

5-2 宇宙天気予報評価技術

5-2 Space Weather Forecast Verification

久保勇樹

KUBO Yūki

自然現象の発生予報には、大きく分けて2通りの方法がある。現象の発生を0/1で予報する決定論的予報と、現象の発生確率を予報する確率予報である。現在のNICTの予報は決定論的予報が中心であるが、将来的な確率予報への移行を見据えて、確率予報モデルの開発なども行われている。本稿では、これまでNICTで行われてきた予報と、現在開発中の確率予報モデルを題材にして、宇宙天気予報評価技術について簡単に述べる。さらに、予報提供者視点での評価法だけでなく、予報利用者視点での評価法についても紹介する。

There are two major methods to forecast natural phenomena; a deterministic forecast, which forecasts an event occurrence as yes or no, and a probabilistic forecast, which forecasts an event occurrence probability. NICT forecast has focused currently on the deterministic forecast. However, we have been developing the probabilistic forecast models of space weather phenomena for future transition from current deterministic forecast to a probabilistic forecast. In this paper, we briefly introduce a forecast verification method for a deterministic and probabilistic space weather forecasting. The verification in terms of forecast users as well as forecast providers is also introduced.

1 はじめに

国立研究開発法人情報通信研究機構(NICT)は、日本で唯一の宇宙天気予報を業務として行っている公的機関である。NICTでは、その前身である郵政省電波研究所が発足した1950年代初めごろから、短波通信障害を事前に察知して利用者に伝える電波警報業務を行ってきた。この業務は、1988年、宇宙天気予報業務と名前を変えると同時に、短波通信障害を引き起こす原因となる太陽活動や地磁気じょう乱にまで予報・警報の範囲を広げていった。現在、NICTの宇宙天気予報業務は、短波通信障害や衛星測位の誤差増大の原因となり得る太陽活動、地磁気じょう乱、電離圏じょう乱などの情報以外に、人工衛星障害の原因となり得る放射線帯電子や、宇宙飛行士や航空機乗務員・乗客の宇宙放射線被ばくの原因となり得る太陽高エネルギー粒子など、様々な情報を発信している[1]。

予報は、その予報精度や特徴に関する情報を伴うことで、より有用な情報となる。そのため、過去に発信した予報からその精度や特徴を正しく評価する技術が必要となってきた。このような技術は、気象予報の分野では1884年にJ.P. Finleyによって発表されたトルネード予報評価[2]以来、長い歴史を持ち、一つの

研究分野として確立しているが、宇宙天気予報ではようやくその重要性が認識されてきたところである。

本稿では、NICTが過去に発信した太陽フレア予報と、NICTで開発されている太陽フレア確率予報モデルを題材として、予報の精度や特徴を評価する技術について述べるとともに、今後の予報の在り方について議論する。

2 NICT 太陽フレア予報

太陽フレアの規模は、米国の気象衛星GOESが観測している太陽フレアに伴って突発的に増加するX線フラックスに基づいて、表1のように決められている。NICTの太陽フレア予報は、ウルシグラムコードの基準(UGEOA)に従っており、予報発令時刻の6:00 UTC(協定世界時)以降24時間以内に発生が予測される最大フレアの規模が4つのカテゴリー(静穏、やや活発、活発、非常に活発)のどこになるかを0か1かで予報する決定論的予報である。この形式での予報は1992年に始まり、現在も続いている。

表1 太陽フレアの規模と予報レベル

最大 X 線量 F_x (Watts m^{-2})	太陽フレアの規模	予報 (UGE OA コード)
$F_x < 10^{-7}$	A	静穏 (0)
$10^{-7} \leq F_x < 10^{-6}$	B	静穏 (0)
$10^{-6} \leq F_x < 10^{-5}$	C	やや活発 (1)
$10^{-5} \leq F_x < 10^{-4}$	M	活発 (2)
$10^{-4} \leq F_x$	X	非常に活発 (3)

表2 2000～2015年の太陽フレア予報と結果

Observation	3	0	19	82	29
	2	23	379	495	43
	1	335	1554	453	11
	0	1979	419	21	2
			0	1	2
		Forecast			

