

光化学大気汚染の予報精度の向上に関する研究 (I)
 予測の検証とカテゴリー予報の判定値設定の検討

深澤秀司, 内田聡子
 (環境情報部)

Note

Research on the improvement of the precision to forecast the photochemical air pollution. (I)
 The verification of predictions and the examination of a setup of decision value for the category forecast.

Shuji FUKASAWA, Satoko UCHIDA
 (Environmental Information Division)

1. はじめに

神奈川県では、光化学大気汚染による緊急時に対応するため、1971年5月に「神奈川県における光化学公害緊急時の暫定措置要綱」を定め、被害防止のため「光化学スモッグ注意報」「光化学スモッグ警報」や「予報」を発令し、ばい煙排出者に対して燃料使用量の削減措置などの対策をとってきた。光化学スモッグ注意報、警報の発令回数は、1972年の31回を最高に1999年までの29年間に387回で、年平均13.3回となっている。¹⁾

また、予報には当日予報と前日予報がある。当日予報は、1972年から行っているが、これまでの27年間の発令回数は114回で、年平均4.2回となっている。当日予報については、その日のオキシダント最高濃度の予測値に基づき光化学スモッグ注意報の発令の可能性が高い場合、朝10時に県内全地域に対して発令を行っており、同時に主要ばい煙排出者に対してあらかじめ指導した計画に沿った燃料使用量の削減措置の実施を要請している。

光化学スモッグの予報は、人に対する被害特に被害の多い小中学生に対する被害の未然防止を目的としていることから、注意報発令の回避が目標であり、削減措置により「空振り」となることはやむを得ないが、「見逃し」をなくすことが理想である。

そこで予報発令の的確性の推移を評価するため、予報を開始した1973年から1999年の予報と注意報（警報を含む）の回数をもとに注意報 (Warning) と予報 (Forecast) の発令回数の比 (W/F) を図1に示した。W/Fは、 $W/F < 1$ の場合は空振りがあることを、 $W/F = 1$ の場合は一致していることを、 $W/F > 1$ の場合は見逃しがあることを示している。W/Fは、1973年～1984年においては2.6以下であったが、その後次第に上昇し1990年～1994年の5年間は62.0と最悪の状態になった。1991年に環境科学センターが発足後、独自の予測式を併用することにより予測精度の向上に努めた結果、最近の5年間に於いて $W/F = 6.3$ となったが、この更なる改善が大きな課題となっている。

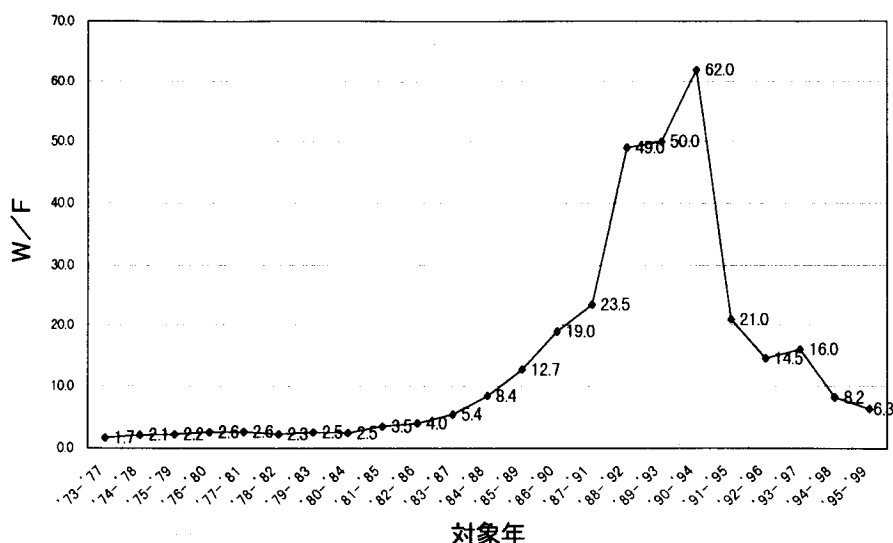


図1 光化学スモッグ注意報・予報比の推移

表1 光化学大気汚染予報判定表

EWS予測値 重回帰予測値	< 80	80~119	≥ 120
< 100	C	C 但し業者判定a又はbの場合は B	B
100~109	C	C 但し業者判定a又はbの場合は B	B
110~119	C	B	B
≥ 120	C	B	A

