

太阳耀斑预报研究进展

李蓉^{①*}, 朱杰^①, 黄鑫^②

① 北京物资学院信息学院, 北京 101149;

② 中国科学院国家天文台, 北京 100012

* 联系人, E-mail: lirong@bao.ac.cn

2014-01-23 收稿, 2014-05-24 接受, 2014-07-18 网络版发表

国家自然科学基金(11273031)和北京市专业建设信息类特色专业建设项目(PXM2014_014214_000017)资助

摘要 太阳耀斑是指发生在太阳表面局部区域中突然和大规模的能量释放过程. 它是空间环境的主要扰动源, 对地球空间环境造成很大影响. 太阳耀斑预报是空间天气预报的重要组成部分, 对其研究具有重要的实用价值和科学意义. 现有的大部分太阳耀斑预报模型是从观测数据提取预报因子, 利用各种统计和数据挖掘技术建立预报因子与耀斑发生之间的关系模型, 利用建立的模型对未来时间的耀斑发生进行预报. 在预报研究中, 预报因子、预报方法和预报模型是 3 个主要研究领域. 其中预报因子的选取和数据处理尤为重要, 是建立预报模型的前期工作. 预报因子主要采用太阳黑子、磁场参量和分形因子等. 预报方法包括统计方法、机器学习方法和数据同化方法. 统计方法在早期的耀斑预报建模中用的较多, 随着数据挖掘技术的发展, 越来越多的机器学习方法应用到预报模型中并取得了较好效果. 而近期发展的数据同化方法有更好的模型修正能力. 预报模型早期基本使用静态模型, 后来发展起来的动态模型具有更强的优势; 而自组织临界模型在物理方面给了耀斑发生更多的解释. 本文分别从这 3 个方面总结了耀斑预报的研究进展, 结合中国科学院国家天文台太阳活动预报中心的工作, 评述了一些重要的研究成果. 最后, 对未来的研究方向进行了总结和展望.

关键词

太阳耀斑
预报因子
机器学习
数据挖掘
预报模型

太阳耀斑是指发生在太阳表面局部区域中突然且大规模的能量释放过程^[1], 是最剧烈的太阳活动之一, 也是空间天气预报需要研究的重要内容之一. 耀斑的爆发主要表现在从射电波段到 X 射线辐射通量的突然增强. 一般分为 3 个过程: 脉冲相、闪相和下降相. 大耀斑在脉冲相之前有时有前兆, 一次大耀斑发生的过程中, 各种电磁辐射和粒子发射都会随时间产生一定变化. 较小的耀斑通常只有可见光、软 X 射线和微波射电辐射增强, 而没有大耀斑爆发时所产生的诸如 γ 射线和软 X 射线, 也没有 II, III, IV 型射电爆发和可测的粒子流.

太阳耀斑对地球空间环境造成很大影响, 是空间环境的主要扰动源^[2~7]. 耀斑爆发时, 发出大量高能粒子到达地球轨道附近时, 将会严重危及宇宙飞

行器内宇航员和仪器的安全^[8]. 当耀斑辐射到地球附近时, 与大气分子发生剧烈碰撞, 破坏电离层, 使其失去反射无线电电波的功能^[9,10]. 无线电通信尤其是短波通信, 以及电视台和电台广播, 会受到干扰甚至中断. 耀斑发射的高能带电粒子流与地球高层大气作用产生极光, 并干扰地球磁场而引起地磁暴. 因此, 太阳耀斑预报的研究不仅是人类空间探索和空间活动的重要保障, 也是理解太阳活动机理的重要环节, 具有重要的实用价值和科学意义.

目前太阳耀斑短期预报主要是利用空间和地面观测得到的太阳图像和观测数据进行预报, 一般可分为经验法、统计法和物理预报法等^[11]. 物理预报是在了解引起空间环境变化的物理过程和机制的基础上, 建立所研究对象的物理模型, 通过给定适当的初

引用格式: 李蓉, 朱杰, 黄鑫. 太阳耀斑预报研究进展. 科学通报, 2014, 59: 2452–2463

Li R, Zhu J, Huang X. Progress in solar flare forecasting (in Chinese). Chin Sci Bull (Chin Ver), 2014, 59: 2452–2463, doi: 10.1360/N972014-00058

始条件和边界条件,利用数值模拟的方法计算预报参量的空间分布和演化.然而,由于太阳耀斑爆发的物理机制仍处于研究阶段^[12-15],迄今为止,还没有建立起可用于预报的物理模型.经验预报和统计预报是根据预报因子和预报对象之间存在的简单统计关系,利用与日地物理过程有关的观测量,给出需要预报的物理量时变曲线、发生概率、位置和强度,不需要深究其中的物理背景,具有很好的实用性.随着各种观测设备的使用,可以得到大量观测数据,这种情况迫切地需要将数据转化成可以使用的预报模型.数据挖掘技术为这种需求提供了一种切实可行的方法.

在太阳耀斑预报研究中,主要关注3个方面:预报因子、预报方法和预报模型.现有的预报系统在预报因子选择上分为几大类:太阳活动区黑子组特征参量、太阳磁场参量、分型维数参量和耀斑自身参量等;预报方法主要采用统计方法、机器学习方法和数据同化方法;预报建模则分为静态模型、动态模型和自组织临界模型.本文主要围绕这3个方面,结合中国科学院国家天文台太阳活动预报组的相关工作,评述耀斑预报的研究进展.

1 预报因子

1.1 太阳黑子参量

太阳黑子和太阳活动有着紧密联系^[16-18].黑子实际上是太阳光球的局部强磁场区,其磁场强度可高达几千高斯,其内部热力学温度比周围区域要低1000 K,因而该区域较暗,呈现为黑子状态.对太阳黑子的观测,主要包括黑子数、黑子群数、黑子面积数和黑子活动区磁场等^[1].一般来说,黑子面积越大,其磁场强度也越大,同时每群黑子所呈现出的形态和极性复杂程度也与太阳耀斑等活动现象密切相关.

太阳物理学家对太阳黑子群提出了各种分类,这些分类都可以作为预报太阳活动的特征参量.1919年美国Wilson山天文台提出黑子群按磁场极性分类的方法,Wilson山磁分类(威尔逊山磁分类)以双极黑子为基本类型,其他类型都看作双极黑子群的变形.在这种分类中,黑子群被分为单极群、双极群和多极群,分别用 α , β , γ 表示,以后又加入新的 δ 位型.

黑子群的形态可用Zurich分型表示.Zurich分型只是黑子群的形态分类,其中“双极群”和“单极群”是

根据黑子群外观上是否由前导和后随两部分组成为判断.基于黑子群的Zurich分型,McIntosh^[19]提出了用3个字母表示的黑子群McIntosh分型.其中第一个字母仍然使用Zurich分型,但把9类合并为7类(A, B, C, D, E, F, H);第二个字母表示黑子群中最大黑子的形状和复杂性,根据半影的情况分为6类(x, r, s, a, h, k);第三个字母表示黑子群中黑子的分布情况,分为4类(x, o, f, c).