3 予報精度評価法

予報と結果に関する情報は、予報(x)と結果(o)を2つの変数とみた場合の同時確率密度関数にすべて含まれていることから、予報精度評価とは予報・結果同時確率密度関数 $p(x,o)$ を解析することと等価である。例えば、2 カテゴリーの決定論的予報と結果は、2行2列の予報結果行列 (contingency table) によって完全に記述される。そして、この contingency table の4つの要素の間には、合計が全予報数になるという1つの束縛条件があるため、contingency table の自由度は3である。そのため、予報・結果同時確率密度関数を余すことなく記述するためには、少なくとも3つの独立な評価指標を示すことが必要であり、一つの評価指標をもってこの予報の精度を議論することはできない [3]。一般的に、2変数の同時確率密度関数は、1変数の周辺確率密度関数と、その1変数を条件づけられた条件付確率密度関数の積に分解することができる。したがって $p(x,o)$ は、予報 x を条件づけられた分解 $p(x,o) = p(o|x) \cdot p(x)$ (calibration-refinement 分解) と、結果 o を条件づけられた分解 $p(x,o) = p(x|o) \cdot p(o)$ (likelihood-base rate 分解) の2種類があり、この2種類の分解それぞれが、評価される予報の性質に対応する。例えば、 $p(o|x)$ は発信した予報の信頼度 (reliability) という性質に対応し、 $p(x|o)$ は現象の発生を区別できていたかどうか (discrimination) という性質に対応する。また、 $p(x)$ と $p(o)$ は、過大・過小予報 (bias) を表す指標に使われる [4]。この考え方は、distribution-oriented approach と呼ばれ、現在の予報評価技術の一般的な考え方となっている [5]。

3.1 決定論的予報評価

本小節で紹介する太陽フレア予報精度評価は、久保ほかの報告 [6] の内容を簡略化したものであり、2000年から2015年に発令された5,844日間の予報データを用いて行った結果である。前節で述べたように、NICT太陽フレア予報は、4カテゴリーの決定論的予報である。決定論的予報評価法は、前述の [2] 以来、様々な評

表3 2カテゴリーの contingency table

結果 (有り)	C	A
結果 (無し)	D	B
	予報 (無し)	予報 (有り)

価指標が提案されているが、それぞれ一長一短がありどの評価指標を選ぶのが良いのかは一概には言えないが、一つのヒントとして、世界気象機関 (WMO) が、サイクロン発生予報 (2 カテゴリー予報) に対する予報評価に関する一つの勧告を出している [7]。サイクロンはまれな現象 ([2] では発生頻度は約 1-2%) であるという点で、太陽フレアと類似していることから、この WMO 勧告に従った評価指標を採用した。それぞれの評価指標の定義や意味、具体的な計算方法などの詳細は、[6] を参照されたい。

表2に2000～2015年の太陽フレア予報と結果の contingency table を示す。この contingency table を、Mクラス以上をイベント、Cクラス以下をイベント無しとして2カテゴリー予報に圧縮したときの、様々な評価指標を算出した結果を図1に示す。代表的な指標を表3の contingency table を用いて説明すると、Probability of detection (検出率) : $POD = A / (A + C)$ 、Probability of false detection (誤検出率) : $POFD = B / (B + D)$ 、Peirce skill score : $PSS = POD - POFD$ (True skill statistic : TSS とも呼ばれる) などがあり、これらは discrimination を評価する指標として用いられる。また、reliability を表す指標としては、False alarm ratio : $FAR = B / (A + B)$ 、bias を表す指標としては、Frequency bias : $FB = (A + B) / (A + C)$ などがある。図1には、NICT予報と、前日結果 = 本日の予報

(Persistence)とした場合、27 日前の結果=本日の予報 (Recurrence)とした場合の、評価指標 (上図)と NICT 予報と Persistence、Recurrence それぞれの評価指標の差 (下図) が 95 % 信頼区間とともに書かれている。信頼区間は、ブートストラップ BCa 法 [8] により計算されている。結果の詳細な議論は [6] を見ていただきたいが、図中の FB を見るとわかるように、NICT 予報は概ね FB~1 であり、過大予報にも過小予報にもなっていないことがわかる。特筆すべきことは、下図の赤い棒グラフを見るとわかるように、多くの評価指標において、NICT 予報は Persistence よりも 95 % の信頼度をもって有意に高いスコアを記録していることであり、Recurrence (青い棒グラフ) に対しては言うに及ばずである。すなわち、NICT 予報は、Persistence や Recurrence に対して、良い予報精度を持っているということを示している。

