

大気汚染の予測に関する基礎的研究

奥谷 巖*

Iwao OKUTANI*

まえがき

近年、地方都市においても交通量が急激に増加しており、特に幹線道路への集中は著しいものがある。自動車の速達性、随時性、戸口性などが、その強い輸送需要となって現われているため、今後ともこの傾向が続くことは避け難いと思われる。しかし、昨今、エネルギー資源の節約、環境保全、省労働力、省空間の面から、自動車を交通体系の一環とした総合的検討の必要性が指摘されている。特に、自動車交通を原因とする大気汚染、いわゆる車社会に対する批判の焦点となっていることは周知の事実である。

本研究は、環境保全の観点から、こうした自動車の負の要因のうち、とくに大気汚染に焦点を絞り、その予測に関して基礎的な研究を行ったものであるが、我が国の排出ガスによる大気汚染状況を見ると、一酸化炭素、硫黄酸化物については改善の傾向がみられるものの、窒素酸化物 (NO_x) については依然として悪い状況にあることに鑑み、汚染物質をとくに窒素酸化物をとり挙げることにした。

研究は大きく分けて2つの立場からなされている。すなわち、1つは環境アセスメントとしての予測という観点であり、空間を系とした場合の、物理的状态量と汚染濃度の相関分析がその骨子となっている。他の1つはダイナミックな短期予測であり、コントロールや警告指示とリンクすることがその最終的なねらいとなるものである。

系の状態量と汚染濃度の相関分析

まずはじめに、大気空間を1つの系と考えた場合の、その系の状態量と NO_x の汚染濃度との間の相関分析を行う。こうした分析が適当に類型化された地域について多数なされるならば、いわゆる環境アセスメントの手段になりうる。本研究でとり挙げた状態量は以下のとおりである。

- | | |
|-----------------|-------------------------------------|
| 1. 交通量 (台/時) | 2. X方向の風速 (m/秒) |
| 3. Y方向の風速 (m/秒) | 4. 気温 ($^{\circ}\text{C}$) |
| 5. 湿度 (%) | 6. 日射量 (cal/cm^2) |
| 7. 雲量 (割) | 8. 大気安定度 |

*信州大学工学部土木工学科 Fac. Eng., Shinshu Univ.

また、説明の便宜上、 NO_x 濃度を第9番目の状態量としておく。なお、A市の場合、大気安定度のデータはない。

分析手法としては重回帰分析法を採用し、データは関東A市、近畿B市、中国C市、九州D市の4地域で実測されたものを用いた。計測地点はそれぞれの地域で3地点であり、その位置関係は図1～図4に示したとおりである。各図は道路の断面図であり、図中の1～3の番号

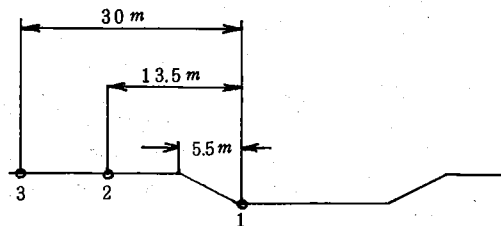


図1. 関東A市の計測地点

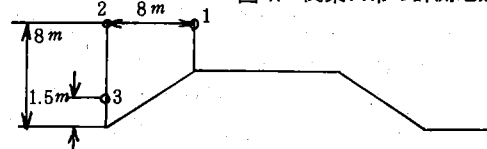


図2. 近畿B市の計測地点

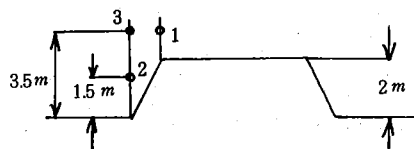


図3. 中国C市の計測地点

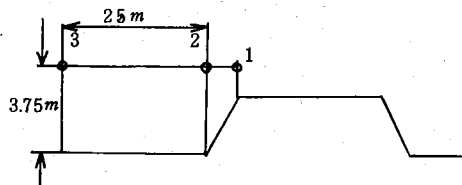


図4. 九州D市の計測地点

が計測地点を表わしている。なお、道路のY方向は交通の流れ方向を、X方向はそれに直角な方向をそれぞれ意味している。

NO_x の汚染濃度は、1時間ごとの平均値であり、合計

10数日分のデータが入手できた。

分析の結果は以下のとおりである。まず表1～表4は各状態量間の相関係数行列を示したものである。

表1. 相関係数行列 (関東A市)