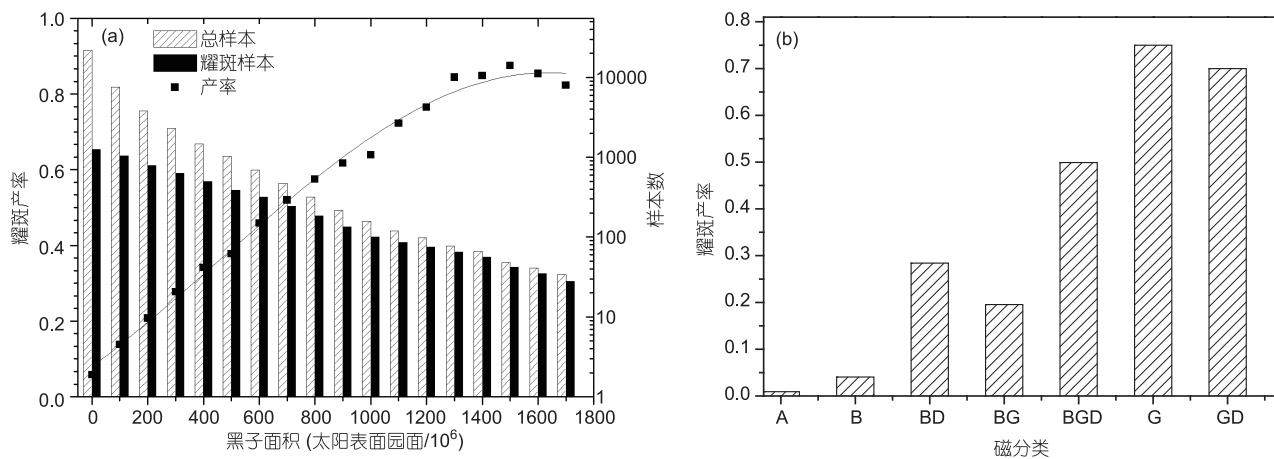
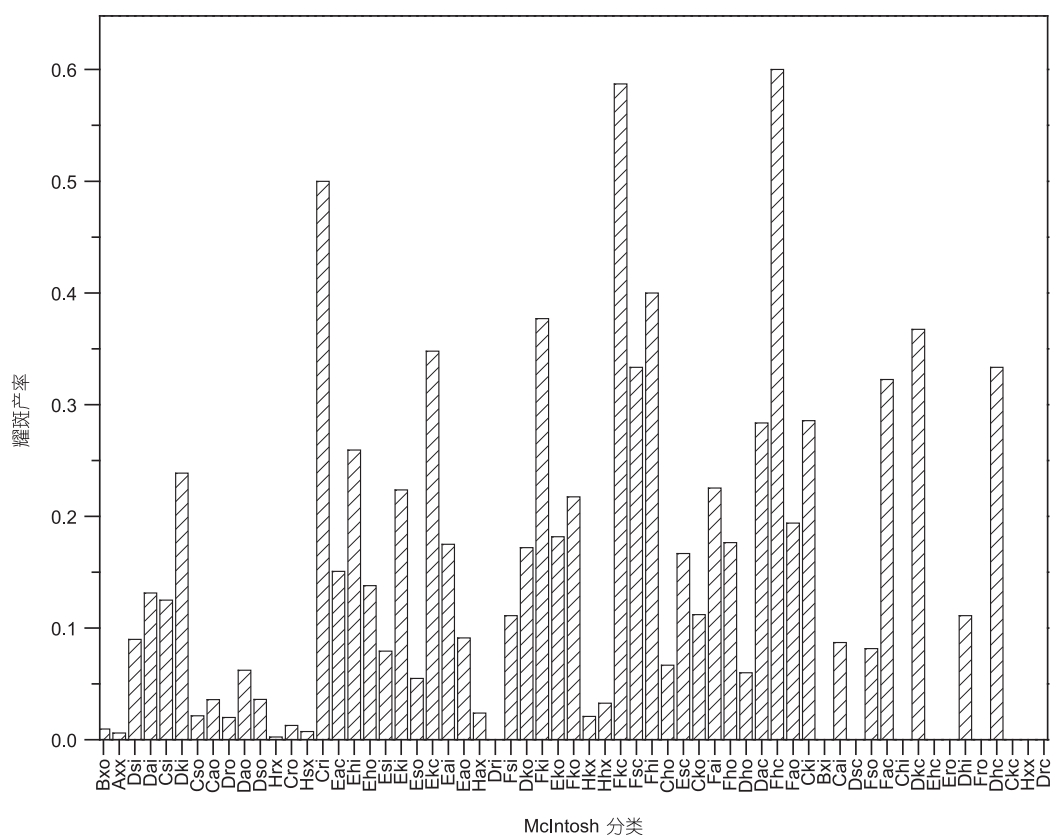
相关研究工作阐述了黑子参量和耀斑之间的关系.Samms等人^[20]研究了Wilson山磁分类和耀斑的关系,8年的活动区数据观测显示,复杂的磁场更容易产生耀斑.虽然大区域更倾向于产生大耀斑,但是磁场分类对耀斑的影响更为明显.McIntosh^[19]指出,McIntosh分型与耀斑的相关性要优于Zurich分型与耀斑的相关性.Atac^[21]综合研究了McIntosh分型和Wilson山磁分型对耀斑的影响,指出具有Dki或Eki,同时具有 δ 分型的黑子且最大磁场强度在1600~2500高斯的活动区更容易产生大耀斑.赵明宇等人^[22]统计了第24太阳活动峰年和第23太阳活动谷年的太阳黑子群和耀斑的活动规律,认为耀斑爆发与黑子群活动具有良好的相关性,峰年期间的黑子数比谷年增多~1.5倍,耀斑产率与黑子群的磁分类密切相关.

国家天文台太阳活动预报组采用大量样本来研究太阳黑子参量和耀斑的相关性.李蓉等人^[23]统计计算了黑子面积磁分类、McIntosh分型与耀斑的产率关系(图1, 2).由图中可以看出,黑子面积和耀斑产率符合Sigmoid函数拟合,而黑子磁分类中 γ 分型黑子和McIntosh分类的Fhc分型的耀斑产率最高.

1.2 太阳磁场参量

太阳大气中等离子体的平衡、运动和稳定性几乎都受到太阳磁场的制约和影响.太阳黑子和宁静日珥、太阳耀斑、爆发日珥和日冕瞬变等活动现象,都直接受太阳磁场的支配.黑子只不过是强磁场的一种表现形式^[24].耀斑的能量来源于磁场,因此研究耀斑与磁场之间的关系对理解耀斑的物理过程和耀斑预报非常重要.

美国大熊湖天文台的Gallagher等人^[25]对耀斑活动和光球磁场的关系进行了详细研究.研究结果表明,大多数活动区典型的磁场梯度 $<0.12 \text{ G km}^{-1}$,只有耀斑产率比较高的活动区才显示了大的、显著的磁场梯度.在磁场梯度显著的活动区中发现,活动区内

图 1 黑子面积(a)与磁分类(b)的耀斑产率^[23]图 2 McIntosh 分类与耀斑产率^[6]

部区域有明显磁性反转线, 并且观测和理论均证实, 大的反向磁场越接近就越容易导致大的磁场梯度. Abramenko 等人^[26]研究了光球磁场 B_z 分量的标度行为及其与耀斑间的关系. B_z 分量耗散谱与活动区耀斑产率具有强的依赖关系, 当耀斑产率很高时, 对应的谱更加平缓. Komm 等人^[27]研究了耀斑和非耀斑活动区的表面磁通量和表面下流量涡度, 研究发现, 虽

然磁通量和涡度间范围比较宽, 但是它们之间或多或少存在着线性关系. 他们还发现, 产生 C 级耀斑的活动区存在大的磁通量或大的涡度; 产生 X 级耀斑的活动区存在大的磁通量和大的涡度; 产生 M 级耀斑的活动区存在大的涡度. 因此, 涡度的测量有利于区分耀斑和非耀斑活动区. Hagyard^[28]指出了向量磁图对太阳活动研究的重要性. Schmieder 等人^[29]的研

究表明,黑子群的分裂导致磁场图像的连续重组,这一过程导致磁力线的剪切,快速的磁流浮现引起磁场复杂拓扑中能量的释放. Wheatland^[30]测试了关于耀斑能量来源于大尺度电流系统的这一假设.

