

# 判別関数法による光化学スモッグ予報

伊藤 政 志   宇田川   満   早 福 正 孝

岡 本 真 一  
(産業公害防止協会)

## 1 はじめに

光化学大気汚染に係る緊急時対策の一環として、各自治体では、緊急時の発令（予報、注意報、警報、重大緊急報）が実施されている。このうち、予報を確実にこなうためには、前日および当日の朝における光化学スモッグ発生の予測が必要となる。予測方法としては、重回帰式や数量化理論を主体とした統計モデルと化学反応式および拡散方程式に基づく物理モデルが開発されている。しかし、現在では、計算量、予測精度、入力データの整備などの問題により、物理モデルは使用されておらず、統計モデルが幾つかの自治体で使用されているに過ぎない。統計モデルの例としては、藤田<sup>1)</sup>、Lin<sup>2)</sup>、関東地方公害対策推進本部<sup>3)</sup>や、大阪府、神奈川県、千葉県で実施されているものがある<sup>4)</sup>。

しかし、現状では、上記の問題点により、簡便なチェックシート方式、フローシート方式が多く用いられている。これらの手法は、高濃度光化学大気汚染をもたらす気象条件として、日射量、最高気温、風向、風速、大気安定度等の気象要素の実況値および予想値を利用し、総合的に判断して予報を行なっている。これらの方式の運用には、熟練した技術者が不可欠である。より容易に、光化学スモッグ予報業務を実施するためには、客観的な定量的予報システムの開発が必要である。このような統計的予測方式としては、重回帰式が多く検討されているが、濃度予測値が、当面は発令基準を越えるか否かを判断することが重要であることを考慮すれば、判別関数法<sup>5),6)</sup>が最も適していると考えられる。

本研究では、光化学オキシダント濃度と気象条件および一次汚染物質濃度との関係についての解析を行ない、説明変数の選択、判別分析による予測モデルの作成、重回帰分析との比較を併せて行なった。

## 2 統計的予測式

オキシダント濃度  $y$  を気象条件などの説明変数  $x_i$  より予測する重回帰式は

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n \quad (1)$$

である。ここで、係数  $b_0, b_1, b_2, \dots, b_n$  は実測値  $y$  と推定値  $\hat{y}$  との差の分散を最小にするように求められる。このときの実測値  $y$  と推定値  $\hat{y}$  の相関係数が重相関係数である。

オキシダント濃度が基準値以上の高濃度日とそれ以下の低濃度を判別するには、

$$f = l_0 + l_1 x_1 + l_2 x_2 + \dots + l_n x_n \quad (2)$$

で計算される  $f$  の値が高濃度日と低濃度日の各グループ間で大きく分離できるように、係数  $l_0, l_1, l_2, \dots, l_n$  を決定する。すなわち、式(2)で計算される変数  $f$  の全分散（全変動）に対する、高濃度日での平均  $\bar{f}_1$  と低濃度日の平均  $\bar{f}_2$  の差の2乗（群間変動）の比率を最大にするような重み係数  $l_i$  を求めることである。このようにして求められる一次式が判別関数である。

なお、以下に述べる判別関数では、説明変数  $x_i$  を基準化（平均が0、分散が1になるように変数変換すること）したものを使用し、各変数間の関係についても検討した。

## 3 気象条件および1次汚染質とオキシダント濃度の関係

まず始めに、学校情報の基準である 0.10 ppm 以上の高濃度日における気象条件などを明らかにするために、表1に示す各変数とオキシダント (Ox) 濃度の関係について検討した。ここで、Ox濃度としては東京都内の総合測定室での1時間値の最大値を使用した。また、最大混合層高度 (MMD) およびNO<sub>2</sub>濃度とノンメタン炭化水素 (NMHC) 濃度については表2のように4ケースを設定した。

表1 統計モデルで使用した変数

<ul style="list-style-type: none"> <li>◦ 風速(日平均)</li> <li>◦ 最高気温</li> <li>◦ 湿度(日平均)</li> <li>◦ 日照時間</li> <li>◦ 最大混合層高度MMD*</li> <li>◦ 前日Ox最高濃度</li> <li>◦ NO<sub>2</sub>濃度*</li> <li>◦ NMHC濃度*</li> </ul>
---