X クラスフレアについても同様の解析を行っている (図 2)。X クラスフレアについても、NICT 予報は Recurrence に対しては有意に良い精度を示しているものの、Persistence に対しては、95 % の信頼度をもって有意に良い精度であるとは言えなかった。これは、X クラスフレアはまれな現象であることからサンプル数が少なく、95 % 信頼区間が大きいため有意な結論を示すことができないということであり、NICT 予報が Persistence と同じ精度であると結論しているわけではないことに注意してほしい。しかしながら、X クラスフレアに対しては、NICT 予報は有意に過小予報で

あることが示されており、これは、X クラスが起こるのであるという予報は、かなり確からしく判断できる場合にしか発令していないと結論できる。

太陽活動は約 11 年の周期で極大、極小を繰り返していることはよく知られたことであり、予報の精度も太陽活動に依存して変化することが容易に推測される。2000~2015 年の 16 年間に太陽活動度に応じて 4 つの期間に分けて、NICT 予報の精度を評価することも行っている。この結果については、紙面の都合上本稿では割愛するが、[6] で詳細に述べているので参考にしてほしい。

3.2 確率予報評価

前述のように、現在の NICT 太陽フレア予報は、いわゆる 0 か 1 かの決定論的予報であるが、自然現象の発生とは本質的に不確実性を伴うものであり、これを決定論的に予報するというのは不確実性を無視した予報ということであり、考え方によっては適切な予報とは言えない。むしろ、自然現象の発生に伴う不確実性を考慮した確率予報のほうが、より合理的な予報であると考えられる。NICT の予報も、今後確率予報に移行していくことが検討されており、AI を用いた太陽フレア確率予測モデル (DeFN-R) が開発されている [9]。本小節では、この AI 太陽フレア予報モデルの結果を用いて、確率予報評価技術について述べる。