耀斑跟光球磁场特征参量的关系已被广泛研究,包括新浮现、强剪切、磁场结构变化、磁通量变化等^[31]. Leka 和 Barnes^[32,33]研究了磁场水平梯度、垂直电流和电流螺度等多个光球磁场物理特征量与耀斑的关系. 他们还用统计方法研究了多个光球磁场特征量,分析了其中发生耀斑的重要参量^[34]. 结果显示,一半的参量对于活动区产生耀斑是重要的,很多参量之间有着很强的联系,被总流量刻画的大活动区也趋向于产生大的垂直电流、自由能和电流螺度,而对于产生大耀斑最重要的参量是自由能 E_e . Georgoulis 等人^[35]定义了有效的连接磁场 B_{eff} , 它是活动区潜在耀斑的度量标准. 结果表明,对于识别耀斑和非耀斑活动区, B_{eff} 具有鲁棒的标准,可有效用于耀

斑预报. Barnes 等人^[36]系统比较了最近研究中提出的 4 个磁场参量,分别为全磁通 Φ_{tot} , 自由能 E_e , 强梯度中心线上的磁通量 R 和有效的连接磁场 B_{eff} , 并且对每个参量采用判别分析方法量化预报耀斑的能力. 尽管文献^[32,34,35]显示他们提出的预报因子具有不同的预报能力,但这里给出这些预报因子的预报能力没有明显差别. Xie 等人^[37]研究了双带耀斑的扩展速度和带位置纵向磁场测量之间的关系,发现它们之间存在着一个幂率关系.

基于大样本, Cui 等人^[38,39]对磁场参量和耀斑的相关性做了统计分析. 分别对从活动区的光球磁场中提取出太阳耀斑的 3 个预报因子(磁场最大水平梯度、中性线长度和孤立奇点个数),从矢量磁场中提取出太阳耀斑的 5 个预报因子(强梯度长度、强剪切长度、强梯度强剪切长度、总电流强度和总电流螺度). 统计结果指出,这些预报因子与耀斑有很强的相关性,可以用 Sigmoid 函数表示(图 3).

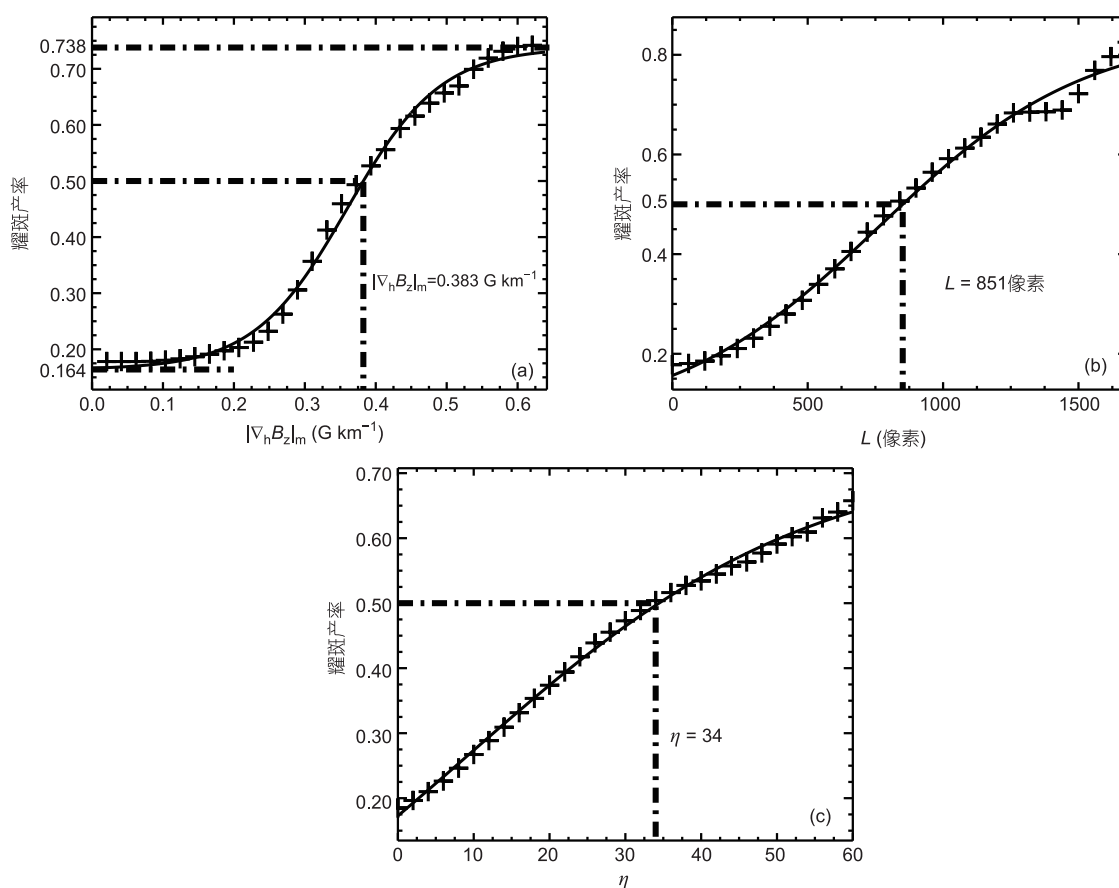


图3 预报因子与耀斑产率的相关性^[38]

(a) 纵向磁场最大水平梯度; (b) 强梯度中性线长度; (c) 孤立奇点数目. 根据与未来 24 小时耀斑产率之间函数映射

1.3 分形因子

混沌和分形是非线性科学的两大主题,混沌运动是一种非常普遍的非线性现象,具有初值放大效应、奇怪吸引子、不可线性迭加、结构自相似及分形几何效应.而分形是刻画混沌运动的直观几何语言,用来描述表明变化莫测而实质上存在有某种内在规律性的几何图像或形状^[40].

太阳活动区是许多能量和太阳活动,包括太阳耀斑和日冕物质抛射的来源.人们尝试了多种方法描述活动区的复杂性,从最初的 Wilson 磁分类到 McIntosh 分类均表明大的、更复杂的活动区产生数量多和大耀斑.然而这些描述方法基于主观形态学,后续研究集中用磁场来描述活动区的复杂性,磁场数据主要基于光球全日面纵向磁图.而这些研究没有把活动区作为一个整体或考虑耀斑的尺寸和频率.由于活动区具有自相似的特征,可以用分形的方法来量化活动区的复杂性^[41].

科研人员围绕用分形方法描述活动区复杂性和太阳耀斑展开了系列研究. Karakatsanis 等人^[42]的研究指出,太阳活动区的演变包含 2 个物理过程:一个是自组织临界,对应于光球下活动和湍动对流区;另一个则对应于光球活动的混沌过程.太阳数据的分形研究分为 4 种:周长面积、线性尺寸面积、计盒方法和多维分形. McAteer 等人^[41]用计盒方法研究了 SOHO 卫星 MDI 磁图每个活动区 Wilson 分类的计算分形维数和耀斑产率,发现每个 Wilson 分类有相似的分形维数频率分布,从而表明了活动区的自相似性.他们还具体给出分形维数一个较低的阈值 1.2~1.5,这个阈值是未来 24 h 产生 M 和 X 级耀斑的一个必要非充分条件. Aschwanden 课题组^[43,44]研究了太阳耀斑的面积分形维数和体积分析维数,结果发现, EUV 波长观测的耀斑面积展现了分形特征,全部的 EUV 流量和面积分形线性相关. Meunier^[45]研究了在太阳活动周不同时间尺度太阳耀斑活动水平的分形维数差异性,还有其他学者采用多维分形的方法研究了太阳磁场和太阳耀斑^[46~48].

1.4 其他预报因子

(i) 耀斑参量. 现存的耀斑预报模型大部分基于太阳活动区特征参量与耀斑之间的联系,采用的

预报因子为活动区的形态参量和磁场参量.这样选择的一个缺点是带有一定程度主观因素,包含了领域专家的知识,如磁分型和 McIntosh 分型.而在耀斑预报中存在一个趋势,即活动区在过去产生大耀斑,那么在将来也产生大耀斑,称之为持续性,这种持续性在未来 24 h 大耀斑预报中被认为是最可靠的预报因子^[49].

