanaka 1991 [132])。

磁剪切角是描述磁场非势性的一个典型参量。Hagyard et al. (1984) [64] 和 Lü et al. (1993) [90] 分别引入了平面磁剪切角和空间磁剪切角。空间剪切角是观测矢量磁场和其对应的势场的夹角；平面剪切角是观测磁场和其势场的方

位角差，可以认为是空间剪切角在天空平面上的投影。Lü et al. (1993) [90] 指出，空间剪切角是实际磁场对其势场偏离的更直接的反映，同时它也与非势磁场自由能量有着直接的关系。

静力学研究使物质维持静止状态的条件。磁流体静力学讨论什么样的磁场位型和导电流体的压力和密度分布关系可使它们保持静止，即在特征时间内物理过程变化非常缓慢。低温高密的日珥、太阳磁流管、太阳黑子等，在他们的典型时间内，物理过程变化得非常缓慢，以致其运动的效应和时间变化的关系不重要，它们可以用静力学模型分析。在静力学近似条件下， $\partial/\partial t \simeq 0$ ， $\mathbf{v} \simeq 0$ 。这时磁流体静力学方程组中的运动方程为

$$\nabla p = \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} + \rho \mathbf{g},$$

其中 $\mu_0 = 4\pi \times 10^{-6} \text{ G km A}^{-1}$ 为真空磁导率。在太阳的色球和低日冕层，磁场强度相对较强，在量级上，磁压力远大于等离子体压力和重力。从而上式近似为无力场形式，洛伦兹力为零，即电流的方向平行于当地磁场的方向，于是有

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = 0,$$

或者写为

$$\nabla \times \mathbf{B} = \alpha(t, \mathbf{r}) \mathbf{B}.$$

α 是时间和空间的函数，称为无力因子。若 α 的分布确定，则无力场可完全确定。对上式取散度，有

$$(\mathbf{B} \cdot \nabla) \alpha = 0.$$

这表明无力因子 α 沿每条磁力线不变。最简单的情况， $\alpha = 0$ ，此时 $\nabla \times \mathbf{B} = 0$ ，即为势场，此时电流密度 $\mathbf{J} = \frac{1}{\mu_0} (\nabla \times \mathbf{B}) = 0$ 、电流螺度密度 $h_c = \mathbf{B} \cdot (\nabla \times \mathbf{B}) = 0$ 。若 $\alpha = \text{常数}$ ，则方程为线性的，称为线性无力场。 α 不为零，则表示当前磁场是偏离势场的。

因此，电流、电流螺度、无力因子 α 对描述磁非势性非常重要。早在上世纪 60 年代，太阳物理学家已经开始从观测上研究活动区的电流 (Severny 1965 [125]) 以及它们与太阳耀斑的关系了 (Moreton & Severny 1968 [99])。Wang et al. (1994) [144] 和 Schrijver (2007) [123] 等指出携带电流的磁场浮现可能导致了大多数的大级别耀斑。Bao et al. (1999) [15] 验证了活动区的电流

螺度 (参见: Seehafer 1990 [124]; Pevtsov et al. 1995 [108]) 在太阳耀斑的演化中起到重要的作用。Nakagawa & Raadu (1972) [102] 从理论上描述了无力因子 α 作为磁力线扭绞的表征; 他们得出, 磁能会随着 α 值的增长而增长。Nindos & Andrews (2004) [103] 讨论了 133 个关联 CME 的大耀斑来陈述耀斑前日冕的磁螺度和无力因子 α 的作用。

活动区与耀斑关联的非势磁场特征的研究已有数十年的历史。耀斑多发活动区的一些非势参量有许多作者做过细致研究。Gary et al. (1987) [53] 研究了 NOAA 编号 2684 的活动区的矢量磁场的非势特征, 并试图回答诸如非势特征与耀斑关系这样的一些问题。Wang et al. (1996) [141] 细致研究了一个耀斑活跃活动区的耀斑出现与非势性发展的关系。Moon et al. 2000 [98] 分析了 NOAA 编号 5747 的活动区的一些磁非势性参量的空间分布和时间演化。Deng et al. (2001) [36] 研究了著名的 NOAA 编号 9077 活动区的磁非势性在巴士底事件大耀斑前后几天的演化情况。Leka & Barnes 2003 [81] 考查了耀斑活动相区别的三个不同类型活动区里大量光球磁参量的强度和时间变化, 这些参量包括水平磁梯度、纵向电流、电流螺度、无力因子 α 、磁剪切角、剩余磁能等, 试图从中寻找直接关系到耀斑产生的重要特征。Dun et al. 2007 [41] 计算了 NOAA 编号 10486 活动区的磁剪切角、纵向电流、电流螺度, 并分析了它们的时间演化及与耀斑的空间关系。Falconer et al. (2002, 2006) [42, 43] 也讨论了一些非势性有关参量和活动区的 CME 产率的关系。还有一些作者则从统计角度来研究活动区光球磁场的物理特征与耀斑的关系。Cui et al. (2006) [33] 研究了最大水平梯度、中性线长度、孤立奇点数; Jing et al. (2006) [74] 则使用了强梯度磁中性线上的磁梯度平均值、强梯度磁中性线长度、总磁能耗散; Cui et al. (2007) [32] 还尝试了强梯度中性线长度、强剪切中性线长度、强梯度强剪切中性线长度; Cui & Wang (2008) [34] 还研究了无符号总电流和无符号总电流螺度。Guo et al. (2006, 2007, 2010) [58–60] 利用从几何角度量化表示活动区磁场复杂性的有效距离参量, 分别研究了活动区磁复杂性的演化、与活动区爆发活动的关系、以及随活动周的演化情况等。大量的文章从观测上研究了活动区磁非势性与爆发活动的紧密关系。Rust et al. (1994) [117] 综述了包括非势特征在内的耀斑前状况的研究进展, 并指出研究从矢量磁像仪得到更多量化结果的重要性。

1.3 太阳耀斑预测

在较长时间尺度上, 太阳对地球有着大致为 11 年的周期性的影响; 在短一些的时间尺度上, 比如一个黑子的演化时间尺度上, 太阳的爆发性活动可以说是随机的, 活动区的演化具有很大的不确定性, 爆发活动的触发因素也是非常不确定的。因为存在人类活动对空间天气预测的实际需求, 太阳活动的预测 (如: 长期的如图 1.6 的太阳周黑子数预测⁶, 图 1.7 的 10.7 cm (2.8 GHz) 太阳射电流量预测⁷ 等, 短期的如耀斑、CME 预测等) 成为太阳物理的研究内容之一。

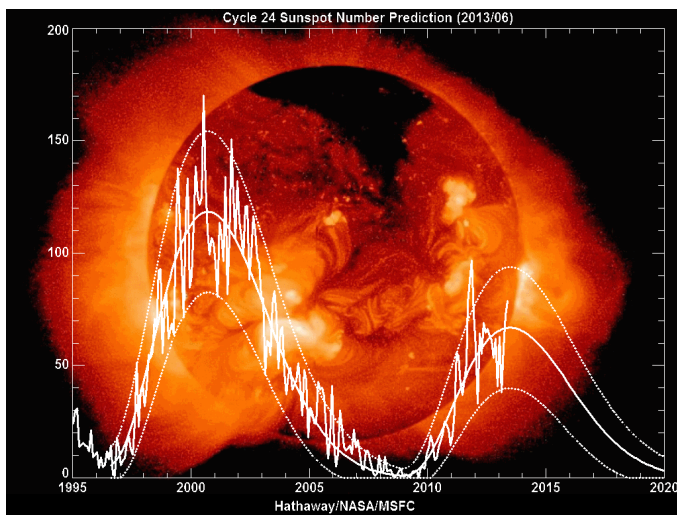


图 1.6: 24 太阳周黑子数预测 (引自 NASA 网站)

太阳活动的预测是空间天气预测的一部分, 太阳活动是空间天气的驱动源, 直接影响着日地空间环境和行星际空间环境。对太阳活动的实时监测、以及耀斑、CME 等爆发事件的预测和监测, 对诸如质子事件、高能粒子事件、地磁暴、电离层暴等灾害性空间天气事件的预测有着重要的参考价值。太阳活动的预测从时间尺度上大致分为长期预测 (一年以上, 如太阳周的预测)、中期预测 (半月至数月, 如黑子数和射电流量月均值预测、活动经度带预测等)、短期预测 (数小时、数天, 如耀斑预测、CME 预测等)。尽管已经发现黑子数和黑子面积、10.7 cm 射电流量等指标有着约 11 年的周期变化, 可以用来表征

⁶http://solarscience.msfc.nasa.gov/images/ssn_predict_l.gif

⁷http://solarscience.msfc.nasa.gov/images/f107_predict.gif

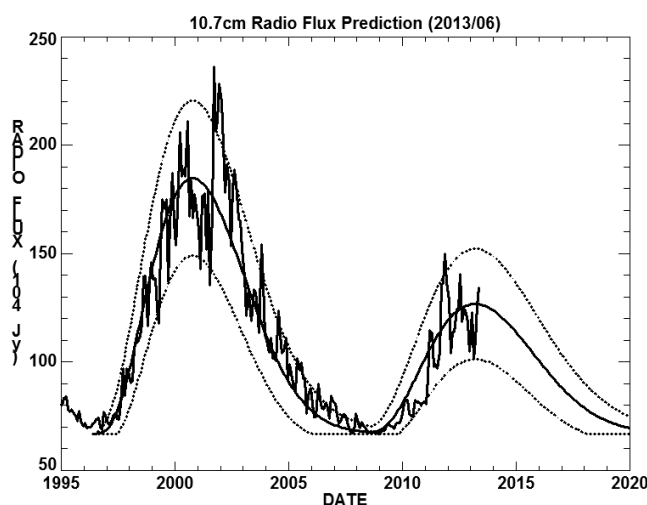


图 1.7: 10.7 cm (2.8 GHz) 太阳射电流量预测 (引自 NASA 网站)

和了解从当下以远的太阳整体活动水平，然而，下一活动周的相对强弱、持续时间、何时达到极大期，这些问题依然是不确定的，也是长周期预测的研究课题，这些课题也与验证和应用太阳发电机理论有着密切的联系。各种时间尺度的太阳活动预报，有赖于对相应尺度的太阳活动特征的研究。中期预测一般较短期和长期预测困难，可喜的是，目前日震学的研究进展已显示出预测活动区出现的能力。Messerotti et al. (2009) [96] 将太阳爆发活动及其引起的空间天气事件用太阳天气事件作为概括，综述了对各种太阳天气事件基于物理和统计两方面的模拟和预测，文中的两个所谓概念图图 2 和图 3 分别勾勒了太阳活动驱动及它们之间的关系、以及太阳磁活动的预测的整体概览。

人们对黑子的观测和认识最早，持续观测的时间也较其他太阳现象长。通过统计不同形态的黑子与耀斑发生之间的关系，人们可以通过黑子形态特征对耀斑发生的概率进行估计。最近十多年，机器学习、数据挖掘等概念和技术逐渐发展起来并应用在许多学科领域后，有一些学者将黑子形态学分类 (如 McIntosh 黑子磁分型, McIntosh 1990 [95]) 数值化，以适应模型的输入和自动演算。

然而，许多研究表明，太阳爆发活动需要可释放的磁能 (磁自由能) 的积累，既然现阶段人们已经具有测量太阳表面尤其是光球矢量磁场的的能力，如果能从磁场中直接计算黑子的磁自由能储存情况和其他表征活动区爆发潜能的物理量，这不仅有助于提高对太阳活动预测能力的认识，也有助于人们进一步理

解太阳活动区磁场与太阳活动的关系。当然，我们也不必放弃对形态学参量的使用，它们同样是对活动区爆发活动潜能的一个反映。目前人们还无法、甚至不可能找到一个单一指标来预测太阳活动的爆发，作者始终认为，多种形式的能量积累和多种形式的触发因素，共同决定着太阳活动的爆发与否和级别大小。尽管有学者发现一些案例中有所谓的前兆现象发生，但是这些前兆现象是否普遍存在，尚未可知。从事观测研究者需要从统计角度对发现的现象予以证认，理论研究者同样也需要统计意义上的观测支持才能使其模型得到广泛的认可。

一些作者或通过黑子形态学参量、或收集不同来源的信息来建立太阳耀斑预测模型 (如: Gallagher et al. 2002b [48]; Qahwaji & Colak 2007 [110]; Li et al. 2007 [86]; Colak & Qahwaji 2009 [30]; Bloomfield et al. 2012 [18])。这些预测因子必须首先经过人工确定，然后才能输入到预测过程中，故而不适于自动操作。另一些耀斑预测工作则是采用从视向磁场中得到的参量 (如: Gallagher et al. 2002b [48]; Georgoulis & Rust 2007 [55]; Yu et al. 2009 [156]; Song et al. 2009 [128]; Mason & Hoeksema 2010 [94]; Yuan et al. 2010 [158]; Ahmed et al. 2013 [2] 等)。Yu 和 Huang 的一系列工作 (如: Yu et al. 2009, 2010a, 2010b [155–157]; Huang et al. 2010 [73]) 使用 Cui et al. (2006) [33] 统计研究的三个视向磁场得到的磁参量作为预测因子，利用机器学习方法做了预报研究。Song et al. (2009) [128] 和 Yuan et al. (2010) [158] 也做了类似的工作，他们则是使用 Jing et al. (2006) [74] 里研究的三个同样得自视向磁场资料的磁参量作为预报因子。

磁场非势能量的积累对于太阳爆发是非常重要的。上一节提到的相关文献都曾表明磁非势性与太阳耀斑的紧密关系。Gallagher et al. (2002a) [47] 很早就提到将矢量磁场得到的参量用于辅助其预测系统的预测。Mason & Hoeksema (2010) [94] 在使用 13 年的 MDI 数据进行耀斑预测后，提到了使用矢量场数据的重要性以得到更有希望预测耀斑的磁场参量。Leka 和 Barnes 的一系列工作 (如: Leka & Barnes 2003a, 2003b, 2007 [81–83]) 在探索耀斑活跃活动区和耀斑宁静活动区之间磁场特性的差别方面做了很重要的工作。然而他们使用的样本数目及样本的时间跨度都受到很大的限制。由于少有长时期的稳定的矢量磁场观测，磁场非势性参量很少被用于太阳耀斑预测。而怀柔太阳观测站 20 余年的矢量磁场观测为这样的预测实验提供了很好的条件。

1.4 本文的研究目的与内容

磁场主导着太阳的演化，磁场的非势性质与诸如耀斑、日冕物质抛射等太阳爆发活动密切相关。尽管反映磁场非势性的物理量的演化并不直接触发太阳爆发，具有强非势性和复杂性的活动区由于拥有足够可供释放的自由能量，相比简单的活动区更容易爆发。由于少有长时期的稳定一致的矢量磁场观测，研究者们大都仅能使用矢量磁场数据进行个例研究，而统计分析只能限于视向磁场资料。从统计角度研究磁非势性的太阳周演化及其与太阳爆发的关系，对于理解太阳活动的长期演化、预测爆发活动等是非常有用的。目前，地面望远镜，由于其维护及操控的便捷，非常适合于长时期不间断、稳定的观测研究，在数据积累方面相比空间望远镜仍有其优势。位于中国科学院国家天文台怀柔太阳观测基地的太阳磁场望远镜 (Solar Magnetic Field Telescope, SMFT)，从 1987 年开始常规观测起已经平稳运行了 20 余年，几乎覆盖了两个完整的太阳活动周 (22 周和 23 周，现在已处在 24 周半程)。

太阳耀斑的短期预报能提供有用的信息以采取及时措施来避免更大的损失。目前领域内大家普遍接受，太阳大气里积累的自由磁能 (非势能量对势场能量的差值) 为太阳表面各种爆发现象提供能量。于是长期以来，太阳物理学家们一直试图弄清楚磁场的非势性质与太阳爆发现象之间是一种什么样的关系，努力理解和发现太阳剧烈爆发的物理机制。这样便可基于这些认识来对爆发做出物理预测，通过对诸如磁场演化、等离子体运动中的各种物理参数随时间的变化进行模拟从而预测其行为。但是在现阶段，这些物理机制人们知道得并不十分清楚。因而相对来讲，基于太阳爆发与其他现象演化的统计关系来进行预测似乎相对更为可行一些。

本文正是围绕上述问题，统计研究了第 22 至 23 太阳活动周光球活动区磁场的非势特性及其与耀斑的关系，试图从统计角度考查活动区磁场非势性的强弱随活动周的演化趋势、以及非势性的积累与耀斑之间的关系 (第三章)。并尝试对这批数据进行统计学习、通过由此建立的耀斑预测模型检验非势性参量用于太阳耀斑预测的性能 (第四章)。通过实验，我们可以仅仅利用一些更全面的磁场特性来预测指定活动区的耀斑的出现与否。在实验中我们指定不同时间窗口和不同的预测耀斑级别进行多组试验，这样做为同时估计耀斑爆发的大致开始时间和强度提供了可能性。此外，我们还讨论了一些经典的二类分类的性能评价指标，希望引起相关人士对预测性能的评价方面更认真的注意 (Doswell et

al. 1990 [39])。

第二章 本论文研究的工具基础

太阳物理是以观测为基础的天体物理学的分支之一，随着观测设备和计算机技术的发展，太阳物理研究也不可避免地面临海量数据和分析处理能力的瓶颈。如何将最新的计算机应用技术和被数据侵袭的太阳物理研究有机结合起来，是天文信息技术(学)需要不断解决的问题，也是本文作者努力的一个方向。本章简单介绍与本论文研究有关的一点工具基础(第 2.1 节)、作者对高性能计算在太阳物理中应用的一点点粗浅的调研和认识(第 2.2 节)、以及对本文数据基础的使用展望(第 2.3 节)。

2.1 数据挖掘概念

数据挖掘、机器学习、人工智能是近半个世纪随着计算机技术的发展兴起且不断发展的三个范畴。三个范畴之间的技术和方法有所交叉，它们在概念上则各自有所侧重。数据挖掘侧重从大量数据中(近些年流行起来所谓海量数据和大数据 Big Data)挖掘有用信息，甚至发现新的知识(所谓知识发现)。机器学习则侧重让机器从给定大量历史数据中学习规律而导出需要的模型，从而可以对新数据做出判断和预测。人工智能则从更宽泛的角度，试图用人工的方法和技术，模仿、延伸和扩展人的智能，让机器智能化。三个范畴各自有着一套涉及面广泛的理论体系。

对于天文学观测获得的数据，从抽象一些的角度讲，我们想从中发现知识和规律；再具体一些，让这些历史数据从自身学习，有监督或者无监督；更具体一些，我们或许需要将它们分类，或者把遇到的问题转换为分类问题。与分类问题最接近的一个概念是模式识别。在模式识别技术中，将样本诸测量值的组合称为模式，对诸测量值经过综合分析找出主要特征称为特征提取和特征选择。运用知识或规律作出判断称为分类判决，这种知识或规律是判决的准则，或称为判决规则。判决结果就是把样本区分成相应的类型。模式是有限的单一样本一系列被测值的综合，模式识别即试图确定某个样本的类别属性，把它归属于多个类型的其中之一(参见 李金宗 1994 [85]; Duda 2001 [40])。

Aschwanden (2010) [9] 很好地综述了处理太阳数据所需要的一些图像处理

技术和自动的特征识别方法。他将这些分为四类：图像预处理过程，空间特征的自动探测，时间特征即事件的自动探测和跟踪，以及如图像可视化、归类、统计、理论建模、预测和预报等的的数据后处理。这个归纳非常有代表性。文章开篇自然首先提到了太阳物理学家在数据爆炸时代所面临的挑战。随着计算机技术的出现和发展，1960 年以后，数据存储和处理能力按照 Moore 定律指数式增长；而 1950 年后，以论文数来衡量的太阳领域的科学产出则是线性增长。那么，科学反馈大致呈现为数据量的一种对数关系。这样便可以估计出对自动数据处理方法的需求，以使科学反馈跟上步伐。换句话说，如果没有自动的数据处理工具的开发和应用，我们无法消化这些指数式增长的数据。

应用图像处理技术自动识别和探测太阳各种静态和动态特征物 (Aschwanden 归纳的中间两类)，为研究特征物的形态学特征、统计特性、指定区域的物理参数研究提供了许多的便利，为研究工作节省了宝贵的时间。关于自动识别和探测特征物，近些年来涌现了一大批工作，如：极紫外和 X 射线图像上的冕洞自动探测 (如：Scholl 2008 [122]; Krista & Gallagher 2009 [79])， $H\alpha$ 太阳单色像上的暗条探测 (如：Shih & Kowalski 2003 [127]; Qu et al. 2005 [113];)，磁图上活动区的识别 (如：Zhang et al. 2010 [163])，运动磁结构的跟踪 (Li & Zhang 2013 [87])， $H\alpha$ 太阳单色像上的耀斑探测 (如：Fernandez Borda et al. 2002 [45]; Qu et al. 2003 [115]; Qu et al. 2004 [112];)，自动识别和分类 CME (如：Robbrecht & Berghmans 2004 [116]; Qu et al. 2006 [114]; Young & Gallagher 2008 [154]; Olmedo et al. 2008 [104]; Yashiro et al. 2008 [153]; Boursier et al. 2009 [21]; Gallagher et al. 2011 [49]; Morgan et al. 2012 [100]; Byrne et al. [22]; Tappin et al. 2012 [133] 比较了自动和手工方法识别 CME 的差别)。利用这些图像处理和自动识别技术，便可很方便地得到一些计算和统计结果，如 EUV 变暗与 CME 之间的关系 (Bewsher et al. 2008 [17])，活动区的统计特征 (Zhang et al. 2010 [163])，运动磁结构的演化特征统计 (Li & Zhang 2013 [87])，太阳大气的湍流表象与诊断 (Georgoulis 2005 [54])，诸如此类，从而为人们发现各种现象背后的物理规律提供观测证据。

数据挖掘技术应用于天文领域已有数十年的历史。在夜天文的研究中，数据挖掘被用于天体分类 (恒星/星系分类，星系形态分类)，图像分类 (数字巡天中的恒星/星系分类)，数据压缩与分类 (光谱压缩和分类)，重建方法 (大尺度巡天中的图像重建)，大尺度结构分析 (有关大尺度结构和微波背景辐射的大尺度巡天) 等。太阳领域里，Calvo et al. [23] 早在 1995 年就利用神经网络

络技术预测太阳年平均黑子数。除上面提到的太阳特征识别使用到模式识别技术外, 太阳爆发活动的预测里也广泛应用到神经网络、机器学习、模式识别等技术。用到的具体技术包括线性判别分析 (即分类或模式识别, 在概率统计中称为判别分析; 如: Leka & Barnes 2007 [83])、决策树 (如: Yu et al. 2009 [156]; Yu et al. 2010a [155])、贝叶斯网络 (如: Yu et al. 2010b [157])、 k 最近邻分类 (如: Li et al. 2007 [86])、多层感知器、径向基函数网络 (如: Qahwaji & Colak 2007 [110])、级联相关神经网络 (如: Cascade Correlation Neural Network, CCNN; Qahwaji et al. 2008 [111])、有序逻辑回归 (如: Song et al. 2009 [128])、以及我们后面将用到的支持向量机等。支持向量机因为其概念清晰、使用轻便而得到广泛应用。

由 Vapnik 引入的支持向量机 (Support Vector Machine, SVM; Boser et al. 1992 [20]; Cortes & Vapnik 1995 [31]; Vapnik 1995 [137]), 是一种基于统计学习理论的通用学习算法, 用于解决分类问题和回归问题。SVM 是数据挖掘的一种新方法, 它借助于最优化方法来解决机器学习问题 (邓乃扬 & 田英杰 2004 [35]), 现在已广泛应用于诸如图像识别、文本识别、生物信息技术等许多领域里。支持向量分类技术已经发展得比较成熟了, 但仍在继续发展和应用中; 支持向量回归技术相比还不十分成熟, 仍在发展。在近十年内, 支持向量机已经被应用于天文领域 (如: Zhang & Zhao 2003 [164]; Woźniak et al. 2004 [148]; Wadadekar 2005 [138]; Gao et al. 2008 [50]; Beaumont et al. 2011 [16]; Peng et al. 2012 [107]), 包括太阳物理方向 (如: Qu et al. 2003 [115]; Qahwaji & Colak 2007 [110]; Li et al. 2007 [86]; Al-Omari et al. 2010 [6]; Labrosse et al. 2010 [80]; Alipour et al. 2012 [7])。

2.2 高性能计算

天文学是一门以观测为主的科学, 天文观测设备的涌现、探测器成像技术以及航天技术的发展使得天文学成为一个数据密集型和计算密集型的学科领域。它对高性能计算的要求, 一方面来源于数据量的急剧增长, 一方面也来源于天文学中的一些大型科学计算本身。太阳物理中的数值模拟、涉及图像数据矩阵的计算、统计研究涉及的相似数据的批量处理等, 使得它对高性能计算的需求日益强烈。以下介绍几个以 GPU 技术为代表的高性能并行计算技术在太阳物理领域的典型应用。

太阳数据采集和处理。太阳磁场的观测需要进行多帧叠加,以获得高信噪比的偏振图像,再加上大面阵 CCD 的使用,需要提高其终端数据处理设备的实时计算能力来保证观测质量、使望远镜的性能尽可能地发挥。目前怀柔太阳观测基地已在其 35 厘米口径太阳磁场望远镜的数据处理终端设备中安装配备有 GPU 显卡,进行图像数据采集系统的高速实时相关计算 (Shen et al. 2013 [126])。

太阳大气磁场反演。磁场测量使人类对太阳的研究不仅仅局限在质量、大小等几何参量,而且扩展到磁场强度、电流、温度、密度、速度等太阳大气的物理参量,大大丰富了人类对太阳的认识,是当今太阳物理学家研究太阳的主要手段。太阳大气的磁场测量,主要依据磁敏谱线在磁场作用下会发生分裂的塞曼效应,由裂距公式推算出太阳大气的磁场强度。而太阳夫琅和费谱线的塞曼效应分裂子线既涉及发射又涉及吸收,解释其强度和偏振状态需要 Stokes 偏振辐射转移方程,将 Stokes 参数和磁场、热力学参数联系起来。太阳磁场望远镜测量特定谱线的偏振信号,由偏振信号推导出矢量磁场的过程称作磁场反演。磁场反演通过辐射转移方程的解生成 Stokes 轮廓,采用非线性最小二乘法来拟合理论和观测的 Stokes 轮廓 ($I(\lambda), Q(\lambda), U(\lambda), V(\lambda)$, 四个 Stokes 分量在观测波长附近随波长变化的轮廓),并输出最优的磁场强度、方位角、磁倾角等物理参量。Borrero et al. 2011 [19] 使用约 50 个 CPU 的计算机集群对反演过程进行并行加速,使其能在十分钟内反演出大小为 4096×4096 像素点的磁图,但这对科研人员使用数据来说仍是个瓶颈,他们必须依赖于这些高性能集群。考虑到 GPU 通用计算的高效, Harker & Mighell 2012 [68] 将 GPU 利用在磁场反演上,速度比串行计算提高约 1 – 2 个量级。

太阳日冕磁场外推。尽管人们已经成功利用塞曼效应测得太阳光球层次的磁场,但由于高温低密的日冕大气造成偏振信号变弱、引发散射偏振等物理效应,日冕磁场的测量仍然存在许多困难。目前太阳物理学家通常采用磁场外推的手段、并结合日冕成像观测,来推测日冕磁场。磁场外推以太阳光球层的观测矢量磁场作为底面边界,基于一定的磁场理论模型假设向上推。利用外推得到的日冕磁场,与软 X 射线、极紫外波段的太阳单色像比对,可以研究太阳高层大气磁场的空间性质及演化规律,进而丰富人们对太阳磁场的产生和发展的认识、另外也为认识太阳活动对日地空间的影响提供线索。在通常的外推过程中,每一层次的计算仅与上一层次的数据有关,本层内各点之间的计算相互独立,便于引入 GPU 加速计算。在 Wang et al. 2013 [142] 的工作中,已利用

了 GPU 技术加速日冕磁场的外推, 并进行日冕磁场的三维重建。他们使用型号为 GeForce GTX 260 的桌面型 GPU 设备进行 64×64 像素大小为底面边界的磁场外推计算, 速度比用四核 CPU 的桌面 PC 提高约 20 倍。