確率予報を評価する際に、ある一つのしきい値確率 (例えば 0.5 など) を設定して、現象の発生確率がそれ

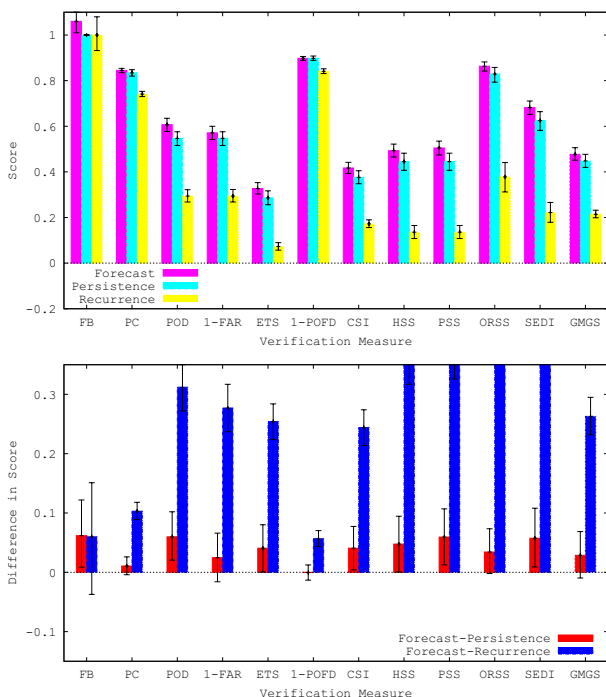


図 1 M クラス以上のフレアに対する評価指標

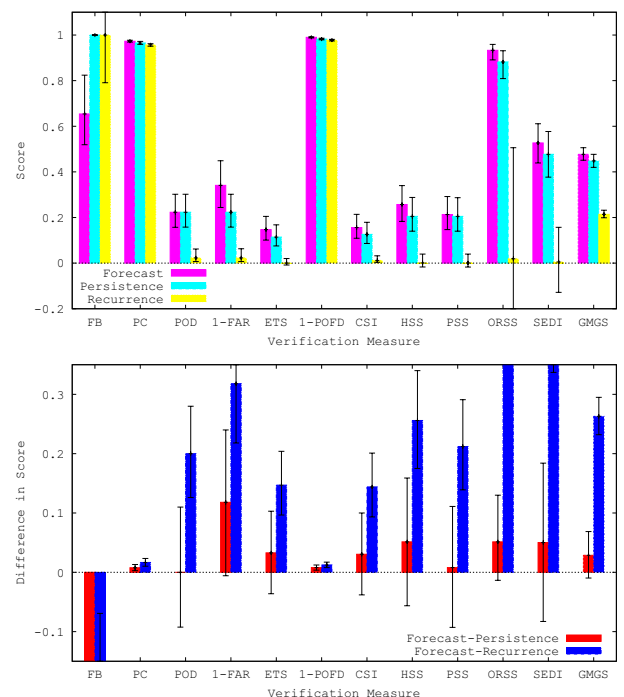


図 2 X クラスフレアに対する評価指標

より大きい小さいかで確率予報を決定論的予報に変換し、それに対して決定論的予報の評価手法を適用するというやり方を、宇宙天気予報の分野で頻繁に見かける。しかし、これでは0から1までの連続値を持つ確率予報を2つのカテゴリに圧縮することになり、予報・結果同時確率密度関数が持つ非常に多くの情報を失ってしまい、十分な評価ができない。また、しきい値確率は作為的に決められることができるため、しきい値確率をいくつにするかというあまり有意義でない議論に陥ってしまうことになる。このようなことから、確率予報としての評価手法は、決定論的予報の評価手法とは異なる手法で行うことが必要である。以下に、確率予報の古典的な評価手法について述べる。なお、以下で述べるようなグラフィカルな評価手法に加えて、Brier skill score とその reliability 成分、resolution 成分や、情報理論を基にした Ignorance score のようなスカラー値として表される評価指標が使われることも多い(例えば [10]) が、紙面の都合でこれらについては本稿では割愛する。

現象発生確率が x であった場合、多数回 x という予報をした時の結果 o (0 か 1 の値をとる) の平均は、当然 x であるのが良いはずである。数学的には、確率 x が与えられたという条件の下での結果 o の条件付き期待値 $E(o|x)$ が、確率 x に等しい、すなわち、 $E(o|x) = x$ となるのが良いということを意味している。これは、予報精度評価の分野で信頼度(reliability)と呼ばれる確率予報が備えるべき重要な性質であり、前述の、calibration-refinement 分解に関連している。これを、縦軸 $E(o|x)$ 、横軸 x としてプロットした場合、高い信頼度を持つ予報であればすべての点対角線に乗ることになり、もし、対角線から外れているならば、その確率予報は信頼度が低いということになる(図3参照)。この図は reliability diagram(信頼度曲線)と呼ばれており、確率予報の評価を行う際に重要な評価手法になる。

一方で、現象が発生した時 ($o=1$) に予報していた現象発生確率は確率1に近い方が良く、現象が発生しなかった時 ($o=0$) に予報していた発生確率は、0に近い方が良い。これは、数学的には結果 o が与えられたという条件の下での確率 x の条件付き確率密度関数 $p_0(x) := p(x|o=0)$ と $p_1(x) := p(x|o=1)$ が、分布としてできるだけ区別されている ($p_0(x)$ と $p_1(x)$ の確率密度関数としての距離が離れている) ほうが良いということの意味している。この性質は、discrimination と呼ばれ、likelihood-base rate 分解に関連した重要な性質である。この discrimination を表す指標として、 $p_0(x)$ と $p_1(x)$ それぞれの累積分布関数 $F_0(x)$ と $F_1(x)$ とが使われる。ここで、 $POD(x) := 1 - F_1(x)$ 、 $POFD(x) := 1 - F_0(x)$

と定義すると、 $POD(x)$ 、 $POFD(x)$ はそれぞれ、しきい値確率を x としたときの、決定論的予報評価の指標である POD 、 $POFD$ となる。そして、しきい値確率 x を0から1まで変化させながら縦軸に $POD(x)$ 、横軸に $POFD(x)$ をプロットしたものは relative operating characteristic (ROC) 曲線と呼ばれている(図4参照)。ここで注意しなくてはならないのは、現象の定義 ($o=0,1$ の定義) は事前に決定され、しきい値確率を変化させる際に変化させてはならないということである。変化させてしまうと、何の現象を区別する能力を評価しているのかわからなくなるからである。もし、 $p_0(x)$ と $p_1(x)$ が完全に区別され、関数として重なり合う部分が全くない、すなわち現象が発生した場合には必ず1に近い高い発生確率を予測、現象が発生しなかった場合には必ず0に近い低い発生確率を予測していたならば、ROC 曲線は、 $POFD-POD$ 平面上を点 (1,1) から $POFD$ 軸と平行に点 (0,1) に移動し、その後、 POD 軸に沿って点 (0,0) と移動する折れ線となる。もし、 $p_0(x)$ と $p_1(x)$ が完全に重なりあう関数であるならば、ROC 曲線は、 $POFD-POD$ 平面上で点 (1,1) と点 (0,0) を結ぶ対角線になる。その場合、確率予報は現象発生を全く区別することができないということである。すなわち ROC 曲線は、その曲線の全体としての形状が対角線から (1,1) → (0,1) → (0,0) を結ぶ折れ線に近づくほど(もしくは ROC 曲線と $POFD$ 軸に挟まれた領域の面積 Area under curve: AUC が1に近いほど) discrimination 性能が高いことを示すことになり、この曲線の形状は、確率予報の評価を行う際に重要な評価手法になる。