基于上述考虑, Wheatland^[49]提出,活动区发生所有耀斑的历史是一个重要因素,决定了将来活动区的发展,应该被用在耀斑预报中.他同时开发了贝叶斯方法预报耀斑发生概率,这个方法使用活动区过去时间序列中的耀斑观测记录和耀斑统计的现象规则作为初始预报,采用贝叶斯公式计算耀斑发生事件的后验概率分布.文献[50]描述了贝叶斯方法在全日面 GOES 事件预报中的应用,对 1976~2003 年的 GOES 数据进行了测试,并和美国国家海洋大气管理局(NOAA)1987~2003 年的预报结果作了详细比较.结果表明,尽管 NOAA 的预报采用了大量其他信息,而耀斑观测记录模型对于预报产生 M-X 级和 X 级耀斑的每日平均数优于 NOAA 方法;在预报 M-X 级的预报准确率略低于 NOAA,但具有更高的预报 X 级事件的准确率.

(ii) 预报因子组. Huang 等人^[51]提出了建立预报因子组的概念,是为了解决预报因子的高维问题.当耀斑系统的预报因子较多或在预报因子中引入时序信息时,耀斑预报因子的维度急剧扩展,预报因子间出现大量冗余信息,会面临数据维度过高的问题.这在机器学习中属于高维数据分类问题,这时求解空间大,很容易出现过拟合和训练时间长的情况.

为了解决这个问题,在粗糙集理论的框架下,对预报因子进行选择,形成预报因子组.具体方法为,以粗糙集中的正域大小作属性重要度的评价,以遗传算法做优化算法进行粗糙集属性约简,找到一系列利用较少属性且尽可能保持原始信息的特征子集,这些找到的特征子集被称为预报因子组.针对预报因子组,采用多模型融合的方法建立预报模型,基于不同的预报因子组建立耀斑预报的基模型后,再对所有结果进行融合,得到一个综合的输出结果.预报因子组的方法简化了所要求解的问题,提高了信息的利用率,取得了较好的结果.

2 预报方法研究现状

2.1 统计方法

耀斑预报的事件统计方法只利用已经观测到的事件对未来情况进行预报. 已观测到的太阳耀斑统计规律对耀斑预报具有重要指导意义.

1939 年 Giovanelli^[52]给出的一个活动区当天发生耀斑的概率表达式为 $P=2.6akf(i)$, 其中 a 为当天的黑子面积, k 和 $f(i)$ 分别代表与黑子群磁型和黑子面积变化率有关的参量. Drake^[53]提出了耀斑峰值流量服从幂率分布, 表示为 $N(S)=\text{const} \times S^{-\gamma}$. 其中 $N(S)$ 为每单位时间每峰值流量 S 中的事件数, const 表示常数. 通常幂指数 γ 比 2 稍小^[54]. Bai^[55]的研究显示, 假设在预报时间段内, 耀斑发生服从泊松分布, X 级耀斑发生的概率为 $\varepsilon_X=1-e^{-\gamma_X T}$, 其中, 幂指数 γ_X 可以用极大似然方法估计.

Wheatland^[49]提出了用贝叶斯方法进行耀斑预报. 这个方法使用一个活动区的耀斑发生记录和耀斑统计的现象规则, 对大耀斑在子序列时间段发生的情况产生一个最初预测, 给定一个发生在 $t_1, t_2, \dots, t_M$ 大于 S_1 的事件序列 $s_1, s_2, \dots, s_M$, 应用贝叶斯事件统计计算 ϵ 的后验分布 $P(\epsilon)$, 再通过现存的耀斑预报方法进行评价.

Bornmann 等人^[56]指出, 空间服务环境中心(space environment services center, SESC)使用活动区的 McIntosh 分型指导耀斑预报, 用多元线性回归法分析 McIntosh 分型参数对耀斑的贡献. 当假设变量间为乘性关系时, 变量和决策间可以达到好的相关性. McIntosh 参数反映了活动区磁场的性质, 第一个参数提供了活动区总磁通量的测量, 第二个参数反映了最大黑子的磁通量, 第三个参数反映了黑子群的总面积.

美国大熊湖天文台的 Gallagher 等人^[25]根据第 22 太阳活动周开始以来约 8 年的 NOAA 资料, 根据 McIntosh 分类的 3 个参量(Z, p, c), 对每一个活动区产生 C 级、M 级或 X 级的 X 射线耀斑的可能性进行预报. 其预报过程如下: 首先利用从 1988 年 11 月到 1996 年 6 月约 8 年的耀斑和活动区数据, 统计 McIntosh 分型中参数的平均日耀斑率. 因为耀斑发生频率符合泊松统计, 所以已知平均耀斑率就可以计算 24 h 内活动区 McIntosh 分型中某类产生一个或多个 C, M 或 X 类耀斑事件的概率. 设 μ 是单位时间

内平均事件数, 利用泊松分布给出 24 h 内观测到一个或多个时间的概率为 $P_\mu(N \geq 1) = 1 - \exp(-\mu)$.

Wheatland^[49]给出了贝叶斯事件统计方法在全日面 GOES 事件预报中的应用, 并对 1976~2003 年的 GOES 数据进行了测试. Bartkowiak 等人^[57]使用多元回归方法预报太阳耀斑活动, 使用 3 种距离基于回归方法, 结果表明, 使用 L1 范数或 Gower 距离度量的性能比使用欧氏距离效果好. Leka 等人^[32]利用 Fisher 判别分析给出了各种磁场参量与耀斑的统计关系, 并给出了初步测试结果.

耀斑预报在我国开展的比较早, 北京天文台的 Zhang 等人^[58]利用多元回归方法建立了太阳耀斑预报模型, 所用预报因子涉及: (1) 太阳活动区光球纵向场常规观测资料; (2) 太阳黑子的光学观测数据, 包括太阳黑子群面积、磁分类和 McIntosh 分类; (3) 10 cm 射电日流量的常规观测数据. 根据 1988 年 4 月到 1990 年 10 月的观测资料, 可预报未来 48 h 的 M 和 X 级耀斑发生概率. 此方法很长一段时间作为太阳活动预报中心的每日预报主要参考工具. Zhu 等人^[59]对采用这种方法的实际预报结果给出了统计, 并和美国 WWA 的预报结果进行了比对, 结果显示, 该方法的预报准确率高于 WWA 的预报结果.

2.2 机器学习方法

近年来, 随着数据挖掘技术的迅速发展和在各个领域的成功应用, 越来越多的机器学习方法应用到太阳活动预报中.

基于 McIntosh 分类, NOAA 空间环境实验室和 Colorado 大学联合开发了一个专家系统 Theo 来预报太阳耀斑^[19]. 利用与 Theo 相同的预报因子, Bradshaw 等人^[60]利用人工神经网络技术建立了耀斑预报模型, 可与 Theo 达到相同的预报精度. 英国 Bradford 大学的 Qahwaji 和 Colak^[61]应用神经网络和支持向量机算法建立了混合耀斑短期预报系统, 以 McIntosh 分型和太阳黑子数为预报因子, 首先利用支持向量机预报是否有耀斑发生, 如果有耀斑发生, 再利用神经网络预报发生耀斑的级别是多大. Baena 等人^[62]应用行为模式学习方法挖掘耀斑产生和活动区数据之间的关系, 这个关系以规则的形式表示出来, 比较容易理解. 此外, Leka 和 Barnes^[34]对从光球磁场提取的特征参量进行了研究, 并应用 Fishers 线型判别预报耀斑发生. Song 等人^[63]选取总磁通、强梯度中心线

长度和总磁场能量分布参量应用 logistic 回归预报耀斑. 基于同样数据, Yuan 等人^[64]采用 logistic 回归和支持向量机相结合建立耀斑预报模型, 首先使用 logistic 回归预报耀斑发生级别的概率, 再将输出的概率作为输入, 应用支持向量机预报耀斑级别. Ahmed 等人^[65]选取了多个和耀斑相关性强的光球磁场参量, 由特征选择算法挑出有效因子, 应用神经网络方法预报耀斑.