太阳背面黑子日震学重建。太阳内部的辐射由于被太阳光球层及以上的气体物质全部吸收, 我们无法观测到太阳的内部。而日震学通过对太阳表面震荡的观测和分析研究太阳内部不同深度处的温度、密度、压强、化学组成、运动等重要信息, 是太阳物理研究不可或缺的一部分。近些年, 日震学的研究除了其对太阳本身认识的巨大贡献外, 对黑子浮现预测、太阳背面活动区监视、空间天气预报也提供了很好的手段。Hartlep et al. 2008 [69] 通过数值模拟应用时距日震学 (Time-distance Helioseismology) 方法反演得到太阳背面的活动区信息, 他们在其工作中使用了 NASA 的 “Columbia” 超级计算机¹ 来完成他们庞大的计算过程。现在, 美国国立太阳天文台的网站上每天都会更新从 GONG (Global Network Oscillation Group, 全球日震观测网)、HMI (Helioseismic and Magnetic Imager, 日震与磁场成像仪, 搭载在太阳动力学天文台 SDO 卫星上的三台仪器之一) 等联合设备或仪器通过日震学方法得到的太阳背面图像²。在时距日震学中使用的计算方法里, 对由时序数据组成的数据立方体进行演算, 大量重复地使用互相关函数 (Cross-Correlation Function) 和三次样条插值, 通常一次运算中主要的时间都消耗在这里, 计算量和数据量巨大, 使用高配的桌面系统计算一幅时距图需要约一周左右的时间, 这已远不能满足研究人员的需要。在合适的层次利用 GPU 加速有望提高计算效率。然而, 由于计算中数据吞吐量也很大, 如何在计算中合理地分配 CPU 并行和 GPU 加速, 还有待测试和精心设计。另外, 日震学研究中还涉及到大量的时域和频域之间的转换, 适合 GPU 加速。

太阳磁流体力学数值模拟。在空间环境中, 99% 以上的宇宙物质呈现出磁化等离子体的状态, 与一般中性介质的流体力学不同, 宇宙物质的运动加入了电磁场的调制, 需要考虑流体与电磁场之间复杂的耦合过程。太阳是一个等离子体球, 其表面大气的运动和演化可以由磁流体力学 (Magnetohydrodynamics, MHD) 基本方程组描述, 然而由于 MHD 方程组的高度非线性, 其解析解只有在少数几种具有对称性或特殊假设下才可求得, 任意边界条件下的三维通解几乎无法得到, 正是这种局限性, 使得数值求解 MHD 方程组成为研究太

¹www.top500.org/system/ranking/7288, www.nasa.gov/hecc/resources/columbia.html

²<http://gong.nso.edu/data/farside>

阳 MHD 过程的一个必要手段。在一定的假设下，可以在数学上证明，当模拟精度不断提高时，数值解逼近真实解。因此合理地利用 MHD 数值模拟，人们可以从中得到一些非常有用的信息，这些信息一定程度上反映了太阳上真实的物理状态，分析这些状态对理解太阳耀斑、日冕物质抛射、暗条爆发等太阳活动的物理机制有重要意义，最终将有助于太阳活动的物理预报。高精度数值求解 MHD 方程组的计算量巨大，主要来自两个方面：一是单点计算量大，MHD 方程组有 8 个方程和 8 个未知数，并且现代的高精度 MHD 算法都要在每点计算 8 个特征速度；二是计算网格数大，典型的三维计算网格数如： $256 \times 256 \times 256$ 。这两个方面造成三维 MHD 数值模拟极其消耗时间，然而大多数 MHD 数值算法在时间上是显式格式的，即 $n + 1$ 时刻的结果只与 n 时刻的结果有关，这就使得并行计算成为可能。因此 GPU 超强的并行计算能力特别适合于大规模的三维 MHD 数值模拟。Zhu et al. 2012 [165] 便利用了 GPU 的计算能力，使得计算时间大幅缩减。Wong et al. 2011 [146] 则详细试验了用 GPU 进行 MHD 模拟，对 1 维、2 维、3 维流体问题加速比分别达到 10、200 和 84 倍，并且对使用单精度和双精度计算进行了比较。

磁螺度数值模拟。由于前面提到的日冕磁场还无法精确测量的原因，一些重要的磁场参量（如磁螺度，反映空间中磁力线的拓扑结构和缠绕形式的物理量）我们也无法直接根据精确测量来推算，同样需要采用数值模拟的方法，在已有观测量的基础上来模拟这些磁场参量在三维空间中的演化。将模拟的结果与用其他方法得到的间接结果，以及其他间接观测结果相比对，可以互相验证一些理论推测或者得到一些新的发现。杨尚斌 (2009) [150] 使用分离变量法计算三维有限空间内的相对磁螺度，并将该方法应用到数据驱动的 MHD 模型中来检验该方法的合理性，通过与边界上螺度传输的比较，验证了其精确程度可以应用到未来的计算中来研究日冕中磁螺度的演化。在计算中需要花费大部分时间在分解电流以求解符合边界条件的矢势上，运算时间主要受限于求解偏微分方程（泊松方程和拉普拉斯方程），而其中用到的快速傅立叶变换很适合使用 GPU 加速，这样将节省大量时间进行更细致的分析。他们也正在使用计算机集群对计算过程做加速和改进。数值模拟日冕中磁螺度的演化将会为研究磁螺度积累与太阳耀斑和日冕物质抛射的关系提供有价值的参考。

数据已成为科学研究的必需品，现在的条件足以使某些领域里的定量分析取代过去的定性分析，且计算精度也在飞速提高。在有限的科研资源条件下，构建服务于科研的具有合理成本和开销的科学计算服务器将是一个有益的选

择。前面简述了太阳物理学中有高性能并行计算需求的几个典型研究案例，这些领域正在或有望使用 GPU 加速计算，希望未来能更好地将 GPU 计算服务于太阳物理研究。林钢华 2011 [88] 描述了信息化建设在诸多方面给太阳物理等天文研究带来的便利，同样，构建服务于太阳物理研究的 GPU 桌面超级计算系统，也将成为构建虚拟太阳天文台的重要组成部分，必定会为太阳物理研究做出可观的贡献。从上世纪八十年代我国太阳磁场望远镜和南京大学多波段太阳光谱仪的研制成功和投入观测，国内的太阳物理观测设备已成功运转了数十年，中国的太阳物理学家在国际上的影响也日渐壮大。进入二十一世纪，国内又有多台地基太阳观测设备投入使用和预研，在不久的将来，中国还要将太阳观测设备送入日地空间，以克服地球大气的扰动和吸收，数据量将会有更加巨幅的增长，科学内容也会继续丰富。在这样的情势下，GPU 超级计算技术对太阳物理领域科学计算的加速作用将会更加鲜明地显现出来。

2.3 本论文研究的数据基础

中国科学院国家天文台怀柔太阳观测基地致力于太阳的观测与研究，近三十年来积累了大量宝贵的太阳光球和色球的单色像、磁场、速度场等观测数据。尽管国外太阳望远镜上天大大推动了太阳物理研究和太阳观测设备的发展，地面观测平台以及地面长时间稳定一致的观测数据的积累对于太阳物理领域来说依然是不可或缺甚至是非常重要的，在地面开展各种观测实验的灵活性方面、在更长周期的统计分析方面，地面望远镜的价值丝毫不逊于空间望远镜，它们的潜力值得挖掘。

近三十年来，计算机技术也发生了非常巨大的变化，数据采集能力、计算机存储能力和计算能力都大大提高。海量数据对于研究人员似乎越来越显得难以驾驭。数据使用平台的设计、数据处理程序的改善，无疑显得非常重要，否则再优质的数据也难以发挥应有的作用、这将是极大的浪费和损失。国际虚拟天文台已经历了十余年的发展，太阳由于其特殊性而独立出来形成了虚拟太阳天文台³，中国现有和未来设备的太阳数据有必要融入这个大家庭。当然，我们首先需要建设好自身的数据使用平台，为科学研究和科学普及提供便利，使这些宝贵的数据获得充分的利用。这将是一件非常有意义的系统工程。近二十年以来空间设备的科学产出明显占据了观测研究的绝大部分，这也是得益于目

³<http://www.virtualsolar.org>

前国际上知名空间设备的数据使用的便利性 (如: Martens et al. 2012 [92] 及其中的文献), 其背后有一大批数据平台支撑人员在默默地奉献。我们完全有必要借鉴其经验和模式, 为未来更多自主数据的使用平台支撑建设储备技术。

第三章 光球活动区磁非势性随活动周的演化及其与耀斑的关系

本章的工作统计研究了太阳光球活动区磁非势性及其与关联耀斑之间的关系。我们选择了怀柔观测站太阳磁场望远镜 1988 – 2008 年间的 1106 个活动区共 2173 张光球矢量磁图，覆盖了 22 周和 23 周的大部分时期。我们对每张矢量磁图计算了平面剪切角均值、空间剪切角均值、纵向电流密度均值、电流螺度密度均值、无力因子、自由磁能密度均值、有效距离参数及经纵场均值加权后的有效距离参数。其中，电流螺度密度均值、自由磁能密度均值和加权后的有效距离参数表现出与活动周演化较高的相关性，它们分别与年均黑子数的相关系数均超过 0.59；其余五个参数显示出与活动周演化相比较弱的变化趋势，但在太阳极大期时的活动区非势性与复杂性仍强于极小期。所有八个参数与活动区耀斑产率之间存在正的相关性，结合这些不同参数将有益于预测活动区发生耀斑的可能性。

3.1 一节中，介绍了 SMFT 的观测和数据处理情况。非势性和复杂性参数的计算在 3.2 一节中展现。与耀斑关联的磁非势性参量的统计分析结果展现在 3.3 一节中。最后在 3.4 节中给出结论。

3.1 观测和数据处理

3.1.1 怀柔太阳观测站的观测

太阳磁场望远镜 SMFT 是一台滤光器型的矢量视频磁像仪，物镜通光口径为 35 厘米，测量太阳光球和色球的磁场和速度场。我们使用的数据由它其中的一条工作谱线 $\text{Fe I } \lambda 5324.19 \text{ \AA}$ 得到，测量的是光球矢量磁场。 $\text{Fe I } \lambda 5324.19 \text{ \AA}$ 线是一条强的宽线，其等效宽度为约 0.334 \AA ，Lande 因子为约 1.5 (Ai et al. 1982 [5]; Wang et al. 1996 [141])。在线翼 -0.075 \AA 处测量圆偏振信号 (Stock 参数 V) 从而获得磁场的纵向分量，在线心处测量线偏振信号 (Stokes 参数 Q 和 U) 从而获得磁场的横向分量。该磁像仪测量一组矢量磁场数据需要耗时约 3 分钟。观测目标主要为黑子所在的活动区。在 1998 年至 2008 年间，主要有三种尺寸相异的 CCD 服役于该磁像仪，所对应的视场大小为以下三种：

2001 年 8 月 25 日之前, 视场大小为 $5.23' \times 3.63'$, 像元分辨率为 $0.61'' \times 0.43''$; 2001 年 8 月 25 日至 2001 年 11 月 30 日, 视场大小为 $4.06' \times 2.77'$, 像元分辨率为 $0.48'' \times 0.32''$; 2001 年 12 月 1 日之后, 视场大小为 $3.75' \times 2.81'$, 像元分辨率为 $0.35'' \times 0.35''$ 。为提高信噪比, 观测系统主要采用 256 帧叠加用于常规观测。分别用 3×5 、 3×5 和 5×5 像元大小的掩模对以上三种数据平滑后, 可以得到更高的信噪比, 空间分辨率变为约 $2'' \times 2''$ 。该矢量磁图的 3σ 噪声约为纵场 20 高斯、横场 150 高斯。

为确认 SMFT 测量磁场的准确性, 从研究特定活动区磁场结构的角度 (Wang et al. 1992 [140]; Bao et al. 2000 [13]; Zhang et al. 2003 [161])、以及从统计分析角度 (Pevtsov et al. 2006 [109]; Xu et al. 2007 [149]) 都有一些工作对不同观测站的不同磁像仪得到的矢量磁图数据做过分析比较。结果显示, 它们的数据基本一致, 包括磁场的纵向分量和横向分量、电流、螺度参数等等。由于观测方法和反演方法不同造成的少量差别对于我们的磁非势性的统计研究是不显著的。

3.1.2 数据选择与预处理

为尽量减小投影效应的影响, 我们挑选出位于距日面中心 30° 以内的活动区样本; 并且每天每个活动区限取一张磁图。图 3.1 所示为挑选出的磁图的一个示例: 左图为定标之前的原始观测数据显示; 右图为定标之后的矢量磁图, 背景为纵向磁场, 箭头指示横场方向和大小。在最后挑选得到的数据中, 大多数的观测时间都在 01:00 UT 至 07:00 UT 之间 (对应于怀柔当地时间 09:00 至 15:00), 这也是怀柔观测站大气状况较好的时段。最后选出成为研究样本的数据共包含了 1106 个活动区的 2173 张磁图, 时间跨度从 1988 年 6 月至 2008 年 3 月。表 3.1 列出了这些样本在约两个活动周内每年的磁图数量和活动区数量的分布。我们在选择活动区时没有偏向性, 这些活动区涵盖了多个种类的活动区, 从简单的单极小黑子, 到复杂的 δ 型黑子。从 1988 年到 2008 年, NOAA 共记录了 6095 个活动区, 这些活动区中有 39.1% 产生过 C 级及以上耀斑, 有 15.0% 产生过 M 级及以上耀斑, 有 1.0% 产生过 X 级耀斑。我们的数据集涵盖的活动区数量占到这段时期活动区总数的 18.1%, 其中分别有 71.3%、20.7% 和 1.0% 的活动区产生过 C 级及以上、M 级及以上和 X 级的耀斑。由于观测目标为日面局部的活动区, 复杂活动区相比于简单活动区会被给予更高的观测频率, 加上我们选择数据时并不对活动区区别对待, 使得我们的数据集更偏向

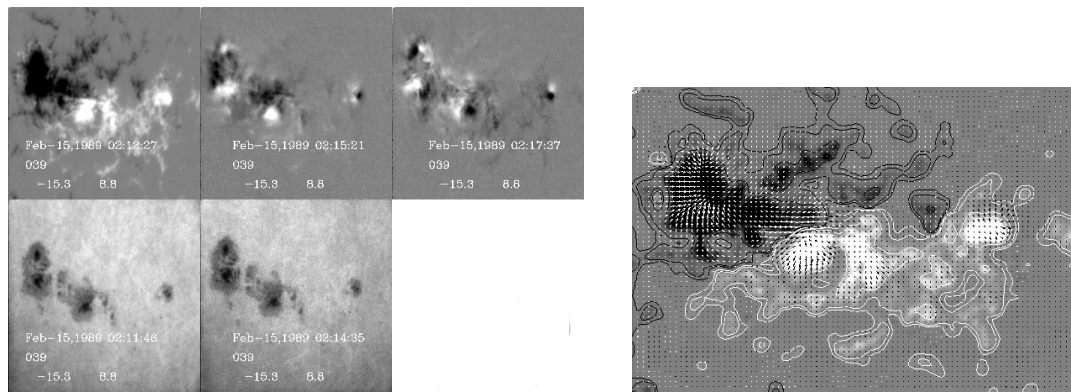


图 3.1: 挑选磁图示例

于更大和更活跃的活动区，但从这些活动区的磁图观测时刻开始计算，向后 48 小时以内所产生的 C 级以上耀斑仅占到这 21 年活动区关联耀斑总数的 18.6%，其中，C 级耀斑占全部 C 级耀斑总数的 18.8%，M 级耀斑占全部 M 级耀斑总数的 17.2%，X 级耀斑占全部 X 级耀斑总数的 19.8%，这和前面提到的活动区数的比例 18.1% 是一致的。因此可以说我们的样本覆盖是合理的和有代表性的。

表 3.1: 数据样本在 1988 年至 2008 年间的数量分布

年份	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
磁图数	10	37	47	108	138	97	74	56	11	54	114
活动区数	8	21	30	58	65	52	39	25	6	32	62
年份	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	总计
磁图数	166	361	285	233	144	67	101	46	20	4	2173
活动区数	88	166	136	123	77	34	50	22	11	2	1106

为分析活动区磁特征与关联耀斑之间的关系，我们将怀柔活动区号与国际 NOAA 活动区号进行了一一对应 (图 3.2)，这样我们便可以利用 GOES 卫星记录到的较完整的软 X 射线耀斑信息。1996 年 2 月之后的磁图可以对照 SolarMonitor¹ 上的 MDI/SoHO 全日面视向磁图进行比对；这之前的数据，

¹<http://www.solarmonitor.org>

Column 1 (Huairou)	Column 2 (Huairou)	Column 3 (NOAA)
3 06016	10862	
4 06019	10865	
5 06020	10866	
6 06021	10867	
7 06023	10869	
8 06025	10871	
9 06029	10875	
10 06030	10876	
11 06031	10878	
12 06035	10880	
13 06037	10885	
14 06042	10892	
15 06043	10893	
16 06048	10897	
17 06049	10898	
18 06058	10908	
19 06059	10909	
20 06060	10910	
21 06063	10914	
22 06064	10913	
23 06075	10930	
		85 00183 09055
		86 00184 09057
		87 00185 09056
		88 00186 09063
		89 00188 09061
		90 00189 09062
		91 00190 09067
		92 00191 09066
		93 00194 09068
		94 00198 09077
		95 00201 09084
		96 00205 09081
		97 00206 09085
		98 00207 09088
		99 00210 09094
		100 00211 09091
		101 00215 09097
		102 00216 09096
		103 00217 09099
		104 00220 09103
		105 00222 09105
		1 91001 06447
		2 91007 06466
		3 91008 06462
		4 91010 06469
		5 91017 06484
		6 91018 06487
		7 91028 06509
		8 91032 06523
		9 91045 06555
		10 91051 06562
		11 91052 06563
		12 91062 06583
		13 91066 06593
		14 91068 06605
		15 91076 06619
		16 91080 06624
		17 91083 06633
		18 91106 06682
		19 91108 06681
		20 91113 06687
		21 91115 06693

图 3.2: 怀柔活动区号与 NOAA 活动区号的一一对应

我们则结合 SOON (Solar Observing Optical Network) 的全日面手描黑子资料² 以及 NOAA.Boulder 的全日面中性线/磁场手绘示意图³ 进行比对; 另外, usaf_mwl (United States Air Force/Mount Wilson) 的活动区详细信息⁴ (含活动区经纬度、类型等) 也同时用作参考以确保无误 (图 3.3)。由于 SMFT 的视场尺寸是固定的, 我们选择的磁图均满足: 包含单独且完整的活动区 (出现多个活动区的磁图和活动区尺寸大于视场尺寸的磁图未被选择), 黑子尽量位于视场中心。这样我们便可以对所有入选的磁图进行统一流程的数据处理。GOES 卫星记录的软 X 射线耀斑事件列表由 NGDC 发布在网上⁵。

一张矢量磁图包含有磁场的纵向分量 B_ℓ 和横向分量 B_t , 磁场强度值由测

²<http://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-imagery/photosphere/sunspot-drawings/soon/>

³<http://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-imagery/composites/full-sun-drawings/boulder/>

⁴ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_REGIONS/USAF_MWL/

⁵<http://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-flares/x-rays/goes/>

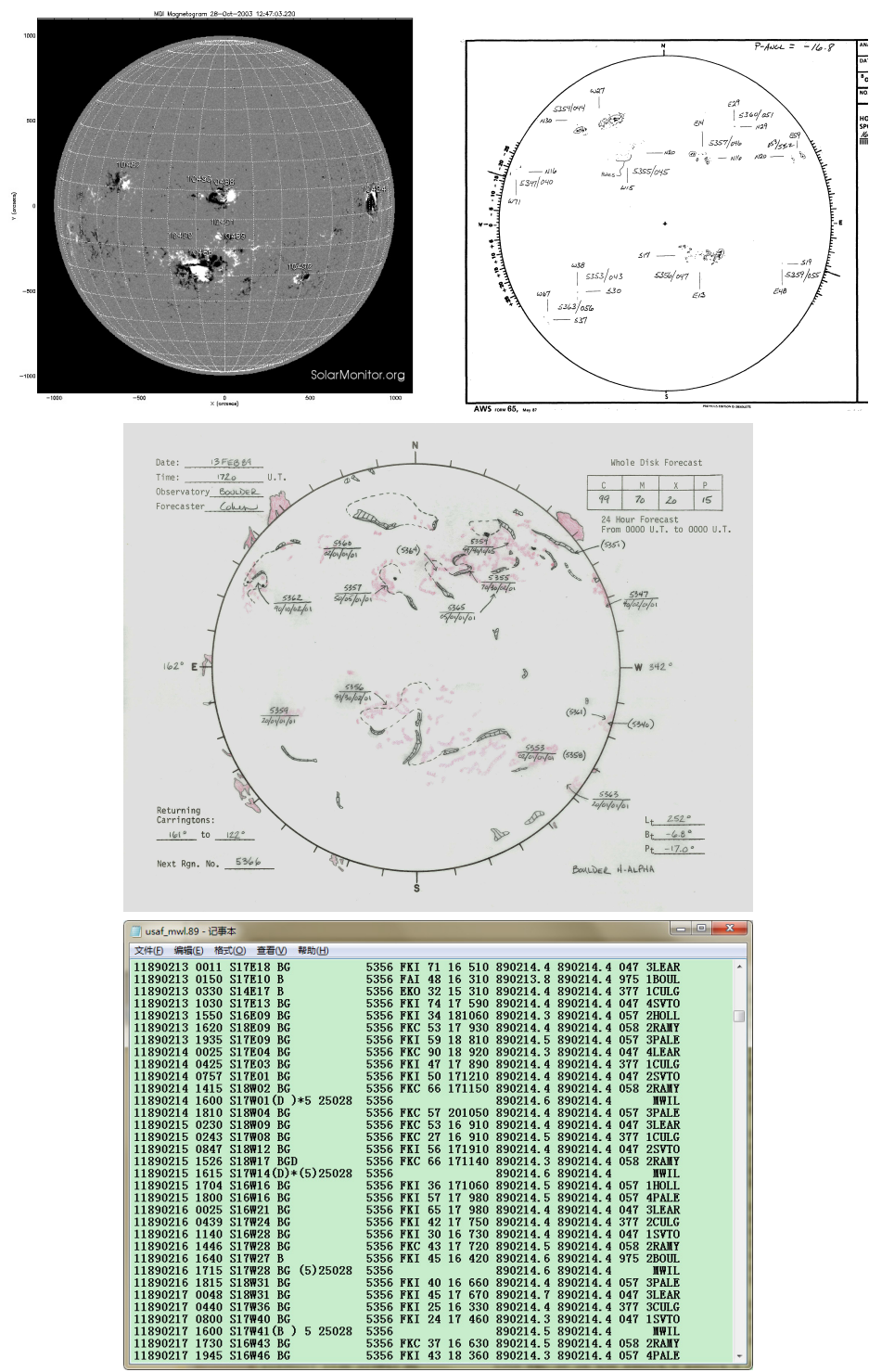


图 3.3: 对应怀柔活动区号与 NOAA 活动区号的参考资料

得的 Stokes 信号经反演得到。SMFT 普遍使用的定标方法基于线性关系：

$$B_\ell = C_\ell \frac{V}{I}, \quad B_t = C_t \sqrt{\left(\frac{Q}{I}\right)^2 + \left(\frac{U}{I}\right)^2}, \quad (3.1)$$

其中, I 、 Q 、 U 、 V 为 Stokes 参数, C_ℓ 和 C_t 是分别对应于 B_ℓ 和 B_t 的定标系数。横场方位角 ϕ 由以下式子得到：

$$\phi = \frac{1}{2} \arctan \left(\frac{U}{Q} \right). \quad (3.2)$$

历史上 SMFT 矢量磁像仪使用几种经不同方法得到的定标结果 (理论定标, Ai et al. 1982 [4]; 观测定标, Wang et al. 1996 [143]; 非线性最小二乘拟合定标, Su & Zhang 2004 [131]), 本工作中使用 $C_\ell = 8381$ 、 $C_t = 6790$ (Su & Zhang 2004 [131])。由于是统计计算, 其结果受定标系数选择的影响可忽略。

横场 180° 不确定性问题我们使用势场近似方法 (Harvey 1969 [70]; Sakurai et al. 1985 [119]) 解决。该方法对于实际磁剪切角小于 90° 的区域比较成功。尽管对于一些偏离势场较多的区域不确定性问题处理得有些粗糙, 但本工作考查的参量都是反映一个活动区的宏观性质和平均量值; 故而在本统计分析中其影响可以忽略。由于所选样本位于日面中心附近, 磁图处理中未做投影改正 (Gary & Hagyard 1990 [52]; Li 2002 [84]; Cui et al. 2007 [32])。

3.2 磁非势性和磁复杂性参量

3.2.1 磁剪切角

在本工作中, 我们把这两个剪切角看作是两个空间三维矢量间的夹角大小, 均为正值, 表示公式如下：

$$\Delta\phi = (\widehat{\mathbf{B}_{to}, \mathbf{B}_{tp}}) = \arccos \left(\frac{\mathbf{B}_{to} \cdot \mathbf{B}_{tp}}{|\mathbf{B}_{to}| |\mathbf{B}_{tp}|} \right), \quad (3.3)$$

$$\Delta\psi = (\widehat{\mathbf{B}_o, \mathbf{B}_p}) = \arccos \left(\frac{\mathbf{B}_o \cdot \mathbf{B}_p}{|\mathbf{B}_o| |\mathbf{B}_p|} \right), \quad (3.4)$$

其中, 下角标 o 和 p 分别表示观测磁场和其对应的势场, t 表示磁场的横向分量。

3.2.2 纵向电流密度，电流螺度密度，无力因子

根据安培定律，并且不考虑位移电流效应，电流密度的纵向分量表示为：

$$J_z = \frac{1}{\mu_0} (\nabla \times \mathbf{B})_z = \frac{1}{\mu_0} \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right), \quad (3.5)$$

这里， J_z 以 A km^{-2} 为单位， $\mu_0 = 4\pi \times 10^{-6} \text{ G km A}^{-1}$ 为真空磁导率， B_x 和 B_y 分别为水平磁场的两个正交分量。

电流螺度定义为：

$$H_c = \int_V \mathbf{B} \cdot (\nabla \times \mathbf{B}) dV,$$

那么电流螺度的纵向分量为：

$$(h_c)_z = B_z (\nabla \times \mathbf{B})_z = B_z \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right). \quad (3.6)$$

目前还没有观测方法能得到电流螺度密度的横向分量 $(h_c)_t$ 。我们使用符号 h_c 来直接表示它的纵向分量。

基于无力场假设，洛伦兹力为零，即，电流的方向平行于当地磁场的方向，于是有：

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = 0,$$

或者写为

$$\nabla \times \mathbf{B} = \alpha \mathbf{B}.$$

如果磁场近似于线性无力场，那么当地的 α 值则可以表示为：

$$\alpha = \frac{(\nabla \times \mathbf{B})_z}{B_z}.$$

我们使用 α_{av} (Hagino et al. 2004 [61]) 来表征一个活动区内的主导扭绞状况：

$$\alpha_{av} = \frac{\sum (\nabla \times \mathbf{B})_z \cdot \text{sign}[B_z]}{\sum |B_z|}. \quad (3.7)$$

3.2.3 自由磁能密度

耀斑和其他爆发性事件释放的能量依赖于自由磁能即非势磁能的积累, 这个自由磁能定义为实际总磁场能量 (E) 与其势场能量 (E_p) 的差值:

$$\Delta E = E - E_p.$$

Hagyard et al. (1981) [63] 引入了源场的概念来描述光球上磁场的非势性:

$$\mathbf{B}_s = \mathbf{B}_o - \mathbf{B}_p,$$

其中, \mathbf{B}_o 是观测得到的矢量磁场, \mathbf{B}_p 是由 \mathbf{B}_o 的纵向分量外推得到的势场, \mathbf{B}_s 就是所谓的表示磁场非势分量的源场。

源场的磁能密度正比于 \mathbf{B}_s^2 :

$$\rho_{\text{free}} = \frac{\mathbf{B}_s^2}{8\pi} = \frac{(\mathbf{B}_o - \mathbf{B}_p)^2}{8\pi}. \quad (3.8)$$

式 (3.8) 用来计算本工作中的光球层次的自由磁能密度, 由该式可以推得:

$$\rho_{\text{free}} = \frac{(B_o - B_p)^2}{8\pi} + \frac{B_o B_p}{2\pi} \sin^2 \left(\frac{\Delta\psi}{2} \right),$$