A：光化学スモッグが発生する恐れがある。光化学スモッグ予報の発令
 B：気象条件によっては、光化学スモッグが発生する恐れがある。
 C：光化学スモッグの発生の恐れはない。

予報精度の向上については、予報の客観的、基礎的データとなる予測式を改良する方法と予報判定のため、予測結果を判断する基準を改良する方法の二つの方法が考えられる。筆者らは、今回後者の方法により1995年から1999年の5年間に得られた複数の予測最高濃度と実測最高濃度のデータを使用して、予報精度の向上についての検討をおこなった。

2. 検討の方法

現在、神奈川県では光化学スモッグの当日予報を行う場合、次の4つの情報の内ア～ウを表1の判定表に配置してA、B、C3段階のカテゴリーにしている。このうちカテゴリーAの場合に当日予報を発令している。

- ア. 民間気象会社が使用している重回帰予測値（以下「重回帰予測」という）
- イ. 県で開発した気象データと大気汚染濃度に基づく統計モデルの予測システムで計算した濃度予測値（以下「EWS予測」という。）
- ウ. 民間気象会社の予測担当者が重回帰予測値に基づき判定するa、b、c3段階の判定（以下「業者判定」という。）
- エ. 民間気象会社が業者判定のために予報担当者が重回帰予測を修正した濃度予想値（以下「人予測」という）

まず、人予測値、重回帰予測値、EWS予測値の3量それぞれの予測精度の検討を行った。精度の確認方法として、鈴木²⁾が示した気象予測の検証方法を参考に行った。鈴木は次の5つの方法を示している。

- ①適中率の計算による方法、
- ②相関係数の計算による方法、
- ③ χ^2 -検定による方法、
- ④情報量の計算による方法、
- ⑤スキル・スコア*とパーセント・スコアによる方法。

これらの方法のうち、①、③、④及び⑤のパーセント・スコアにはそれぞれ次のような課題があることから採用をしなかった。

①の適中率の計算による方法は、発生すると予測して発生した場合だけでなく、発生しないと予測し発生しなかった場合も適中とする（以下、単純適中率*という。）ことから、発生頻度の小さな現象に対しては、予測精度の評価法として適切ではないと言われている。³⁾従ってこうした単純適中率の代わりに、発生頻度の少ない現象の適中率を示すものとしてスレット・スコア*を使うことにした。

また、③の χ^2 -検定による方法及び④の情報量の計算による方法は、予報効果の評価法であり、予報の適中度を示すものではない。⑤のパーセント・スコアは、単純適中率に対して起こりにくい事象を当てた場合を起こりやすい事象を当てた場合よりもウエイトを高く評価することを考慮したものである。しかしウエイトを経験的にそれぞれに0から1の範囲の設定する必要があるため取扱いが難しい。

こうしたことから、3つの予測値それぞれの検証にあたっては、①スレット・スコアによる方法、②相関係数による方法、③スキル・スコアによる方法を使用した。

次に実況値が注意報発令基準である120ppb以上となる場合、各予測値の判定基準値を検討するため、判定値を変化させ、スレット・スコアとバイアス・スコアにより評価を行った。²⁾

現在、予報の精度を上げるための手法として、2つの予測値を組み合わせて判断を行っている。そこで、重回帰予測と人予測とはその関係が独立ではないことから組合せの対象とはしないで、従来のガイダンスの組み合わせとしてEWS予測と重回帰予測と新たな組み合わせであるEWS予測と人予測による予測精度について、W/Fを計算し比較検討を行った。

3. 結果と考察

3.1 各予測値の予測精度

表2～4は、実況値（実測最高濃度）と各予測値の閾値を120ppbにし、2値のカテゴリー問題に変換した5年間の予測と実況（実測最高濃度）の対比表である。

そこで、これに基づきスレット・スコア、スキルスコアを計算した結果と、予測として技術性がない前日の最高濃度を当日の予報とする場合を「前日継続予測」として表5に示した。

