

# 機械学習を用いたリアルタイム太陽フレア予測

○西塚直人・久保勇樹・杉浦孔明・田光江・亘慎一・石井守  
情報通信研究機構 (NICT)

宇宙天気現象の源は太陽活動であり、太陽フレアのより早く正確な予報は宇宙天気予報にとって重要な課題である。近年、様々な太陽衛星観測によって、太陽表面の詳細なデータがリアルタイムで取得可能になってきた。我々は、これらのデータをもとに黒点磁場の複雑性や彩層発光現象を判別し、機械学習を用いて今後24時間以内に発生する太陽フレアの予測モデルを開発した。本講演では我々の予測モデルの紹介と将来展望について議論する。

## 1. はじめに

太陽面爆発現象フレアは、宇宙天気の源である。太陽フレアが発生すると、地球にも大量のX線や紫外線、有害な高エネルギー粒子が到達し、地磁気嵐を起こす。その結果、衛星故障や大停電を引き起こすことがある。特に太陽フレアの発生と同時に起こるデリンジャー現象や極冠吸収は航空通信の阻害ともなるが、フレアの予測なしには回避不能である。また、コロナ質量放出(CME)到来の場合も2・3日しかインフラ対応のための時間がない。従って、太陽フレアの発生をより早く正確に予測する必要性があり、かつ宇宙利用が増大する現代において、宇宙天気予報は必要なリスク管理といえる。

既存の宇宙天気予報(フレア予測)の手法には大きく分けて2タイプある。1つは、物理法則に基づいた数値シミュレーションによる予測である。もう1つは、観測量から物理的知見に基づいた経験的・確率的な予測であり、数百年間研究されてきた太陽物理学の知識や経験則から重要だと思われる具体的な物理量を用いている。

太陽フレアの発生は100年以上の謎であり、まだその発生機構は完全に理解されていない。太陽内部から表面に磁場が浮上してくることで、太陽黒点は形成される。さらに黒点の成長と拡大・複雑化に伴って、黒点上空の

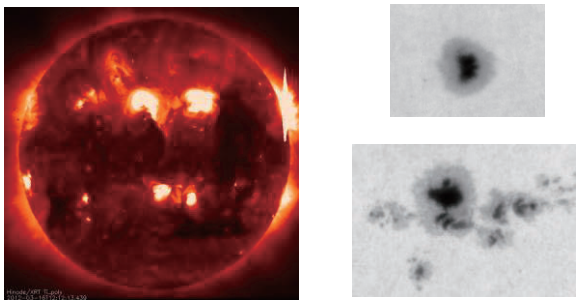


図1: 太陽軟X線全面像(Hinode衛星)と白色光による黒点観測(京都大学飛騨天文台 SMART)

コロナではX線放射を伴うフレアが観測されている(図1)。さらに近年は白色光による黒点観測だけでなく、Hinode衛星やSDO衛星による光球ベクトル磁場の定常観測が可能により、大規模(Xクラス)フレアが起きる直前の黒点磁場の複雑化と磁気シアの増大がより詳細に明らかになってきた。

黒点は常にN極とS極とが対になった構造をしている。図2の背景の白黒カラーは正負の視線方向磁場(S極・N極)を表している。白黒の境界線を磁気中性線と呼ぶ。磁気中性線は地震という断層面に相当し、ここに磁気シアとよばれる歪みが蓄積されるのが観測される(図中の赤矢印が水平磁場を表す)。さらに磁気中性線付近に浮上磁場が現われると、それまで安定的であった黒点磁場構造が不安定化してフレアが起きる。磁気中性線付近で局所的に時間変化する構造を監視することで、フレアのトリガー機構を捉えられる可能性がでてきた。