其中, $B_o = |\mathbf{B}_o|$, $B_p = |\mathbf{B}_p|$ 。可以看到, 空间剪切角 $\Delta\psi$ 与非势磁能有着直接的关系 (Lü et al. 1993 [90]; Wang et al. 1996 [141])。

3.2.4 有效距离

有效距离 (effective distance, d_E) 是 Chumak & Chumak (1987) [26] 提出的一个活动区的结构参量, 在耀斑少的活动区和耀斑活跃的活动区呈现出一定的区别 (Chumak et al. 2004 [27])。作为一个量化的磁复杂性参量, d_E 从几何意义上描述了活动区两极间的孤立或互相渗透的程度 (Guo et al. 2006 [60])。 d_E 的计算式为:

$$d_E = \frac{R_p + R_n}{R_{pn}}, \quad (3.9)$$

其中

$$R_p = \sqrt{\frac{A_p}{\pi}}, \quad R_n = \sqrt{\frac{A_n}{\pi}}.$$

A_p 和 A_n 分别为正极和负极的总面积, R_p 和 R_n 分别为正负极区域的等效半径, R_{pn} 为两极的磁流中心间的距离 (Guo et al. 2006, 2007, 2010 [58–60]; Guo & Zhang 2007 [57])。

考虑到有效距离参量 d_E 单独表征一个活动区复杂性存在一定的局限性, 我们将 d_E 乘上一个因子 $\overline{|B_z|}$ 。修改后的有效距离 d_{Em} 表示为:

$$d_{Em} = d_E \overline{|B_z|} = \frac{R_p + R_n}{R_{pn}} \overline{|B_z|}. \quad (3.10)$$

d_E 是无量纲的, $\overline{|B_z|}$ 从某种程度上为 d_{Em} 提供了一些实际物理意义。这里, d_{Em} 的量纲与磁场强度一致, 以高斯为单位。 d_{Em} 也反映了活动区的复杂程度, 可以认为是经黑子纵场平均强度加权过的复杂程度。

3.3 统计分析结果

3.3.1 22 – 23 太阳周非势性的强度分布

根据上节介绍的磁非势性和磁复杂性参量, 使用表达式 (3.3) – (3.10) 来计算每张磁图的以下参量的值: 平面剪切角均值 $\overline{\Delta\phi}$ 、空间剪切角均值 $\overline{\Delta\psi}$ 、纵向电流密度绝对值的均值 $\overline{|J_z|}$ 、电流螺度密度绝对值的均值 $\overline{|h_c|}$ 、自由磁能密度均值 $\overline{\rho_{free}}$ 、平均无力因子的绝对值 $|\alpha_{av}|$ 、纵向磁场的有效距离 d_E 、修改后的有效距离 d_{Em} 。对于 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $|\alpha_{av}|$ 、 $\overline{\rho_{free}}$ 、 d_{Em} , 满足纵场大于 20 G 的区域参与计算。对于 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$, 除了纵场大于 20 G 外, 还要求横场大于 200 G, 即, 将计算区域仅限制在黑子半影和沿磁中性线。计算 d_E 时, 纵场的下限设为 80 G。图 3.4 为一个磁图样本的例子, 该磁图由 SMFT 采集于 1989 年 2 月 15 日 02:12 UT, NOAA 活动区编号为 5356。白色和黑色区域分别对应纵向磁场为正和为负。箭头指示横向磁场方向, 箭头长度标示出横场相对大小, 横场强度小于 50 G 的区域未予标示。白色和黑色实等强度线指示纵场为 ± 20 G, 标识出计算 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $|\alpha_{av}|$ 、 $\overline{\rho_{free}}$ 和 d_{Em} 时的区域边界。白色和黑色虚等强度线指示纵场为 $B_z = \pm 80$ G, 标识出计算 d_E 区域边界。灰色 (电子版中为红色) 实等强度线指示横场为 200 G, 标识出计算 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ 时的进一步边界限制。磁图视场大小为约 $5.23' \times 3.63'$ 。

图 3.5 展示了 1988 年至 2008 年的活动区样本的以上八个参量的分布情况。我们画出这个时段内的月平均黑子数, 以作参照比较。定义活跃样本 (耀

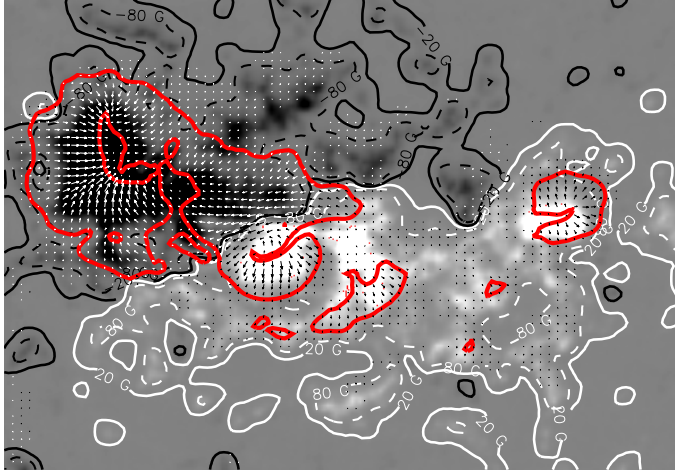


图 3.4: 一张矢量磁图示例 — NOAA 活动区 5356, 观测于 1989 年 2 月 15 日 02:12 UT

斑活跃活动区) 为在一定的向后时间内等效耀斑级别 (见 3.3.2 一节中定义的等效耀斑指数) 超过一定阈值的活动区样本。没能超过一定阈值的则为宁静样本 (耀斑宁静活动区)。图 3.5 中菱形小框 (电子版中为红色) 标识出活跃样本, 黑色圆点标识宁静样本。同时, 全部样本的年均值、以及活跃样本和宁静样本各自的年均值也分别计算出来, 分别画在图 3.6 和图 3.7 中。误差棒给出相应年均值的标准误差。同样, 叠加上了月平均黑子数的连线以供对照, 如图 3.6 和图 3.7 中虚线所示。

在图 3.5 中, 可以看到 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ 有着相似的分布。二者的线性相关系数为 0.875。它们没有随活动周的起伏而变化的趋势。其值绝大多数分布于区间 $[10^\circ, 40^\circ]$ 内。 $\overline{\Delta\psi}$ 的分布更加集中于这个区间。样本数据的 $\overline{\Delta\psi}$ 的均值为 22.0° 、标准差为 6.6° , $\overline{\Delta\phi}$ 的均值为 23.4° 、标准差为 7.6° 。

结合图 3.5 – 3.7, $|\overline{J_z}|$ 随活动周没有明显的起伏, 大多数样本的值在 $1.0 - 3.0 \times 10^3 \text{ A km}^{-2}$ 内浮动。在太阳活动峰年有一些活动区有着较大的 $|\overline{J_z}|$ 值。 $|\overline{J_z}|$ 值大于 $2.5 \times 10^3 \text{ A km}^{-2}$ 的样本多为活跃样本。在 2001 年至 2003 年间, 也即第 23 太阳活动周的下降期的开始阶段, $|\overline{J_z}|$ 值大多 (91.7%) 低于其整体的平均水平 $2.0 \times 10^3 \text{ A km}^{-2}$ 。2003 年后, 其值恢复到平均水平。图 3.7 显示出的 $|\overline{J_z}|$ 年均值演化的形状与 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ 非常相似。

对于 $|\alpha_{av}|$, 其值大都集中在 $0 - 1.5 \times 10^{-5} \text{ km}^{-1}$ 区间内, 随活动周也没有

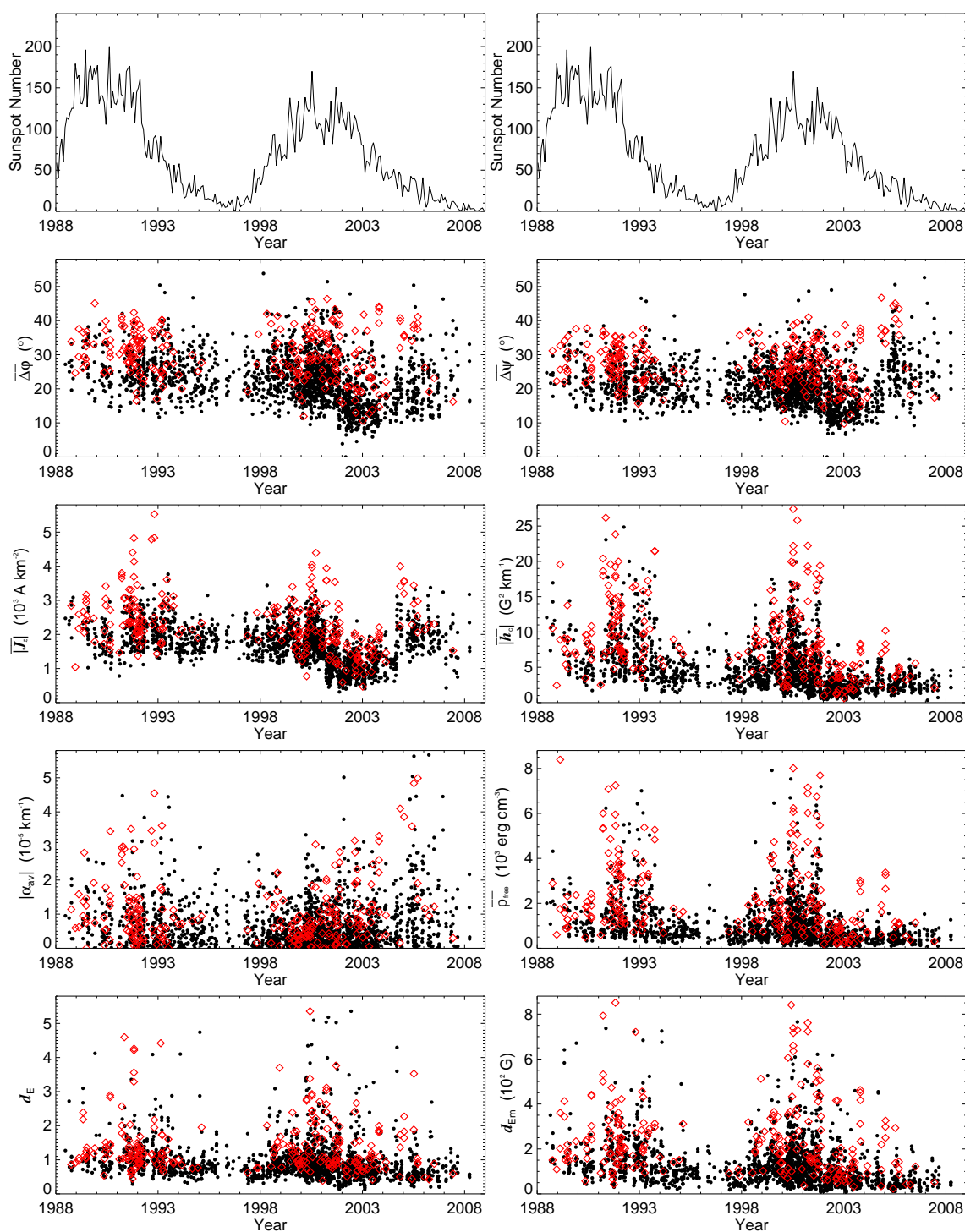


图 3.5: 1988 – 2008 年活动区样本的 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $\overline{|\alpha_{av}|}$ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的散点图

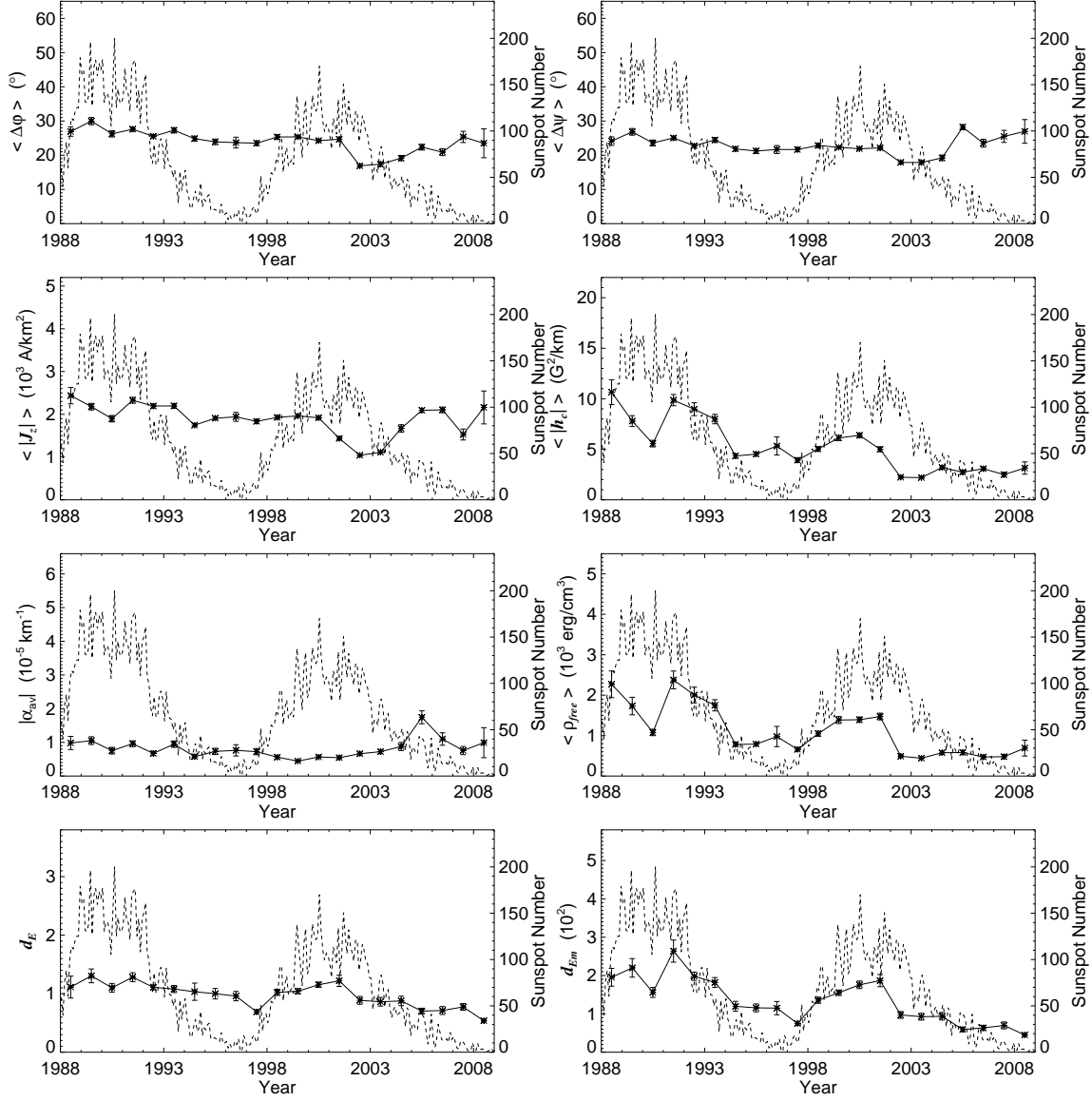


图 3.6: 1988 – 2008 年活动区样本的 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $\overline{|\alpha_{av}|}$ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的年均值

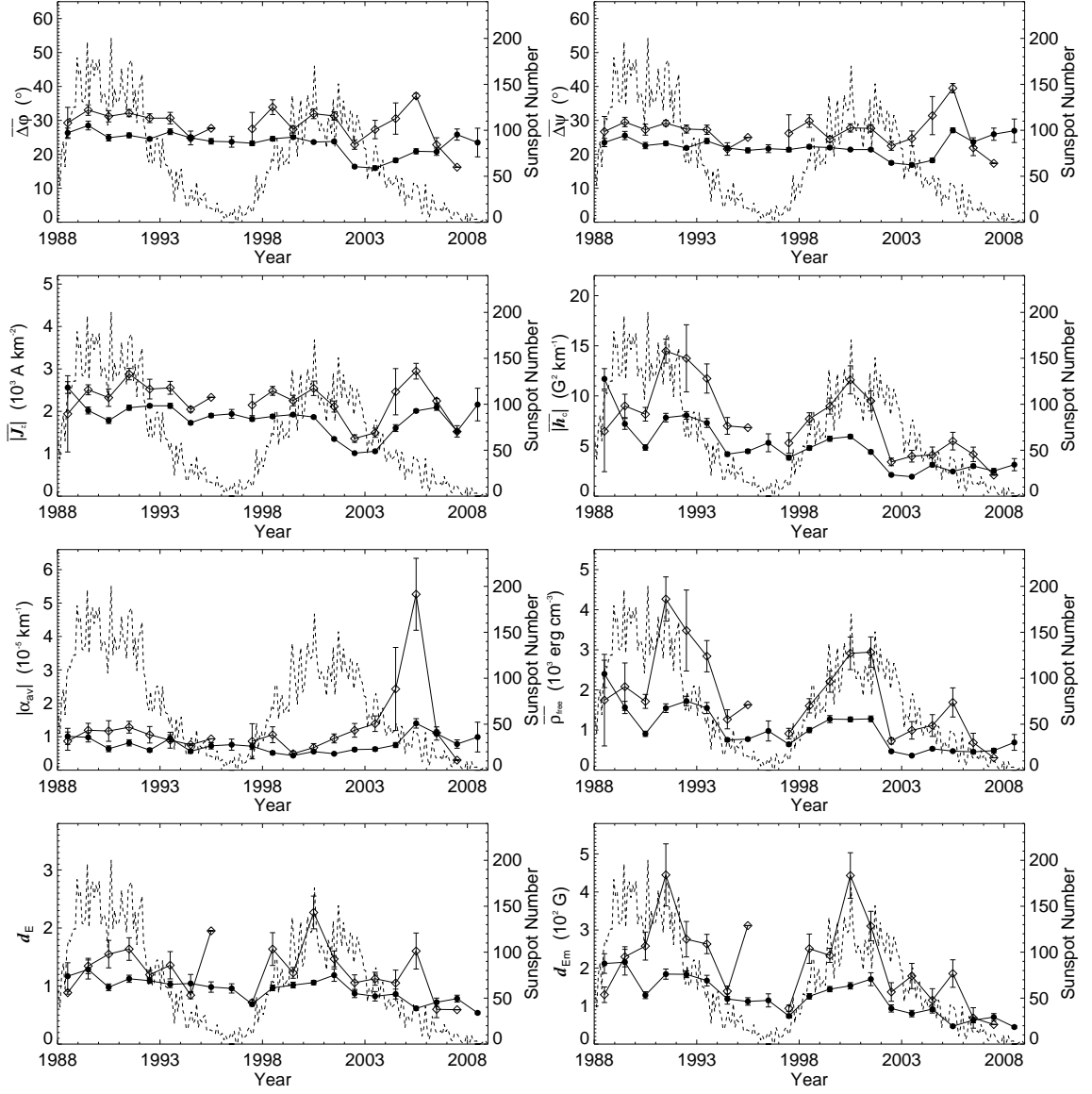


图 3.7: 1988 – 2008 年活动区样本中活跃样本与宁静样本分别的 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $\overline{|\alpha_{av}|}$ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em} 的年均值

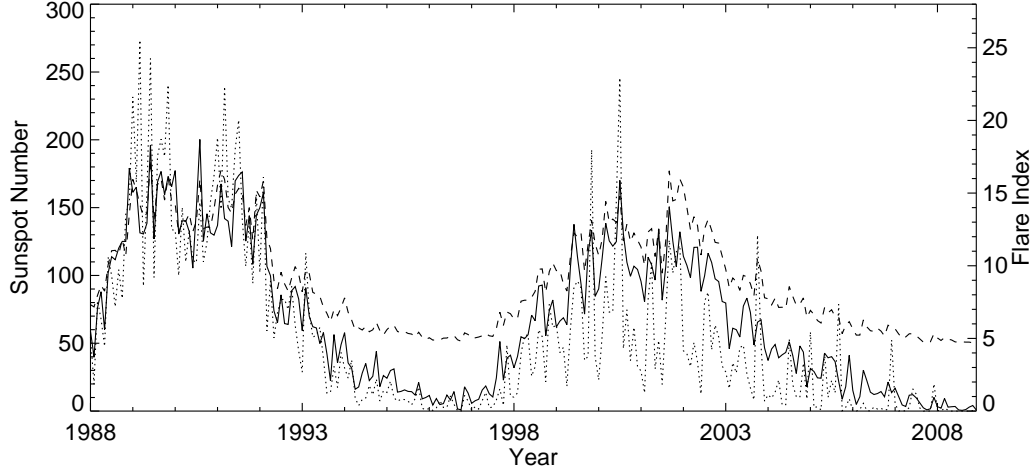


图 3.8: 1988 – 2008 年月均黑子数、耀斑指数、10.7 cm 射电流量

明显的变化。除了 2004 年和 2005 年，它在活跃样本与宁静样本之间的差别也不显著。

$\overline{|h_c|}$ 、 $\overline{\rho_{\text{free}}}$ 和 d_{Em} 与活动区的变化吻合得较好。两个活动周内都能看到两个明显的钟形结构。三个参量的整体水平随着活动峰年的来临而增长、随着活动低年的来临而下降。与月均黑子数相比较，1989 – 1990 年有一个低谷，这可能是由于可获得样本的缺少，或者是一些未知的不确定性或者未知的演化过程所致。这一点在 Bao & Zhang (1998) [14] 研究 22 太阳周的平均电流螺度演化时也得到了，但原因未知。表 3.2 列出了这八个参量年 (月) 均值分别与年 (月) 均黑子数的线性相关系数。从中也能看出平均黑子数与 $\overline{|h_c|}$ 、 $\overline{\rho_{\text{free}}}$ 和 d_{Em} 三者有着比较紧密的关系。

表 3.2: 八个参量年 (月) 均值分别与年 (月) 均黑子数的线性相关系数

	$\overline{\Delta\phi}$	$\overline{\Delta\psi}$	$\overline{ J_z }$	$\overline{ h_c }$	$ \alpha_{\text{av}} $	$\overline{\rho_{\text{free}}}$	d_{E}	d_{Em}
年均黑子数	0.379	0.010	0.093	0.594	-0.176	0.666	0.799	0.814
月均黑子数	0.242	0.085	0.054	0.448	-0.067	0.474	0.367	0.530

比较表示磁场结构复杂性的两个参量 d_{E} 和 d_{Em} ，在太阳活动峰年的一些

活动区的 d_E 值比在太阳活动低年的活动区的值要大,但是样本中那些有着较大 d_E 值的活动区的数目并不多,大多数样本的 d_E 值分布在一个狭窄的区间内,均值为 1.04,标准差为 0.93。结合了平均磁场强度, d_{Em} 更确切地反映了活动区的磁复杂程度,而 d_E 仅有限地与活动区的形态有关、缺少了些物理内容。二者都随活动周的发展而变化,这意味着在太阳活动峰年有着更多的复杂活动区,而在太阳活动低年有着较少的复杂活动区。

使用 SoHO/MDI 和 NSO/Kitt 峰的综合磁图, Guo et al. (2010) [59] 发现,在 23 周的下降期, $d_E > 1$ 的复杂活动区所占全部活动区的比例有所降低,尽管有一些大的起伏,他们所提到的大的起伏有一部分是对应于 2005 年的一个小峰,在其文中的图里能清楚地看到。在研究 23 周磁螺度的长期演化时, Yang & Zhang (2012) [151] 发现磁螺度的积累在 2005 年时两个半球有着相同的趋势。Zhang et al. (2010) [162] 和 Tiwari et al. (2009) [134] 也发现 22 周和 23 周的下降相时活动区并不遵守一般的半球螺度法则,类似的结果在 Hao & Zhang (2011) [67] 对 Hinode/SP 数据的分析研究工作中也体现出来。在图 3.5 和 3.7 中,每张小图里都能看到 2005 年的一个小的凸起。于是,我们也检查了 23 周的月均黑子数和每月 10.7 厘米射电流量的变化情况,发现在 2005 年附近并无异常情况,仅仅是在 2004 年末和 2005 年耀斑稍稍多了一些(图 3.8)。表 3.3 给出了一个 22 周下降期与 23 周下降期的一个比较。我们选择相比较的一组两个年份有着大致相同的年均黑子数,也即它们有着大致相似的太阳活动水平。一个例外是,将 2002 年和 2003 年合在一起与 1992 年进行比较。表 3.3 中列出的值为 23 周相比 22 周的相对差值(即 $\frac{x_{23}-x_{22}}{x_{22}}$, x_{23} 和 x_{22} 分别表示 23 周年份的指定样本与 22 周年份的全部样本所比较的相应参量的年均值)。总体活动水平 23 周的下降相比 22 周的下降相要低一些,与其一致,表 3.3 中的各参量也反映出相同的趋势,也即这些参量 23 周的值要普遍低于 22 周。然而,活跃样本的所有这些参量在 2005 年相比 1994 年有着正的相对差值。这显示出活跃样本的非势性在 2005 年稍强。23 周这种非单调的下降或许可以看作是 23 周和 24 周之间长而深的极小期的一种前兆。

3.3.2 非势性与其关联耀斑

在第一章中我们已陈述过软 X 射线耀斑分类。为方便分析非势性与耀斑间的关系,我们设定时间窗口 τ 表示从磁图观测时刻起向后的一段时间。在确定的时间窗口 τ 内,软 X 射线等效耀斑指数 (Flare Index, FI) 根据实际发生

表 3.3: 22 周下降期 (1992 – 1996) 与 23 周下降期 (2002 – 2007) 各参量年均值比较

	年份	活跃样本	宁静样本	全部样本		年份	活跃样本	宁静样本	全部样本
$\overline{\Delta\phi}$	02&03 vs. 92	-0.057	-0.398	-0.329	$\overline{\Delta\psi}$	02&03 vs. 92	0.009	-0.269	-0.215
	04 vs. 93	0.082	-0.335	-0.298		04 vs. 93	0.244	-0.259	-0.214
	05 vs. 94	0.407	-0.183	-0.101		05 vs. 94	0.747	0.218	0.291
	06 vs. 95	0.165	-0.170	-0.126		06 vs. 95	0.372	0.068	0.108
	07 vs. 96	0.295	0.031	0.070		07 vs. 96	0.450	0.129	0.177
$ \overline{J_z} $	02&03 vs. 92	-0.358	-0.560	-0.509	$ \overline{h_c} $	02&03 vs. 92	-0.603	-0.906	-0.753
	04 vs. 93	0.092	-0.272	-0.240		04 vs. 93	-0.505	-0.609	-0.600
	05 vs. 94	0.591	0.136	0.199		05 vs. 94	0.074	-0.442	-0.370
	06 vs. 95	0.212	0.086	0.102		06 vs. 95	-0.208	-0.336	-0.320
	07 vs. 96	-0.032	-0.244	-0.212		07 vs. 96	-0.493	-0.538	-0.531
$ \alpha_{av} $	02&03 vs. 92	0.891	-0.135	0.032	$\overline{\rho_{free}}$	02&03 vs. 92	-0.609	-1.001	-0.768
	04 vs. 93	1.576	-0.236	-0.074		04 vs. 93	-0.415	-0.696	-0.671
	05 vs. 94	6.398	1.325	2.028		05 vs. 94	0.687	-0.407	-0.256
	06 vs. 95	1.980	0.280	0.502		06 vs. 95	0.027	-0.467	-0.402
	07 vs. 96	0.192	-0.042	-0.008		07 vs. 96	-0.565	-0.502	-0.511
d_E	02&03 vs. 92	0.165	-0.282	-0.212	d_{Em}	02&03 vs. 92	-0.105	-0.661	-0.521
	04 vs. 93	-0.099	-0.197	-0.189		04 vs. 93	-0.425	-0.491	-0.485
	05 vs. 94	0.454	-0.444	-0.320		05 vs. 94	0.319	-0.637	-0.505
	06 vs. 95	-0.350	-0.278	-0.288		06 vs. 95	-0.432	-0.455	-0.452
	07 vs. 96	-0.267	-0.180	-0.193		07 vs. 96	-0.441	-0.384	-0.393

耀斑的不同级别的系数分别加权后相加得到，以此来量度一个活动区的耀斑爆发能力：

$$FI = 100 \sum_{\tau} I_X + 10 \sum_{\tau} I_M + \sum_{\tau} I_C, \quad (3.11)$$