NICT で開発している AI 太陽フレア確率予報モデル DeFN-R の結果に対して、reliability diagram と ROC 曲線を描いてみると以下のようなになる。

図3から、DeFN-R は非常に良い reliability 特性を持っていることがわかる。図3の対角線上の誤差棒は、

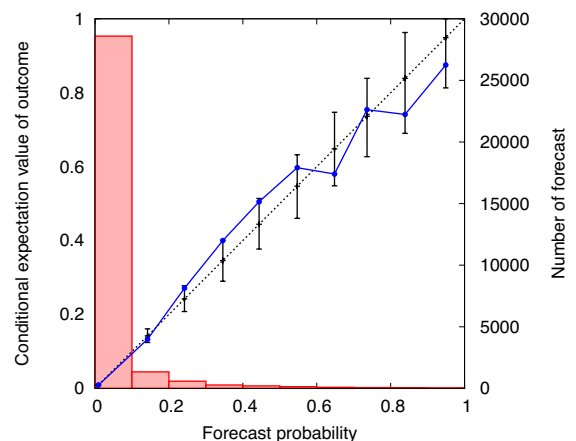


図3 Mクラス以上のフレア発生確率予報に対する reliability diagram

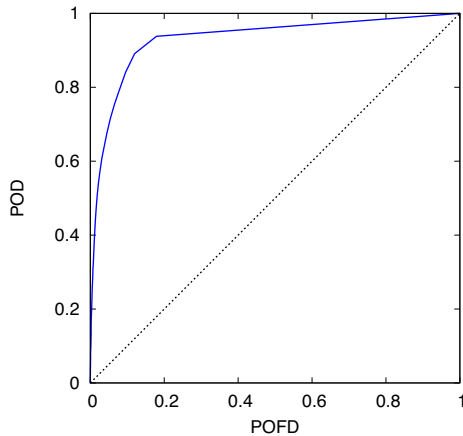


図4 Mクラス以上のフレア発生確率予報に対するROC曲線

95 % consistency bar と呼ばれ、5 % 棄却域を持った統計的仮説検定に関する量であるが、詳細は久保の報告 [11] を参照されたい。図4からは、ROC 曲線が (1,1) → (0,1) → (0,0) の折れ線にかなり近く、AUC が 0.93 と 1 に非常に近い値であることから、DeFN-R は discrimination 特性についても高い性能を持っていることがわかる。このように、DeFN-R は reliability と discrimination 両方について高い性能を兼ね備えた確率予報モデルであるということが出来る [9]。

最後に、ROC 曲線は、ある一つのしきい値確率に関しての記述ではなく、曲線全体の形状及び AUC が確率予報全体としての精度に関する情報を記述しているということを強調しておく。ROC 曲線を解釈する際に、点 (0,1) に最も近い点を与えるしきい値確率が、確率予報を決定論的予報に圧縮する際の最適なしきい値確率であるという議論が、宇宙天気予報の分野でよくされている。しかしこの議論は正しくない。点 (0,1) に最も近い ROC 曲線上の点 P というのは、すなわち、 $L^2 := (1-POD)^2 + POFD^2$ が最も小さい点を意味しているが、 L^2 という量は予報精度に関する評価指標としての意味を持った量とはみなされていない。さらに、一般的には点 P では POD 最大でもなく、POFD 最小でもない。すなわち L^2 が最小であるということが、確率予報を決定論的予報に圧縮する際の、最適なしきい値確率を与えるという意味にはなっていないということである。このように、ROC 曲線は、その全体としての形状が評価指標としての意味を持つのであって、ROC 曲線上のある特定の一点を与えるしきい値確率に関してはほとんど意味を持たない。ROC 曲線が特定の一点について述べている唯一のことは、 $L^1 := |1-POD| + |POFD|$ が最小の点、すなわち POD-POFD が最大の点についてである。この点は、ROC 曲線上で、POD=POFD を表す対角線からの距離が最大になる点であり、これは決定論的予報に対する discrimination