	1	2	3	4	5	6	7	9
1	1.000							
2	-0.003	1.000						
3	-0.153	0.646	1.000					
4	0.503	0.030	-0.219	1.000				
5	-0.477	0.217	0.327	-0.595	1.000			
6	0.362	-0.026	-0.067	0.184	-0.345	1.000		
7	0.493	0.008	0.182	-0.409	0.392	-0.224	1.000	
9	0.613	0.121	0.169	0.006	0.085	0.156	-0.228	1.000

表2. 相関係数行列 (近畿B市)

	1	2	3	4	5	6	7	8	9
1	1.000								
2	0.223	1.000							
3	-0.043	0.152	1.000						
4	0.613	0.064	-0.002	1.000					
5	-0.732	-0.199	-0.111	-0.657	1.000				
6	0.593	0.623	-0.229	0.537	-0.729	1.000			
7	0.077	-0.113	-0.223	0.072	0.022	-0.009	1.000		
8	-0.361	-0.021	0.090	-0.193	0.287	-0.301	-0.194	1.000	
9	0.136	-0.088	-0.552	-0.049	0.126	0.185	0.262	-0.178	1.000

表3. 相関係数行列 (中国C市)

	1	2	3	4	5	6	7	8	9
1	1.000								
2	0.137	1.000							
3	-0.346	0.007	1.000						
4	0.624	0.189	-0.449	1.000					
5	-0.655	-0.306	0.410	-0.759	1.000				
6	0.490	0.147	-0.429	0.559	-0.718	1.000			
7	0.087	0.014	0.072	0.005	0.323	-0.169	1.000		
8	-0.497	-0.265	0.319	-0.396	0.537	-0.616	-0.007	1.000	
9	0.576	0.041	-0.710	0.561	-0.495	0.308	-0.064	-0.241	1.000

表4. 相関係数行列 (九州D市)

	1	2	3	4	5	6	7	8	9
1	1.000								
2	-0.145	1.000							
3	-0.269	0.236	1.000						
4	0.522	-0.274	-0.205	1.000					
5	-0.435	0.318	0.314	-0.555	1.000				
6	0.1399	-0.394	-0.253	0.619	-0.649	1.000			
7	0.065	0.159	0.036	0.178	0.118	-0.231	1.000		
8	-0.376	0.035	0.172	-0.351	0.240	-0.211	-0.240	1.000	
9	0.152	0.180	0.013	-0.124	0.189	-0.266	0.052	0.023	1.000

これらの表より次のようなことがいえる。すなわち、説明変数となる各状態量間では、気温、湿度、日射量の間にある程度の相関がみられるが、それも強い相関ではなく、また他の変量間には各地域全体を通じて相関が高いものはない。各表の最下行はNOxの汚染濃度(ppb)と他の状態量との単相関係数を示しているが、注目する程に相関の高いものはみあたらない。このことは、汚染濃度を単独の説明変数で説明することの困難性を示唆するものであろう。

このことを踏まえ、次にNOxの汚染濃度を外部基準とした場合の各状態量の標準重みベクトルを求めてみた。結果は表5に示したとうりであるが、本表中の1～8行

の各値が標準重みを、最下行の値が重相関係数(=ρ)をそれぞれ表わしている。標準重みの符号が地域によらず一定であるのは交通量のみであり、地域の特性(地形や建物などの条件)によって、各状態量が汚染濃度に与える影響の仕方はまちまちであることがわかる。また、重相関係数が0.7を超えるのは関東(A市)と中国(C市)の2地域のみであり、本研究でとり挙げた状態量だけで一般的な汚染濃度の説明式を得ることはむずかしいことがわかる。

表5. 各状態量の標準重み係数

	関東	近畿	四国	九州
1	1.125	0.860	0.471	0.543
2	-0.138	0.212	0.005	-0.118
3	0.289	-0.084	-0.719	-0.546
4	-0.115	-0.038	0.249	-0.202
5	0.535	0.288	-0.041	0.763
6	0.006	-0.717	-0.295	0.505
7	0.048	-0.125	-0.103	0.208
		-0.122	0.113	-0.094
	0.7714	0.4138	0.8304	0.6559

カルマンフィルターによる汚染濃度の予測

以上みてきたように、ここで扱っている状態量で汚染濃度を予測することは必ずしも容易ではないが、次にわれわれは近代制御理論の中で開発されたカルマンフィルター理論を応用して、その動的な短期予測の可能性について調べてみることにする。