中国科学院国家天文台太阳活动预报中心在应用机器学习方法预报太阳耀斑方面做了一系列的工作. 选取文献[59]主要的预报因子, Li 等人^[66]应用支持向量机-k 近邻方法建立了一个耀斑短期预报系统, 并得到良好的预报精度. Wang 等人^[67]利用从光球磁图中提取的 3 个预报因子, 纵向磁场最大水平梯度、中性线长度和孤立奇点个数, 使用应用多层感知器建立了一个耀斑短期预报系统, 预报精度能够达到有长期预报经验的预报人员的水平. Huang 等人^[68-71]应用几种机器学习方法建立了耀斑预报模型, 结果显示了磁场数据时间序列演化信息在耀斑预报中的有效性. Li 等人^[72]在耀斑预报建模中采用神经网络和无监督聚类相结合的方法, 结果表明了聚类算法在耀斑预报建模中的重要性. 另一项工作是结合光球磁场参量和黑子参量建立了耀斑预报模型, 模拟预报结果显示采用综合预报因子比单独使用光球磁场参量和黑子参量作为预报因子更为有效^[73].

此外, 太阳活动在经度和纬度的分布上存在一定规律^[74], 太阳活动经度本质上给出太阳活动更加频繁的区域. 考虑到选取太阳活动经度带数据作为预报因子可以进一步提高太阳耀斑的预报精度, Huang 等人^[75,76]在除了选取太阳光球磁场参量外, 添加活动区中心到活动经度带中心的距离作为新的预报因子, 并验证了其有效性.

2.3 数据同化方法

数据同化起源于 20 世纪中期, 随着 20 世纪 80 年代全球对地观测能力的提高和全球环境变化研究需求而日益引起关注, 当前数据同化已经发展为地球系统科学研究的重要方法之一, 被广泛应用于大气、海洋、陆面和生态等多个领域^[77].

数据同化的根本思想是用同化数据引入对模型进行更正, 常表现为数值模拟和物理现象, 典型应用是用误差数据的更正来产生预报. 数据同化分为两

种, 连续数据同化算法和顺序同化算法. 连续数据同化算法定义一个同化的时间窗口 T , 利用该同化窗口内的所有观测数据和模型状态值进行最优估计, 通过迭代而不断调整模型初始场, 最终将模型轨迹拟合到在同化窗口周期内获取的所有观测上, 如三维变分和四维变分算法等. 顺序数据同化算法基于卡尔曼滤波器, 用模型误差和观测误差统计来发现模型和观测数据的最优组合.

Béanger 等人^[78]探讨应用四维变分数据同化技术来预报太阳耀斑. 变分算法构建代价函数描述状态量预报与观测之间的差异, 利用变分思想把数据同化问题转为一个极值求解问题, 在满足动态约束条件下, 最小化代价函数, 使得误差最小的状态量即为最优状态估计量. 最小化代价函数通常运行最小化算法, 包括最优迭代步长法、共轭梯度法和牛顿法. 使用变分方法需要引入一个模型, 在 Lu 等人^[79]研究基础上, 将耀斑发生视为自组织临界模型, 具体化为时间序列能量释放的雪崩模型. 将时间序列的能量释放作为数据同化目标, 建立相应的代价函数, 使用共轭梯度算法最小化代价函数得到最优的状态变量, 作为自组织临界模型的初始值. 实验比较了基于 SOC 模型的 DNS 预报和利用四维变分法进行数据同化的预报. 两种方法虽然使用了相同的最初始条件, 但不同自组织临界模型的初始条件驱动产生不同的时间序列能量释放, 实验结果四维变分法的预报和观测基本符合, 而 DNS 漏报了一个大的雪崩事件. 由此可以看到, 四维变分法成功产生雪崩模型的最优初始条件, 从而再生了耀斑雪崩模型的时间序列的能量释放, 为耀斑预报铺垫了基础.

3 耀斑预报模型研究

3.1 静态模型

目前太阳耀斑预报模型大致上为两类: 一类是基于专家系统的模型, 另一类是利用观测数据建立的模型^[24]. 基于专家系统的预报方法通过领域专家提供耀斑预报知识来建立预报规则, 该方法的预报精度理论上能够达到领域专家的预报水平. 专家系统的预报精度在很大程度上受制于专家经验的丰富程度, 缺乏客观性. 目前, 关于耀斑预报专家系统的研究呈现出逐步减少的趋势. 从数据中建模的利用耀斑爆发观测数据, 提取预报因子, 然后利用各种数

据统计和数据挖掘技术建立预报因子与耀斑发生之间的关系模型, 利用建立的模型对耀斑的发生进行预报. 随着数据挖掘方法的发展, 这类方法在耀斑预报中得到了广泛应用.

耀斑预报问题可以作为一个建模问题. 模型输入时从数据中选取的预报因子, 可以是太阳黑子参量和磁场参量. 选取和计算出来的参量大都是计算出的原始参量, 不适合直接建立预报模型, 需要对其进行规范化. 以图 4 为例, 采用黑子参量和太阳射电流量作为预报因子, 首先利用这些参量与耀斑直接的产率关系, 把原始数据规范化, 再带入机器学习算法建立预报模型. 图中预报因子组(Area, Wils, McIn, Flux)分别为黑子面积、磁分型、McIntosh 分类和射电流量, (Area', Wils', McIn', Flux')是规范后的数据.

现阶段大部分耀斑预报模型基本都是静态模型, 系统输出只与当前时刻的输入有关. 即从数据中建立当前时刻的观测值与预报时段内耀斑情况的关系, 用公式可以表示为

$$y = f(p(t)), \quad (3)$$

其中, y 是预报时段内耀斑的情况, f 为输入和输出间的关系, 可以具体化为一种机器学习算法; $p(t)$ 为当前时刻的观测向量, 即预报因子.

3.2 动态模型

考虑到静态耀斑预报模型没有包含活动区观测的动态演化信息, 近年来提出了建立太阳耀斑预报的动态模型. 建立模型不仅需要考虑当前活动区的观测情况, 而且需要考虑之前活动区的影响, 即在当前观测下判断是否产生耀斑还需要看之前观测的情况. 平滑窗技术是建立动态模型的主要方法, 可以表

示为

$$y = f(p(t), p(t - \Delta t), \dots, p(t - w\Delta t)), \quad (4)$$

其中, t 为当前时刻的观测量, 随着时间的推移, $t - w\Delta t$ 到 t 间的观测随时间滑动, 称为滑动窗, w 为滑动窗的尺寸. 在不同时刻 t 分别提取活动区参量 $p(t)$, 然后按照时间排开, 可形成预报因子流($p(t)$, $p(t - \Delta t)$, $\dots$, $p(t - w\Delta t)$). 这些预报因子流能够反映活动区的演化信息.

动态模型介绍如图 5 所示, 同样以黑子数据作为预报因子为例, 首先将原始数据带入耀斑产率拟合得到规范化数据, 然后利用平滑窗技术得到数据的时间序列演化数据, 如黑子面积 Area' 变为预报因子流 $\{Area'(t), Area'(t - \Delta t), \dots, Area'(t - w\Delta t)\}$, 最后将得到的时间序列数据带入机器学习算法建立预报模型. 由图 5 看出, 模型输入数据的维数会因为引入演化数据而成倍增加.