表2 MMDとNO<sub>2</sub>及びNMHC濃度

	MMD	NO <sub>2</sub> 及びNMHC濃度*
データ1	当日9時の 高層観測	7~9時の 3時間平均
" 2	前日21時	"
" 3	当日9時	日平均
" 4	前日21時	"

\*国設, 江戸川, 梶谷の3測定室の平均値

\*表2参照

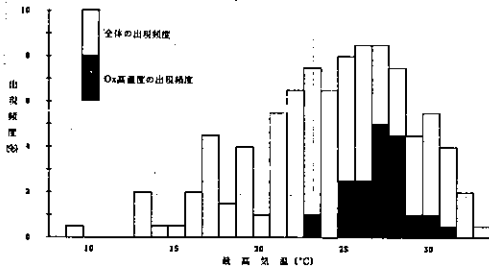


図1 最高気温とオキシダント高濃度(0.10 ppm以上)の出現日数の関係

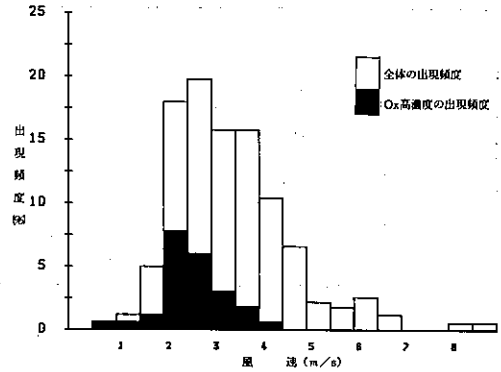


図2 日平均風速とオキシダント高濃度(0.10 ppm以上)の出現日数の関係

この解析に使用したデータは昭和57年4月より9月までの6か月間のものである。この期間中の最高気温と日平均風速の出現頻度のヒストグラムを図1, 図2に示す。この図には、Ox濃度が0.10 ppm以上となった高濃度日も図示してあるが、この結果より、最高気温が23℃以下および日平均風速4.5 m/s以上の日には高濃度汚染が生じていないことがわかる。

次に、表1, 表2に示す説明変数を用いて、重回帰分析および判別分析により、Ox高濃度汚染の推定を行なった。推定結果の評価には、表3に示すような分割表が便利である。また、重回帰式と判別関数の推定精度を式(3), 式(4)によって計算される正解率1および正解率2によって評価した。

$$\text{正解率1} = \frac{a + d}{a + b + c + d} \quad (3)$$

$$\text{正解率2} = \frac{d}{c + d} \quad (4)$$

ここで、a, b, c, dは表3に示す分割表中での度

表3 分割表

		推定値	
		低濃度	高濃度
実測値	低濃度	a	b
	高濃度	c	d

数であり、正解率1は全体の精度、正解率2は高濃度日に注目した判別精度の尺度である。

前日21時の館野高層気象観測データと当日の最高気温より求めたMMDと当日朝のNO<sub>2</sub>およびNMHC濃度を用いたケース2(表2におけるデータ2)での判別分析と重回帰分析の結果を表4, 表5に示す。この結果を見ると、全体の推定誤差を小さくするという観点からは重回帰式の精度が高いが、このために推定値が平均値の周囲に偏る傾向が見られる(図3参照)。したがって、重回帰分析では低濃度日より出現日数

の少ない高濃度日のみの判別率（正解率2）は低くなり、高濃度日に注目した推定精度では判別分析の方が良好である。

判別関数の値（スコア）と実測Ox濃度の散布図を図4に示す。この図を見ると、散布図のパターンは図3の重回帰式と類似であるが、高濃度日のプロットが図中の右側に寄っており、判別能力は重回帰式よりも高いことがわかる。

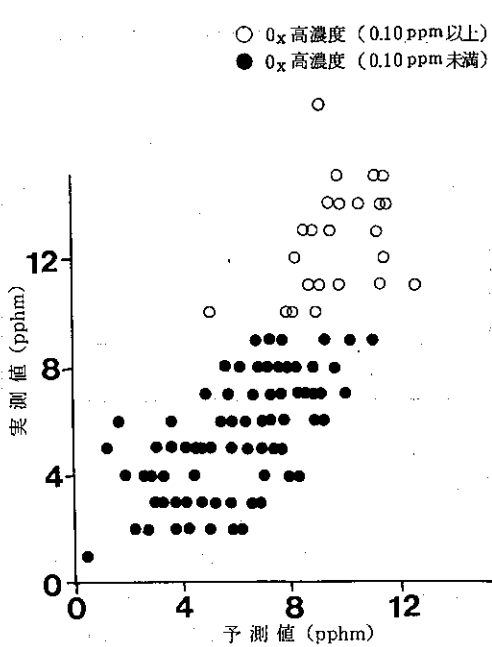
表4 テータ2を用いた場合の判別分析と重回帰分析の比較（昭和57年4～9月）

重回帰		推定		判別				
		低	高					
実測	低	112	4	実測	低	90	26	合計
	高	18	10		高	4	24	
合計		130	14	合計		94	50	144