(予測方法I)

予測のために使うm個の状態量を $x_1(t) \sim x_m(t)$ とし、これらを要素とするベクトルを $x(t)$ とする。ここに、 t は時刻を表わすものとする。現時点をいま t として、それ以前の状態量により、 k 時点先の汚染濃度を予測する式を次のように与える。すなわち、一般性を失うことなく $x_1(t)$ をNOxの汚染濃度としたとき

$$x_1(t+k) = H^0(t)x(t) + H^1(t)x(t-1) + \dots + H^k(t)x(t-r) + U(t) \quad (1)$$

ここに、 $H^l(t)$ は m 次元のパラメータベクトル ($l=0 \sim r$) であり、 $U(t)$ はノイズ項である。

いま、

$$A(t) = \begin{bmatrix} x^T(t) & & & & \\ & x^T(t) & 0 & & \\ & & x^T(t-1) & & \\ & & & x^T(t-1) & 0 \\ 0 & & & & x^T(t-r) & 0 \\ & & & & & x^T(t-r) \end{bmatrix}$$

$$h(t) = (H^0(t), H^1(t), \dots, H^k(t))$$

$$y(t) = x(t+k)$$

とすると、式(1)は

$$y(t) = A(t)h(t) + U(t)$$

表9. 方法IIによる予測精度 (近畿B市)

d	A			B			C		
	1	2	3	1	2	3	1	2	3
3	0.981 0.739	1.091 0.845	1.169 0.920	1.146 0.918	1.274 1.031	1.349 1.101	2.843 1.943	3.018 2.288	3.063 2.637
4	0.474 0.389	0.432 0.340	0.461 0.362	0.538 0.408	0.519 0.378	0.517 0.368	1.715 0.762	1.800 0.717	1.643 0.649
5	0.335 0.349	0.378 0.407	0.422 0.458	0.458 0.466	0.504 0.531	0.535 0.579	0.714 1.015	0.783 1.056	0.776 1.088
6	0.443 0.346	0.350 0.390	0.445 0.478	0.595 0.582	0.485 0.553	0.602 0.616	1.546 2.250	1.504 4.495	1.517 0.767
7	0.689 0.543	0.841 0.606	0.847 0.675	0.785 0.672	0.995 0.822	1.002 0.801	2.250 1.376	4.495 2.122	3.767 1.709
8	0.353 0.317	0.447 0.317	0.501 0.404	0.483 0.353	0.588 0.455	0.656 0.546	1.880 0.868	2.287 1.132	2.335 1.163
9	0.800 0.572	0.975 0.662	1.151 0.782	0.869 0.617	1.188 0.810	1.299 0.897	2.784 1.159	4.926 5.513	4.457 1.382
10	0.693 0.667	0.717 0.684	0.818 0.733	0.835 0.847	1.012 1.013	1.101 1.000	1.494 1.562	2.377 1.982	2.476 2.052
11	0.350 0.191	0.366 0.250	0.322 0.235	0.541 0.336	0.587 0.395	0.412 0.286	2.463 1.177	2.512 1.208	1.242 0.545
12	0.613 0.295	0.653 0.338	0.726 0.393	0.536 0.318	0.574 0.388	0.630 0.411	3.716 0.850	3.322 1.000	3.966 0.835
13	1.131 0.727	1.199 0.825	1.270 0.927	1.237 0.924	1.216 1.059	1.144 1.172	6.088 3.362	6.362 3.530	6.403 3.556
14	0.415 0.304	0.468 0.324	0.513 0.374	0.492 0.352	0.577 0.404	0.559 0.410	1.479 0.665	1.914 0.911	1.366 0.794
15	0.928 0.251	0.300 0.249	0.377 0.285	0.344 0.306	0.361 0.312	0.416 0.336	0.948 0.611	0.627 0.538	0.996 0.539

それぞれ表わしており、また、1~3の番号は1時間先~3時間先の予測を示している。各欄中、上段の値は真値を生データとした場合の誤差率を、下段の値は真値を平滑値とした場合の誤差率をそれぞれ表わしている。

関東では、予測方法IIの方がどちらかといえばよい精度を与えているのに対し、近畿では逆に方法Iの方がよりよい精度を示している。また日によって精度に大きな差があることがわかる。予測時点が何時間先かによって精度は著しく異なることはない。いずれにしても、一般的に予測精度はあまりよいとはいえず、今後何らかの方法によってデータ処理を行い、より安定した予測結果が得られるような検討をしてゆかねばならないと思われる。