この結果、3つの予測値の予報精度は、単純適

中率では、EWS予測が一番高いが、高濃度に着目した適中率であるスレットスコアは、人予測、EWS予測、重回帰予測の順に落ちている。特に重回帰予測は、前日継続予測とあまり変わらない精度となっていることがわかった。このことは、重回帰予測の前日最高濃度の影響が過大であることも考えられる。

相関係数（決定係数）による検証結果は、散布図を図2に示した。重回帰予測は、ばらつきが大きく、予測値の実測値に対する説明力が低いことが分かった。また、回帰直線の傾きから、平均値としてのズレは人予測、重回帰予測において実況より4～5%高めなっており、逆にEWS予測は4%程度低めになっている。

以上の検討から、予報発令のためには、これまでの重回帰予測とEWS予測の組合せに代えて人予測とEWS予測を組み合わせとすることがあることがわかった。

3.2 各予測の判断基準に関する検討

3.1では、予測値により予報を発令する基準を実況値に合わせて120ppbに置いて検討した。しかし、図2の散布図を見ると予測値100ppb以上でのばらつきが大きいことがわかる。そこで、予報発令を行う予測値の基準を検討するため、判定値を変化させスレット・スコアとバイアスコアを求めたものが表6である。

スレットスコアは、適中率の一種であることから値が高い方が、適中率が良いことを示す。またバイアスコアは、1に近いほど見逃しと空振りのバランスがとれていることを示している。

表2 人予測と実況の対比表

予測 \ 実況	120ppb以上	120ppb未満	合計
	120ppb以上	25	
120ppb未満	31	959	990
合計	56	989	1045

表3 重回帰予測と実況の対比表

予測 \ 実況	120ppb以上	120ppb未満	合計
	120ppb以上	15	
120ppb未満	47	952	999
合計	62	991	1053

表4 EWS予測と実況の対比表

予測 \ 実況	120ppb以上	120ppb未満	合計
	120ppb以上	17	
120ppb未満	44	907	951
合計	61	920	981

表5 予測の検証結果

方法 \ 予測	人予測	EWS予測	重回帰予測	前日継続予測
単純適中率	0.941	0.942	0.918	
スレット・スコア	0.29	0.23	0.14	0.12
スキルスコア	0.42	0.34	0.22	0.18

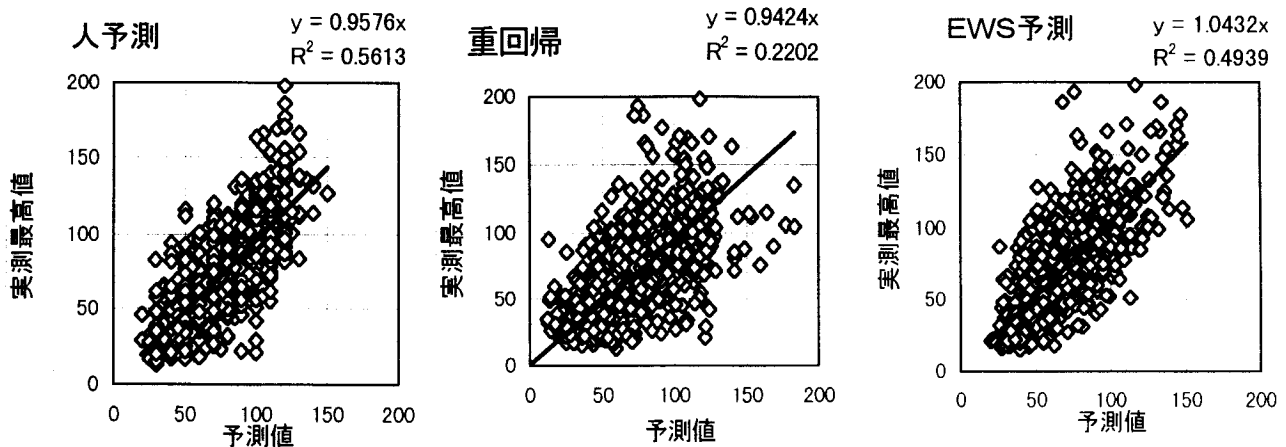


図2 各予測値と実測値（最高濃度）の散布図と回帰式

表6 予測判定値によるスレットスコアとバイアスコアの変化

予測の種類 判定値 (ppb)	スレットスコア			バイアスコア		
	人予測	重回帰予測	EWS予測	人予測	重回帰予測	EWS予測
1 2 0	0.29	0.14	0.20	1.0	0.9	0.4
1 1 5	0.35	0.17	0.22	1.1	1.2	0.5
1 1 0	0.33	0.16	0.26	1.7	1.7	0.8
1 0 5	0.34	0.19	0.22	2.1	2.6	1.0
1 0 0	0.25	0.18	0.25	3.5	3.1	1.4
9 5	0.23	0.18	0.25	3.8	3.5	1.9
9 0	0.17	0.17	0.23	5.6	4.0	2.6