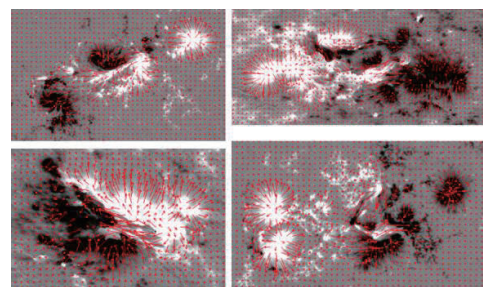


図2: Xフレア発生前のベクトル磁場観測(SDO衛星)。背景白黒は視線方向磁場、赤矢印は水平磁場を表す。

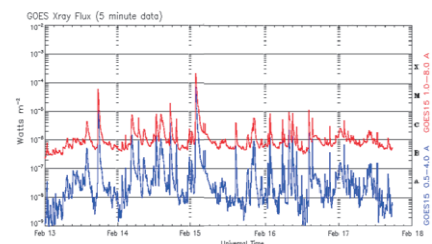


図3: GOES衛星によるX線放射量観測。放射量が1桁増加する毎にクラスがA, B, C, M, Xと増加。

## 2. 宇宙天気予報の現状

現在、NICT では毎日午後 2 時半に宇宙天気予報会議を開催し、予報情報を Web 配信している。国際宇宙環境サービス (ISES) 加盟国の地域警報センター (RWC) が予報情報配信の役割を担い、日本では NICT が担当している。太陽フレア予報の他、高エネルギー粒子、地磁気嵐の予報などを行っている。従来の人手を介する予報では太陽フレアの的中率は約 50-80% 程度であるが、最大規模 (X クラス) のフレアを予測することは依然難しい。TSS (True Skill Statistics) というスキルスコアで評価すると TSS~0.5 となる。ここで TSS は -1 から 1 の間の値を取り、全正解が 1、完全ランダムで 0 の値になる。

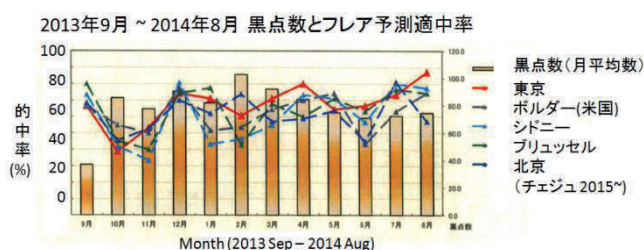


図 4: ISES 加盟国の太陽フレア予報的中率の月変化

表 1: 人手による太陽フレア予測でのチェック点

① 白色光:	黒点面積、形状 ( $\alpha \beta \delta \gamma$ )
② 軟 X 線:	フレア実績、背景値
③ 光球磁場:	磁気中性線の勾配・長さ 磁場構造の複雑性 磁気シア角、浮上磁場
④ 彩層低部:	1600 Å 連続光での増光
⑤ リム観測:	東端領域からの廻り込み

太陽は毎日多波長で観測されており、太陽フレアの予測には、SDO 衛星・STEREO 衛星によって撮られた表 1 のような観測データ (白色光・軟 X 線・光球ベクトル磁場・彩層低部発光・リム観測) を参考としながら、画像や動画を目視確認して、人手による予測を行っている。太陽衛星観測によって、昼夜にも天候にも左右されない高精度・高解像度の観測が可能になった。さらに SDO 衛星は 1 日に約 1.5TB のデータを取得しており、10 年前に比べて約 1000 倍の準リアルタイムデータが利用できるようになってきた。

その一方で、近年の太陽観測データ量は膨大で、人の処理能力を超えてしまっている。また、この大量データから緻密な経験則を人手で導き出すのも困難である。そ

こで我々は、近年社会応用が進みつつある機械学習の宇宙天気予報への応用に着目し、今までの予報精度を上回することを検討した。さらに機械学習を用いることで予測の自動化やリアルタイム予報も可能であり、日々の予報結果からより効率良くフィードバックを得ることもできる。さらに最終的には、学術的にも統計的にフレア発生を決める物理機構が何かかわかると有意義である。