由于动态耀斑预报模型采用了活动区的演化数据, 预报精度得到明显提高. Yu 等人^[68]采用平滑窗技术建立了耀斑动态模型, 通过互信息函数推出最佳的平滑窗尺寸. 应用决策树 C4.5 和学习矢量量化(LVQ)技术建立预报模型, 比较了静态模型和动态模型的预报性能. 结果表明, 采用动态模型对耀斑发生的报准率和耀斑未发生的报准率都有大幅度显著提高. 而他们后续研究中, 为解决因引入时间序列数据引起的模型输入维数增加问题, 应用小波变换将太阳活动区光球磁场序列数据进行多尺度分析, 在每个频段对分解后的序列进行特征提取^[69]. 结果发现, 高频段和低频段的预报因子包含了预报因子的主要信息量. 采用这两个频段的数据建立模型看可起到降维作用, 解决了高维数据建模速度慢的问题, 同时提高了预报精度. Li 等人^[80]对太阳黑子数据建立了动态模型, 其结果验证了黑子活动区的时间演化信息对于太阳耀斑预报同样重要.

3.3 自组织临界模型

自组织临界理论(self-organized criticality, SOC)是美国物理学 Bak 和 Tang^[81]于 1987 提出的理论. 由大量相互作用成分组成的系统会自然地自组织临界态发展, 当系统达到自组织临界态时, 即使小的干扰事件也可引起系统发生一系列巨变. “自组织”是指该状态的形成主要是由系统内部组织间的相互作用产生, 而不是由任何外界因素控制或主导所致. “临

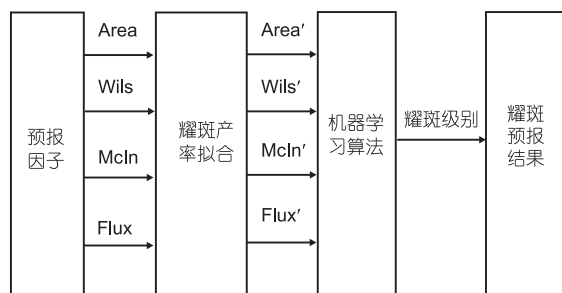


图4 太阳耀斑的静态预报模型结构

预报因子序列{Area, Wils, McIn, Flux}分别对应太阳黑子、磁分类、McIntosh 分型、射电流量的观测初始值; {Area', Wils', McIn', Flux'}是耀斑产率拟合后的{0,1}之间的规范化值

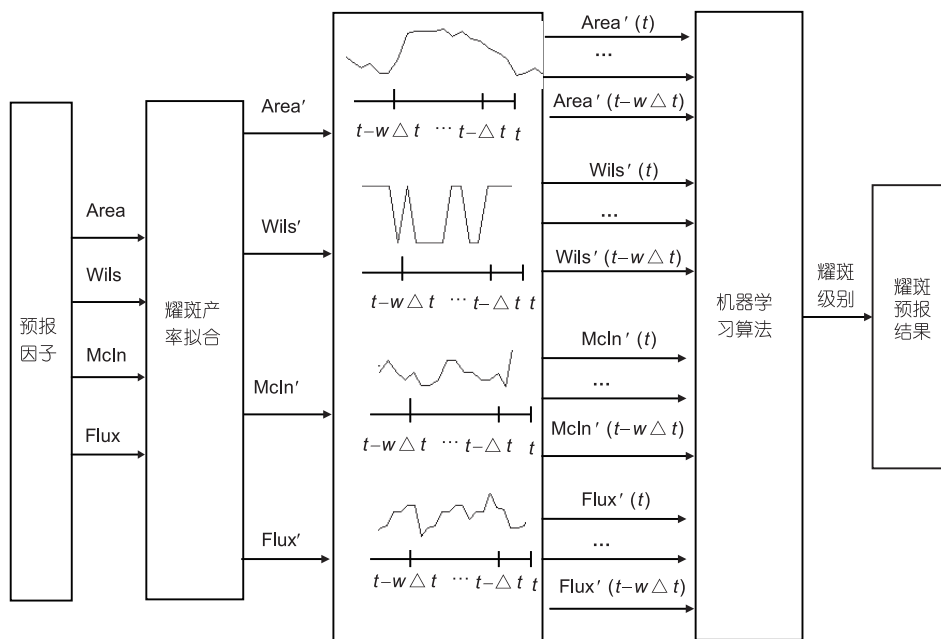


图 5 太阳耀斑的动态预报模型结构

预报因子序列{Area, Wils, McIn, Flux}分别对应太阳黑子、磁分类、McIntosh 分型、射电流量的观测初始值; {Area', Wils', McIn', Flux'}是耀斑产率拟合后的{0,1}之间的规范化值. 通过平滑窗技术, 预报因子添加了时间序列数据, 如黑子面积 Area'形成了预报因子流 Area'(t), Area'(t-Δt), ..., Area'(t-wΔt)

界态”是指系统处于一种特殊敏感状态, 微小的局部变化可以不断放大并扩延至整个系统. 也就是说, 系统在临界态时, 其所有成分的行为都相互关联, 系统内事件大小与其频率之间是幂函数关系^[78].

一个简单的例子是沙堆模型, 当向沙堆添加沙子时, 沙堆的平均斜坡增加, 直到达到近似常量的状态. 一旦自组织临界状态达到了, 额外的沙砾会引起雪崩, 导致沙堆发生整体性的连锁改变或重新组合, 沙堆的结构将随每粒新沙落下而变得脆弱, 最终发生结构性失衡. 临界态时, 沙崩规模的大小与其出现的频率呈幂函数关系.

众所周知, 太阳耀斑发生频率与其峰值强度之间存在幂率谱的关系, 此规律与太阳周无关. Lu 等人^[79]认为, 日冕磁场类似于沙堆, 由于光球对流运动产生缓慢, 随机扭转起着沙堆中额外沙砾的作用. 当磁场参量在电流层相反面的不连续夹角小于临界角, 由于日冕等离子体的传导性, 磁重联缓慢进行, 当夹角大于临界角, 磁重联爆发性的进行. 临界磁场的强度和拓扑也随之改变, 而周边磁场电流层的不连续夹角会增加, 大于临界角时引发额外的磁重联事件. 这时日冕磁场整个区域达到不稳定的边缘状

态, 称之为磁重联雪崩.

耀斑由许多小事件组成, 也可以说日冕被许多小的非热能事件加热. 通常认为存储于日冕磁场的磁能是耀斑的能量源. 每个磁重联事件释放磁场能量, 通过能量释放, 磁场的梯度小于临界状态值. 当另外一个磁场向量加入时, 这个过程被重复, 磁场被驱动到一个临界状态.

4 展望

综上所述, 太阳耀斑预报研究技术, 包括预报因子、预报方法和预报模型, 近年来得到了很大发展. 现有的预报因子描述了活动区的非势性和复杂性, 具有一定的预报能力. 但由于预报因子之间存在一定的相关性, 组合预报模型的预报能力有限. 考虑到 CME 与耀斑之间的紧密联系^[82], 今后研究将不局限于从磁场中提取因子, 还要从日冕中提取有效因子. 预报模型刻画了耀斑的产生, 现有 SOC 模型和数据同化方法没有实际预报能力, 可以对其进一步研究. 而随着人工智能技术的发展, 会不断出现新的机器学习方法, 这些方法将具有更好的泛化能力, 从而建立具有更强预报能力的预报模型.

另一方面, 依据物理规律的太阳活动预报模式处于探索阶段. 运用磁流体力学方程并利用观测数据为边界条件对太阳活动进行数值模拟或仿真, 对提高太阳活动预报能力具有十分重要的意义. 在未来 20 年, 太阳耀斑预报模式将从现在的以统计预报模式为主, 过渡到基于物理模型的预报为主. 随着太阳物理学不断取得进展, 发电机模型将会更加有

效地描述太阳内部物质运动, 日冕物质抛射和耀斑模型将会更加准确描述日冕中的爆发过程. 随着国内外更为先进的太阳观测仪器投入使用^[83,84], 通过时序观测数据和数值模拟技术, 可实现太阳活动区演化仿真和太阳爆发前物理演变过程模拟, 由此寻求太阳爆发临界点, 进而形成基于数据驱动模拟的太阳耀斑预报技术.