表5 判別分析と重回帰分析の比較

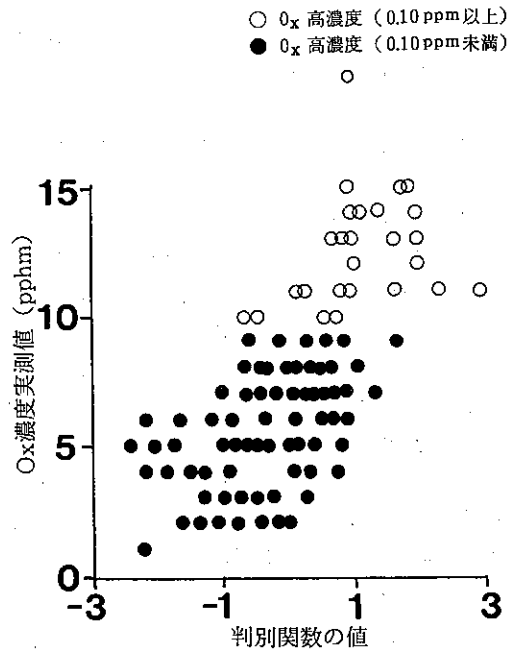
使用データ	重回帰分析			判別分析		
	重相関係数	正解率1	正解率2	マハラノビス距離	正解率1	正解率2
データ1	0.73	85%	36%	63.8	80%	86%
" 2	0.73	85	36	64.3	79	86
" 3	0.74	88	46	73.7	81	86
" 4	0.74	88	43	74.3	83	86

\*マハラノビス距離は基準化しないデータの分散共分散行列より計算した値



重回帰分析 CASE 2  
MMD 21時  
NO<sub>2</sub>, NMHC 7～9時

図3 重回帰式による計算値と実測値の比較



判別分析 CASE 2  
MMD 21時  
NO<sub>2</sub>, NMHC 7～9時

図4 判別関数の値（スコア）と実測Ox濃度の関係

判別関数の重み係数および判別関数の値とOx濃度との相関係数を表6に示す。この判別関数は基準化された説明変数に基づくものであり、各説明変数の相対的な重要性は重み係数の絶対値の大小で評価できる。この結果を見ると、風速、最高気温、日照時間、前日Ox、NO<sub>2</sub>、NMHC濃度が重要であり、MMDと湿度はあまり重要な要因ではないことがわかる。NO<sub>2</sub>とNMHC濃度の重みは同程度であるが、日平均値を用いた場合には、NMHC濃度の方が重みが大きくなっている。