機械学習は、正式には“統計的機械学習”と呼ばれ、人工知能と統計学の融合分野である。コンピュータでアルゴリズムを構築し、学習データを読み込ませることにより、自動的に今あるデータを分類し、まだ見ぬデータを予測させることができる。さらにヒトの情報処理能力を超えるような複雑かつ大量のデータを分類 & 予測する時に、力を発揮する。そのため、100 枚程度の観測画像と 2,3 個の特徴量のみを扱った予測であれば、頭の良い人の予測と機械学習との間に差はないが、次の章で述べるような観測画像 30 万枚、特徴量 60 個という場合においては、圧倒的に機械学習の方が有利となる。

## 3. モデル概要

我々が新たに開発した太陽フレア予測モデルの概要を図 5 に示す。①過去 2010-2015 年の SDO 衛星による 1 時間毎の太陽全面画像 (光球磁場・紫外線: 計 30 万枚) をダウンロードして、観測データベースを作成する。②

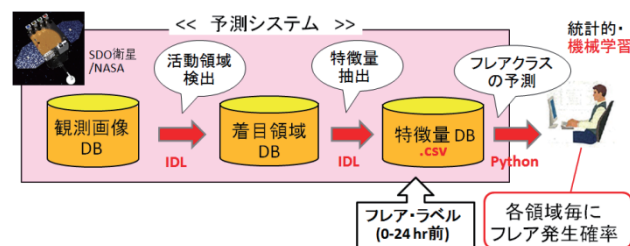


図 5: 太陽フレア予測システムの処理フロー

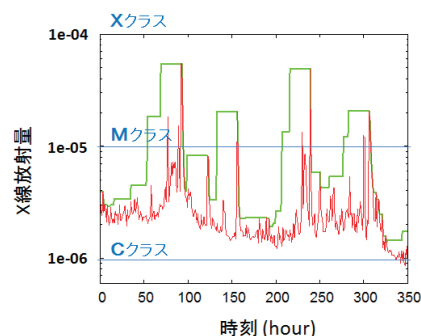


図 6: GOES 衛星が観測した軟 X 線放射量 (赤) と、ある時刻から 24 時間後までの X 線の最大値 (緑)

全面磁場画像から磁場の強い領域(黒点・活動領域)を自動検出する。③各領域にて物理特徴量を抽出する。④特徴量データベースにフレア発生のラベル(教師データ)を添付する。⑤これらをもとに機械学習によって24時間以内に発生するフレアの最大規模の予測を行う。使用データ期間中にXクラスフレアは40例、Mクラスフレアは460例観測された。フレア判定にはGOES衛星のX線放射量観測データを用いる。

各活動領域を自動検出した後、物理特徴量を抽出する。フレア予測に有効であると過去の論文にて指摘されている物理量を中心に特徴量として計算した。SDO衛星HMI望遠鏡の視線方向磁場・ベクトル磁場、AIA望遠鏡1600ÅフィルターによるUV連続光(彩層低部)画像、そしてGOES衛星X線放射量から、表2に示すような特徴量を計算した。さらに2時間・12時間・24時間差分による物理量の時間変化も計算することで、エネルギー蓄積やトリガー機構の出現などを捉えられるようにした。特徴量は数を多くすると過学習の原因ともなるため、有効なもののみを残した。

抽出した特徴量の内、4つの量のみを2次元特徴量空間に示した結果を図7に示す。赤はXフレア24時間前、

表2: 抽出した太陽活動領域の特徴量リスト

面積	視線方向磁場(HMI)	彩層発光面積 (AIA1600Å)
磁束量絶対値の和		彩層発光の最大輝度
正味の磁束量		彩層発光の輝度総量
視線方向Bzの最大値/平均値		
Bz勾配の最大値/平均値	ベクトル磁場(HMI)	
磁気中性線の最大長	カレント・ヘリシティ ( $\Sigma B_z \cdot J_z$ )	
磁気中性線の長さの和	ローレンツ力 ( $\Sigma B^2$ )	
磁気中性線の本数	垂直電流 ( $J_z$ )	
	自由磁気エネルギー	
	平均シア角	
	ツイスト変数	
同一領域のフレア履歴(X, Mクラス)		
前日のフレア履歴		
軟X線の2時間/4時間平均値		
前日の軟X線最大値 (GOES)		
	時間微分 (2, 12, 24 hrs)	