式中 I_X 、 I_M 和 I_C 分别表示 X 级、M 级和 C 级耀斑类别里各自字母后面的数字 (Antalova 1996 [8]; Abramenko 2005 [1])。由于在太阳活动峰年附近，X 射线流量的背景比较高，B 级耀斑淹没在背景里难以探测 (参见 Feldman et al. 1997 [44]; Joshi et al. 2010 [76])，我们在计算中没有包含 B 级耀斑。在 3.3.1 一节中我们设定 $FI = 10.0$ (即等效于一个 M1.0 耀斑) 来划分活跃样本和宁静样本，在后面的分析中我们也使用同样的设定。

图 3.7 中，活动峰年附近每个参量的活跃样本的年均值要高于宁静样本的年均值，尤其对于 $|\overline{h_c}|$ 、 $\overline{\rho_{free}}$ 和 d_{Em} 三者。然而，无力因子 $|\alpha_{av}|$ 在活跃样本与宁静样本间显示的差别不显著，除了 2005 年差别比较明显，当然这个较大差别也有着较高的不确定性。对于同样的 τ 值，活跃样本的各参量的年均值随

FI 阈值的提高而变大, 活跃样本与宁静样本的差距也随之逐渐略为增大。如果确定 FI 阈值, 活跃样本的各参量的年均值随 τ 值的增加而变小, 活跃样本与宁静样本的差距相应是略为减小的。设定不同的 τ 值和不同的 FI 阈值, 各参量的年均值连线都有着基本相似的演化趋势 (图 3.7)。然而, 当 FI 阈值升高到一定程度 (比如 24 小时内的 M3.0, 48 小时内的 M5.0), 某些极小年附近年份的活跃样本就没有了, 尤其是在 1994 – 1996 年和 2006 – 2008 年。这意味着活跃样本更可能具有相对强的非势性和显著的复杂性, 于是这些表征非势性和复杂性的参量可以作为预测因子应用在太阳耀斑的预测上。

另外, 为考查有着一定非势性和复杂性的活动区它们的耀斑爆发可能性情况, 我们还计算了耀斑产率 (Cui et al. 2006, 2007 [32, 33]; Cui & Wang 2008 [34]; Park et al. 2010 [105]), 它被定义为:

$$P(X) = \frac{N_A(\geq X)}{N_T(\geq X)}, \quad (3.12)$$

其中 $N_T(\geq X)$ 为满足所研究参量的值大于其某一阈值 X 的活动区样本的总个数, $N_A(\geq X)$ 为这些样本中的活跃样本个数。

各参量的耀斑产率图见图 3.9。为保证统计尽可能可靠, 在图中显示的结果里总样本个数都不超过 50。图中的加号、叉乘号、菱形框、星号分别表示设置的时间窗口分别为 6 小时、12 小时、24 小时和 48 小时。用 Boltzmann sigmoid 函数拟合每组数值:

$$f(x) = a_2 + \frac{a_1 - a_2}{1 + e^{\frac{x - x_0}{W}}},$$

a_1 、 a_2 分别是 x 趋于大值方向和小值方向时 $f(x)$ 的两个渐近值, W 是曲线在两个渐近值处分别对应的两个 x 之间的宽度, x_0 位于该宽度的中心点, 曲线在 x_0 处有最大斜率。 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ 曲线形状非常相似。 $|\overline{J_z}|$ 与前面两者也很相像, 但当其值在 $2.5 \times 10^3 \text{ A km}^{-2}$ 附近时有一个很陡很快速的上扬。而 $|\overline{h_c}|$ 和 $|\alpha_{av}|$ 在这个计算里的增长趋势相似。 $\overline{\rho_{free}}$ 和 d_{Em} 的拟合曲线形状相像。当 d_E 的值大于 1.0 时, 活动区的耀斑产率几乎没有明显变化了。这些图反映出各量不同的耀斑产率水平及其它们的变化趋势, 这也是图 3.5 的散点分布图和图 3.7 的演化图的很自然的结果。图 3.9 中, 活跃样本与宁静样本相接近的参量比差异更明显的那些参量的拟合曲线要更趋平坦一些。而在每张小图里 X 最大值对应的耀斑产率可能反映了各量能对耀斑预测的贡献, 这些值在实际的预测模型中可以被赋予相应的权重。

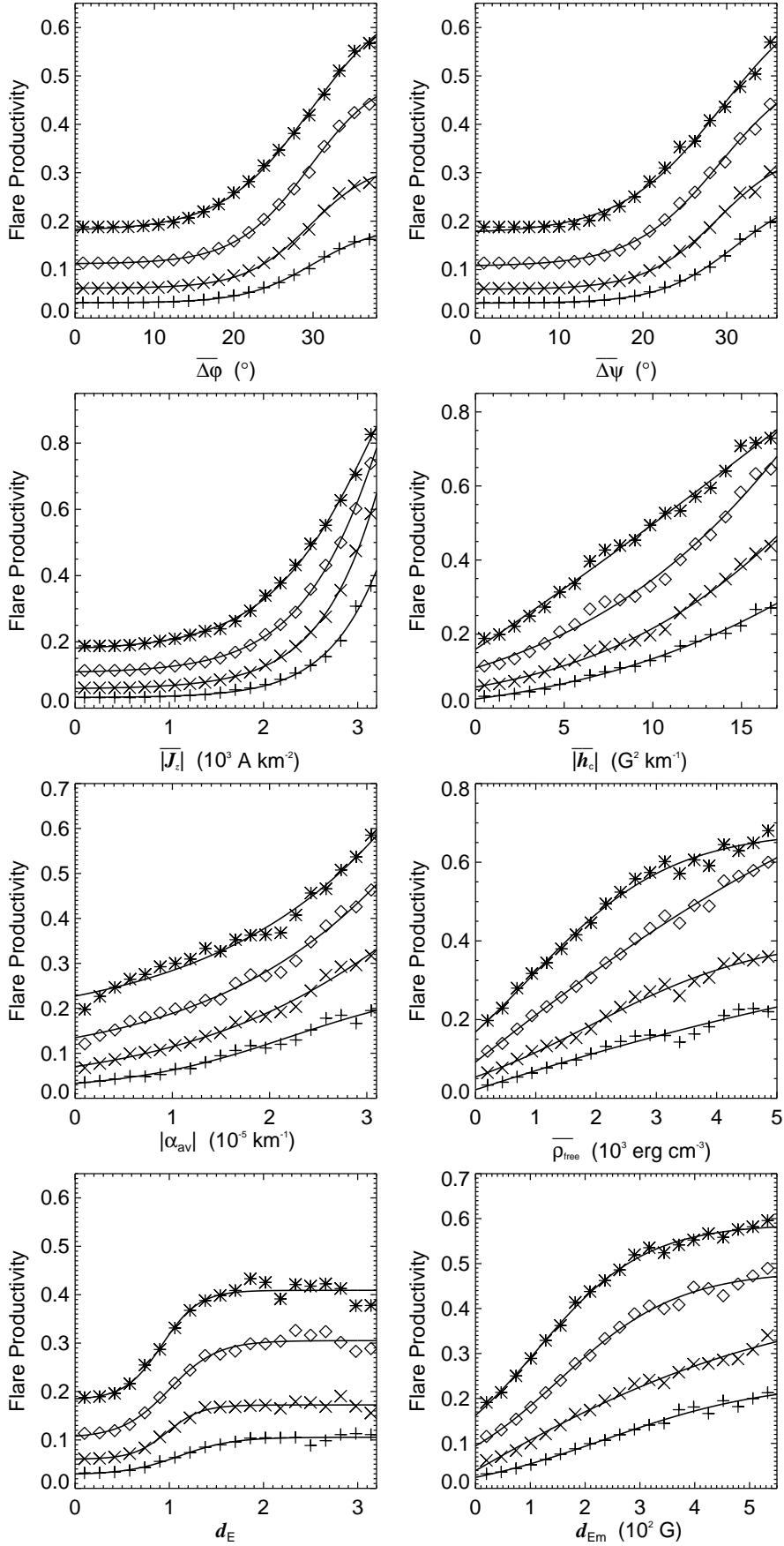


图 3.9: $\Delta\phi$ 、 $\Delta\psi$ 、 $|J_z|$ 、 $|h_c|$ 、 $|\alpha_{av}|$ 、 ρ_{free} 、 d_E 和 d_{Em} 的耀斑产率

3.3.3 非势参量与有效距离的比较

表 3.4: 八个参量间的线性相关系数

	$\overline{\Delta\phi}$	$\overline{\Delta\psi}$	$\overline{ J_z }$	$\overline{ h_c }$	$ \alpha_{av} $	$\overline{\rho_{free}}$	d_E	d_{Em}
$\overline{\Delta\phi}$	1.000	—	—	—	—	—	—	—
$\overline{\Delta\psi}$	0.875	1.000	—	—	—	—	—	—
$\overline{ J_z }$	0.528	0.499	1.000	—	—	—	—	—
$\overline{ h_c }$	0.386	0.274	0.737	1.000	—	—	—	—
$ \alpha_{av} $	0.306	0.425	0.285	0.107	1.000	—	—	—
$\overline{\rho_{free}}$	0.372	0.313	0.566	0.911	0.101	1.000	—	—
d_E	0.259	0.186	0.162	0.220	0.076	0.242	1.000	—
d_{Em}	0.330	0.241	0.357	0.566	0.069	0.618	0.856	1.000

表 3.4 列出了本章所研究的八个参量两两之间的线性相关系数。与前两节展示的结果一致, d_{Em} 与 $\overline{|h_c|}$ 、 $\overline{\rho_{free}}$ 二者有着相对紧密的联系。与 d_E 相比较, d_{Em} 和非势性参量有着更高一些的正相关, 而 d_E 几乎与非势性参量没有多少关系。直接相联系的参量 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$ 、 d_E 和 d_{Em} 、以及 $\overline{|h_c|}$ 和 $\overline{|J_z|}$, 它们两两之间高相关。另外 $\overline{|h_c|}$ 和 $\overline{\rho_{free}}$ 相关性也很高。图 3.10 显示了相关系数最高的这四对的散点图, 其相关系数由高到低为 0.911、0.875、0.865 和 0.737。图中矩形框 (电子版中为红色) 标识活跃样本, 黑色叉乘号标识宁静样本。对活跃样本和宁静样本分别考查两两参量的相关性, 除 d_E 和 d_{Em} 外, 其余两两组合的线性相关系数均是活跃样本高于宁静样本。各小图里的曲线为对其所有点的二项式最小二乘拟合。

3.4 结论与讨论

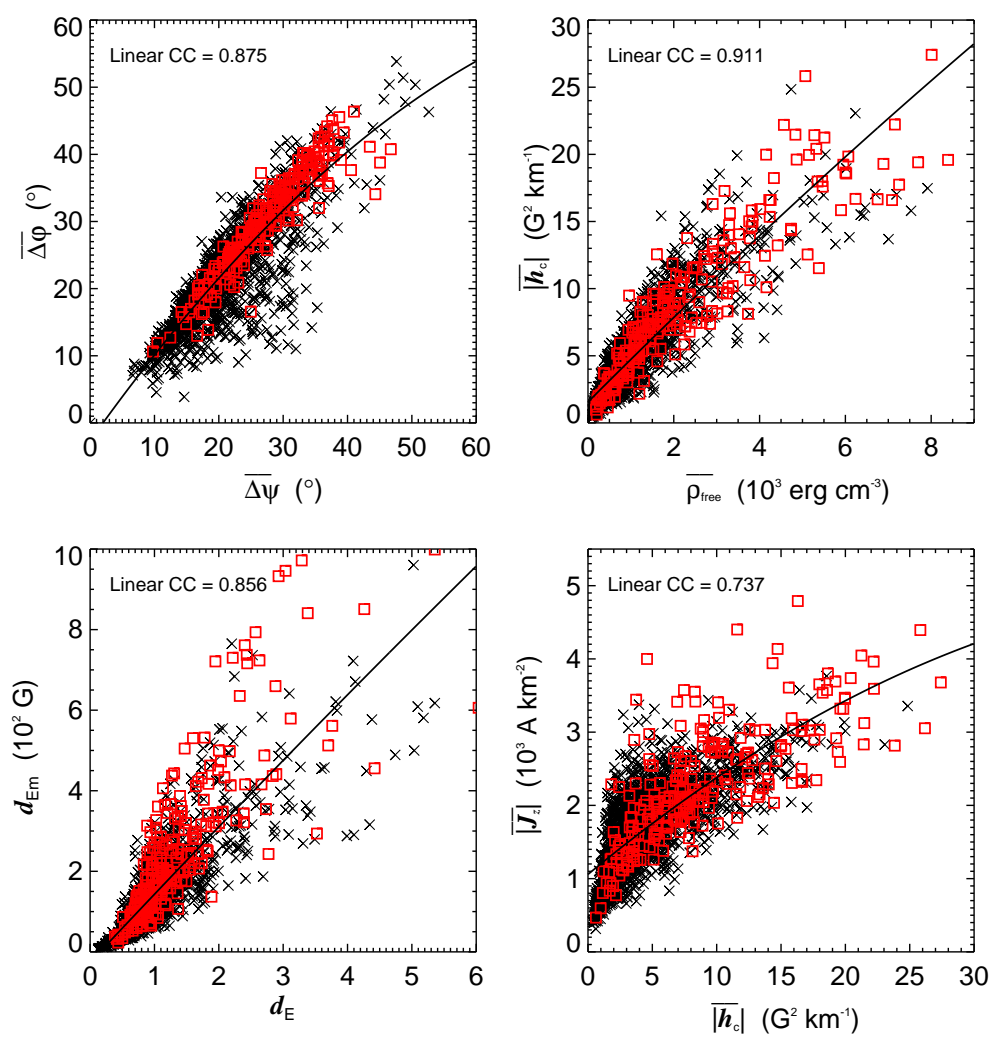
通过计算怀柔基地太阳磁场望远镜观测的 1106 个活动区共 2173 张磁图的表征活动区非势性和复杂性的八个参量 ($\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $|\alpha_{av}|$ 、 $\overline{\rho_{free}}$ 、 d_E 和 d_{Em}) 随活动周的演化及与它们的关联耀斑的关系, 我们得出以下主要结果:

- (1) 总体上讲, 活动区的两个磁剪切角均值 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$, 纵向电流密度绝对值的均值 $|\overline{J_z}|$, 平均无力因子的绝对值 $|\alpha_{av}|$, 以及有效距离 d_E 随整体太阳活动水平变化不显著。然而, 这些参量仍能反映出活动区在太阳活动峰年较低年具有更强的非势性。
- (2) 电流螺度密度绝对值的均值 $|\overline{h_c}|$ 、自由磁能密度均值 $\overline{\rho_{free}}$ 、纵场加权后的有效距离 d_{Em} 与平均黑子数显现出高的正相关, 且三者之间也有着相对紧密的关系。三者与年均黑子数的线性相关系数都高于 0.59。它们可以和传统的黑子数一样用于表征太阳活动水平。
- (3) 本工作中分别考查活跃样本和宁静样本的磁非势性和磁复杂性, 对于理解耀斑活跃活动区的演化及其与活动周磁活动水平的关系是有用的。这些磁非势性和磁复杂性参数可以作为预测因子用于太阳耀斑预测中 (下一章我们对此做了实际的实验)。
- (4) 有效距离参数 d_E 里缺少磁场强度信息, 经纵场加权后的有效距离 d_{Em} 则由于加入了磁场强度信息而能更好地反映活动区的磁活动特征。

活动区磁场里积累的非势能量提供了爆发太阳耀斑的能量, 并且耀斑的触发也同磁非势性有着一定的关系。本章讨论的这些参量将从不同的角度提供太阳活动区非势性的基本信息。不同非势性参量的综合分析能为预测太阳耀斑提供非势磁场与太阳耀斑关系的一个客观基础。

尽管地面设备长期观测的积累是不可缺少的, 但是诸如 Hinode、SDO 等一些空载太阳设备得到的太阳矢量磁场数据, 由于不受大气视宁度等的影响, 能得到更准确、更佳空间分辨率的数据, 可以与地面数据互相补充, 也能做些比较研究。由于空间观测数据不受天气和昼夜交替的影响, 连续性也较地面观测数据更好, 数据积累足够长时间后可以对不同活动区的磁场演化做统计研究。

在本工作中, 临近太阳边缘和高纬区域的活动区并未包含。但我们相信, 如果研究它们, 基本的结论应该是一样的。当然, 我们可以估计这些活动区的爆发耀斑的可能性, 但是由于垂直于视线方向的磁场测量有着相对较大的噪声 (横场测量误差约为 100 G), 对处于这些位置的活动区的预测会相对粗糙和不可靠一些。

图 3.10: $\Delta\phi$ 和 $\Delta\psi$ 、 $|h_c|$ 和 $\overline{\rho_{\text{free}}}$ 、 d_E 和 d_{Em} 、 $|h_c|$ 和 $|J_z|$ 的相关

第四章 基于光球活动区磁非势性的太阳耀斑预测

通过上一章我们得到的活动区磁非势性与太阳耀斑的统计关系，我们期望将磁非势性实际用于太阳耀斑预测中，以考查磁非势性参量作为耀斑预测因子的性能。本章的工作即是检验磁非势性参量用于太阳耀斑预测的性能。

关于预测，我们需要明确观测目标是什么，是预测事件的有无、事件发生的概率、还是事件某个特定指标的数值。然后我们需要有预测因子和预测方法。与预测目标密切相关的特征会很自然地入选预测因子，这需要有大量经验的积累和统计分析过程；预测方法同样预测类型有关系，不同的预测类型发展有一类预测方法。在预测模型构建出来以后，我们需要对模型的预测性能进行评估。性能符合要求后，预测模型则可进入实际应用中。我们这里的工作是一个最基本的二类分类的确定性预测。

基于怀柔太阳观测基地太阳磁场望远镜跨两个太阳周的光球活动区矢量磁场观测，我们从中计算得到这些矢量磁场数据的一些磁非势性参量，由此建立了一个机器学习模型来预测指定活动区在一定时间内的耀斑发生情况。预测模型由支持向量分类器 (Support Vector Classifier) 构造，这是一种成熟且应用广泛的通用分类器。我们讨论了一些经典的二分预测评价指标，来综合评估预测的性能或称质量。我们考查了不同耀斑级别和不同时间窗口的多种组合，对于设置不同耀斑阈值和研究时间窗口的数据集，分别训练和测试。通过这种间接的方法我们可以估计一个特定活动区爆发耀斑的大致时间和级别。我们试验了不同组合的预报因子。在我们的 100 组实验中，评价指标 Heidke Skill Score 值的范围在 0.23 至 0.48 之间，97% 的 True Skill Statistics 指标值高于 0.36。仅从纵向磁场得到的预报因子组合表现的确不差，然而，它们在预测更高级别的耀斑时却似乎不太敏感。而采用了从矢量场中得到的非势性参量作为预报因子，则能提高对更高级别耀斑 (如 M5.0 级以上耀斑) 的预测能力。

4.1 数据与方法

4.1.1 数据及其预处理

本工作中使用的观测数据全部来自怀柔太阳观测站 SMFT 望远镜观测得

到的光球活动区矢量磁图。所使用的数据集与前一工作完全一样，数据选取、矢量磁场定标、横场 180° 不确定性改正、投影改正等均与前一章一致，这里不再重复。软 X 射线耀斑信息亦同上章一样，从 NASA 的 NGDC 网站上获取。

4.1.2 作为预报因子的磁非势性参量

作为预报因子的磁场参量从以下参量中选取，它们将作为预测模型的输入。它们是：平面磁剪切角均值 $\overline{\Delta\phi}$ ，空间磁剪切角均值 $\overline{\Delta\psi}$ ，纵向电流密度绝对值的均值 $\overline{|J_z|}$ ，电流螺度密度绝对值的均值 $\overline{|h_c|}$ ，平均无力因子的绝对值 $|\alpha_{av}|$ ，自由磁能密度均值 $\overline{\rho_{free}}$ ，纵场有效距离 d_E ，纵场加权后的有效距离 d_{Em} (Yang et al. 2012 [152])，纵场水平梯度均值 $\overline{\nabla_h B_z}$ ，纵场水平梯度最大值 $(\nabla_h B_z)_m$ ，强梯度 ($>0.05 \text{ G km}^{-1}$) 中性线长度 L_{gnl} ，纵向磁能耗散密度均值 $\overline{\varepsilon(B_z)}$ (Cui et al. 2006 [33]; Jing et al. 2006 [74]; Cui et al. 2007 [32])。将以上参量的有关表达式分列如下：

- 平面磁剪切角

$$\Delta\phi = \arccos \left(\frac{\mathbf{B}_{to} \cdot \mathbf{B}_{tp}}{|\mathbf{B}_{to}| |\mathbf{B}_{tp}|} \right),$$

- 空间磁剪切角

$$\Delta\psi = \arccos \left(\frac{\mathbf{B}_o \cdot \mathbf{B}_p}{|\mathbf{B}_o| |\mathbf{B}_p|} \right),$$

- 纵向电流密度

$$J_z = \frac{1}{\mu_0} (\nabla \times \mathbf{B})_z = \frac{1}{\mu_0} \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right),$$

- 电流螺度密度的纵向分量

$$h_c = B_z (\nabla \times \mathbf{B})_z = B_z \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right),$$

- 无力因子

$$\alpha_{av} = \frac{\sum (\nabla \times \mathbf{B})_z \cdot \text{sign}[B_z]}{\sum |B_z|},$$

- 自由磁能密度

$$\rho_{\text{free}} = \frac{\mathbf{B}_s^2}{8\pi} = \frac{(\mathbf{B}_o - \mathbf{B}_p)^2}{8\pi},$$

- 有效距离参量

$$d_E = \frac{R_p + R_n}{R_{pn}},$$

- 纵场强度加权的有效距离参量

$$d_{Em} = d_E \overline{|B_z|} = \frac{R_p + R_n}{R_{pn}} \overline{|B_z|},$$

- 纵场水平梯度

$$\nabla_h B_z = \left[\left(\frac{\partial B_z}{\partial x} \right)^2 + \left(\frac{\partial B_z}{\partial y} \right)^2 \right]^{1/2},$$

- 纵向磁能耗散密度

$$\varepsilon(B_z) = 4 \left[\left(\frac{\partial B_z}{\partial x} \right)^2 + \left(\frac{\partial B_z}{\partial y} \right)^2 \right] + 2 \left(\frac{\partial B_z}{\partial x} + \frac{\partial B_z}{\partial y} \right)^2.$$

从而得到 $\overline{\Delta\phi}$ 、 $\overline{\Delta\psi}$ 、 $\overline{|J_z|}$ 、 $\overline{|h_c|}$ 、 $|\alpha_{av}|$ 、 $\overline{\rho_{\text{free}}}$ 、 d_E 、 d_{Em} 、 $\overline{\nabla_h B_z}$ 、 $(\nabla_h B_z)_m$ 、 L_{gnl} 和 $\overline{\varepsilon(B_z)}$ ，这些参量都是活动区整体磁非势性或磁复杂性的宏观和平均度量值。每张磁图计算出的这些参量的值构成预测因子，作为学习机器的输入向量，并且这些用作训练或测试的数据都有确定的类别标号。在计算中，各磁图都将表示为 (\mathbf{x}_i, y_i) ，这里 $\mathbf{x}_i \in \mathbf{R}^n$ 是预测因子向量， $y_i \in \{1, -1\}$ 是其相应的类别标号 ($y_i = 1$ 表示磁图为活跃样本， $y_i = -1$ 表示磁图为宁静样本，类别标号规则将在下节中说明)。 \mathbf{x}_i 的向量长度依赖于预测因子的数目。我们在实验中使用了以下五个预测因子组合：

$$\text{V06 } (\overline{\Delta\psi}, \overline{|J_z|}, \overline{|h_c|}, |\alpha_{av}|, \overline{\rho_{\text{free}}}, d_{Em}),$$

$$\text{V08 } (\overline{\Delta\phi}, \overline{\Delta\psi}, \overline{|J_z|}, \overline{|h_c|}, |\alpha_{av}|, \overline{\rho_{\text{free}}}, d_E, d_{Em}),$$

$$\text{L05 } (d_{Em}, \overline{\nabla_h B_z}, (\nabla_h B_z)_m, L_{\text{gnl}}, \overline{\varepsilon(B_z)}),$$

$$\text{A10 } (\overline{\Delta\psi}, \overline{|J_z|}, \overline{|h_c|}, |\alpha_{av}|, \overline{\rho_{\text{free}}}, d_{Em}, \overline{\nabla_h B_z}, (\nabla_h B_z)_m, L_{\text{gnl}}, \overline{\varepsilon(B_z)}),$$

$$\text{A12 } (\overline{\Delta\phi}, \overline{\Delta\psi}, \overline{|J_z|}, \overline{|h_c|}, |\alpha_{av}|, \overline{\rho_{\text{free}}}, d_E, d_{Em}, \overline{\nabla_h B_z}, (\nabla_h B_z)_m, L_{\text{gnl}}, \overline{\varepsilon(B_z)}).$$

4.1.3 预测方法：支持向量分类

预测一个活动区在指定的时间段内是否会发生耀斑，可以被抽象为一个分类问题。一个用于分类的机器学习系统能够从有确定类别标号的训练样本中学习、并从中构建一个分类模型，该模型可以较好地将训练样本分类，之后便能利用该模型来对一个未知类别标号的样本进行分类了。支持向量机的两个核心概念是最大间隔原则和核函数。通过求解一个最优化问题，得到一个最优的分隔超平面（如果训练样本仅有两个特征为分隔线，有三个特征为分隔平面，有三个以上特征则为分隔超平面），该超平面使两类数据拥有最大的间隔。如果数据是线性不可分的，则利用核函数将训练向量映射到一个更高维的向量特征空间中，在那个空间里数据是线性可分的。根据最大间隔原则，支持向量分类机利用一个核函数将原始特征空间变换到一个更高维的特征空间，这样可以使得一个线性不可分问题转化为一个线性可分问题，从而其相应的最优化问题得到解决，也就得到一个支持向量分类机。

原始优化问题可以表示为：

$$\begin{aligned} \min_{\mathbf{w} \in \mathcal{H}, b \in \mathbf{R}, \xi \in \mathbf{R}^l} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \xi_i, \\ \text{subject to} \quad & y_i(\mathbf{w} \cdot \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, i = 1, \dots, l, \\ & \xi_i \geq 0, i = 1, \dots, l. \end{aligned}$$

这里 $C > 0$ 为对松弛变量 ξ_i 之和的惩罚参数。对应于两类数据间的距离，取倒数、平方以及 $1/2$ 因子是为了数学计算上的方便。 $\phi(\mathbf{x}_i)$ 表示应用了核函数之后在高维空间中的训练向量。将核函数表示为 $K(\mathbf{x}_i, \mathbf{x}_j)$ ，则上述优化问题的对偶问题为

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j=1}^l \alpha_j, \\ \text{subject to} \quad & \sum_{i=1}^l y_i \alpha_i = 0, \\ & 0 \leq \alpha_i \leq C, i = 1, \dots, l, \end{aligned}$$

α_i 是拉格朗日乘子。从对偶问题中求解出最优超平面的系数 α_i^* 。 $\alpha_i^* \neq 0$ 的训

练向量 \mathbf{x}_i 被称为支持向量，这些向量对最后的决策函数产生贡献：

$$f(\mathbf{x}) = \text{sgn}(\mathbf{w}^* \cdot \phi(\mathbf{x}) + b^*) = \text{sgn} \left(\sum_{i=1}^l \alpha_i^* y_i K(\mathbf{x}, \mathbf{x}_i) + b^* \right),$$

其中 \mathbf{w}^* 和 b^* 为原始问题的解 (取任一正 α_j^* 求出 $b^* = y_j - \sum_{i=1}^l y_i \alpha_i^* K(\mathbf{x}_i, \mathbf{x}_j)$)。 $f(\mathbf{x})$ 的正负则对应于两个不同的类别。

经常使用的核函数有：多项式核，高斯径向基核，Sigmoid 核等等。试验了一些核后，我们采用高斯径向基核来进行计算，它在其他的一些应用中也是效果比较好的。高斯径向基核的数学表示为 $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2)$ ，其中 σ 为核的一个参数。实验中我们使用 SVM 软件 LIBSVM (Chang & Lin 2011 [24])，我们也用另一常用 SVM 软件 SVM^{light} (Joachims 1999 [75]) 做了对照计算，结果完全一致。由于前者相对更加易用，全部的计算由它完成。

需要指出的是，这种分类或预测模型是基于统计关系的，没有明显的物理含义；然而，与耀斑爆发联系紧密的物理量将会对模型的预测性能产生正面的影响。这也是我们采用磁非势性和磁复杂性参量作为耀斑预测因子的原因。