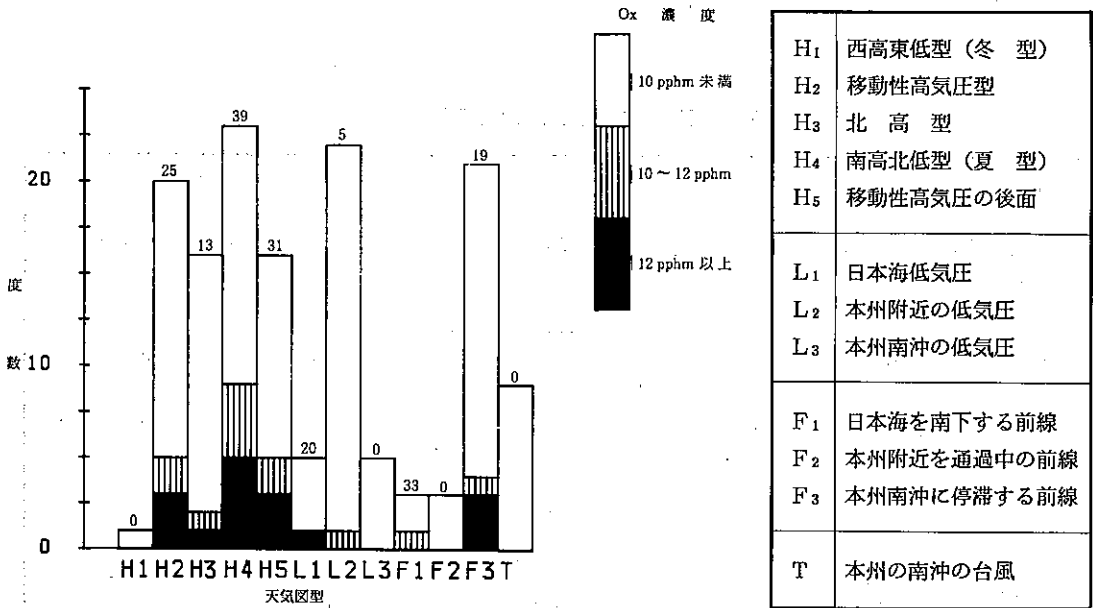
4 当日予報モデルの検討

光化学スモッグの当日予報では、当日午前9時前後までに入手できる気象要素等の測定値および十分な精度の得られる予測値を説明変数として、予測式を構成することが必要である。最高気温、風速、日照時間などの気象要素について、予測値と実況値の比較は環境庁<sup>4)</sup>が実施している。この結果によれば、最高気温は

表6 判別関数の重みベクトル

変数	ケース1	ケース2	ケース3	ケース4
1. 風速	-0.38	-0.36	-0.12	-0.12
2. 最高気温	-0.36	0.39	0.35	0.36
3. 湿度	-0.11	-0.09	-0.08	-0.10
4. 日照時間	0.32	0.35	0.32	0.36
5. MMD	-0.12	-0.15	-0.01	0.11
6. 前日Ox濃度	0.30	0.28	0.35	0.33
7. NO <sub>2</sub> 濃度	0.28	0.32	0.20	0.22
8. NMHC濃度	0.21	0.21	0.50	0.48
全変動と群間変動の比率	0.56	0.56	0.58	0.59
判別スコアとOx濃度との相関係数	0.68	0.68	0.68	0.68

計算条件のケース1~4は表2におけるデータ1~4に対応する。



数値は0.10 ppm以上の高濃度日の出現比率(%) (全体で19.7%)

図5 気圧配置別出現頻度とオキシダント高濃度汚染の出現頻度

よく一致しているが、日照時間はバラツキが大きい。したがって、当日予報の作成時に正確な予測値を得ることが困難な日平均風速、湿度、日照時間を削除し、気圧配置型および台風発生の有無を説明変数として追

加した。

気圧配置型別の出現日数とOx高濃度汚染出現頻度を図5に示す。ここで、Ox高濃度汚染の基準は学校情報を発令する0.10 ppm以上と注意報を発令する

0.12 ppm 以上の 2 つの場合について検討した。この結果を見ると、H<sub>4</sub> 型（夏型）及び H<sub>5</sub> 型（移動性高気圧の後面）で高濃度汚染の頻度が高くなっている。海風発生の有無による O<sub>x</sub> 高濃度汚染の出現頻度を見ると図 6 に示すように、海風が発生した日に高濃度汚染の頻度が高くなっている。

この結果に基づいて、気圧配置型及び海風発生の有無を数量化して、判別関数による予測式に導入した。気圧配置型については、O<sub>x</sub> 最高濃度が 0.10 ppm を越える割合が 20% 以上である H<sub>2</sub>, H<sub>4</sub>, H<sub>5</sub>, L<sub>1</sub>, F<sub>1</sub> 型（図 5 参照）の場合「1」、その他の場合は「0」とした。また、海風の有無に関しては、海風の発生した場合を「1」、ない場合を「0」として数量化した。

以上の説明変数を用いて、0.10 ppm 以上及び 0.12 ppm 以上の出現を判別関数により解析した結果を表 7 に示す。マハラノビスの距離、全分散に対する群間変動の比率、正解率などを見ると、海風の有無及び気圧配置型を導入した方が、判別能力が高くなっている。また判別関数の重みベクトルの値により、海風の有無、及び NO<sub>2</sub> 濃度が重要な要因であることがわか