注 人予測、重回帰予測、EWS予測と実況値の全てがそろったデータ (n=950) を使用。

* スレットスコアの白抜き文字は最高値を示す

* バイアスコアのグレーのセルは空振り領域を示す

これまでの予報発令判定基準は、重回帰予測120ppb以上かつEWS予測120ppb以上としていたが、予測判定値120ppbにおけるバイアスコアは、重回帰予測で0.9、EWS予測で0.4といずれも1.0未満であった。従って、従来の判定が「見逃し」をしやすい判定値設定であることが確認できた。

表6のスレットスコアの最高値は、人予測が115ppb、重回帰予測が105ppb、EWS予測が110ppbである。人予測に付いては、115ppb～105ppbの範囲ではスレットスコアが0.33以上でありあまり差がない。しかし、105ppbは、バイアスコアが2.1であり、予報回数のうち半分以上が空振りとなる可能性がある。また、人予測では、数値が5単位に丸められて提供されるので、その5年間の予測回数をみると、125ppb以上が16回、120ppbが39回、115ppbが7回、110ppbが33回、105ppbが19回となっている。これは5単位より10単位のほうが出しやすいという予報担当者の心理から起こる偏りと考えら

れるので、バイアスコアが、115ppbよりも悪くなるが、110ppbを判断基準と設定することで、予報発令回数の増加が期待できる。

EWS予測については、100ppbのスレットスコアも0.25と110ppbと大きな差がない、しかし、人予測が「空振り」の傾向があることから、バイアスコアは0.8で「見逃し」がわずかに卓越する110ppbに判断基準の設定を設定することも考えられる。

EWS予測の予測回数は、114～110ppbが15回、109～105ppbが12回、104～100ppbが21回となっており、予測回数からも判断基準の優劣は判断できなかった。

そこで次に、人予測について115ppbと110ppbを判断基準として、EWS予測値の判断基準を110～100ppbで変化させ予測値の組合せによる精度の検討を行った。

3. 3 予測値の組合せによる精度の検討

5年間のデータに、人予測とEWS予測の判断基準を組合せて比較したのが表7である。

これを見ると、EWS予測の判断基準が105ppbの場合は、適中率では判断基準110ppbと同じになり、空振り率が高くなる。従って、判断基準として110ppbのほうが適切であることがわかる。

人予測115、110ppbとEWS予測110、100ppbの選択は、適中率重視か空振り率重視かの違いで選択できることがわかった。

なお、空振りについては、その内容も重要である。120ppb以上の高濃度になるとして予報を出して、120ppbからはるかに低い濃度となれば、予報の信頼を失うことになる。そこで、空振りの場合の実況値について、判断基準ごとに平均値と分散を求めたが、有意な差は認められなかった。

以上から、空振り率を少なくすることを重視する場合は、人予測115ppb、EWS予測110ppbを判断基準とし、適中率を重視する場合は、人予測110ppb、EWS予測100ppbを基準とすることになる。この両者の中間であれば、人予測115ppb、EWS予測100ppbか人予測110ppb、EWS予測110ppbのいずれでも良いことがわかった。

光化学オキシダントについては、この5年間で高濃度日（最高濃度120ppb以上を記録した日）62日のうち、光化学スモッグ注意報の発令が38日であった。しかし、当日予報の発令日数は6日と著しく低い状態であった。

そこで、人予測110ppb、EWS予測110ppbを判断基準として設定し、5年間のデータに当てはめたところ、予報発令は31日になった。W/Fで評価すると、現在の6.3から $38/31=1.2$ へと見かけ上の