支持向量分类机的一个缺点是，不易寻找合适的分类模型参数，在构建模型的过程中不可避免地需要一个搜索的过程。然而这个问题在其他机器学习方法中也直接或间接地存在，同样不可避免。

4.2 实验与结果

4.2.1 实验设计

根据各样本在指定时间窗口内的爆发耀斑的级别，将每一个样本标记为正例 (爆发了超过指定级别的耀斑) 和负例 (没有爆发超过指定级别的耀斑)。本章中的时间窗口的开始时间均为每张磁图各自的观测时间。我们设定耀斑级别阈值分别为 C1.0、C5.0、M1.0、M5.0 和 X1.0，设定时间窗口分别为 6 小时、12 小时、24 小时和 48 小时。表 4.1 列出了按特定耀斑级别阈值和时间窗口划分的各组合的正例和负例数量分布。可以看到，对于不同的耀发阈值和时间窗口组合，正例与负例的样本比例是不同的 (从接近平衡到非常不平衡)。在每个标记集里，我们再将全集划分为训练集和测试集，通过在训练集上训练以得到分类器，然后在测试集上测试来评估分类器的性能。

表 4.1: 耀发 (f) 和无耀发 (n-f) 的样本分布

时间窗口	类别	耀发阈值				
		C1.0	C5.0	M1.0	M5.0	X1.0
48 hr	f	918	427	252	71	42
	n-f	1255	1746	1921	2102	2131
24 hr	f	697	291	167	39	25
	n-f	1476	1882	2006	2134	2148
12 hr	f	475	181	95	22	17
	n-f	1698	1992	2078	2151	2156
6 hr	f	309	101	58	12	9
	n-f	1864	2072	2115	2161	2164

k -折交叉确认被用来避免过拟合。将全集随机分为大致相同数量的 k 份, 保持每份中的正负例比例与原始全集相同。用其中的 $(k - 1)$ 份作为训练集, 剩下的 1 份用于测试。每份都要作为一次测试集, 共进行训练-测试 k 次。这里, 我们对大多数集合使用惯用的 $k = 10$; 对于正例数小于 50 的集合, 10 折后正例数小于 5, 我们采用 $k = 5$ 。

4.2.2 性能评价

对上述二类分类的测试结果中成功与否的计数可以构造出一个 2×2 混淆矩阵 (也称为 contingency table), 见表 4.2。一些用于评估这种二类预测性能的评价指标将由此表得出。简便起见, 我们使用符号 (x, y, z, w) 来表示它的四个元素, 其中 x 为预测正确的正例数 (称为 True Positive 真正或 Hit 命中), y 为预测错误的正例数 (称为 False Negative 假负或 Miss 漏报), z 为预测错误的负例数 (称为 False Positive 假正或 False Alarm 误报), w 为预测正确的负例数 (称为 True Negative 真负或 Correct Rejection 正确判负)。显然地, x 和 w 会对预测评估产生正面的影响, y 和 z 则产生负面的影响。 N_0 和 N_1 分别表示负例总数和正例总数。

由表 4.2, 我们能得到最直接的八个与各元素的边际和相关的比例 (POD¹

¹Probability of Detection

表 4.2: 2×2 混淆矩阵的定义

实测	预测		总计
	是	否	
是	x	y	$N_1 = x + y$
否	z	w	$N_0 = z + w$
			$N = x + y + z + w$

$= x/(x + y)$, $\text{FOH}^2 = x/(x + z)$, $\text{FAR}^3 = z/(x + z)$, $\text{POFD}^4 = z/(z + w)$, $\text{FOM}^5 = y/(x + y)$, $\text{DFR}^6 = y/(y + w)$, $\text{PON}^7 = w/(z + w)$, $\text{FOCN}^8 = w/(y + w)$; Doswell et al. 1990 [39]). 它们之间有着非常简单的关系: $\text{POD} + \text{FOM} = 1$, $\text{FOH} + \text{FAR} = 1$, $\text{PON} + \text{FOFD} = 1$, $\text{FOCN} + \text{DFR} = 1$. 其中, 以 x 或 w 为分子的四个比例 POD 、 FOH 、 PON 和 FOCN , 我们希望它们的值越大越好, 剩下的四个则希望它们的值越小越好. 此外, 还可以得到一些其他的评价指标, 如 F_1 (F_β , $\beta = 1$; Van 1979 [136]; Chinchor 1992 [25]), Heidke Skill Score (HSS; Doolittle 1888 [38]; Heidke 1926 [71]; 参见 Woodcock 1976 [147] 和 Doswell et al. 1990 [39]), True Skill Statistic (TSS; 又称为 Peirce Skill Score, Peirce 1884 [106] 或 Hanssen-Kuipers' discriminant, Hanssen & Kuipers [66]; 参见 Woodcock 1976 [147] 和 Doswell 1990 [39]), Critical Success Index (CSI; 又称为 threat score, Gilbert 1884 [56]; Donaldson et al. 1975 [37]; 参见 Schaefer 1990 [120]), Gilbert Skill Score (GSS; Gilbert 1884 [56]; 参见 Schaefer 1990 [120]), Clayton Skill Score (CSS; Clayton 1889, 1934 [28, 29]; 参见 Wandishin & Brooks 2002 [139]) 等等, 表 4.3 给出了对它们的一个小的总结. 其中 GSS 的最小值为 $-1/3$ 而非 -1 并不特别要紧, 对于负的 GSS 只需注意负得越多预测性能越差就足够了 (Schaefer 1990 [120]). 完美的预测对应于它们的上界 1, 当然在实际中很难真正达到完美. 尽管由 “Finley 事件” (该事件由文章 Finley 1884 [46] 引起, 故

²Frequency of Hits, 或称 True Positive Rate

³False Alarm Ratio

⁴Probability of False Detection

⁵Frequency of Misses

⁶Detection Failure Ratio

⁷Probability of a Null event

⁸Frequency of Correct Null forecasts, 或称 True Negative Rate

名, 参见 Murphy 1996 [101]) 引发的热议已经过去了一个多世纪, 对这个看似简单的 2×2 问题的研究依然在继续 (Stephenson 2000 [129])。本章中, 我们仅考虑几个经典指标, 它们在实际应用中更直观、更容易理解和使用。

正确预测的比例即通常说的报准率 $(x + w)/N$ (后面用 ACC 表示), 虽说最简单最容易被人想到, 但对于评价预测来说也常常是一个误导人们的指标, 尤其对于预测的双方中有一方的数量具有压倒性优势, 比如罕见事件预测 (全部报无即全部漏报时的 ACC 会很高, 但这样做没有任何意义, 我们的目标是希望在尽可能报对正例的同时减少对负例的误报, 但实际中真正要做到这一点很难, 很多时候迫不得已需要用误报来换取尽量少的漏报)。正是考虑到往往会出现这种情况, 在预测罕见事件时, 去掉式中的 w 而得到 $x/(x + y + z)$, 这正是 CSI。然而, ACC、CSI 和 F_1 没有排除掉正确预测里通过随机预测就可正确预测的那一部分。而所谓的 skill score 则正是用来表示预测的一个相对的准确率, 相比于一些标准的参考预测, 如随机预测、常数预测等。Skill score (SS) 的一般形式表示为:

$$SS = \frac{S - S_{\text{ref}}}{S_{\text{perfect}} - S_{\text{ref}}} \times 100\%,$$

其中 S 是实际预测的正确数量, S_{ref} 为参考预测的正确数量, S_{perfect} 为完美预测的正确数量 (即全部预测正确)。一个无技巧 (no skill) 预测其值为 0, 正值表示预测结果好于参考预测, 完美预测的值为 1。HSS 为这样的 skill score, 它将 ACC 与随机预测相比较得到 (从 x 中减去了正确报正的期望 $(x + z)(x + y)/N$, 从 w 中减去了正确报负的期望 $(y + w)(z + w)/N$)。GSS 则是技巧修改 (skill-modified) 于 CSI, 从正确正例预测里减去了期望的正确正例预测值 (对于 CSI 和 GSS 更详细的分析可见 Schaefer 1990 [120])。 F_1 为 POD 和 FOH 的调和平均, HSS 恰好是技巧修改后的 POD 和技巧修改后的 FOH (Schaefer 1990 [120] 中的 POD_s 和 RS_s) 的调和平均。技巧修改后的量总是小于未经修改的量。GSS 总是小于 CSI, 事件比例决定了二者有多接近。HSS 总比 F_1 小, 而 TSS 通常比 POD 小 (当 $z = 0$ 时, $\text{TSS} = \text{POD}$), CSS 通常比 FOH 小 (当 $y = 0$ 时, $\text{CSS} = \text{FOH}$)。

表 4.3: 评价指标

评价指标	推导	公式	w 主导的情形	范围
GSS	$GSS = \frac{x-E_1}{(x-E_1)+y+z}$	$GSS = \frac{x \cdot w - y \cdot z}{(y+z)(x+y+z+w)+x \cdot w - y \cdot z}$	\rightarrow CSI	$[-1/3, 1]$
HSS	$HSS = \frac{(x-E_1)-(w-E_0)}{(x-E_1)+y+z+(w-E_0)}$	$HSS = \frac{2(x \cdot w - y \cdot z)}{(x+y)(y+w)+(x+z)(z+w)}$	$\rightarrow F_1$	$[-1, 1]$
TSS	$TSS = \text{POD} - \text{POFD}$	$TSS = \frac{x \cdot w - y \cdot z}{(x+y)(z+w)}$	\rightarrow POD	$[-1, 1]$
CSS	$CSS = \text{FOH} - \text{DFR}$	$CSS = \frac{x \cdot w - y \cdot z}{(x+z)(y+w)}$	\rightarrow FOH	$[-1, 1]$
CSI	$CSI = \frac{x+w-w}{x+y+z+w-w}$	$CSI = \frac{x}{x+y+z}$	CSI	$[0, 1]$
F_1	$F_1 = 2(\text{POD}^{-1} + \text{FOH}^{-1})^{-1}$	$F_1 = \frac{2x}{2x+y+z}$	F_1	$[0, 1]$

这些评价指标通过 x 、 y 、 z 和 w 而彼此间互相关联。HSS、GSS、TSS 和 CSS 的一个共同点是，它们的分子里都含有因子 $(x \cdot w - y \cdot z)$ 。在随机预测里，该因子的值为 0，这样它们四个 skill score 的值也为 0，即意味着没有技巧。在常数预测时（预测全部为正 $y = w = 0$ ，或全部为负 $x = z = 0$ ），这个因子也为 0；此时 CSS 分母为 0 没有意义，这也许是 CSS 并不常用的原因。也就是说，如果这些指标的值大于 0，则表示预测结果优于随机预测和常数预测。而 CSI、 F_1 和 ACC 三者的值在随机情况下则依赖于正负例的比例。四个 skill scores 的另一共同点为，它们对正例和负例是公平对待的。即，如果我们将关注点放在预测事件不发生上，同时交换 x 和 w 、以及 y 和 z ，它们四个的值是不变的。这一点是 CSI 和 F_1 所不具有的。

另外来考虑一下四个 skill scores 的大小关系，保持它们的分子完全一致，比较它们的分母：

$$\begin{aligned} D_{\text{HSS}} - D_{\text{TSS}} &= \frac{1}{2}(y - z)(y - z + x - w), \\ D_{\text{TSS}} - D_{\text{CSS}} &= (z - y)(x - w), \\ D_{\text{GSS}} - D_{\text{HSS}} &= \frac{1}{2}(y + z)(x + y + z + w), \\ D_{\text{GSS}} - D_{\text{TSS}} &= y(x + y) + z(z + w). \end{aligned}$$

GSS 总是小于 TSS 和 HSS，除了 $y = z = 0$ 即完美预测时。TSS 与 CSS、HSS 与 TSS 的大小关系不确定。在罕见事件预测时， $w \gg x$ 绝对主导，优化 TSS 时经常会导致 $z > y$ ，故此时 HSS 小于 TSS。而在 w 不明显主导时，HSS 与 TSS 的值非常接近，在我们的结果表 4.5 中是能看出来的。对混淆矩阵和评价指标更详细的介绍可见于 Wilks (2006) [145]。

Woodcock (1976) [147] 综述了部分 skill scores，主要为 TSS 和 HSS。并强调在优化 skill score 的过程中需要注意的四点，即正负比例、不同指标反映不同属性、一个数字难免会丢失信息、多次计算 skill score 得标准差以确认统计显著，这也是我们在评价或优化一个预测结果时需要引起注意的地方。尽管 TSS 被认为是与实验无关的，即它不依赖于实验的正负例之比，只要保持 x/y 和 z/w 不变，TSS 的值不会发生变化，但这并不意味着这些结果的预测效果是相同的（从下节 Woodcock 1976 [147] 的例子中可以看出）。在某一指标的值相近的情况下，我们考查其他的指标，会发现，值更高者预测效果更好。但这也仅是对相同环境下的实验（同一批数据的实验）而言，对不同的实验一味地

比较它们的指标其实是没有多少意义的, 说明不了什么问题。我们除了可以简单从直观上判断结果是否较好之外 (即与标准预报如随机预报或常数预报相比), 需要实现的只能是在相同环境下使预测结果更优, 比较两种预测方法也应该基于这个相同环境的条件。

4.2.3 实验结果

同时优化所有的评价指标几乎是不可能的 (Manzato 2005 [91]; 见 Bloomfield et al. 2012 [18] 的结果所示)。我们计算了几个最关心的评价指标 (POD, FOH, TSS, HSS, GSS, CSI, F_1 , $\sqrt[3]{\text{POD} \cdot \text{FOH} \cdot \text{FOCN}}$) 的几何平均值来优化模型, 这样做的目的是让这些指标之间互相制衡、以防偏倚。我们使用了网格搜索以得到最终相对较优的支持向量分类器的 (C, σ^2) 的值。表 4.5 中给出的是预测因子组合 A12 的各项预测结果中该几何平均值为最高的一组值。给出了它们各自在 k 次测试里相应值的平均值及其标准误差。在每行的最后给出了负例所占的比例以作参照。本章最后的表 4.6 – 4.10 则给出了完整的结果。除了当 $x = 0$ 或 $y = z = 0$ 时外, F_1 总是大于 CSI。在罕见事件情形下, HSS 接近于 F_1 , 故 HSS 一般会大于 CSI。我们的结果中仅有两组的 HSS 小于 CSI (C1.0, 48 小时; C1.0, 24 小时), 这也是正例所占比例最高的两组, 它们的 w 并非极端主导。预测级别为 $\geq \text{C1.0}$ 和 $\geq \text{C5.0}$ 的耀斑时, 由纵场得到的预测因子组合 (L05) 的预测表现稍稍好于主要包含横场分量的预测因子组合 (V06, V08)。然而, 这个优势随着预测耀斑级别的增大而逐渐失去。当预测级别为 $\geq \text{M5.0}$ 和 $\geq \text{X1.0}$ 的耀斑时, 仅由纵场得到的预测因子组合变得反不如其他了。这看上去仅来自纵场的预测因子似乎对预测大级别耀斑不那么敏感。整体来看, 采用包含了纵场预测因子和横场分量的组合 (A10, A12), 其预测性能或多或少会有所提高。

HSS 和 TSS 是在二类预测评价领域里最常讨论和应用的两个指标 (如: Woodcock 1976 [147]; Doswell et al. 1990 [39]; Manzato 2005 [91])。当 $y = z$ 时, HSS 与 TSS 的值相等; 当 $N_1 = N_0$ 即正负例平衡时, HSS 恒等于 TSS。Bloomfield et al. 2012 [18] 在其文中呼吁使用 TSS 替代 HSS 来更可靠地比较耀斑预报。然而, 没有一个标量指标是能完全覆盖全部的预测结果信息的。甚至无偏向的 TSS (它与事件发生的频度无关), 会在评价罕见事件的预测时失效 (Doswell et al. 1990 [39])。罕见事情的情形下, TSS 接近于 POD, 这样 w 和 z 对结果的贡献很有限。经验上讲, 如果 x 的比重提高了 ($N_1 = x + y$ 的值在

一个实验环境下是固定不变的), z 的值也会变大。偏向 $((x+z)/(x+y) \neq 1)$ 可能在优化一个指标时无意被引入 (Manzato 2005 [91])。追求更高的 POD 或者 TSS, 将会导致更高的 FAR 和更低的 FOH。更少漏报的代价将是经受更多的误报, 然而“狼来了”也同样是我們不愿看到的。另外, 相同的 TSS 也并不表示有相同的预测性能。举一个例子, 在表 4.4 中我们重列了 Woodcock (1976) [147] 文中一个例子的一部分。预测 P1 有 75% 的正例预测准确率和 50% 的负例预测准确率 (POD=75%、PON=50%), 预测 P2 有 50% 的正例预测准确率和 75% 的负例预测准确率 (POD=50%、PON=75%)。在四种情形下, TSS 保持不变, 但预测结果的好坏却是的确不一样的。因此, 仅仅一个指标可能会误导预测评价, 多指标综合评价是非常需要的。这个观点在 Schaefer (1990) [120]、Doswell et al. (1990) [39]、Marzban (1998) [93] 等文中也有表述和讨论。我们认为对不同实验数据进行预测比较是不甚合适的, 不同的数据可能有着各自不同的内禀特性。

表 4.4: 两种情形下的两种预测结果举例

情形 1: $N_1 = x + y = 140$, $N_0 = z + w = 60$						
(x, y, z, w)	TSS	HSS	GSS	CSS	CSI	F_1
预测 1: (105, 35, 30, 30)	0.250	0.244	0.139	0.239	0.617	0.764
预测 2: (70, 70, 15, 45)	0.250	0.198	0.110	0.215	0.452	0.622
情形 2: $N_1 = x + y = 60$, $N_0 = z + w = 140$						
(x, y, z, w)	TSS	HSS	GSS	CSS	CSI	F_1
预测 1: (45, 15, 70, 70)	0.250	0.198	0.110	0.215	0.346	0.514
预测 2: (30, 30, 35, 105)	0.250	0.244	0.139	0.239	0.316	0.480

关于高相关因子之间的冗余问题, 由于非势性质在包括我们考查的参量在内的不同非势性参量上均有所体现, 我们希望所使用的因子能从不同的侧面反映实际磁场对稳定状态的偏离。对高相关性因子的使用确实并未对预测产生明显积极的影响, 在实际操作中可择一参与计算。另外, 大多数机器学习算法本身可以对特征即输入的预测因子进行选择。

表 4.5: 测试 SVC 的评价结果

预测设定		评价指标										
耀斑级别	时间窗口	POD	FOH	FOCN	CSI	F ₁	TSS	CSS	HSS	GSS	ACC	N ₀ /N
≥C1.0	48 hr	0.707±0.011	0.690±0.013	0.782±0.008	0.538±0.013	0.698±0.011	0.474±0.020	0.472±0.020	0.473±0.020	0.312±0.018	0.742±0.010	0.578
	24 hr	0.677±0.019	0.617±0.013	0.840±0.008	0.478±0.016	0.645±0.014	0.478±0.022	0.458±0.020	0.466±0.021	0.306±0.018	0.761±0.010	0.679
	12 hr	0.653±0.028	0.508±0.010	0.895±0.007	0.399±0.013	0.569±0.014	0.475±0.024	0.404±0.014	0.430±0.016	0.275±0.013	0.786±0.006	0.781
	6 hr	0.560±0.026	0.423±0.022	0.923±0.004	0.317±0.017	0.479±0.020	0.430±0.026	0.346±0.025	0.378±0.024	0.235±0.018	0.826±0.009	0.858
≥C5.0	48 hr	0.627±0.026	0.549±0.019	0.906±0.005	0.415±0.020	0.584±0.020	0.500±0.027	0.455±0.024	0.474±0.025	0.313±0.021	0.825±0.008	0.803
	24 hr	0.626±0.028	0.457±0.022	0.939±0.004	0.357±0.018	0.524±0.018	0.507±0.026	0.396±0.024	0.437±0.022	0.282±0.019	0.847±0.008	0.866
	12 hr	0.485±0.044	0.400±0.028	0.952±0.004	0.283±0.028	0.435±0.033	0.419±0.044	0.352±0.031	0.379±0.035	0.239±0.029	0.896±0.006	0.917
	6 hr	0.595±0.064	0.250±0.029	0.979±0.003	0.219±0.029	0.351±0.038	0.508±0.065	0.229±0.032	0.306±0.041	0.187±0.029	0.898±0.006	0.954
≥M1.0	48 hr	0.642±0.028	0.433±0.021	0.950±0.004	0.350±0.021	0.516±0.023	0.531±0.030	0.383±0.024	0.438±0.026	0.284±0.021	0.860±0.007	0.884
	24 hr	0.550±0.039	0.419±0.023	0.962±0.003	0.314±0.023	0.474±0.028	0.486±0.039	0.381±0.026	0.424±0.030	0.273±0.023	0.907±0.004	0.923
	12 hr	0.554±0.056	0.344±0.021	0.979±0.002	0.266±0.025	0.415±0.030	0.505±0.052	0.323±0.022	0.382±0.031	0.240±0.024	0.934±0.004	0.956
	6 hr	0.523±0.067	0.225±0.028	0.986±0.002	0.191±0.028	0.312±0.039	0.474±0.067	0.211±0.030	0.286±0.040	0.173±0.028	0.939±0.004	0.973
≥M5.0	48 hr	0.634±0.048	0.319±0.020	0.987±0.002	0.268±0.020	0.419±0.025	0.587±0.047	0.306±0.021	0.393±0.026	0.247±0.020	0.942±0.004	0.967
	24 hr	0.329±0.114	0.434±0.094	0.988±0.002	0.231±0.075	0.353±0.093	0.321±0.114	0.422±0.095	0.343±0.094	0.224±0.075	0.980±0.003	0.982
	12 hr	0.460±0.104	0.275±0.067	0.994±0.001	0.213±0.056	0.338±0.075	0.447±0.105	0.269±0.067	0.329±0.076	0.207±0.056	0.982±0.002	0.990
	6 hr	0.667±0.139	0.220±0.038	0.998±0.001	0.202±0.042	0.327±0.058	0.654±0.139	0.218±0.039	0.322±0.059	0.198±0.042	0.985±0.002	0.994
≥X1.0	48 hr	0.522±0.070	0.487±0.141	0.991±0.001	0.326±0.072	0.474±0.082	0.507±0.070	0.477±0.141	0.462±0.084	0.316±0.073	0.976±0.005	0.981
	24 hr	0.480±0.102	0.375±0.094	0.994±0.001	0.262±0.061	0.401±0.078	0.469±0.102	0.369±0.095	0.392±0.079	0.256±0.061	0.983±0.003	0.982
	12 hr	0.533±0.062	0.278±0.023	0.996±0.001	0.214±0.010	0.353±0.013	0.522±0.061	0.274±0.023	0.346±0.013	0.210±0.010	0.985±0.002	0.990
	6 hr	0.700±0.200	0.169±0.055	0.999±0.001	0.167±0.056	0.270±0.084	0.688±0.199	0.167±0.056	0.265±0.084	0.164±0.056	0.987±0.003	0.996

4.3 结论与讨论

基于 SMFT 长期稳定一致的光球矢量磁场观测, 我们采用了一些仅从视向场数据中所无法获得的非势性参量, 来研究太阳耀斑的预测。实时处理和无需人工干预是我们的预测系统的两个优势。预测模型的输入由本地获得, 用于模型中的关键参量可以直接由计算机自动计算得到而无需人工操作。

从我们的实验中发现, 来自纵场的预测因子组合, 确实在预测耀斑中表现不错; 然而, 在预测大级别耀斑时, 它们可能不如矢量磁场得到的预测因子敏感。来自横场的信息尽管在预测低级别耀斑时贡献有限, 但在预测诸如 $\geq M5.0$ 和 $\geq X1.0$ 这样大级别的耀斑时确实提高了预测性能。因此, 在耀斑预测时包含进横场信息是一个合理的选择。

为避免单个指标对优化过程的误导以及对结果的误用, 预测结果应该被仔细地评估。关注多个评价指标可更公平地评价预测结果。分开训练集和测试集、使用 k -折交叉确认等过程在提高预测模型的泛化能力方面是非常必要的。另外, 不同数据集的内禀特性会使得同样的预测工具表现得非常不同, 故而在同样的数据环境下比较预测因子或预测方法应该才是有意义的。

一些研究者已经开始试验使用 SDO 卫星上搭载的 HMI 仪器观测到的矢量磁场数据进行耀斑预测了。然而, 建立在统计信息上的预测方法受限于当前 HMI 数据有限的时间跨度。统计预测结果不仅依赖于所使用的预测方法, 同时也依赖于历史数据集。在将来, 使用优质的 HMI 数据的太阳爆发活动预测还将有很长的路要走。

我们不选择对各评价指标作加权处理 (如对关心部分的奖励及对误报和漏报等的惩罚等), 因为这样处理是依赖于用户或者说是基于用户的, 有着随意性。用户目的不同, 评价自然也会不同。在该领域的纯研究中, 通常不会考虑经济、实际应用时的特殊性, 因为需要有用户意向, 这是依用户而变的, 无法在研究中绝对化。对不同的用户, 成功和不成功的预测会产生不同的经济影响, 故而评价过程中应该对不同的用户给预测施予不同的权重。而对于科学目的, 对成功和失败的等权考虑是普遍接受和采用的 (Woodcock 1976 [147])。如果考虑代价, 有些事件的漏报会产生很大的代价, 所以宁可误报; 然而, 误报多到一定程度也会有很大代价, 这里面有很多不确定的因素需要考虑, 无法绝对地估量。Stephenson 2000 [129] 指出, skill 和 value 是有区别的, skill 衡量的是预测和实际的一般关联, 而预测 value 则给出的是不同的信息, 它依赖于

预测用户的特定需求。

预测的目标有很多种，比如概率预测、确定性的预测、值的预测。我们的预测是确定预测里的二类预测。我们探讨了耀斑有无预测的 2×2 的问题，虽然可以通过建立多组预测模型来间接给出耀斑的级别和发生时间，但我们仍可沿此思路继续探讨多类分类问题。虽然多类分类显然会更复杂一些，但能更直接地通过一组模型来给出级别预测， $k \times k$ 的问题也早已纳入相关研究人员的视野。直接的多类预测、以及更直接的值的预测即回归问题（实际上分类也是根据值进行的，将连续量离散化）将在未来考虑。