参考文献

- 1 林元章. 太阳物理学导论. 北京: 科学出版社, 2000
- 2 Zhao P, Zhou X J, Liu G. Decadal-centennial-scale change in Asian-Pacific summer thermal contrast and solar activity. *Chin Sci Bull*, 2011, 56: 3012–3018
- 3 Wu J Y, Shao X H, Kong X G, et al. Imprint of solar activity on Nanjing stalagmite annual layer thickness sequence during the Last Glacial Maximum. *Chin Sci Bull*, 2006, 51: 441–447
- 4 Liao D H, Liao X H. New evidence for possible impact of solar activity on long-term fluctuation of the earth rotation. *Chin Sci Bull*, 2001, 46: 905–908
- 5 郑玲, 傅绥燕, 宗秋刚, 等. 极光半球能量季节变化与太阳活动周期的关联. *科学通报*, 2013, 58: 853
- 6 王跃, 翦知湔, 赵平, 等. 全新世太阳活动驱动的太平洋上层热力结构的瞬变演化. *科学通报*, 2013, 58: 379–384
- 7 冯博, 韩延本. 太阳活动对黄帝陵 500 年侧柏年轮变化的可能影响. *中国科学 G 辑: 物理学 力学 天文学*, 2009, 39: 776–784
- 8 陈刘成, 胡彩波, 谢廷峰, 等. 太阳风暴对卫星导航系统的影响分析. *中国科学: 物理学 力学 天文学*, 2011, 41: 556–563
- 9 刘立波, 万卫星, 陈一定, 等. 电离层与太阳活动性关系. *科学通报*, 2011, 56: 477–487
- 10 Liu L B, Wan W X, Chen Y D, et al. Solar activity effects of the ionosphere: A brief review. *Chin Sci Bull*, 2011, 56: 1202–1211
- 11 赵海娟. 太阳活动预报. 硕士学位论文. 北京: 中国科学院研究生院, 2004
- 12 Ning Z J. Power conversion factor in solar flares. *Chin Sci Bull*, 2012, 57: 1397–1404
- 13 Wang J X. Solar activity studies: From a magnetohydrodynamics description to a plasma perspective. *Chin Sci Bull*, 2012, 57: 1362–1368
- 14 Lin J. Studies of solar flares and CMEs related to the space solar missions in the future. *Sci China Ser G-Phys Mech Astron*, 2009, 52: 1646–1654
- 15 Ning Z J. The investigation of the Neupert effect in two solar flares. *Sci China Ser G-Phys Mech Astron*, 2009, 52: 1686–1690
- 16 Yin Z Q, Ma L H, Han Y B, et al. Long-term variations of solar activity. *Chin Sci Bull*, 2007, 52: 2737–2741
- 17 杨若文, 曹杰, 黄玮, 等. 太阳常数与太阳黑子数关系的交叉小波分析. *科学通报*, 2009, 54: 871–875
- 18 王家龙. 第 24 太阳周将是一个低太阳周? *科学通报*, 2009, 54: 3664–3668
- 19 McIntosh P S. The classification of sunspot groups. *Solar Phys*, 1990, 125: 251–267
- 20 Sammis I, Tang F, Zirin H. The dependence of large flare occurrence on the magnetic structure of sunspots. *Astrophys J*, 2000, 540: 583–587
- 21 Atac T. Statistical relationship between sunspots and major flares. *Astrophys Space Sci*, 1987, 129: 203–208
- 22 赵明宇, 陈军权, 刘煜, 等. 太阳活动峰年和谷年期间黑子群与耀斑的统计分析. *中国科学: 物理学 力学 天文学*, 2014, 44: 109–120
- 23 李蓉, 崔延美. 应用数据挖掘技术的短期太阳耀斑预报模型. *中国科学: 物理学 力学 天文学*, 2011, 11: 1342–1350
- 24 黄鑫. 基于序列数据的太阳耀斑预报方法研究. 博士学位论文. 哈尔滨: 哈尔滨工业大学, 2010
- 25 Gallagher P T, Moon Y, Wang H M. Active-region monitoring and flare forecasting I. Data processing and first results. *Solar Phys*, 2002, 209: 171–183
- 26 Abramenko V, Yurchyshyn V, Wang H, et al. Scaling behavior of structure functions of the longitudinal magnetic field in active regions on the sun. *Astrophys J*, 2002, 577: 487–495
- 27 Komm R, Hill F. Solar flares and solar subphotospheric vorticity. *J Geophys Res Space Phys*, 2009, 114: A06105
- 28 Hagyard M. The significance of vector magnetic field measurements. *Soc Astron Ital*, 1990, 61: 337–357
- 29 Schmieder B, Hagyard M, Ai G, et al. Relationship between magnetic field evolution and flaring sites in AR 6659 in June 1991. *Solar Phys*, 1994, 150: 199–219
- 30 Wheatland M. A test to confirm the source of energy for solar flares. *Pub Astron Soc Austr*, 2001, 18: 351–354
- 31 崔延美. 太阳光球磁场特性与耀斑相关性研究. 博士学位论文. 北京: 中国科学院研究生院, 2007