改善が見られた。しかし、高濃度日62日のうち18日が適中したが、44日見逃しであり、また予報の空振りは13日となった。そこで、実質的なW/Fは $38/18=2.1$ と考えられる。

今回の判断基準で、従来の基準より予報発令が5倍の31日となり、3倍にあたる18日が高濃度日となり、予報の空振りがあるが、見逃しを20%改善できることがわかった。

しかし、見逃しがまだ44日あり、その原因を分析すると、人予測によるものが2日、EWS予測によるものが19日、両予測とも見逃しが14日、どちらかの予測データが得られなかったのが9日であった。このことから、EWS予測式の改善が、精度の向上には重要といえる。

4. まとめ

光化学スモッグの当日予報の精度を向上させるため、気象予測の検証方法を使って検討したところ、次の結果が得られた。

(1) これまで予報に使用してきた民間気象会社提供の重回帰予測は、単純適中率では92%と人予測、EWS予測の94%と大差がないが、相関係数及び発生頻度の少ない現象の評価法であるスレットスコア及びスキルスコアでの評価は著しく低いため、これを予報の判断基準として使用することは問題があり、人予測とEWS予測を使うことが適切であることがわかった。

(2) 予測結果を濃度ポテンシャルとして考え、当日予報の判断基準を検討したところ、これまでの判断基準値である120ppbから110ppbに変更し、人予測とEWS予測を組み合わせることにより、従前より見逃しを21%改善できることがわかった。

表7 EWS予測の判断基準による予測精度の変化

	人予測判断基準値	EWS予測の予報発令		
		110ppb	105ppb	100ppb
予報回数 (A)	115ppb	25	28	33
適中回数 (B)		17	17	19
適中率 (B/55)		30.9%	30.9%	34.5%
空振り回数 (C)		8	11	14
空振り率 (C/A)		32.0%	39.3%	42.4%
予報回数 (A)	110ppb	31	36	44
適中回数 (B)		18	18	22
適中率 (B/55)		32.7%	32.7%	40%
空振り回数 (C)		13	18	22
空振り率 (C/A)		41.9%	50.0%	50.0%

* 5年間の最高濃度 120ppb 以上で2つの予測値データが揃っている 55 日について検討をした。

5. おわりに

光化学スモッグの当日予報は、被害の未然防止ということから非常に重要なものである。今回、予測値をポテンシャルとして捉え、実況の濃度との関係を整理することで改善ができることがわかった。しかし、適中率（スレットスコア）は30%程度でありさらに二つの予測値の AND、OR の検討などによる改善が必要である。

また、EWS予測の見逃しの事例を詳細に解析し、気象予測に起因するもの、予測式に起因するもの、汚染小気塊発生などの発生機構に起因するものなどの原因別に分け、精度向上の手法を検討することになっている。

参考文献

- 1) 神奈川県環境農政部大気水質課：神奈川県における光化学大気汚染の現状と対策，（2000）
- 2) 鈴木栄一：気象統計学，地人書館（1981）
- 3) 立平良三：気象予報による意思決定，東京堂出版（1999）

注記

*単純適中率、スレットスコア、バイアスコア及びスキルスコアは、次のような関係になっている。

実況 \ 予報	現象あり	現象なし
予報あり	a回 適中	b回 空振り
予報なし	c回 見逃し	d回 適中

総回数： $N = a + b + c + d$

単純適中率 = $(a + d) / N$

スレット・スコアは、Critical Success Indexともいわれ、予報で現象ナシかつ実況で現象ナシを情報価値がないものとして除いた回数のうち適中した回数を比で表したもので次式により計算する。

スレットスコア = $a / (a + b + c)$

バイアス・スコアは、予測が「見逃し」と「空振り」のどちらかに偏っているかを示す数で次式より計算する。これが1より小さい場合「見逃し」卓越、1より大きいとき「空振り」卓越であることを示している。

バイアスコア = $(a + b) / (a + c)$

スキルスコアは、予測技術が低い方法での偶然的な適中と予測による適中の差から予測技術の高さを評価するもので次式により計算する。

スキルスコア = $(a + b - r) / (N - r)$

今回 r は、偶然適中度の代わりに持続性予報適中数 $r = \{(a + b)(a + c) + (c + d)(c + b)\} / N$ とした。