表 4.6: C1.0 级以上耀斑的预测结果评价

MC1-48h ($w_{10} = 1.0$, $w_{11}=2.5$, $w_{step} = 0.5$, $k = 10$) $N_0/N = 0.578$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	16.0	2.0	0.478	0.729 \pm 0.011	0.596 \pm 0.009	0.763 \pm 0.010	0.488 \pm 0.010	0.655 \pm 0.009	0.366 \pm 0.019	0.358 \pm 0.018	0.218 \pm 0.014	0.692 \pm 0.010	0.676 \pm 0.010
V08	1.0	64.0	1.5	0.486	0.619 \pm 0.014	0.664 \pm 0.011	0.735 \pm 0.008	0.472 \pm 0.012	0.640 \pm 0.011	0.389 \pm 0.018	0.399 \pm 0.018	0.246 \pm 0.014	0.671 \pm 0.010	0.706 \pm 0.009
L05	1.0	16.0	2.0	0.556	0.681 \pm 0.014	0.707 \pm 0.009	0.773 \pm 0.008	0.531 \pm 0.012	0.693 \pm 0.010	0.474 \pm 0.016	0.480 \pm 0.016	0.314 \pm 0.014	0.719 \pm 0.009	0.746 \pm 0.008
A10	1.0	8.0	2.0	0.557	0.706 \pm 0.010	0.691 \pm 0.012	0.781 \pm 0.008	0.537 \pm 0.013	0.698 \pm 0.011	0.474 \pm 0.020	0.472 \pm 0.020	0.312 \pm 0.018	0.725 \pm 0.010	0.742 \pm 0.010
A12	1.0	8.0	2.0	0.558	0.707 \pm 0.011	0.690 \pm 0.013	0.782 \pm 0.008	0.538 \pm 0.013	0.698 \pm 0.011	0.474 \pm 0.020	0.472 \pm 0.020	0.312 \pm 0.018	0.725 \pm 0.010	0.742 \pm 0.010
MC1-24h ($w_{10} = 1.5$, $w_{11}=3.5$, $w_{step} = 0.5$, $k = 10$) $N_0/N = 0.679$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	32.64	2.5	0.440	0.644 \pm 0.015	0.525 \pm 0.012	0.812 \pm 0.008	0.407 \pm 0.012	0.578 \pm 0.012	0.368 \pm 0.021	0.337 \pm 0.019	0.211 \pm 0.014	0.650 \pm 0.011	0.698 \pm 0.010
V08	1.0	16.0	2.0	0.446	0.567 \pm 0.013	0.581 \pm 0.017	0.797 \pm 0.006	0.402 \pm 0.012	0.572 \pm 0.012	0.370 \pm 0.018	0.378 \pm 0.022	0.231 \pm 0.015	0.640 \pm 0.010	0.728 \pm 0.010
L05	1.0	1.0	2.5	0.520	0.627 \pm 0.019	0.643 \pm 0.013	0.826 \pm 0.008	0.465 \pm 0.015	0.634 \pm 0.014	0.462 \pm 0.020	0.469 \pm 0.019	0.304 \pm 0.017	0.693 \pm 0.012	0.768 \pm 0.008
A10	1.0	4.0	2.5	0.532	0.676 \pm 0.020	0.622 \pm 0.013	0.841 \pm 0.008	0.479 \pm 0.015	0.647 \pm 0.014	0.481 \pm 0.022	0.462 \pm 0.020	0.309 \pm 0.018	0.706 \pm 0.013	0.763 \pm 0.009
A12	1.0	8.0	2.5	0.530	0.677 \pm 0.019	0.617 \pm 0.013	0.840 \pm 0.008	0.478 \pm 0.016	0.645 \pm 0.014	0.478 \pm 0.022	0.458 \pm 0.020	0.306 \pm 0.018	0.705 \pm 0.013	0.761 \pm 0.010
MC1-12h ($w_{10} = 2.5$, $w_{11}=4.5$, $w_{step} = 0.5$, $k = 10$) $N_0/N = 0.781$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	16.0	3.0	0.410	0.558 \pm 0.030	0.467 \pm 0.020	0.870 \pm 0.008	0.341 \pm 0.018	0.506 \pm 0.021	0.379 \pm 0.030	0.337 \pm 0.026	0.217 \pm 0.019	0.608 \pm 0.019	0.763 \pm 0.010
V08	1.0	32.0	2.5	0.417	0.505 \pm 0.024	0.521 \pm 0.020	0.863 \pm 0.006	0.343 \pm 0.015	0.510 \pm 0.017	0.373 \pm 0.022	0.384 \pm 0.023	0.233 \pm 0.015	0.608 \pm 0.015	0.788 \pm 0.008
L05	1.0	16.0	4.5	0.473	0.689 \pm 0.033	0.479 \pm 0.012	0.902 \pm 0.009	0.394 \pm 0.017	0.563 \pm 0.017	0.479 \pm 0.030	0.381 \pm 0.019	0.412 \pm 0.022	0.666 \pm 0.017	0.768 \pm 0.008
A10	1.0	4.0	3.5	0.476	0.636 \pm 0.033	0.513 \pm 0.012	0.892 \pm 0.008	0.395 \pm 0.016	0.564 \pm 0.017	0.466 \pm 0.017	0.405 \pm 0.017	0.427 \pm 0.020	0.661 \pm 0.016	0.788 \pm 0.007
A12	1.0	32.0	3.5	0.480	0.653 \pm 0.028	0.508 \pm 0.010	0.895 \pm 0.007	0.399 \pm 0.013	0.569 \pm 0.014	0.475 \pm 0.024	0.404 \pm 0.014	0.275 \pm 0.013	0.666 \pm 0.013	0.786 \pm 0.006
MC1-06h ($w_{10} = 3.5$, $w_{11}=5.5$, $w_{step} = 0.5$, $k = 10$) $N_0/N = 0.858$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	16.0	5.0	0.355	0.563 \pm 0.031	0.341 \pm 0.015	0.919 \pm 0.005	0.270 \pm 0.016	0.423 \pm 0.019	0.382 \pm 0.031	0.260 \pm 0.020	0.300 \pm 0.023	0.178 \pm 0.016	0.559 \pm 0.019
V08	1.0	8.0	4.5	0.359	0.521 \pm 0.032	0.366 \pm 0.018	0.915 \pm 0.005	0.275 \pm 0.017	0.428 \pm 0.022	0.371 \pm 0.032	0.281 \pm 0.023	0.314 \pm 0.025	0.189 \pm 0.018	0.557 \pm 0.020
L05	1.0	0.125	5.0	0.404	0.528 \pm 0.025	0.436 \pm 0.021	0.919 \pm 0.004	0.311 \pm 0.014	0.473 \pm 0.016	0.411 \pm 0.022	0.354 \pm 0.022	0.375 \pm 0.020	0.233 \pm 0.015	0.593 \pm 0.015
A10	1.0	0.125	5.5	0.410	0.582 \pm 0.037	0.407 \pm 0.022	0.926 \pm 0.006	0.314 \pm 0.020	0.475 \pm 0.023	0.439 \pm 0.036	0.333 \pm 0.025	0.370 \pm 0.027	0.230 \pm 0.021	0.600 \pm 0.021
A12	1.0	0.25	5.0	0.412	0.560 \pm 0.026	0.423 \pm 0.022	0.923 \pm 0.004	0.317 \pm 0.017	0.479 \pm 0.020	0.430 \pm 0.026	0.346 \pm 0.025	0.378 \pm 0.024	0.235 \pm 0.018	0.600 \pm 0.017

表 4.7: C5.0 级以上耀斑的预测结果评价

MC5-48h ($w_{10} = 2.5, w_{11}=4.5, w_{step} = 0.5, k = 10$) $N_0/N = 0.803$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	4.0	3.5	0.431	0.606 \pm 0.020	0.454 \pm 0.018	0.895 \pm 0.005	0.352 \pm 0.017	0.519 \pm 0.017	0.427 \pm 0.024	0.349 \pm 0.022	0.379 \pm 0.023	0.236 \pm 0.018	0.626 \pm 0.015	0.778 \pm 0.010
V08	1.0	32.0	3.0	0.444	0.576 \pm 0.024	0.492 \pm 0.023	0.892 \pm 0.006	0.362 \pm 0.021	0.529 \pm 0.021	0.428 \pm 0.028	0.384 \pm 0.028	0.402 \pm 0.028	0.255 \pm 0.023	0.631 \pm 0.018	0.798 \pm 0.011
L05	1.0	0.5	4.5	0.500	0.684 \pm 0.021	0.504 \pm 0.019	0.915 \pm 0.005	0.409 \pm 0.018	0.579 \pm 0.018	0.516 \pm 0.025	0.419 \pm 0.023	0.455 \pm 0.024	0.297 \pm 0.021	0.680 \pm 0.015	0.804 \pm 0.011
A10	1.0	1.0	4.0	0.506	0.688 \pm 0.019	0.510 \pm 0.017	0.917 \pm 0.004	0.415 \pm 0.016	0.585 \pm 0.016	0.524 \pm 0.021	0.427 \pm 0.021	0.463 \pm 0.021	0.304 \pm 0.018	0.685 \pm 0.013	0.807 \pm 0.009
A12	1.0	64.0	3.0	0.504	0.627 \pm 0.026	0.549 \pm 0.019	0.906 \pm 0.005	0.415 \pm 0.020	0.584 \pm 0.020	0.500 \pm 0.027	0.455 \pm 0.024	0.474 \pm 0.025	0.313 \pm 0.021	0.677 \pm 0.017	0.825 \pm 0.008
MC5-24h ($w_{10} = 2.5, w_{11}=5.5, w_{step} = 0.5, k = 10$) $N_0/N = 0.866$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	4.0	4.0	0.398	0.523 \pm 0.026	0.421 \pm 0.019	0.923 \pm 0.004	0.304 \pm 0.016	0.464 \pm 0.019	0.410 \pm 0.026	0.345 \pm 0.022	0.371 \pm 0.022	0.230 \pm 0.017	0.586 \pm 0.017	0.838 \pm 0.007
V08	1.0	0.125	4.5	0.398	0.550 \pm 0.023	0.404 \pm 0.017	0.926 \pm 0.004	0.303 \pm 0.013	0.463 \pm 0.016	0.422 \pm 0.022	0.330 \pm 0.019	0.365 \pm 0.019	0.225 \pm 0.014	0.589 \pm 0.014	0.829 \pm 0.007
L05	1.0	0.5	5.0	0.453	0.619 \pm 0.030	0.443 \pm 0.015	0.937 \pm 0.004	0.347 \pm 0.016	0.513 \pm 0.017	0.497 \pm 0.027	0.380 \pm 0.018	0.424 \pm 0.020	0.271 \pm 0.017	0.634 \pm 0.015	0.844 \pm 0.006
A10	1.0	1.0	3.0	0.466	0.519 \pm 0.021	0.550 \pm 0.026	0.926 \pm 0.003	0.363 \pm 0.017	0.531 \pm 0.019	0.451 \pm 0.021	0.477 \pm 0.028	0.460 \pm 0.022	0.301 \pm 0.018	0.640 \pm 0.016	0.877 \pm 0.006
A12	1.0	2.0	4.5	0.464	0.626 \pm 0.028	0.457 \pm 0.022	0.939 \pm 0.004	0.357 \pm 0.018	0.524 \pm 0.018	0.507 \pm 0.026	0.396 \pm 0.024	0.437 \pm 0.022	0.282 \pm 0.019	0.643 \pm 0.016	0.847 \pm 0.008
MC5-12h ($w_{10} = 3.5, w_{11}=8.5, w_{step} = 0.5, k = 10$) $N_0/N = 0.917$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	64.0	6.5	0.344	0.541 \pm 0.044	0.300 \pm 0.015	0.955 \pm 0.004	0.240 \pm 0.017	0.384 \pm 0.023	0.428 \pm 0.042	0.256 \pm 0.018	0.311 \pm 0.024	0.186 \pm 0.017	0.534 \pm 0.024	0.858 \pm 0.004
V08	1.0	32.0	6.0	0.351	0.508 \pm 0.044	0.326 \pm 0.018	0.953 \pm 0.004	0.248 \pm 0.019	0.394 \pm 0.025	0.413 \pm 0.042	0.279 \pm 0.021	0.327 \pm 0.027	0.198 \pm 0.019	0.537 \pm 0.025	0.872 \pm 0.005
L05	1.0	1.0	4.0	0.374	0.442 \pm 0.039	0.410 \pm 0.022	0.950 \pm 0.003	0.273 \pm 0.024	0.424 \pm 0.029	0.385 \pm 0.039	0.359 \pm 0.025	0.370 \pm 0.031	0.231 \pm 0.024	0.553 \pm 0.027	0.902 \pm 0.003
A10	1.0	0.125	7.0	0.390	0.590 \pm 0.038	0.339 \pm 0.014	0.960 \pm 0.003	0.275 \pm 0.016	0.429 \pm 0.020	0.486 \pm 0.036	0.300 \pm 0.016	0.362 \pm 0.021	0.223 \pm 0.016	0.575 \pm 0.019	0.871 \pm 0.004
A12	1.0	1.0	4.5	0.388	0.485 \pm 0.044	0.400 \pm 0.028	0.952 \pm 0.004	0.283 \pm 0.028	0.435 \pm 0.033	0.419 \pm 0.044	0.352 \pm 0.031	0.379 \pm 0.035	0.239 \pm 0.029	0.565 \pm 0.029	0.896 \pm 0.006
MC5-06h ($w_{10} = 5.0, w_{11}=11.0, w_{step} = 1.0, k = 10$) $N_0/N = 0.954$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	1.0	10.0	0.293	0.446 \pm 0.061	0.246 \pm 0.040	0.972 \pm 0.003	0.197 \pm 0.037	0.316 \pm 0.048	0.376 \pm 0.065	0.217 \pm 0.043	0.271 \pm 0.051	0.167 \pm 0.038	0.464 \pm 0.047	0.908 \pm 0.008
V08	1.0	0.125	10.0	0.298	0.426 \pm 0.060	0.265 \pm 0.044	0.971 \pm 0.003	0.203 \pm 0.035	0.324 \pm 0.049	0.365 \pm 0.064	0.236 \pm 0.047	0.282 \pm 0.052	0.174 \pm 0.036	0.465 \pm 0.050	0.915 \pm 0.007
L05	1.0	8,16,32,64	7.0	0.337	0.486 \pm 0.049	0.295 \pm 0.027	0.974 \pm 0.003	0.226 \pm 0.024	0.363 \pm 0.032	0.429 \pm 0.049	0.269 \pm 0.029	0.324 \pm 0.034	0.198 \pm 0.024	0.512 \pm 0.033	0.921 \pm 0.005
A10	1.0	1.0	10.0	0.336	0.585 \pm 0.055	0.253 \pm 0.024	0.978 \pm 0.003	0.216 \pm 0.022	0.351 \pm 0.031	0.500 \pm 0.055	0.231 \pm 0.026	0.306 \pm 0.033	0.184 \pm 0.023	0.519 \pm 0.034	0.900 \pm 0.006
A12	1.0	0.125	11.0	0.338	0.595 \pm 0.064	0.250 \pm 0.029	0.979 \pm 0.003	0.219 \pm 0.029	0.351 \pm 0.038	0.508 \pm 0.065	0.229 \pm 0.032	0.306 \pm 0.041	0.187 \pm 0.029	0.519 \pm 0.040	0.898 \pm 0.006

表 4.8: M1.0 级以上耀斑的预测结果评价

MM1-48h ($w_{10} = 3.5, w_{11}=7.0, w_{step} = 0.5, k = 10$) $N_0/N = 0.884$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	4.0	4.5	0.436	0.599 \pm 0.025	0.418 \pm 0.021	0.944 \pm 0.003	0.328 \pm 0.019	0.491 \pm 0.022	0.488 \pm 0.028	0.362 \pm 0.024	0.410 \pm 0.025	0.261 \pm 0.020	0.856 \pm 0.007
V08	1.0	32.0	4.0	0.443	0.591 \pm 0.028	0.436 \pm 0.019	0.944 \pm 0.003	0.335 \pm 0.018	0.499 \pm 0.020	0.489 \pm 0.028	0.379 \pm 0.021	0.422 \pm 0.023	0.270 \pm 0.019	0.862 \pm 0.007
L05	1.0	1.0	6.5	0.440	0.658 \pm 0.028	0.394 \pm 0.016	0.951 \pm 0.004	0.327 \pm 0.016	0.491 \pm 0.017	0.524 \pm 0.027	0.345 \pm 0.018	0.404 \pm 0.021	0.255 \pm 0.017	0.841 \pm 0.008
A10	1.0	0.25	4.5	0.463	0.599 \pm 0.030	0.461 \pm 0.028	0.945 \pm 0.004	0.353 \pm 0.024	0.518 \pm 0.026	0.504 \pm 0.031	0.407 \pm 0.030	0.445 \pm 0.030	0.291 \pm 0.025	0.870 \pm 0.008
A12	1.0	0.5	5.0	0.463	0.642 \pm 0.028	0.433 \pm 0.021	0.950 \pm 0.004	0.350 \pm 0.021	0.516 \pm 0.023	0.531 \pm 0.030	0.383 \pm 0.024	0.438 \pm 0.026	0.284 \pm 0.021	0.860 \pm 0.007
MM1-24h ($w_{10} = 3.0, w_{11}=9.0, w_{step} = 1.0, k = 10$) $N_0/N = 0.923$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	32.0	7.0	0.380	0.621 \pm 0.038	0.310 \pm 0.019	0.966 \pm 0.003	0.262 \pm 0.019	0.412 \pm 0.025	0.505 \pm 0.039	0.276 \pm 0.022	0.345 \pm 0.027	0.212 \pm 0.019	0.864 \pm 0.007
V08	1.0	32,64	6.0	0.381	0.567 \pm 0.039	0.332 \pm 0.021	0.962 \pm 0.003	0.268 \pm 0.022	0.419 \pm 0.027	0.472 \pm 0.041	0.294 \pm 0.024	0.356 \pm 0.030	0.220 \pm 0.022	0.879 \pm 0.005
L05	1.0	4.0	5.0	0.425	0.533 \pm 0.029	0.423 \pm 0.018	0.960 \pm 0.002	0.310 \pm 0.018	0.471 \pm 0.022	0.472 \pm 0.030	0.383 \pm 0.020	0.421 \pm 0.023	0.270 \pm 0.019	0.908 \pm 0.003
A10	1.0	8.0	5.0	0.425	0.544 \pm 0.021	0.415 \pm 0.016	0.961 \pm 0.002	0.309 \pm 0.014	0.470 \pm 0.017	0.480 \pm 0.021	0.376 \pm 0.017	0.419 \pm 0.018	0.267 \pm 0.014	0.906 \pm 0.003
A12	1.0	0.25	5.0	0.430	0.550 \pm 0.039	0.419 \pm 0.023	0.962 \pm 0.003	0.314 \pm 0.023	0.474 \pm 0.028	0.486 \pm 0.039	0.381 \pm 0.026	0.424 \pm 0.030	0.273 \pm 0.023	0.907 \pm 0.004
MM1-12h ($w_{10} = 5.0, w_{11}=12.0, w_{step} = 1.0, k = 10$) $N_0/N = 0.956$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	2.0	10.0	0.322	0.577 \pm 0.036	0.237 \pm 0.018	0.979 \pm 0.002	0.202 \pm 0.018	0.334 \pm 0.022	0.490 \pm 0.035	0.216 \pm 0.019	0.290 \pm 0.024	0.172 \pm 0.018	0.898 \pm 0.006
V08	1.0	0.25	11.0	0.326	0.598 \pm 0.036	0.237 \pm 0.015	0.980 \pm 0.002	0.204 \pm 0.014	0.337 \pm 0.019	0.507 \pm 0.033	0.218 \pm 0.015	0.293 \pm 0.020	0.173 \pm 0.014	0.896 \pm 0.007
L05	1.0	16,32,64	8.0	0.386	0.587 \pm 0.053	0.319 \pm 0.026	0.981 \pm 0.002	0.258 \pm 0.024	0.405 \pm 0.028	0.527 \pm 0.049	0.300 \pm 0.027	0.370 \pm 0.030	0.231 \pm 0.024	0.925 \pm 0.005
A10	1.0	1.0	5.0	0.378	0.472 \pm 0.048	0.381 \pm 0.027	0.976 \pm 0.002	0.261 \pm 0.025	0.408 \pm 0.030	0.435 \pm 0.046	0.356 \pm 0.028	0.379 \pm 0.031	0.238 \pm 0.025	0.941 \pm 0.005
A12	1.0	8.0	6.0	0.391	0.554 \pm 0.056	0.344 \pm 0.021	0.979 \pm 0.002	0.266 \pm 0.025	0.415 \pm 0.030	0.505 \pm 0.052	0.323 \pm 0.022	0.382 \pm 0.031	0.240 \pm 0.024	0.934 \pm 0.004
MM1-06h ($w_{10} = 9.0, w_{11}=15.0, w_{step} = 1.0, k = 10$) $N_0/N = 0.973$														
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC
V06	1.0	0.5	11.0	0.272	0.437 \pm 0.079	0.211 \pm 0.038	0.984 \pm 0.002	0.172 \pm 0.033	0.282 \pm 0.050	0.395 \pm 0.079	0.195 \pm 0.040	0.256 \pm 0.051	0.155 \pm 0.033	0.944 \pm 0.004
V08	1.0	1.0	10.0	0.290	0.437 \pm 0.071	0.236 \pm 0.040	0.984 \pm 0.002	0.188 \pm 0.034	0.304 \pm 0.051	0.399 \pm 0.072	0.220 \pm 0.042	0.280 \pm 0.052	0.172 \pm 0.034	0.948 \pm 0.005
L05	1.0	0.125	9.0	0.341	0.370 \pm 0.055	0.414 \pm 0.089	0.982 \pm 0.002	0.237 \pm 0.043	0.366 \pm 0.054	0.351 \pm 0.055	0.397 \pm 0.089	0.348 \pm 0.056	0.224 \pm 0.044	0.965 \pm 0.004
A10	1.0	0.25	9.0	0.304	0.457 \pm 0.066	0.247 \pm 0.032	0.985 \pm 0.002	0.195 \pm 0.031	0.317 \pm 0.042	0.419 \pm 0.066	0.232 \pm 0.034	0.293 \pm 0.044	0.179 \pm 0.031	0.948 \pm 0.004
A12	1.0	0.125	11.0	0.307	0.523 \pm 0.067	0.225 \pm 0.028	0.986 \pm 0.002	0.191 \pm 0.028	0.312 \pm 0.039	0.474 \pm 0.067	0.211 \pm 0.030	0.286 \pm 0.040	0.173 \pm 0.028	0.939 \pm 0.004

表 4.9: M5.0 级以上耀斑的预测结果评价

MM5-48h ($w_{10} = 8.0$, $w_{11}=14.0$, $w_{step} = 1.0$, $k = 10$) $N_0/N = 0.967$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	4.0	11.0	0.429	0.704 \pm 0.040	0.322 \pm 0.027	0.990 \pm 0.001	0.282 \pm 0.022	0.435 \pm 0.028	0.649 \pm 0.039	0.312 \pm 0.027	0.408 \pm 0.030	0.260 \pm 0.023	0.602 \pm 0.024	0.938 \pm 0.006
V08	1.0	0.25	11.0	0.426	0.677 \pm 0.036	0.327 \pm 0.033	0.989 \pm 0.001	0.284 \pm 0.029	0.435 \pm 0.035	0.624 \pm 0.038	0.316 \pm 0.034	0.408 \pm 0.037	0.263 \pm 0.029	0.596 \pm 0.029	0.939 \pm 0.007
L05	1.0	0.5	12.0	0.340	0.648 \pm 0.064	0.232 \pm 0.029	0.987 \pm 0.002	0.209 \pm 0.028	0.338 \pm 0.037	0.572 \pm 0.065	0.219 \pm 0.030	0.304 \pm 0.039	0.185 \pm 0.028	0.521 \pm 0.037	0.915 \pm 0.007
A10	1.0	0.25	8.0	0.401	0.577 \pm 0.057	0.341 \pm 0.034	0.985 \pm 0.002	0.270 \pm 0.029	0.418 \pm 0.036	0.536 \pm 0.056	0.326 \pm 0.035	0.393 \pm 0.037	0.251 \pm 0.029	0.569 \pm 0.033	0.947 \pm 0.005
A12	1.0	0.125	10.0	0.407	0.634 \pm 0.048	0.319 \pm 0.020	0.987 \pm 0.002	0.268 \pm 0.020	0.419 \pm 0.025	0.587 \pm 0.047	0.306 \pm 0.021	0.393 \pm 0.026	0.247 \pm 0.020	0.580 \pm 0.024	0.942 \pm 0.004
MM5-24h ($w_{10} = 5.0$, $w_{11}=23.0$, $w_{step} = 1.0$, $k = 5$) $N_0/N = 0.982$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.25	12.0	0.324	0.479 \pm 0.105	0.260 \pm 0.044	0.991 \pm 0.002	0.208 \pm 0.044	0.335 \pm 0.063	0.455 \pm 0.103	0.251 \pm 0.046	0.320 \pm 0.063	0.197 \pm 0.044	0.487 \pm 0.068	0.968 \pm 0.002
V08	1.0	0.125	15.0	0.314	0.529 \pm 0.118	0.227 \pm 0.037	0.991 \pm 0.002	0.193 \pm 0.039	0.316 \pm 0.057	0.498 \pm 0.116	0.219 \pm 0.039	0.299 \pm 0.058	0.181 \pm 0.039	0.481 \pm 0.068	0.961 \pm 0.002
L05	1.0	1.0	21.0	0.293	0.657 \pm 0.101	0.174 \pm 0.032	0.994 \pm 0.002	0.162 \pm 0.031	0.274 \pm 0.048	0.598 \pm 0.102	0.168 \pm 0.034	0.253 \pm 0.049	0.148 \pm 0.031	0.477 \pm 0.056	0.936 \pm 0.006
A10	1.0	8,16,32,64	5.0	0.326	0.329 \pm 0.114	0.428 \pm 0.086	0.988 \pm 0.002	0.224 \pm 0.068	0.348 \pm 0.086	0.321 \pm 0.114	0.416 \pm 0.087	0.338 \pm 0.086	0.218 \pm 0.067	0.493 \pm 0.081	0.980 \pm 0.002
A12	1.0	16.0	5.0	0.330	0.329 \pm 0.114	0.434 \pm 0.094	0.988 \pm 0.002	0.231 \pm 0.075	0.353 \pm 0.093	0.321 \pm 0.114	0.422 \pm 0.095	0.343 \pm 0.094	0.224 \pm 0.075	0.495 \pm 0.087	0.980 \pm 0.003
MM5-12h ($w_{10} = 9.0$, $w_{11}=36.0$, $w_{step} = 1.0$, $k = 5$) $N_0/N = 0.990$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.125	11.0	0.357	0.550 \pm 0.059	0.272 \pm 0.036	0.995 \pm 0.001	0.224 \pm 0.032	0.362 \pm 0.042	0.535 \pm 0.060	0.268 \pm 0.036	0.353 \pm 0.043	0.218 \pm 0.032	0.526 \pm 0.039	0.980 \pm 0.002
V08	1.0	0.125	9.0	0.350	0.500 \pm 0.032	0.287 \pm 0.042	0.995 \pm 0.000	0.223 \pm 0.031	0.361 \pm 0.041	0.487 \pm 0.033	0.282 \pm 0.042	0.353 \pm 0.041	0.217 \pm 0.032	0.518 \pm 0.034	0.982 \pm 0.002
L05	1.0	0.25	11.0	0.333	0.410 \pm 0.078	0.319 \pm 0.042	0.994 \pm 0.001	0.220 \pm 0.040	0.354 \pm 0.055	0.401 \pm 0.078	0.313 \pm 0.043	0.346 \pm 0.055	0.215 \pm 0.040	0.498 \pm 0.054	0.985 \pm 0.002
A10	1.0	1.0	10.0	0.321	0.460 \pm 0.104	0.264 \pm 0.068	0.994 \pm 0.001	0.206 \pm 0.054	0.329 \pm 0.074	0.447 \pm 0.105	0.258 \pm 0.068	0.320 \pm 0.075	0.200 \pm 0.055	0.479 \pm 0.074	0.981 \pm 0.002
A12	1.0	16.0	10.0	0.328	0.460 \pm 0.104	0.275 \pm 0.067	0.994 \pm 0.001	0.213 \pm 0.056	0.338 \pm 0.075	0.447 \pm 0.105	0.269 \pm 0.067	0.329 \pm 0.076	0.207 \pm 0.056	0.486 \pm 0.074	0.982 \pm 0.002
MM5-06h ($w_{10} = 10.0$, $w_{11}=60.0$, $w_{step} = 1.0$, $k = 5$) $N_0/N = 0.994$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.25	10,11	0.362	0.667 \pm 0.139	0.236 \pm 0.056	0.998 \pm 0.001	0.218 \pm 0.060	0.343 \pm 0.076	0.654 \pm 0.140	0.234 \pm 0.056	0.338 \pm 0.077	0.214 \pm 0.060	0.525 \pm 0.075	0.986 \pm 0.002
V08	1.0	0.25	10,12,13	0.362	0.667 \pm 0.139	0.236 \pm 0.056	0.998 \pm 0.001	0.218 \pm 0.060	0.343 \pm 0.076	0.654 \pm 0.140	0.234 \pm 0.056	0.338 \pm 0.077	0.214 \pm 0.060	0.525 \pm 0.075	0.986 \pm 0.002
L05	1.0	0.125	17.0	0.306	0.367 \pm 0.186	0.273 \pm 0.186	0.996 \pm 0.001	0.257 \pm 0.188	0.300 \pm 0.184	0.360 \pm 0.186	0.270 \pm 0.187	0.295 \pm 0.185	0.255 \pm 0.189	0.368 \pm 0.184	0.990 \pm 0.003
A10	1.0	0.125	26,27	0.363	0.833 \pm 0.105	0.204 \pm 0.040	0.999 \pm 0.001	0.199 \pm 0.042	0.324 \pm 0.058	0.814 \pm 0.106	0.203 \pm 0.041	0.318 \pm 0.058	0.195 \pm 0.042	0.544 \pm 0.056	0.980 \pm 0.002
A12	1.0	0.5	14.0	0.347	0.667 \pm 0.139	0.220 \pm 0.038	0.998 \pm 0.001	0.202 \pm 0.042	0.327 \pm 0.058	0.654 \pm 0.139	0.218 \pm 0.039	0.322 \pm 0.059	0.198 \pm 0.042	0.517 \pm 0.065	0.985 \pm 0.002