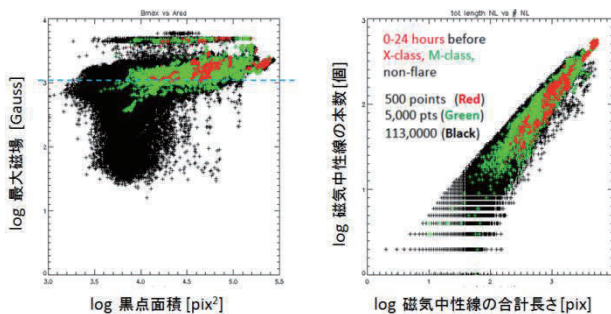


図7: 2次元特徴量空間上の分布。Xフレア発生24時間前(赤)、Mフレア発生24時間前(緑)、フレアなし(黒)

緑はMフレア24時間前、黒はX,Mフレアを伴わないイベントを示す。分布から、面積、最大磁場、磁気中性線の本数・合計長が大きい程、大規模のフレアが起きる傾向がわかる。さらに機械学習ではこれらの高次元の分布を学習させて、新たなデータ点の予測分類を行う。

#### 4. 予測と評価

完成した特徴量データベースをランダムに学習用と試験用とに7対3に分割した。そして学習用データをもとに訓練を行い、試験用データを用いて予測評価を行った。ここで、一般にどの機械学習アルゴリズムが一番フレア予測に最良であるのかを調べた先行研究はない。また実際に実行して比較してみない限りどれが良いかはわからない。従って、我々は以下の3つの機械学習手法の性能比較を行った。①サポートベクターマシン(SVM)、②k近傍法(kNN)、③アンサンブル学習(Extremely Randomized Trees: ERT)である。

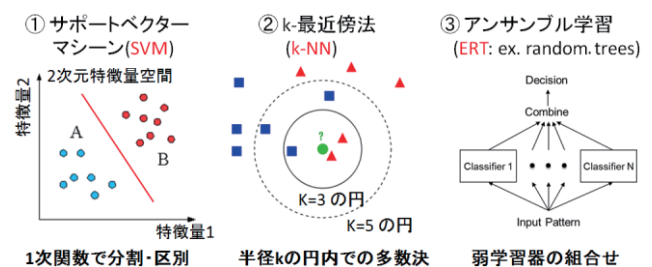


図8: 使用した機械学習の3つのアルゴリズム

SVM	実際	flare	no
予測 flare	144	15	
予測 no flare	18	54439	
	TSS= 0.89		

kNN	実際	flare	no
予測 flare	146	6	
予測 no flare	16	54448	
	TSS= <b>0.93</b>		

ERT	実際	flare	no
予測 flare	130	1	
予測 no flare	32	54453	
	TSS= 0.80		

図9: 機械学習によるフレア予測の結果比較

太陽フレア予測の結果を表9に示す。評価にはTSS(True Skill Statistics)というスキルスコアを用いた。これは太陽フレアのようにイベント発生数が非常に少ないインバランスな予測問題を評価するのに適している。予測の結果、我々はどの手法でもTSS=0.8以上を達成し、かつ最大でTSS=0.93を記録した。これは人手の



表 3: 機械学習で計算した特徴量の重要度ランキング

順位	特徴量	重要度(重み)	
1.	Xhis	0.0519	・フレア履歴 (合計, 前日のみ) ・前日の最大X線強度
2.	Xmax1d	0.0495	
3.	Mhis	0.0365	
4.	TotNL	0.0351	・磁気中性線の総合長 ・磁気中性線の本数 ・磁束の絶対値の和 ・Bzの平均値/最大値
5.	Mhis1d	0.0342	
6.	NumNL	0.0341	
7.	Usflux	0.0332	
8.	CHArea	0.0235	・彩層発光の面積
9.	Bave	0.0230	
10.	Xhis1d	0.0224	・ローレン力の和 ・磁場の垂直方向からの傾き ・極性毎の正味の電流の和
11.	TotBSQ	0.0199	
12.	VUSflux	0.0196	
13.	Bmax	0.0193	
14.	MeanGAM	0.0179	
15.	dt24SavNCPP	0.0171	