- 32 Leka K D, Barnes G. Photospheric magnetic field properties of flaring versus flare-quiet active regions. I. Data, general approach, and sample results. *Astrophys J*, 2003, 595: 1277–1295
- 33 Leka K D, Barnes G. Photospheric magnetic field properties of flaring versus flare-quiet active regions. II. Discriminant analysis. *Astrophys J*, 2003, 585: 1296–1306
- 34 Leka K D, Branes G. Photospheric magnetic field properties of flaring versus flare-quiet active regions. IV. A statistically significant sample. *Astrophys J*, 2007, 656: 1173–1186
- 35 Georgoulis M K, Rust D M. Quantitative forecasting of major solar flares. *Astrophys J*, 2007, 661: L109–L112
- 36 Barnes G, Leka K D. Evaluating the preformance of solar flare forecasting methods. *Astrophys J*, 2008, 688: L107–L110
- 37 Xie W B, Wang H M, Jing J, et al. The correlation between expansion speed and magnetic field in solar flare ribbons. *Sci China Ser G-Phys Mech Astron*, 2009, 52: 1754–1759
- 38 Cui Y M, Li R, Zhang L Y, et al. Correlation between solar flare productivity and photospheric magnetic field properties. 1. Maximum horizontal gradient, length of neutral line, number of singular points. *Solar Phys*, 2006, 237: 45–59
- 39 Cui Y M, Li R, He H, et al. Correlation between solar flare productivity and photospheric magnetic field properties. II Magnetic gradient and magnetic Shear. *Solar Phys*, 2007, 242: 1–8
- 40 申学会, 陈举华. 分形与混沌理论在湍流研究中的应用. *河南科技大学学报(自然科学版)*, 2005, 26: 27–30
- 41 McAteer R T, Gallagher P T, Ireland J. Statistics of active region complexity: A large-scale fracta dimension survey. *Astrophys J*, 2005, 631: 628–635
- 42 Karakatsanis L P, Pavlos G P. SOC and chaos into the solar activity. *Nonlin Phenom Compl System*, 2008, 11: 280–284
- 43 Aschwanden M J, Aschwanden P D. Solar flare geometries. I. The area fractal dimension. *Astrophys J*, 2008, 674: 530–543
- 44 Aschwanden M J, Aschwanden P D. Solar flare geometries. II. The volume fractal dimension. *Astrophys J*, 2008, 674: 544–553
- 45 Meunier N. Complexity of magnetic structures: Flares and cycle phase dependence. *Astron Astrophys*, 2004, 420: 333–342
- 46 Lawrence J K, Ruzmaikin A A, Candavid A C. Multifractal measure of the solar magnetic field. *Astrophys J*, 1993, 417: 805–811
- 47 Conlon P A, Gallagher P T, McAteer R T J, et al. Multifractal properties of evolving active regions. *Solar Phys*, 2008, 248: 297–309
- 48 Yu Z G, Anh V, Eastes R, et al. Multifractal analysis of flare indices and their horizontal visibility graphs. *Nonlin Proc Geophys*, 2012, 19: 657–665
- 49 Wheatland M S. A bayesian approach to solar flare prediction. *Astrophys J*, 2004, 609: 1134–1139
- 50 Wheatland M. A statistical solar flare forecast method. *Space Weath*, 2005, 3: S07003
- 51 Huang X, Yu D, Hu Q. Short-term solar flare prediction using predictor teams. *Solar Phys*, 2010, 263: 175–184
- 52 Giovanelli R G. The relations between eruptions and sunspots. *Astrophys J*, 1939, 89: 2555–567
- 53 Drake J. Characteristics of soft solar X-ray bursts. *Solar Phys*, 1971, 16: 152–185
- 54 Aschwanden M, Dennis B, Benz A. Logistic avalanche processes, elementary time structures, and frequency distributions in solar flares. *Astrophys J*, 1998, 497: 972–993
- 55 Bai T. Variability of the occurrence frequency of solar flares as a function of peak hard X-ray rate. *Astrophys J*, 1993, 404: 805–809
- 56 Bornmann P, Shaw D. Flare rates and the McIntosh active-region classifications. *Solar Phys*, 1994, 150: 127–146
- 57 Bartkowiak A, Jakimiec M. Distance-based regression in prediction of solar flare activity. *Questiio*, 1994, 18: 7–12
- 58 Zhang G Q, Wang J L. A new scheme used for the short-term prediction of X-xary flares. *Progr Geophys*, 1994, 9: 54–58
- 59 Zhu C L, Wang J L. Verification of short-term prediction of solar X-ray bursts for the maximum phase (2000–2001) of solar cycle 23. *Chin J Astron Astrophys*, 2003, 3: 563–568
- 60 Bradshaw G, Fozzard R, Ceci L. A connectionist expert system that actually works. *Adv Neu Inform Proc Sys*, 1989, 1: 248–255
- 61 Qahwaji R, Colak T. Automatic short-term solar flare prediction using machine learning and sunspot associations. *Solar phys*, 2007, 241: 195–211
- 62 Baena M, Morales R. The influence of active region information on the prediction of solar flares: An empirical model using data mining. *Annales Geophys*, 2005, 3129–3138
- 63 Song H, Tan C, Jing J, et al. Statistical assessment of photospheric magnetic features in imminent solar flare predictions. *Solar Phys*, 2008, 254: 101–125
- 64 Yuan Y, Shih F Y, Jing J, et al. Automated flare forecasting using a statistical learning technique. *Res Astron Astrophys*, 2010, 10: 785–796
- 65 Ahmed O W, Qahwaji R, Colak T, et al. Solar flare prediction using advanced feature extraction, machine learning, and feature selection. *Solar Phys*, 2013, 283: 157–175
- 66 Li R, He H, Cui Y M, et al. Support vector machine combined with K-nearest neighbors for solar flare forecast. *Chin J Astron Astrophys*, 2007, 7: 441–447

- 67 Wang H N, Cui Y M, Li R, et al. Solar flare forecasting model supported with artificial neural network techniques. *Adv Space Res*, 2007, 42: 1464–1468
- 68 Yu D R, Huang X, Wang H N, et al. Short-term solar flare prediction using a sequential supervised learning method. *Solar Phys*, 2009, 255: 91–105
- 69 Yu D R, Huang X, Wang H N, et al. Short-term solar flare prediction using multi-resolution predictors. *Astrophys J*, 2010, 709: 321–326
- 70 Yu D, Huang X, Wang H N, et al. Short-term solar flare level prediction using a bayesian network approach. *Astrophys J*, 2010, 710: 869–877
- 71 Huang X, Wang H N, Dai X H. Influences of misprediction costs on solar flare prediction. *Sci China-Phys Mech Astron*, 2012, 55: 1956–1962
- 72 Li R, Wang H N, Huang X, et al. Solar flare forecasting using learning vector quantity and unsupervised clustering techniques. *Sci China-Phys Mech Astron*, 2011, 54: 1367–1554
- 73 李蓉, 朱杰, 崔延美. 结合活动区光球磁场参量和黑子参量的太阳耀斑预报模型. *科学通报*, 2013, 58: 1845–1850
- 74 Gao P X, Li K J, Li Q X. Latitude migration of solar activity at high latitudes. *Chin Sci Bull*, 2008, 53: 8–11
- 75 Huang X, Zhang L Y, Wang H N. Improving the performance of solar flare prediction using active longitudes information. *Astron Astrophys*, 2012, 549: A127
- 76 Huang X, Wang H N. Solar flare prediction using highly stressed longitudinal magnetic field parameters. *Res Astron Astrophys*, 2013, 13: 351–358
- 77 马建文. 数据同化算法研发与是实验. 北京: 科学出版社, 2013
- 78 Bélanger E, Vincent A, Charbonneau P. Predicting solar flares by data assimilation in avalanche models. I. Model design and validation. *Solar Phys*, 2007, 245: 141–165
- 79 Lu E D, Hamilton R J. Avalanches and the distribution of solar flares. *Astrophys J*, 1991, 380: L89–L92
- 80 Li R, Zhu J. Solar flare forecasting based on sequential sunspot data. *Res Astron Astrophys*, 2013, 9: 1118–1126
- 81 Bak P, Tang C, Wiesenfeld K. Self organized criticality: An explanation of $1/f$ noise. *Phys Rev Lett*, 1987, 59: 381–384
- 82 陈耀. 日冕动力学研究进展: 冕流、日冕物质抛射及其相互作用. *科学通报*, 2013, 58: 1620–1650
- 83 张双南. 我国空间天文发展的现状和展望. *中国科学: 物理学 力学 天文学*, 2012, 42: 1308–1320
- 84 甘为群, 黄宇, 颜毅华. 太阳空间探测的过去与未来. *中国科学: 物理学 力学 天文学*, 2012, 42: 1274–1281

Progress in solar flare forecasting

LI Rong¹, ZHU Jie¹ & HUANG Xin²

¹*School of Information, Beijing Wuzi University, Beijing 101149, China;*

²*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China*

Solar flare is an abrupt and scale of energy release process in solar surface local region. It is a main disturbance resource of space environment and has tremendous influence. Existed solar flare forecasting models include obtaining predictor from observed data, constructing relation model of predictors and flare occurrence by using statistic or data mining method, and predicting future flare occurrence with this model. In solar flare forecasting research, the predictor, the forecasting method and model are three main sides. As an important part, predictor and data process are preprocessing work. Predictors usually select solar sunspot parameter, magnetic parameters and fractal dimension and so on. Forecasting methods include statistic method, machine learning method and data assimilation method. Statistic method is mainly used in early model. With the development of data mining technique, more and more machine learning methods are concerned and receive satisfied result. Data assimilation method has good model correction ability. In forecasting model, most of them are static model. Recently, the time revolution dynamic model is developed and has better performance. Besides, self-organized model developed recently gives a physical description about burst mechanism of solar flare. This paper summarizes the research progress in these three sides of flare forecasting. Connected with the work in solar activity prediction center of National Astronomical Observatories, important research progresses are reviewed. Finally, future research trend is prospected.

solar flare, predictors, machine learning, data mining, forecasting model

doi: 10.1360/N972014-00058