表 4.10: X1.0 级以上耀斑的预测结果评价

MX1-48h ($w_{10} = 5.0, w_{11}=20.0, w_{step} = 1.0, k = 5) N_0/N = 0.981$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.125	10.0	0.413	0.567±0.080	0.362±0.107	0.991±0.001	0.285±0.076	0.422±0.089	0.542±0.081	0.353±0.108	0.408±0.091	0.274±0.077	0.569±0.077	0.967±0.006
V08	1.0	0.125	7.0	0.420	0.522±0.070	0.405±0.096	0.991±0.001	0.294±0.058	0.441±0.074	0.504±0.072	0.396±0.096	0.428±0.076	0.284±0.059	0.579±0.068	0.973±0.005
L05	1.0	2.4	10.0	0.356	0.519±0.047	0.293±0.038	0.990±0.001	0.229±0.030	0.369±0.041	0.493±0.047	0.283±0.038	0.353±0.042	0.217±0.031	0.527±0.035	0.965±0.005
A10	1.0	64.0	6.0	0.441	0.525±0.085	0.468±0.143	0.991±0.002	0.311±0.071	0.457±0.081	0.508±0.084	0.458±0.143	0.445±0.084	0.301±0.072	0.598±0.075	0.974±0.005
A12	1.0	4.0	6.0	0.453	0.522±0.070	0.487±0.141	0.991±0.001	0.326±0.072	0.474±0.082	0.507±0.070	0.477±0.141	0.462±0.084	0.316±0.073	0.609±0.074	0.976±0.005
MX1-24h ($w_{10} = 5.0, w_{11}=38.0, w_{step} = 1.0, k = 5) N_0/N = 0.982$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.125	8.0	0.400	0.520±0.150	0.346±0.100	0.994±0.002	0.289±0.090	0.415±0.119	0.509±0.151	0.341±0.101	0.407±0.121	0.283±0.090	0.519±0.139	0.983±0.003
V08	1.0	64.0	7.0	0.412	0.520±0.150	0.364±0.119	0.994±0.002	0.305±0.104	0.426±0.131	0.509±0.152	0.359±0.120	0.418±0.133	0.299±0.105	0.525±0.146	0.983±0.004
L05	1.0	0.125	15.0	0.318	0.440±0.117	0.262±0.077	0.993±0.001	0.212±0.071	0.328±0.093	0.425±0.119	0.255±0.079	0.318±0.095	0.205±0.072	0.468±0.091	0.978±0.003
A10	1.0	1.0	7.0	0.387	0.480±0.102	0.379±0.092	0.994±0.001	0.264±0.060	0.403±0.076	0.469±0.102	0.373±0.092	0.395±0.077	0.258±0.060	0.546±0.073	0.983±0.003
A12	1.0	0.5	8.9	0.385	0.480±0.102	0.375±0.094	0.994±0.001	0.262±0.061	0.401±0.078	0.469±0.102	0.369±0.095	0.392±0.079	0.256±0.061	0.544±0.076	0.983±0.003
MX1-12h ($w_{10} = 10.0, w_{11}=50.0, w_{step} = 1.0, k = 5) N_0/N = 0.990$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.125,0.25	10.0	0.305	0.467±0.062	0.239±0.027	0.996±0.000	0.185±0.020	0.310±0.029	0.455±0.061	0.235±0.027	0.303±0.029	0.180±0.020	0.475±0.031	0.984±0.002
V08	1.0	0.125	10.0	0.319	0.467±0.062	0.257±0.027	0.996±0.000	0.198±0.024	0.328±0.034	0.456±0.062	0.253±0.027	0.322±0.034	0.194±0.024	0.488±0.036	0.985±0.001
L05	1.0	8.0	11.0	0.369	0.533±0.062	0.307±0.049	0.996±0.001	0.233±0.027	0.375±0.035	0.523±0.061	0.303±0.048	0.369±0.035	0.229±0.027	0.537±0.031	0.986±0.001
A10	1.0	0.125	10.0	0.335	0.467±0.062	0.289±0.064	0.996±0.000	0.213±0.041	0.343±0.056	0.456±0.063	0.285±0.064	0.337±0.057	0.208±0.041	0.499±0.052	0.985±0.002
A12	1.0	64.0	10.0	0.349	0.533±0.062	0.278±0.023	0.996±0.001	0.214±0.010	0.353±0.013	0.522±0.061	0.274±0.023	0.346±0.013	0.210±0.010	0.521±0.014	0.985±0.002
MX1-06h ($w_{10} = 10.0, w_{11}=40.0, w_{step} = 1.0, k = 5) N_0/N = 0.996$															
γ	C	w_1	MInd	POD	POH	FOCN	CSI	F_1	TSS	CSS	HSS	GSS	GM3	ACC	
V06	1.0	0.25	15.0	0.315	0.700±0.200	0.179±0.059	0.999±0.001	0.178±0.060	0.283±0.091	0.688±0.200	0.178±0.060	0.279±0.090	0.175±0.060	0.457±0.128	0.987±0.003
V08	1.0	64.0	18.0	0.314	0.700±0.200	0.179±0.057	0.999±0.001	0.177±0.058	0.283±0.088	0.688±0.199	0.177±0.058	0.279±0.088	0.174±0.058	0.458±0.126	0.987±0.003
L05	1.0	0.5	38.0	0.289	0.800±0.200	0.145±0.041	0.999±0.001	0.145±0.041	0.243±0.067	0.781±0.198	0.144±0.042	0.238±0.068	0.141±0.042	0.449±0.114	0.980±0.004
A10	1.0	2.0	10.0	0.316	0.700±0.200	0.180±0.051	0.999±0.001	0.177±0.052	0.286±0.081	0.689±0.199	0.179±0.052	0.282±0.081	0.174±0.052	0.462±0.123	0.988±0.002
A12	1.0	0.5	13,14	0.303	0.700±0.200	0.169±0.055	0.999±0.001	0.167±0.056	0.270±0.084	0.688±0.199	0.167±0.056	0.265±0.084	0.164±0.056	0.449±0.123	0.987±0.003

第五章 总结与展望

5.1 总结

本文通过处理怀柔太阳观测站太阳磁场望远镜共 21 年、约跨两个完整的太阳活动周 (第 22 和第 23 太阳周) 的活动区光球矢量磁场数据 (共涉及 1106 个活动区的 2173 张磁图, 样本的选取合理、有代表性, 详见第 3.1.2 节), 统计计算了反映磁场非势性和复杂性的八个参量各自在整个活动区的平均强度随太阳活动周的变化, 考查了这些非势性的平均表征与耀斑之间的统计关系, 还利用支持向量分类器在这些数据的基础上建立预测模型、考查了活动区磁非势性作为耀斑预测因子的预测性能。

得到的主要结论有: (1) 活动区的两个磁剪切角均值 $\overline{\Delta\phi}$ 和 $\overline{\Delta\psi}$, 纵向电流密度绝对值的均值 $\overline{|J_z|}$, 平均无力因子的绝对值 $|\alpha_{av}|$, 以及有效距离 d_E 随整体太阳活动水平变化不显著。然而, 这些参量仍能反映出活动区在太阳活动峰年较低年具有更强的非势性。(2) 电流螺度密度绝对值的均值 $\overline{|h_c|}$ 、自由磁能密度均值 $\overline{\rho_{free}}$ 、纵场加权后的有效距离 d_{Em} 与平均黑子数显现出高的正相关, 且三者之间也有着相对紧密的关系。三者与年均黑子数的线性相关系数都高于 0.59。它们可以和表征太阳活动水平。(3) 经纵场加权后的有效距离 d_{Em} 由于加入了磁场强度信息而能更好地反映活动区的磁活动特征。(4) 仅从纵向磁场得到的磁参量在预测较大级别耀斑时不如矢量场得到的参量敏感。将横场信息融入到耀斑预测因子里是有益的。

工作的特点是: 由于怀柔观测站的活动区数据所使用的怀柔编号并未形成与 NOAA 活动区编号相对应的完整记录, 使得研究活动区矢量磁场与耀斑关系的统计工作难以开展, 这也是相关研究者所遗憾的, 我们工作中所整理的数据和记录将会有益于将来更多统计和比较工作的开展; 基于长期稳定的矢量磁场观测数据预测耀斑; 利用同一批数据进行多种实验, 由于数据样本覆盖全面、考查的内容能够比较全面完整, 这在现有的相似工作中是不常见的; 利用多个指标评估和优化模型避免了单个评价指标的误导, 分开训练和测试集、使用 k -折交叉确认使模型具有很好的泛化能力 (其实这两点在预测建模中是不应该被忽视的); 基于磁场导出参量的预测可以直接应用到实际的观测中进行实

时预测，为观测提供很好的参考，具有实用价值。

值得一提的是，第二部分工作的审稿人说到，“这篇文章为读者介绍了预测耀斑或者任何能表示为两类的现象（事件发生或事件不发生）的一种精良的技术。文章应该被发表，它将教育耀斑预报工作者应用合适的技术来评估数据。文中提到的有关‘Finley 事件’的参考文献也非常有意思。”，充分肯定了我们的工作。

5.2 展望

质子耀斑是指产生质子事件的耀斑爆发。如果耀斑引起高能质子事件，质子的探测会晚于电磁辐射的探测。因此对于质子事件的预测，大多在发生了耀斑之后再来判断是否会发生质子事件。或者，将整个耀斑预测分为两部分，先判断是否发生耀斑，如果发生耀斑再判断是否为质子耀斑。在历史上发生的耀斑中，质子耀斑所占比重并不大，通常较大级别的耀斑容易产生质子事件，那么对应这些质子耀斑的光球磁场位型和无质子事件的爆发是否有着显著的区别，统计结果是怎样的，我们可以基于已有数据继续进行探讨。

活动区上方日冕的温度与活动区光球非势性的关系如何，也是研究磁能加热日冕的一个值得探讨的课题。

SDO 上天仅 3 年多，由于统计预测对历史数据的依赖，它对活动周的覆盖尚不够、可能还不足以用来进行成功率可以接受的统计预测，能否通过虽不同但类似的仪器之间做归一处理来利用怀柔历史数据，这或许也是值得思考和尝试的一个课题。不同仪器测量数据的比较已有一些同事已经或正在开展，这是一些很好的基础工作。

对于爆发事件有无预测即二类分类还可以继续研究和寻找技术途径，另外多类预测和回归预测恐怕是太阳物理和空间科学研究者更感兴趣的。继续关注这方面的问题，并关注相应评价机制的使用，也是本人的一个兴趣点所在。

参考文献

- [1] ABRAMENKO, V. I. Relationship between Magnetic Power Spectrum and Flare Productivity in Solar Active Regions. *Astrophys. J.* 629 (Aug. 2005), 1141–1149.
- [2] AHMED, O. W., QAHWAJI, R., COLAK, T., HIGGINS, P. A., GALLAGHER, P. T., AND BLOOMFIELD, D. S. Solar Flare Prediction Using Advanced Feature Extraction, Machine Learning, and Feature Selection. *Solar Phys.* (Nov. 2011), 404.
- [3] AI, G. X., AND HU, Y. F. Propose for a solar magnetic field telescope and its working theorem. *Publ. Beijing Astron. Obs.* 8 (1986), 1–10.
- [4] AI, G. X., LI, W., AND ZHANG, H. Q. FeI λ 5324.19 Å line forms in the solar magnetic field and the theoretical calibration of the solar magnetic field telescope. *Acta Astron. Sinica* 23 (Mar. 1982), 39–48.
- [5] AI, G. X., LI, W., AND ZHANG, H. Q. Formation of the FeI 5324.19 line in the sun and theoretical calibration of solar magnetic telescope. *Chinese Astron. Astrophys.* 6 (June 1982), 129–136.
- [6] AL-OMARI, M., QAHWAJI, R., COLAK, T., AND IPSON, S. Machine Learning-Based Investigation of the Associations between CMEs and Filaments. *Solar Phys.* 262 (Apr. 2010), 511–539.
- [7] ALIPOUR, N., SAFARI, H., AND INNES, D. E. An Automatic Detection Method for Extreme-ultraviolet Dimmings Associated with Small-scale Eruption. *Astrophys. J.* 746 (Feb. 2012), 12.
- [8] ANTALOVA, A. Daily Soft X-Ray Flare Index (1969=1972). *Contrib. Astron. Obs. Skalnaté Pleso* 26 (Oct. 1996), 98–120.

- [9] ASCHWANDEN, M. J. Image Processing Techniques and Feature Recognition in Solar Physics. *Solar Phys.*262 (Apr. 2010), 235–275.
- [10] BABCOCK, H. D. The Sun’s Polar Magnetic Field. *Astrophys. J.*130 (Sept. 1959), 364.
- [11] BABCOCK, H. D., AND LIVINGSTON, W. C. Changes in the Sun’s Polar Magnetic Field. *Science* 127 (May 1958), 1058.
- [12] BABCOCK, H. W., AND BABCOCK, H. D. Mapping the Magnetic Fields of the Sun. *Publ. Astron. Soc. Pac.*64 (Dec. 1952), 282.
- [13] BAO, S. D., PEVTSOV, A. A., WANG, T. J., AND ZHANG, H. Q. Helicity Computation Using Observations From two Different Polarimetric Instruments. *Solar Phys.*195 (July 2000), 75–87.
- [14] BAO, S. D., AND ZHANG, H. Q. Patterns of Current Helicity for the Twenty-second Solar Cycle. *Astrophys. J. Lett.*496 (Mar. 1998), L43.
- [15] BAO, S. D., ZHANG, H. Q., AI, G. X., AND ZHANG, M. A survey of flares and current helicity in active regions. *A&AS* 139 (Oct. 1999), 311–320.
- [16] BEAUMONT, C. N., WILLIAMS, J. P., AND GOODMAN, A. A. Classifying Structures in the Interstellar Medium with Support Vector Machines: The G16.05-0.57 Supernova Remnant. *Astrophys. J.*741 (Nov. 2011), 14.
- [17] BEWSHER, D., HARRISON, R. A., AND BROWN, D. S. The relationship between EUV dimming and coronal mass ejections. I. Statistical study and probability model. *Astron. Astrophys.*478 (Feb. 2008), 897–906.
- [18] BLOOMFIELD, D. S., HIGGINS, P. A., MCATEER, R. T. J., AND GALLAGHER, P. T. Toward Reliable Benchmarking of Solar Flare Forecasting Methods. *Astrophys. J. Lett.*747 (Mar. 2012), L41.

- [19] BORRERO, J. M., TOMCZYK, S., KUBO, M., SOCAS-NAVARRO, H., SCHOU, J., COUVIDAT, S., AND BOGART, R. VFISV: Very Fast Inversion of the Stokes Vector for the Helioseismic and Magnetic Imager. *Solar Phys.*273 (Oct. 2011), 267–293.
- [20] BOSER, B. E., GUYON, I. M., AND VAPNIK, V. N. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (New York, NY, USA, 1992), COLT '92, ACM, pp. 144–152.
- [21] BOURSIER, Y., LAMY, P., LLEBARIA, A., GOUDAIL, F., AND ROBELUS, S. The ARTEMIS Catalog of LASCO Coronal Mass Ejections. Automatic Recognition of Transient Events and Marseille Inventory from Synoptic maps. *Solar Phys.*257 (June 2009), 125–147.
- [22] BYRNE, J. P., MORGAN, H., HABBAL, S. R., AND GALLAGHER, P. T. Automatic Detection and Tracking of Coronal Mass Ejections. II. Multiscale Filtering of Coronagraph Images. *Astrophys. J.*752 (June 2012), 145.
- [23] CALVO, R. A., CECCATO, H. A., AND PIACENTINI, R. D. Neural network prediction of solar activity. *Astrophys. J.*444 (May 1995), 916–921.
- [24] CHANG, C.-C., AND LIN, C.-J. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:1–27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [25] CHINCHOR, N. Muc-4 evaluation metrics. In *Proceedings of the 4th conference on Message understanding* (Stroudsburg, PA, USA, 1992), MUC4 '92, Association for Computational Linguistics, pp. 22–29.
- [26] CHUMAK, O. V., AND CHUMAK, Z. N. A stochastic method of describing the evolution of solar active regions. *Kinematika i Fizika Nebesnykh Tel* 3 (June 1987), 7–13. (in Russian).

- [27] CHUMAK, O. V., ZHANG, H. Q., AND GUO, J. Integral properties of the magnetic fields of solar active regions under quiet and flare activity conditions. *Astron. Astrophys. Trans.* 23 (June 2004), 525–531.
- [28] CLAYTON, H. H. Verification of weather forecasts. *American Meteorological Journal* 6 (1889), 211–219.
- [29] CLAYTON, H. H. Rating weather forecasts. *Bull. Amer. Meteor. Soc.* 15 (1934), 279–283.
- [30] COLAK, T., AND QAHWAJI, R. Automated Solar Activity Prediction: A hybrid computer platform using machine learning and solar imaging for automated prediction of solar flares. *Space Weather* 7 (June 2009), 6001.
- [31] CORTES, C., AND VAPNIK, V. Support-vector networks. *Mach. Learn.* 20 (September 1995), 273–297.
- [32] CUI, Y. M., LI, R., WANG, H. N., AND HE, H. Correlation between Solar Flare Productivity and Photospheric Magnetic Field Properties II. Magnetic Gradient and Magnetic Shear. *Solar Phys.* 242 (May 2007), 1–8.
- [33] CUI, Y. M., LI, R., ZHANG, L. Y., HE, Y. L., AND WANG, H. N. Correlation Between Solar Flare Productivity and Photospheric Magnetic Field Properties. 1. Maximum Horizontal Gradient, Length of Neutral Line, Number of Singular Points. *Solar Phys.* 237 (Aug. 2006), 45–59.
- [34] CUI, Y. M., AND WANG, H. N. Correlation between solar flare productivity and photospheric vector magnetic fields. *Adv. Space Res.* 42 (Nov. 2008), 1475–1479.
- [35] DENG, N. Y., AND TIAN, Y. J. 数据挖掘中的新方法 — 支持向量机. 北京: 科学出版社, 2004.
- [36] DENG, Y. Y., WANG, J. X., YAN, Y. H., AND ZHANG, J. Evolution of Magnetic Nonpotentiality in NOAA AR 9077. *Solar Phys.* 204 (Dec. 2001), 11–26.

-
- [37] DONALDSON, R. J., DYER, R. M., AND KRAUS, M. J. An Objective Evaluator of Techniques for Predicting Severe Weather Events. *Amer. Meteor. Soc., Boston* (1975), 321–326.
- [38] DOOLITTLE, M. H. Association Ratios. *Bulletin of the Philosophical Society, Washington* 7 (1888), 122–127.
- [39] DOSWELL, III, C. A., DAVIES-JONES, R., AND KELLER, D. L. On Summary Measures of Skill in Rare Event Forecasting Based on Contingency Tables. *Wea. and Forecasting* 5 (Dec. 1990), 576–585.
- [40] DUDA, R. O., HART, P. E., AND STORK, D. G. *Pattern Classification*, 2nd ed. John Wiley & Sons, Inc., 2001.
- [41] DUN, J. P., KUROKAWA, H., ISHII, T. T., LIU, Y., AND ZHANG, H. Q. Evolution of Magnetic Nonpotentiality in NOAA AR 10486. *Astrophys. J.* 657 (Mar. 2007), 577–591.
- [42] FALCONER, D. A., MOORE, R. L., AND GARY, G. A. Correlation of the Coronal Mass Ejection Productivity of Solar Active Regions with Measures of Their Global Nonpotentiality from Vector Magnetograms: Baseline Results. *Astrophys. J.* 569 (Apr. 2002), 1016–1025.
- [43] FALCONER, D. A., MOORE, R. L., AND GARY, G. A. Magnetic Causes of Solar Coronal Mass Ejections: Dominance of the Free Magnetic Energy over the Magnetic Twist Alone. *Astrophys. J.* 644 (June 2006), 1258–1272.
- [44] FELDMAN, U., DOSCHEK, G. A., AND KLIMCHUK, J. A. The Occurrence Rate of Soft X-Ray Flares as a Function of Solar Activity. *Astrophys. J.* 474 (Jan. 1997), 511.
- [45] FERNANDEZ BORDA, R. A., MININNI, P. D., MANDRINI, C. H., GÓMEZ, D. O., BAUER, O. H., AND ROVIRA, M. G. Automatic Solar Flare Detection Using Neural Network Techniques. *Solar Phys.* 206 (Apr. 2002), 347–357.

- [46] FINLEY, J. P. Tornado Predictions. *Amer. Meteor. J.* 1 (1884), 85–88.
- [47] GALLAGHER, P. T., DENKER, C., YURCHYSHYN, V., SPIROCK, T., QIU, J., WANG, H. M., AND GOODE, P. R. Solar activity monitoring and forecasting capabilities at Big Bear Solar Observatory. *Annales Geophysicae* 20 (July 2002), 1105–1115.
- [48] GALLAGHER, P. T., MOON, Y.-J., AND WANG, H. M. Active-Region Monitoring and Flare Forecasting I. Data Processing and First Results. *Solar Phys.* 209 (Sept. 2002), 171–183.
- [49] GALLAGHER, P. T., YOUNG, C. A., BYRNE, J. P., AND MCATEER, R. T. J. Coronal mass ejection detection using wavelets, curvelets and ridgelets: Applications for space weather monitoring. *Advances in Space Research* 47 (June 2011), 2118–2126.
- [50] GAO, D., ZHANG, Y. X., AND ZHAO, Y. H. Support vector machines and kd-tree for separating quasars from large survey data bases. *MNRAS* 386 (May 2008), 1417–1425.
- [51] GARCIA, H. A. Temperature and emission measure from GOES soft X-ray measurements. *Solar Phys.* 154 (Oct. 1994), 275–308.
- [52] GARY, G. A., AND HAGYARD, M. J. Transformation of vector magnetograms and the problems associated with the effects of perspective and the azimuthal ambiguity. *Solar Phys.* 126 (Mar. 1990), 21–36.
- [53] GARY, G. A., MOORE, R. L., HAGYARD, M. J., AND HAISCH, B. M. Nonpotential features observed in the magnetic field of an active region. *Astrophys. J.* 314 (Mar. 1987), 782–794.
- [54] GEORGOULIS, M. K. Turbulence In The Solar Atmosphere: Manifestations And Diagnostics Via Solar Image Processing. *Solar Phys.* 228 (May 2005), 5–27.
- [55] GEORGOULIS, M. K., AND RUST, D. M. Quantitative Forecasting of Major Solar Flares. *Astrophys. J. Lett.* 661 (May 2007), L109–L112.

-
- [56] GILBERT, G. F. Finley's Tornado Predictions. *American Meteorological Journal* 1 (1884), 166–172.
- [57] GUO, J., AND ZHANG, H. Q. A statistical study of quantifying magnetic complexity of solar active regions. *Adv. Space Res.*39 (2007), 1773–1776.
- [58] GUO, J., ZHANG, H. Q., AND CHUMAK, O. V. Magnetic properties of flare-CME productive active regions and CME speed. *A&A* 462 (Feb. 2007), 1121–1126.
- [59] GUO, J., ZHANG, H. Q., CHUMAK, O. V., AND LIN, J. B. The evolution of the quantified magnetic complexity of active regions in the 23rd solar cycle. *Mon. Not. Roy. Astron. Soc.*405 (June 2010), 111–114.
- [60] GUO, J., ZHANG, H. Q., CHUMAK, O. V., AND LIU, Y. A Quantitative Study on Magnetic Configuration for Active Regions. *Solar Phys.*237 (Aug. 2006), 25–43.
- [61] HAGINO, M., AND SAKURAI, T. Latitude Variation of Helicity in Solar Active Regions. *Publ. Astron. Soc. Japan*56 (Oct. 2004), 831–843.
- [62] HAGYARD, M. J., CUMINGS, N. P., WEST, E. A., AND SMITH, J. E. The MSFC Vector Magnetograph. *Solar Phys.*80 (Sept. 1982), 33–51.
- [63] HAGYARD, M. J., LOW, B. C., AND TANDBERG-HANSEN, E. On the presence of electric currents in the solar atmosphere. I - A theoretical framework. *Solar Phys.*73 (Oct. 1981), 257–268.
- [64] HAGYARD, M. J., SMITH, J. B., TEUBER, D., AND WEST, E. A. A quantitative study relating observed shear in photospheric magnetic fields to repeated flaring. *Solar Phys.*91 (Mar. 1984), 115–126.
- [65] HALE, G. E. On the Probable Existence of a Magnetic Field in Sun-Spots. *Astrophys. J.*28 (Nov. 1908), 315.

- [66] HANSSEN, A. W., AND KUIPERS, W. J. A. On the Relationship Between the Frequency of Rain and Various Meteorological Parameters. *Mededeelingen en Verhandelingen 81* (1965), 2–15.
- [67] HAO, J., AND ZHANG, M. Hemispheric Helicity Trend for Solar Cycle 24. *Astrophys. J. Lett.* 733 (June 2011), L27.
- [68] HARKER, B. J., AND MIGHELL, K. J. A GPU-computing Approach to Solar Stokes Profile Inversion. *Astrophys. J.* 757 (Sept. 2012), 8.
- [69] HARTLEP, T., ZHAO, J., MANSOUR, N. N., AND KOSOVICHEV, A. G. Validating Time-Distance Far-Side Imaging of Solar Active Regions through Numerical Simulations. *Astrophys. J.* 689 (Dec. 2008), 1373–1378.
- [70] HARVEY, J. W. ‘*Magnetic Fields Associated with Solar Active-Region Prominences*’. PhD thesis, University of Colorado, Boulder, 1969.
- [71] HEIDKE, P. Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst. *Geogr. Ann. Stockh.* 8 (1926), 301–349.
- [72] HENNEY, C. J., KELLER, C. U., AND HARVEY, J. W. SOLIS-VSM Solar Vector Magnetograms. In *Astronomical Society of the Pacific Conference Series* (Dec. 2006), R. Casini and B. W. Lites, Eds., vol. 358 of *Astronomical Society of the Pacific Conference Series*, p. 92.
- [73] HUANG, X., YU, D. R., HU, Q. H., WANG, H. N., AND CUI, Y. M. Short-Term Solar Flare Prediction Using Predictor Teams. *Solar Phys.* 263 (May 2010), 175–184.
- [74] JING, J., SONG, H., ABRAMENKO, V., TAN, C. Y., AND WANG, H. M. The Statistical Relationship between the Photospheric Magnetic Parameters and the Flare Productivity of Active Regions. *Astrophys. J.* 644 (June 2006), 1273–1277.

-
- [75] JOACHIMS, T. Making large-scale SVM learning practical. In *Advances in Kernel Methods - Support Vector Learning*, B. Schölkopf, C. Burges, and A. Smola, Eds. MIT Press, Cambridge, MA, 1999, ch. 11, pp. 169–184. Software available at <http://svmlight.joachims.org/>.
- [76] JOSHI, N. C., BANKOTI, N. S., PANDE, S., PANDE, B., UDDIN, W., AND PANDEY, K. Statistical analysis of Soft X-ray solar flares during solar cycles 21, 22 and 23. *New Astron.*15 (Aug. 2010), 538–546.
- [77] KAHLER, S. W., AND KREPLIN, R. W. The NRL SOLRAD X-ray detectors - A summary of the observations and a comparison with the SM-S/GOES detectors. *Solar Phys.*133 (June 1991), 371–384.
- [78] KRALL, K. R., SMITH, JR., J. B., HAGYARD, M. J., WEST, E. A., AND CUMMINGS, N. P. Vector magnetic field evolution, energy storage, and associated photospheric velocity shear within a flare-productive active region. *Solar Phys.*79 (July 1982), 59–75.
- [79] KRISTA, L. D., AND GALLAGHER, P. T. Automated Coronal Hole Detection Using Local Intensity Thresholding Techniques. *Solar Phys.*256 (May 2009), 87–100.
- [80] LABROSSE, N., DALLA, S., AND MARSHALL, S. Automatic Detection of Limb Prominences in 304 Å EUV Images. *Solar Phys.*262 (Apr. 2010), 449–460.
- [81] LEKA, K. D., AND BARNES, G. Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active Regions. I. Data, General Approach, and Sample Results. *Astrophys. J.*595 (Oct. 2003), 1277–1295.
- [82] LEKA, K. D., AND BARNES, G. Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active Regions. II. Discriminant Analysis. *Astrophys. J.*595 (Oct. 2003), 1296–1306.