予報を上回り、かつ既存の論文に比べても良い結果である。特に kNN が他の SVM や ERT より良い精度を出すことも分かった。

さらに機械学習によって導いた特徴量の重要度のランキングを表 3 に示す。上位からフレア履歴や前日の X 線最大強度、磁気中性線の合計長・本数、磁束の絶対値の和、視線方向(垂直方向)磁場の平均値/最大値、彩層発光面積、そしてローレン力の和などのベクトル磁場特徴量の順に並んでいる結果が得られた。これら上位の特徴量はフレア発生の前兆現象として深い関わりを持っていると考えられ、フレアの予測により効果的である。

## 5. まとめと今後の計画

我々は太陽フレアの予測をするために、機械学習を用いたフレア予測モデルを開発した。2010 年～2015 年の 6 年分の観測データを用い、軟 X 線や光球ベクトル磁場、彩層低部発光の観測画像をもとにした特徴量データベースを作成した。そしてこれをもとに、これから 24 時間以内に発生する最大フレアのクラスを予測するモデルを開発した。3 つの機械学習アルゴリズムを用いて予測と精度比較を行い、人手を上回る世界でもトップクラスの予測精度を達成することに成功した。

現行モデルは X, M の 2 クラス予測しか出せないが、今後 C クラスフレアのラベル付加を行うことで、太陽活動領域毎における X/M/C クラスフレアの発生確率をそれぞれ出せるモデルへと拡張する。活動領域毎にフレアの発生確率予報を出すというのは世界的トレンドである。CME 噴出やプロトン現象の予測にフレア予測を利用する時、どの活動領域でフレアが起こるかを知ることが重要になるためである。

また近年は、欧米を中心に宇宙天気予報モデルの国際比較が盛んに行われている。NASA の共同コミュニティ・モデリングセンター(CCMC)には、フレアスコアボードというサイトがあり、世界各機関によるリアルタイム・フレア予測結果の Web 公開比較を行っている。公開予測情報は、人手を介した経験予測モデルによるものも、機械学習を用いた自動予測によるものも機関によって様々である。我々 NICT モデルも近々リアルタイム運用と実証実験を行い、CCMC 等国際比較に参加予定である。また予測期間を 2-3 日に延ばしたり、逆に 6 時間・12 時間以内の直近予測を試したりも検討する。

最後に機械学習を用いた太陽フレア予測は、他の宇宙天気現象の予測にも応用可能である。まず周囲の状況を捉える観測データとイベント発生のラベル付けさえできれば、機械学習させることはできる。まずは太陽フレアの予測から CME 予測、プロトン現象予測へと拡張可能であり、最終的には太陽と地球磁気圏・電離圏をつないだ包括的な予報モデルを構築したい。また短期的には、機械学習を用いることで予測精度の向上を図り、かつ予測に有効な特徴量探しをすることができる。その一方で長期的には、得られた知見を数値モデリングに活かし、太陽フレア機構や宇宙天気現象のさらなる理解へとつないでいきたい。

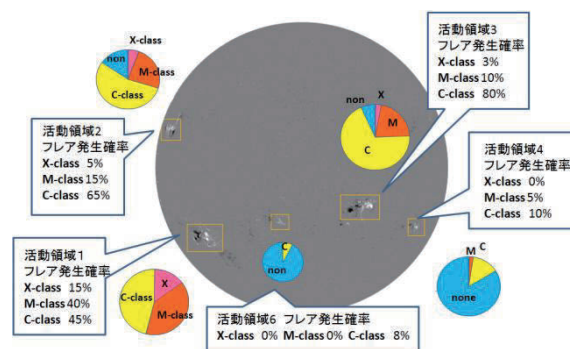


図 10: 領域毎の太陽フレア確率予報のイメージ

## 参考文献

- 1) Nishizuka, N., Sugiura, K., Kubo, Y., Den, M., Watari, S., & Ishii, M., 2016, *Astrophysical Journal*, 受理