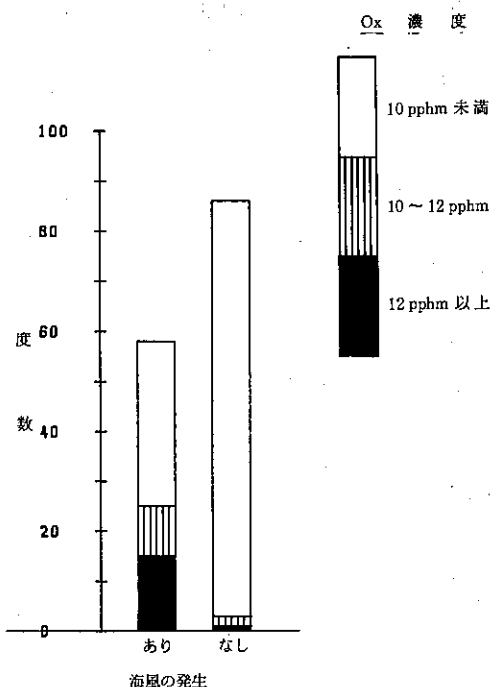


表 7 判別分析による O<sub>x</sub> 高濃度汚染予報の評価

計算条件	ケース 5		ケース 6	
	10pphm	12pphm	10pphm	12pphm
高濃度日の基準 (O <sub>x</sub> 濃度)	10pphm	12pphm	10pphm	12pphm
マハラノビス距離**	5.08	4.34	10.55	7.05
全変動と群間変動の比率	0.51	0.48	0.65	0.58
判別スコアと O <sub>x</sub> 濃度との相関係数	0.58	0.53	0.70	0.67
判別関数の重みベクトル				
1. 最高気温	0.56	0.49	0.38	0.35
2. 海風の有無	-	-	0.62	0.56
3. M M D*	0.08	- 0.09	- 0.15	- 0.25
4. 前日 O <sub>x</sub> 濃度	0.40	0.56	0.19	0.37
5. NO <sub>2</sub> 濃度*	0.43	0.71	0.24	0.53
6. NMHC 濃度*	0.29	- 0.08	0.21	- 0.09
7. 天気図型	-	-	0.18	0.10
評価				
正解率 1	78 %	78 %	83 %	87 %
正解率 2	79	69	89	94

\*印は表 2 におけるデータ 2 を使用

\*\*マハラノビス距離は基準化しないデータの分散共分散行列より計算した値

る。表6との比較では、NMHC濃度に係る重み係数が小さくなっている。これは、その他の説明変数との相関々係によって生じた差であると思われるが、条件によって重み係数の変動が大きい要因は、他の説明変数との相関が高いことも考えられるので、より長い期間のデータを解析し、説明変数の選択について再検討の必要があると思われる。

### 5 ま と め

判別関数によるオキシダント濃度の高濃度汚染を予報する数式モデルについて検討を行なった。この結果、当日予報に利用することを目的として、当日10時頃までに入手できる説明変数により、O<sub>x</sub>最高濃度が基準値0.10 ppmおよび0.12 ppmを超えるか否かを判別できる数式モデルを開発した。内挿時の正解率は80%程度であり、推定精度は良好である。しかし、説明変数の選択方法により、重み係数が変動する場合もあり、より安定した予測精度を保証するためには、より長期間のデータに基づく検証が不可欠であると思われる。

### 参 考 文 献

- 1) 藤田真一ほか：大気汚染の短期予測—オキシダントの予測手法を中心として—, 環境技術, 11, 31, (1982)。
- 2) Lin, G.: Oxidant Prediction by Discriminant Analysis in the South Coast Air Basin of California, Atmos. Environ., 16, 135, (1982)。
- 3) 関東地方公害対策推進本部, 大気部会, 光化学共同解析委員会：環境庁委託業務結果報告書(関東地域光化学オキシダント広域解析調査), 116, (1978)。
- 4) 環境庁：光化学大気汚染に係る統計解析モデル検討調査, (1982)。
- 5) 芝祐 順：行動科学における相関分析法, 3版, 東京大学出版会, 222, (1970)。
- 6) Anderson, T. W.: Introduction to Multivariate Statistical Analysis, Wiley, (1958)。
- 7) 岡本真一ほか：判別関数法による光化学スモッグ予報, 第24回大気汚染学会講演要旨集, (1983)。