性能を評価する PSS を最大にするしきい値確率を与える点である [12]。すなわち、この点は、PSS によって測られる、discrimination 性能を最大化するという意味において、最適なしきい値確率を与えることになる。

なお、PSS を最大化するしきい値確率は、確率予報の信頼度曲線の値が現象発生頻度の値に等しくなる確率であることが [11] によって示されている。

3.3 利用者視点での予報評価

これまで述べてきた予報評価手法は、予報を提供する側の視点での評価であった。しかし、予報を利用する側、すなわち予報利用者にとっての良い予報と、予報提供者にとっての良い予報は必ずしも一致しない。また、ある利用者にとって最適な予報が、すべての利用者にとって最適な予報であるとは限らない。したがって、利用者視点で見た時の予報評価が重要になる。本小節では、簡単な意思決定理論に基づいた予報評価手法について述べる。

予報利用者は、大規模な現象(太陽フレアなど)発生の予報に基づいて回避行動を実施すると仮定する。もし、回避行動を実施せずに大規模な現象が発生した場合には、損失 L が発生するが、経費 C をかけて回避行動を起こすことで、この損失は回避できるとする。この場合、行動の有無と現象の有無の組合せにより、表4のように4つの支出のパターンが考えられる。これに対して表5のように実際の予報結果行列(ただし、各要素を全要素の合計で割って、合計が1になるように規格化)を T とすると、予報に従って回避行動を起こした場合の平均支出 E は、 $E = \text{Tr}(A^t \cdot T)$ とあらわせる。ここで A は表4に示される Cost-Loss 行列である。完全に正しい予報に対する平均支出を E_p とすると、これは $E_p = s(C+L)$ となる ($s=a+c$)。これに対して、予報に関わらず常に回避行動をとる、もしくは常に回避行動をとらない場合の平均支出 E_0 は、 $E_0 = \min(C, sL)$ となる。したがって、予報に従った行動のほうが予報と関係なく常に同じ行動をとるよりも平均支出が小さい (E

表4 Cost-Loss 行列 A

	現象有り	現象無し
行動有り	C	C
行動無し	L	0

表5 予報結果行列 T (a+b+c+d=1)

	現象有り	現象無し
予報有り	a	b
予報無し	c	d

5 定常業務

が E_0 よりも小さい) ならば、利用者はこの予報から利益を得ることができる。ここで、Relative economic value を以下のように定義する。

$$V := \frac{E_0 - E}{E_0 - E_p} = \frac{\min(\alpha, s) - s + s(1 - \alpha)\text{POD} - (1 - s)\alpha\text{POFD}}{\min(\alpha, s) - s\alpha}$$

この時、 V は $\alpha=C/L$ (Cost-Loss 比) の関数であり、 V が正の値をとる Cost-Loss 比を持つ利用者にとって、この予報は価値のあるものとなる (例えば [13])。

以下では、上記の議論を NICT 太陽フレア予報、AI 太陽フレア予報モデルに適用し、利用者視点での予報の価値について議論する。図 5 は、NICT 太陽フレア予報 (M クラス以上) に対する、Relative economic value 曲線 (以後 REV 曲線という) を表す。解析期間に