- [83] LEKA, K. D., AND BARNES, G. Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active Regions. IV. A Statistically Significant Sample. *Astrophys. J.* 656 (Feb. 2007), 1173–1186.
- [84] LI, H. Projection Effects on Physical Parameters Obtained from Solar Vector Magnetograms. *Chinese J. Astron. Astrophys.* 2 (Apr. 2002), 174–182.
- [85] LI, J. Z. 模式识别导论. 北京: 高等教育出版社, 1994.
- [86] LI, R., WANG, H. N., HE, H., CUI, Y. M., AND DU, Z. L. Support Vector Machine combined with K-Nearest Neighbors for Solar Flare Forecasting. *Chinese J. Astron. Astrophys.* 7 (June 2007), 441–447.
- [87] LI, X. B., AND ZHANG, H. Q. Moving Magnetic Features Around AR 10930 from High-resolution Data Observed by Hinode/SOT. *Astrophys. J.* 771 (July 2013), 22.
- [88] LIN, G. H. 太阳物理研究的信息化发展. 科研信息化技术与应用 2 (2011), 74–81.
- [89] LOW, B. C., AND NAKAGAWA, Y. Dynamics of solar magnetic fields. VI. *Astrophys. J.* 199 (July 1975), 237–246.
- [90] LÜ, Y. P., WANG, J. X., AND WANG, H. N. Shear angle of magnetic fields. *Solar Phys.* 148 (Nov. 1993), 119–132.
- [91] MANZATO, A. An Odds Ratio Parameterization for ROC Diagram and Skill Score Indices. *Wea. and Forecasting* 20 (2005), 918.
- [92] MARTENS, P. C. H., ATTRILL, G. D. R., DAVEY, A. R., ENGELL, A., FARID, S., GRIGIS, P. C., KASPER, J., KORRECK, K., SAAR, S. H., SAVCHEVA, A., SU, Y., TESTA, P., WILLS-DAVEY, M., BERNASCONI, P. N., RAOUAFI, N.-E., DELOUILLE, V. A., HOCHEDÉZ, J. F., CIRTAİN, J. W., DEFOREST, C. E., ANGRYK, R. A., DE MOORTEŁ, I., WIEGELMANN, T., GEORGOULIS, M. K., MCATEER, R. T. J., AND

- TIMMONS, R. P. Computer Vision for the Solar Dynamics Observatory (SDO). *Solar Phys.*275 (Jan. 2012), 79–113.
- [93] MARZBAN, C. Scalar Measures of Performance in Rare-Event Situations. *Wea. and Forecasting* 13 (Sept. 1998), 753–763.
- [94] MASON, J. P., AND HOEKSEMA, J. T. Testing Automated Solar Flare Forecasting with 13 Years of Michelson Doppler Imager Magnetograms. *Astrophys. J.*723 (Nov. 2010), 634–640.
- [95] MCINTOSH, P. S. The classification of sunspot groups. *Solar Phys.*125 (Feb. 1990), 251–267.
- [96] MESSEROTTI, M., ZUCCARELLO, F., GUGLIELMINO, S. L., BOTHMER, V., LILENSTEN, J., NOCI, G., STORINI, M., AND LUNDSTEDT, H. Solar Weather Event Modelling and Prediction. *Space Sci. Rev.*147 (Nov. 2009), 121–185.
- [97] MICKEY, D. L., CANFIELD, R. C., LABONTE, B. J., LEKA, K. D., WATERTON, M. F., AND WEBER, H. M. The Imaging Vector Magnetograph at Haleakala. *Solar Phys.*168 (Oct. 1996), 229–250.
- [98] MOON, Y.-J., YUN, H. S., CHOI, G., PARK, Y. D., AND MICKEY, D. L. Nonpotential Parameters of Solar Active Region AR 5747. *J. Korean Astron. Soc.* 33 (Apr. 2000), 47–55.
- [99] MORETON, G. E., AND SEVERNY, A. B. Magnetic Fields and Flares in the Region CMP 20 September 1963. *Solar Phys.*3 (Feb. 1968), 282–297.
- [100] MORGAN, H., BYRNE, J. P., AND HABBAL, S. R. Automatically Detecting and Tracking Coronal Mass Ejections. I. Separation of Dynamic and Quiescent Components in Coronagraph Images. *Astrophys. J.*752 (June 2012), 144.
- [101] MURPHY, A. H. The Finley Affair: A Signal Event in the History of Forecast Verification. *Wea. and Forecasting* 11 (Mar. 1996), 3–20.

- [102] NAKAGAWA, Y., AND RAADU, M. A. On Practical Representation of Magnetic Field. *Solar Phys.*25 (July 1972), 127–135.
- [103] NINDOS, A., AND ANDREWS, M. D. The Association of Big Flares and Coronal Mass Ejections: What Is the Role of Magnetic Helicity? *Astrophys. J. Lett.*616 (Dec. 2004), L175–L178.
- [104] OLMEDO, O., ZHANG, J., WECHSLER, H., POLAND, A., AND BORNE, K. Automatic Detection and Tracking of Coronal Mass Ejections in Coronagraph Time Series. *Solar Phys.*248 (Apr. 2008), 485–499.
- [105] PARK, S.-H., CHAE, J., AND WANG, H. M. Productivity of Solar Flares and Magnetic Helicity Injection in Active Regions. *Astrophys. J.*718 (July 2010), 43–51.
- [106] PEIRCE, C. S. The Numerical Measure of the Success of Prediction. *Science* 4 (1884), 453–454.
- [107] PENG, N. B., ZHANG, Y. X., ZHAO, Y. H., AND WU, X. B. Selecting quasar candidates using a support vector machine classification system. *Mon. Not. Roy. Astron. Soc.*425 (Oct. 2012), 2599–2609.
- [108] PEVTSOV, A. A., CANFIELD, R. C., AND METCALF, T. R. Latitudinal variation of helicity of photospheric magnetic fields. *Astrophys. J. Lett.*440 (Feb. 1995), L109–L112.
- [109] PEVTSOV, A. A., DUN, J. P., AND ZHANG, H. Q. Helicity Measurements from Two Magnetographs. *Solar Phys.*234 (Apr. 2006), 203–212.
- [110] QAHWAJI, R., AND COLAK, T. Automatic Short-Term Solar Flare Prediction Using Machine Learning and Sunspot Associations. *Solar Phys.*241 (Mar. 2007), 195–211.
- [111] QAHWAJI, R., COLAK, T., AL-OMARI, M., AND IPSOON, S. Automated Prediction of CMEs Using Machine Learning of CME - Flare Associations. *Solar Phys.*248 (Apr. 2008), 471–483.

-
- [112] QU, M., SHIH, F., JING, J., AND WANG, H. Automatic Solar Flare Tracking Using Image-Processing Techniques. *Solar Phys.*222 (July 2004), 137–149.
- [113] QU, M., SHIH, F. Y., JING, J., AND WANG, H. Automatic Solar Filament Detection Using Image Processing Techniques. *Solar Phys.*228 (May 2005), 119–135.
- [114] QU, M., SHIH, F. Y., JING, J., AND WANG, H. Automatic Detection and Classification of Coronal Mass Ejections. *Solar Phys.*237 (Sept. 2006), 419–431.
- [115] QU, M., SHIH, F. Y., JING, J., AND WANG, H. M. Automatic Solar Flare Detection Using MLP, RBF, and SVM. *Solar Phys.*217 (Oct. 2003), 157–172.
- [116] ROBBRECHT, E., AND BERGHMANS, D. Automated recognition of coronal mass ejections (CMEs) in near-real-time data. *Astron. Astrophys.*425 (Oct. 2004), 1097–1106.
- [117] RUST, D. M., SAKURAI, T., GAIZAUSKAS, V., HOFMANN, A., MARTIN, S. M., PRIEST, E. R., AND WANG, J. X. Preflare state. *Solar Phys.*153 (Aug. 1994), 1–17.
- [118] SAKURAI, T., ICHIMOTO, K., NISHINO, Y., SHINODA, K., NOGUCHI, M., HIEI, E., LI, T., HE, F., MAO, W., LU, H., AI, G., ZHAO, Z., KAWAKAMI, S., AND CHAE, J.-C. Solar flare telescope at Mitaka. *Publ. Astron. Soc. Japan*47 (Feb. 1995), 81–92.
- [119] SAKURAI, T., MAKITA, M., AND SHIBASAKI, K. Observation of magnetic field vector in solar active regions. In: Schmidt, H.U. (ed.), *Theoretical Problems in High Resolution Solar Physics, MPA-212*, Max Planck Institute for Astrophysics, Munich (1985), 312–315.
- [120] SCHAEFER, J. T. The Critical Success Index as an Indicator of Warning Skill. *Wea. and Forecasting* 5 (Dec. 1990), 570–575.

- [121] SCHERRER, P. H., SCHOU, J., BUSH, R. I., KOSOVICHEV, A. G., BOGART, R. S., HOEKSEMA, J. T., LIU, Y., DUVAL, T. L., ZHAO, J., TITLE, A. M., SCHRIJVER, C. J., TARBELL, T. D., AND TOMCZYK, S. The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO). *Solar Phys.*275 (Jan. 2012), 207–227.
- [122] SCHOLL, I. F., AND HABBAL, S. R. Automatic Detection and Classification of Coronal Holes and Filaments Based on EUV and Magnetogram Observations of the Solar Disk. *Solar Phys.*248 (Apr. 2008), 425–439.
- [123] SCHRIJVER, C. J. A Characteristic Magnetic Field Pattern Associated with All Major Solar Flares and Its Use in Flare Forecasting. *Astrophys. J. Lett.*655 (Feb. 2007), L117–L120.
- [124] SEEHAFFER, N. Electric current helicity in the solar atmosphere. *Solar Phys.*125 (Feb. 1990), 219–232.
- [125] SEVERNY, A. B. The Nature of Solar Magnetic Fields (The Fine Structure of the Field). *Soviet Astron.*9 (Oct. 1965), 171.
- [126] SHEN, Y. B., LIN, J. B., JI, K. F., AND DENG, Y. Y. New real-time correlation solar observing system based on GPU for acquiring the deep-integration magnetogram. *New Astron.*25 (Dec. 2013), 32–37.
- [127] SHIH, F. Y., AND KOWALSKI, A. J. Automatic Extraction of Filaments in H α Solar Images. *Solar Phys.*218 (Dec. 2003), 99–122.
- [128] SONG, H., TAN, C. Y., JING, J., WANG, H. M., YURCHYSHYN, V., AND ABRAMENKO, V. Statistical Assessment of Photospheric Magnetic Features in Imminent Solar Flare Predictions. *Solar Phys.*254 (Jan. 2009), 101–125.
- [129] STEPHENSON, D. B. Use of the “Odds Ratio” for Diagnosing Forecast Skill. *Wea. and Forecasting* 15 (Apr. 2000), 221–232.
- [130] SU, J. T. 太阳矢量磁场测量. PhD thesis, 中国科学院研究生院, 2004.

-
- [131] SU, J. T., AND ZHANG, H. Q. Calibration of Vector Magnetogram with the Nonlinear Least-squares Fitting Technique. *Chinese J. Astron. Astrophys.* 4 (Aug. 2004), 365–376.
- [132] TANAKA, K. Studies on a very flare-active delta group - Peculiar delta SPOT evolution and inferred subsurface magnetic rope structure. *Solar Phys.* 136 (Nov. 1991), 133–149.
- [133] TAPPIN, S. J., HOWARD, T. A., HAMPSON, M. M., THOMPSON, R. N., AND BURNS, C. E. On the autonomous detection of coronal mass ejections in heliospheric imager data. *Journal of Geophysical Research (Space Physics)* 117 (May 2012), 5103.
- [134] TIWARI, S. K., VENKATAKRISHNAN, P., AND SANKARASUBRAMANIAN, K. Global Twist of Sunspot Magnetic Fields Obtained from High-Resolution Vector Magnetograms. *Astrophys. J. Lett.* 702 (Sept. 2009), L133–L137.
- [135] TSUNETA, S., ICHIMOTO, K., KATSUKAWA, Y., NAGATA, S., OTSUBO, M., SHIMIZU, T., SUEMATSU, Y., NAKAGIRI, M., NOGUCHI, M., TARBELL, T., TITLE, A., SHINE, R., ROSENBERG, W., HOFFMANN, C., JURCEVICH, B., KUSHNER, G., LEVAY, M., LITES, B., ELMORE, D., MATSUSHITA, T., KAWAGUCHI, N., SAITO, H., MIKAMI, I., HILL, L. D., AND OWENS, J. K. The Solar Optical Telescope for the Hinode Mission: An Overview. *Solar Phys.* 249 (June 2008), 167–196.
- [136] VAN RIJSBERGEN, C. *Information retrieval*. London: Butterworths, 1979. <http://www.dcs.gla.ac.uk/Keith/Preface.html>.
- [137] VAPNIK, V. N. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [138] WADADEKAR, Y. Estimating Photometric Redshifts Using Support Vector Machines. *Publ. Astron. Soc. Pac.* 117 (Jan. 2005), 79–85.

- [139] WANDISHIN, M. S., AND BROOKS, H. E. On the relationship between Clayton's skill score and expected value for forecasts of binary events. *Meteorol. Appl.* 9 (Dec. 2002), 455–459.
- [140] WANG, H. M., VARSIK, J., ZIRIN, H., CANFIELD, R. C., LEKA, K. D., AND WANG, J. X. Joint vector magnetograph observations at BBSO, Huairou Station and Mees Solar Observatory. *Solar Phys.* 142 (Nov. 1992), 11–20.
- [141] WANG, J. X., SHI, Z. X., WANG, H. N., AND LÜ, Y. P. Flares and the Magnetic Nonpotentiality. *Astrophys. J.* 456 (Jan. 1996), 861.
- [142] WANG, R., YAN, Y. H., AND TAN, B. L. Three-dimensional nonlinear force-free field reconstruction of solar active region 11158 by direct boundary integral equation. *Solar Physics* (2013), 1–23.
- [143] WANG, T. J., AI, G. X., AND DENG, Y. Y. Calibration of Nine-channel Solar Magnetic field Telescope. II. The results of the observational calibration. *Astrophys. Rep. Publ. Beijing Astron. Obs.* 28 (June 1996), 41–48.
- [144] WANG, T. J., XU, A. A., AND ZHANG, H. Q. Evolution of vector magnetic fields and vertical currents and their relationship with solar flares in AR 5747. *Solar Phys.* 155 (Nov. 1994), 99–112.
- [145] WILKS, D. S. *Statistical Methods in the Atmospheric Sciences*. International Geophysics Series. Elsevier Academic Press, 2006. Chapter 7.
- [146] WONG, H.-C., WONG, U.-H., FENG, X., AND TANG, Z. Efficient magnetohydrodynamic simulations on graphics processing units with cuda. *Computer Physics Communications* 182, 10 (2011), 2132–2160.
- [147] WOODCOCK, F. The Evaluation of Yes/No Forecasts for Scientific and Administrative Purposes. *Mon. Wea. Rev.* 104 (1976), 1209.
- [148] WOŹNIAK, P. R., WILLIAMS, S. J., VESTRAND, W. T., AND GUPTA, V. Identifying Red Variables in the Northern Sky Variability Survey. *Astron. J.* 128 (Dec. 2004), 2965–2976.

-
- [149] XU, H. Q., GAO, Y., ZHANG, H. Q., SAKURAI, T., PEVTSOV, A. A., AND SOKOLOFF, D. Helicity comparison among three magnetographs. *Adv. Space Res.*39 (2007), 1715–1722.
- [150] YANG, S. B. 浮现活动区的磁螺度研究. PhD thesis, 中国科学院研究生院, 2009.
- [151] YANG, S. B., AND ZHANG, H. Q. Large-scale Magnetic Helicity Fluxes Estimated from MDI Magnetic Synoptic Charts over the Solar Cycle 23. *Astrophys. J.*758 (Oct. 2012), 61.
- [152] YANG, X., ZHANG, H. Q., GAO, Y., GUO, J., AND LIN, G. H. A Statistical Study on Photospheric Magnetic Nonpotentiality of Active Regions and Its Relationship with Flares During Solar Cycles 22 – 23. *Solar Phys.*280 (Sept. 2012), 165–181.
- [153] YASHIRO, S., MICHALEK, G., AND GOPALSWAMY, N. A comparison of coronal mass ejections identified by manual and automatic methods. *Annales Geophysicae* 26 (Oct. 2008), 3103–3112.
- [154] YOUNG, C. A., AND GALLAGHER, P. T. Multiscale Edge Detection in the Corona. *Solar Phys.*248 (Apr. 2008), 457–469.
- [155] YU, D. R., HUANG, X., HU, Q. H., ZHOU, R., WANG, H. N., AND CUI, Y. M. Short-term Solar Flare Prediction Using Multiresolution Predictors. *Astrophys. J.*709 (Jan. 2010), 321–326.
- [156] YU, D. R., HUANG, X., WANG, H. N., AND CUI, Y. M. Short-Term Solar Flare Prediction Using a Sequential Supervised Learning Method. *Solar Phys.*255 (Mar. 2009), 91–105.
- [157] YU, D. R., HUANG, X., WANG, H. N., CUI, Y. M., HU, Q. H., AND ZHOU, R. Short-term Solar Flare Level Prediction Using a Bayesian Network Approach. *Astrophys. J.*710 (Feb. 2010), 869–877.

- [158] YUAN, Y., SHIH, F. Y., JING, J., AND WANG, H. M. Automated flare forecasting using a statistical learning technique. *Research in Astron. and Astrophys.* 10 (Aug. 2010), 785–796.
- [159] ZEEMAN, P. Doubles and triplets in the spectrum produced by external magnetic forces. *Phil. Mag.* 44 (1897), 55.
- [160] ZEEMAN, P. On the Influence of Magnetism on the Nature of the Light Emitted by a Substance. *Astrophys. J.* 5 (May 1897), 332.
- [161] ZHANG, H. Q., LABONTE, B., LI, J., AND SAKURAI, T. Analysis of Vector Magnetic Fields in Solar Active Regions by Huairou, Mees and Mitaka Vector Magnetographs. *Solar Phys.* 213 (Mar. 2003), 87–102.
- [162] ZHANG, H. Q., SAKURAI, T., PEVTSOV, A. A., GAO, Y., XU, H. Q., SOKOLOFF, D. D., AND KUZANYAN, K. A new dynamo pattern revealed by solar helical magnetic fields. *Mon. Not. Roy. Astron. Soc.* 402 (Feb. 2010), L30–L33.
- [163] ZHANG, J., WANG, Y., AND LIU, Y. Statistical Properties of Solar Active Regions Obtained from an Automatic Detection System and the Computational Biases. *Astrophys. J.* 723 (Nov. 2010), 1006–1018.
- [164] ZHANG, Y. X., AND ZHAO, Y. H. Classification in Multidimensional Parameter Space: Methods and Examples. *Publ. Astron. Soc. Pac.* 115 (Aug. 2003), 1006–1018.
- [165] ZHU, X. S., WANG, H. N., FAN, Y. L., DU, Z. L., AND HE, H. A method of magnetosonic characteristics to correct the “rarefaction shocks” problem arising in ZEUS. *Science in China G: Physics and Astronomy* 55 (Sept. 2012), 1720–1726.

发表文章目录

- [1] **YANG Xiao**, ZHANG HongQi, GAO Yu, GUO Juan, LIN GangHua. A Statistical Study on Photospheric Magnetic Nonpotentiality of Active Regions and Its Relationship with Flares During Solar Cycles 22 – 23. *Solar Physics*, vol. 280(1), pp.165-181, 2012.
- [2] **YANG Xiao**, LIN GangHua, ZHANG HongQi, MAO XinJie. Magnetic Nonpotentiality in Active Regions as a Predictor of Solar Flares. *The Astrophysical Journal Letters*, vol. 774(2), L27 (6pp), 2013.
- [3] **YANG Xiao**, ZHANG HongQi, LIN GangHua, GAO Yu, GUO Juan. A Statistical Study on Photospheric Active-Region Magnetic Nonpotentiality and Associated Flares During Solar Cycles 22 – 23. In: Kosovichev, A. G., de Gouveia Dal Pino E. M., Yan, Y. H. (ed.) *Solar and Astrophysical Dynamos and Magnetic Activity, IAUS-294*, pp.587-588, 2013.
- [4] 杨潇, 林钢华, 白先勇, 朱小帅. GPU 计算在太阳物理中的应用. 科研信息化技术与应用, vol. 3(3), pp.69-76, 2012.
- [5] LIU Suo, SU JiangTao, ZHANG HongQi, DENG YuanYong, GAO Yu, **YANG Xiao**, MAO XinJie. A Statistical Study on Force-Freeness of Solar Magnetic Fields in the Photosphere. *Publications of the Astronomical Society of Australia*, vol. 30(1), e005 (7 pages), 2013.

简 历

基本情况

杨潇，女，汉族，祖籍湖北省应城市，1984 年 9 月出生于湖北省江汉油田，现为中国科学院国家天文台怀柔基地在读博士研究生。

教育状况

2002 年 9 月至 2006 年 7 月，武汉大学，软件工程专业，本科学历，获工学学士学位。

2007 年 9 月至 2009 年 9 月，中国科学院国家天文台，天文技术与方法专业，硕士研究生在读，导师为林钢华高级工程师，后转博。

2009 年 9 月 至今，中国科学院国家天文台，天体物理专业，博士研究生在读，导师为张洪起研究员和林钢华研究员。

研究兴趣

太阳物理，天文信息技术

联系方式

通讯地址：北京市朝阳区大屯路甲 20 号，中国科学院国家天文台

邮编：100012

E-mail: yangx@nao.cas.cn, yxisswu@gmail.com

主页: <http://sun.bao.ac.cn/~yangx>

致 谢

一路走来，并不容易，尽管完成的工作非常微不足道，但总算没让答卷空白。要感谢的人太多太多。

感谢我的导师张洪起研究员和林钢华研究员。张老师是前任怀柔基地首席科学家，四年前他组织的国际螺度会议为我打开了工作思路，文章的顺利接收直接得益于他的把关。他在太阳物理领域默默地耕耘和奉献三十余载仍孜孜不倦，让人由衷地敬服。林老师是我硕士阶段的导师，是我进入天文领域的领路人，在我的学习工作中给予了我许多启发。她旺盛的精力，对科研的执着和奋斗精神，一直感染着我，给我前进的动力。二位老师给我创造了许多不可多得的锻炼机会，受益匪浅。

感谢怀柔基地主任邓元勇研究员，他睿智、高屋建瓴、知人善任，是怀柔基地强有力的主心骨，对我的工作多有画龙点睛式的点拨。感谢首席科学家张枚研究员，她是所有女同胞的榜样，也是所有男同胞的偶像。感谢怀柔基地站长王慧老师，她的气场和亲和力凝聚着这个大家庭，她对我们的关爱体现在每一个细节里，让我们倍感幸福。

感谢北京师范大学毛信杰教授及其夫人吕正女士。毛老师几年来一直坚持给怀柔基地的同学和同事们讲授太阳磁流体力学课程，精益求精；对我文章写作中语言文字的润色方面给予了非常重要的支持，并且对我的工作提出了许多建议；另外，他健康快乐的生活方式，是我们的标杆。吕师母总是热情地招待晚辈们的不时到访，对我们如子女般爱护有加。

感谢我研一时的几位老师：国家天文台林元章研究员、曹惠来副研究员，北京大学徐仁新教授，虚拟经济与数据科学研究中心田英杰研究员，计算所史忠植研究员，软件所范植华研究员。研一是我入门天文的第一年，集中教学期间不仅为我打下相关的知识基础，还对我后来的工作启发良多。

先生之风，山高水长。感谢成长路上诸多良师益友的提携和鼓励。

怀柔基地是个充满热情、积极向上、干劲十足的团队，团队里每一个人身上的闪光点都在潜移默化地影响着我。感谢怀柔基地的所有同事和战友，你们永远是我学习的榜样。感谢我的合作者郭娟博士、高裕博士、刘锁博士、白先

勇给我的最有力的支持。感谢郝娟博士、王晓帆博士对我观测的启蒙，感谢汪国萍老师对我观测中的指导。感谢姜杰博士在工作、生活诸多方面对我的关心和照顾。感谢苏江涛研究员、杨尚斌博士、包星明博士、陈洁博士、赵辉博士、穆军博士、滕飞博士对我研究工作中的解惑和探讨。感谢王东光研究员、胡柯良高工、林佳本高工、王丙祥工程师、孙英姿高工、孙文君、朱晓明、曾真、沈洋斌。感谢张志勇高工，你那里的光学“玩具”让我爱不释手。感谢李小波博士，你的奇思异想、妙语连珠，趣味横生。感谢在研究生办公室一起奋斗过的师兄弟姐妹们：谢文彬博士、王栋博士、王传宇博士、赵明宇博士、王硕博士、徐海清博士、侯俊峰博士、赵翠工程师、于佳、李静、宋永亮。感谢王绚对我组内杂务上的诸多帮助。感谢李威老师、张鼎波、杨永田老师，默默地在为怀柔站贡献力量。感谢怀柔站阮文东、支坤香、凌林、包亚东、王立东、荆帅对我在站期间各个方面的照顾。感谢在怀柔站上一同执行观测任务的CAPS 组蔡昊、刘涛、房冉、刘毅、马兴华，和太阳射电团组的诸位。

感谢太阳活动预报团组王华宁研究员允许我旁听他的组会。感谢汪景琇研究员在我改稿过程中的答疑解惑。感谢云南天文台林隼研究员、屈中权研究员、刘煜研究员，南京大学陈鹏飞教授组织的暑期班课程。感谢太阳磁场与磁活动团组张宇宗博士、金春兰博士、宋乔博士、肖俊敏博士、姜法宇，太阳活动预报组黄鑫博士、郭晓博、阎岩、戴幸华，石家庄学院刘继宏博士，山东大学威海分校阮桂萍博士，云南天文台田占军、闫晓理博士，无论是交流工作还是闲聊娱乐，和你们相处都非常轻松愉快。

感谢研究生办杜红荣老师、艾华老师、马怀宇老师对我读研期间学习生活上的照顾和帮助。

感谢我的同学：李楠、姜云英、程利娜、张悦旻、谢艳霞、王萌、杜薇、肖媛、韩晓红、徐振鹏、郭志清、王康、朱小帅、范玉良、彭南博、侯立刚、朱明、杨培强、张春光、刘东浩、陈如荣、付晓辉、刘昊、马杰、马锦、刘俊、胡会东、巍巍。感谢物理学院 2007304 班的每一位同学、物理学院排球队的每一位队友。张弛有道，感谢一路上有你们，让我的研究生生涯里充满了欢声笑语。一并还要感谢的是我考研路上对我关照有加的：武汉大学文学院孙冬妮老师、朱姝师姐，历史学院雷蕾师姐，生科院张翼师姐，水电学院倪晓峰师兄；还有和我一起并肩战斗并最终如愿进入中科院的唯一国软老同学王勤思。

感谢我的母亲肖道芸老师、父亲杨利斌老师、姨父邓有伍工程师，三位对

我的知识和性格上的影响是我一生的财富。感谢我所有的亲友，感谢你们在远方默默的支持。还有我来不及报答的外婆和姨妈，如今你们的温暖只能来自于梦里，醒后百感交集。

本人攻读博士期间的工作得到国家自然科学基金 (编号 60940030、10921303、41174153、10903015、11003025、11103037、11103038、11203036 和 11221063)、中国科学院国家天文台青年人才基金、中国科学院知识创新工程 (编号 KJCX2-EW-T07)、科技部国家基础研究项目 (编号 2011CB811401)、以及中国科学院太阳活动重点实验室的资助；在工作中使用到 Linux 和 MS 操作系统、IDL 程序语言、C/C++ 程序语言、Python 脚本程序、CT_EX 套装编译环境等一系列计算机工具，以及 Google 搜索引擎、ADS 天文文献系统等网络工具。在此一并表示感谢。

时光飞逝，然所识尚浅，未来的路途依然艰辛。惟愚志难摧，望能以微薄之力，助能者们开天辟地。附上本人 2009 年 10 月 21 日在怀柔值班时所作《浪淘沙》一首，与吾辈天文界后生们共勉：

呆影乱拨弦，思绪延绵。
重滩远涉梦魂牵。
云淡风清留恋处，北望南天。

思寄宇寰间，漫想无边。
心涛澎湃撼山巅。
阔步同行平祝愿，志定心虔。

谨以此文献给我敬爱的外婆赵三秀女士 (1917 – 2010) 和姨妈肖道兰医生 (1937 – 2008)。