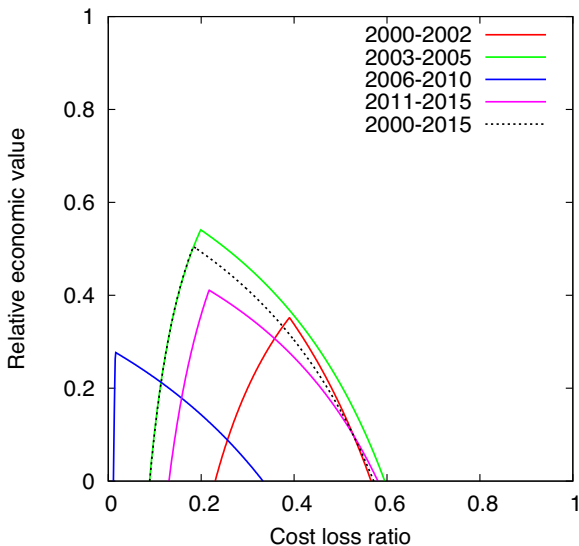


図 5 NICT 太陽フレア予報 (M クラス以上) に対する REV 曲線

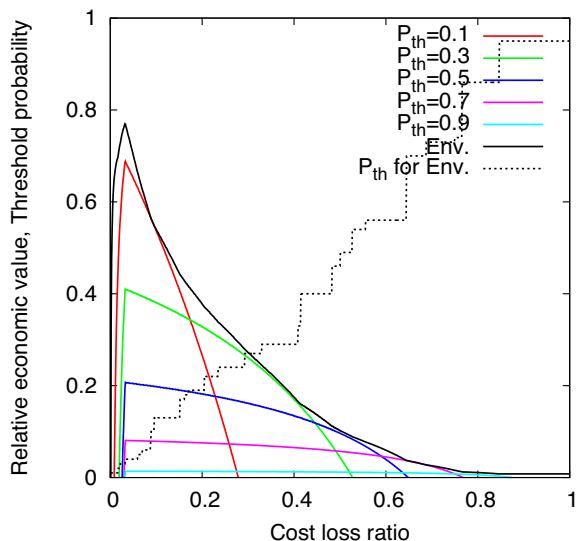


図 6 DeFN-R (M クラス以上) に対する REV 曲

よって多少違いはあるが、大まかに言って、Cost-Loss 比が 0.1 ~ 0.6 くらいを持つ利用者にとっては有益な予報となっているが、0.6 より大きい Cost-Loss 比の利用者にとっては、この予報に従って回避行動を起こしても利益を得ることはできない。

次に DeFN-R に対する REV 曲線を図 6 に示す。DeFN-R は確率予報モデルであるので、0 から 1 まで 0.01 刻みでしきい値確率 P_{th} を取り、そのすべてについて REV 曲線を計算した。図中の色付きの曲線はそれぞれ $P_{th}=0.1, 0.3, 0.5, 0.7, 0.9$ に対する REV 曲線で、黒実線はすべての P_{th} に対する REV 曲線の包絡線である。黒点線は、ある Cost-Loss 比に対して、REV 曲線の包絡線を与えるしきい値確率を書いてある。例としても、予報提供者が、しきい値確率 0.1 を用いて (図 6 赤線) 確率予報から決定論的予報に変換した後の情報を発信していたとしたら、その決定論的予報から利益を得ることができるのは、Cost-Loss 比 0.3 程度以下を持った利用者のみであり、それより大きな Cost-Loss 比の利用者にとっては、この予報に従って意思決定をしても利益は得られない。しかし、もし、予報提供者が決定論的予報に変換せずに確率予報のまま発信していたならば、利用者自身がそれぞれ持つ Cost-Loss 比に応じた最適なしきい値確率に従って意思決定することで、ほぼすべての Cost-Loss 比を持つ利用者が確率予報から利益を得ることができる。その REV 曲線が図中の黒実線になる。そして、利用者自身がそれぞれ持つ Cost-Loss 比に応じた最適なしきい値確率は、図の黒点線になることがわかるだろう。利用者が最適な意思決定をするために必要なしきい値確率の情報を、確率予報とともに提供することで、利用者にとって確率予報は、決定論的予報よりも高い価値を持つことになる。そして、この利用者の最適な意思決定のためのしきい値確率は、実は確率予報の信頼度曲線の逆関数になっていることが数学的に示される。このことから、予報提供者視点での予報評価を行うことで、利用者の意思決定を最適化するしきい値確率の情報も得られることになる。

4 まとめ

今まで述べてきた予報評価手法は、現象が発生するかしないかの 2 カテゴリー予報としての評価である。しかし、前述のように NICT 太陽フレア予報は 4 カテゴリー予報であることから、多カテゴリー予報としての評価も必要である。例えば、決定論的予報の評価指標としては、Gandin-Murphy-Gerrity Score [14][15] などが提唱されている。確率予報の評価手法としては、Ranked Probability Score (例えば [10]) が良く知られて

いるが、それ以外に多カテゴリー Reliability diagram[16]、多カテゴリー ROC 解析 [17] などがある。しかし、多カテゴリーの予報評価手法は、2カテゴリーの予報評価手法に比べると発展途上であり、今後更なる研究が必要である。4カテゴリー太陽フレア予報の評価については、[6]に簡単に述べられているのでそちらを参照されたい。

最後に、今後の NICT 宇宙天気予報の方向性について簡単に述べて、本稿のまとめにする。NICT の宇宙天気予報は、長年にわたり決定論的予報を行ってきた。しかし、自然現象は本質的に不確実性を伴い、その不確実性を伴った現象の予測が決定論的であるというのは不自然であり、不確実性を反映した確率予報のほうが合理的であると考えられるのは前述のとおりである。自然現象の予測という点では宇宙天気予報よりはるかに進んでいる気象予報では、基本的に確率予報が一般的である。例えば降雨予報などには必ず確率の情報が付帯されている。気象予報は数値シミュレーションという決定論的手法によって予測されているが、アンサンブルという手法を用いることで、数値シミュレーションによる予測から確率予報を作り出している。このような予報技術の進化に伴って、気象予報は決定論的予報から確率予報へ移行してきた。宇宙天気予報においても、決定論的予報から確率予報への移行が現実的になりつつある時期と考えている。さらに、3.3 で述べたように、利用者の視点から考えても、確率予報は決定論的予報に比べて多くの利用者が利益を享受できる、より価値の高い情報であると考えられる。このようなことを考慮すると、今後の NICT 宇宙天気予報の進むべき方向は確率予報への移行であり、太陽フレア確率予報のみならず宇宙天気現象の確率予報モデルの開発が極めて重要である。

謝辞

本業務の一部は、総務省委託業務「0155-0099 電波伝搬の観測・分析等の推進」によって行われたものである。また、西塚直人テニユアトラック研究員から DeFN-R モデルの予報結果の提供を受けた。ここに感謝する。

【参考文献】

- 1 大辻賢一, 久保勇樹, “宇宙天気予報業務,” 情報通信研究機構研究報告, 本特集号, 5-1, 2021.
- 2 J. P. Finley, “Tornado Predictions,” Amer. Meteor. J., vol.1, pp.85-88, 1884.
- 3 A. H. Murphy, “Forecast Verification: Its Complexity and Dimensionality,” Mon. Wea. Rev., vol.119, pp.1590-1601, 1991.
- 4 A. H. Murphy, “What is a good forecast? An essay on the nature of goodness in weather forecasting,” Wea. Forecasting, vol.8, pp.281-293, 1993.

- 5 A. H. Murphy and R. L. Winkler, “A General Framework for Forecast Verification,” Mon. Wea. Rev., vol.115, pp.1330-1338, 1987.
- 6 Y. Kubo, M. Den, and M. Ishii, “Verification of operational solar flare forecast: Case of Regional Warning Center Japan,” J. Space Weather and Space Clim., vol.7, A20, 2017.
- 7 WMO, Verification Methods for Tropical Cyclone Forecasts, WWRP 2013-7, 2014. https://www.wmo.int/pages/prog/arep/wwrp/new/documents/WWRP_2013_7_TC_verification_15_Nov_en.pdf
- 8 T. J. DiCiccio and B. Efron. “Bootstrap confidence intervals,” Stat.Sci., vol.11, pp.189-228, 1996.
- 9 N. Nishizuka, Y. Kubo, K. Sugiura, M. Den, and M. Ishii, “Reliable Probability Forecast of Solar Flares: Deep Flare Net-Reliable (DeFN-R),” Astrophys. J., vol.899, 150, 2020.
- 10 I. T. Jolliffe and D. B. Stephenson, Forecast verification: A practitioner's guide in atmospheric science, 2nd edn., John Wiley and Sons Ltd., Chichester, UK., 2012.
- 11 Y. Kubo, “Why do some probabilistic forecasts lack reliability?,” J. Space Weather and Space Clim., vol.9, A17, 2019.
- 12 A. Manzato, “A Note on the Maximum Peirce Skill Score,” Wea. Forecasting, vol.22, pp.1148-1154, 2007.
- 13 D. S. Richardson, “Skill and relative economic value of the ECMWF ensemble prediction system,” Q. J. R. Meteor. Soc., vol.126, pp.649-667, 2000.
- 14 L. S. Gandin and A. H. Murphy, “Equitable Skill Score for Categorical Forecasts,” Mon. Wea. Rev., vol.120, pp.361-370, 1992.
- 15 J. P. Gerrity, “A Note on Gandin and Murphy's Equitable Skill Score,” Mon. Wea. Rev., vol.120, pp.2709-2712, 1992.
- 16 T. M. Hamill, “Reliability Diagrams for Multicategory Probabilistic Forecasts,” Wea. Forecasting, vol.12, pp.736-741, 1997.
- 17 M. S. Wandishin and S. J. Mullen, “Multiclass ROC Analysis,” Wea. Forecasting, vol.24, pp.530-547, 2009.



久保勇樹 (くぼ ゆうき)

電磁波研究所
電磁波伝搬研究センター
宇宙環境研究室
研究マネージャー/
宇宙天気予報グループ
グループリーダー
博士(学術)
太陽物理学、予報評価技術