Box-Jenkins法による汚染濃度の予測

カルマンフィルターを用いた予測では、汚染濃度を予測するために、交通量、風速をはじめとする9つの状態量を用いたが、それらの状態量をうまく整理し切っていないがために、満足ゆく予測精度が得られなかった。そこで、われわれは次に汚染濃度のみの時系列を用いてk時点先の汚染濃度を予測するという最もシンプルな方式を考えてみた。

ある予測変量を、その変量のみの過去の時系列から予測する方法として、最も柔軟性があるものはBox-Jenkins法といわれており、本研究でもその方法を採用する。本方法の最も一般的な予測方式は通常ARIMA(p;d,q)と表記されるが、ここではとくにIMA(d,q)のモデルを考える。

いま、予測地点の時刻tにおける汚染濃度を新たにZ_tと表わし、Bを後退作用素としたとき、IMA(d,q)モデルは最も一般的には次のような式によって表現することができる。

$$(1-B)^d Z_t = \theta(B) A_t$$

ここに

$$BZ_t = Z_{t-1}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

A_t: ホワイトノイズ

まず、dの決定は次のようにして行う。d = 0, 1, 2, …… に対し、(1-B)^dZ_tの時系列をつくり、自己相関係数のグラフを作成してゆく。そして、(1-B)^dZ_tのグラフで自己相関係数が急激に減衰していたならばそのときのdをもってdとする。また、qについては、W_t = (1-B)^dZ_tと似た時系列に対する自己相関係数のグラフにおいて、時間ずれqに対応する点まである程度の大きさがあり、q+1の時間ずれに対応するもので急激に0に近い値まで値が落ち込んでいたならば、このW_tに対しこのqを用いる。なお、上述の判断過程で自己相関係数を事実上0か否かを判定するためには、Bartlettの公式を用いる。すなわち、いまnをデータ数としたとき

$$\sigma(r_s) \div \left(\frac{1}{n}\right) \{1 + 2(r_1^2 + r_2^2 + \dots + r_q^2)\}^{\frac{1}{2}} \quad s > q$$

r_j: j時間のずれに対する自己相関係数
なるr_sの標準偏差に対し、2σ(r_s)なる変動範囲を考え、その中にr_sが入っていたならば事実上0とみなすのである。普通、dもqも1~2の範囲にある。

次に、θ₁~θ₂のパラメータ決定法について述べる。まず、時間ずれsに対してW_tの自己共分散r_sを求める。このとき、われわれは次のような非線形方程式を得る。

$$\begin{aligned} \sigma_a^2 &= r_0 / (1 + \theta_1^2 + \dots + \theta_q^2) \\ \theta_j &= -(r_k / \sigma_a^2) + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots \\ &\quad + \theta_{q-j} \theta_q \quad (j=1 \sim q-1) \end{aligned}$$

ここに、r_k = r_k / r₀

ここに示した移動平均過程が定常であるためには、まずq=1の場合

$$-0.5 < r_1 < 0.5$$

$$-1 < \theta_1 < 1$$

でなければならない。また、q=2の場合

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$-1 < \theta_2 < 1$$

でなければならない。したがって、非線形連立方程式の解もこうした範囲から選ぶようにする。

(d,q)の組み合わせは、上で説明したようにして理論的には決められるのであるが、実際のデータを扱った場合には、必ずしも判断が明確でないことが多い。そこで

以下で述べる適用結果では、 $d \leq 2$ 、 $q \leq 2$ の範囲で、定常過程となるすべての組み合わせについて、検討している。

表10～表13が、先に示した表6等に対応する予測精度を示したものである。先と同様に各欄の上段の値は対生

表10. 予測精度 (関東1地点)

(d,q)	A			B			C		
	1	2	3	1	2	3	1	2	3
(1,1)	0.354 0.117	0.496 0.242	0.605 0.353	0.393 0.148	0.533 0.287	0.634 0.404	2.661 0.550	4.525 0.887	5.701 1.765
(1,2)	0.347 0.113	0.486 0.233	0.597 0.344	0.387 0.142	0.523 0.278	0.629 0.397	2.529 0.516	4.366 0.866	5.628 1.735
(2,1)	0.302 0.128	0.467 0.296	0.686 0.703	0.357 0.158	0.547 0.353	0.791 0.584	2.371 0.634	3.255 1.150	4.182 1.979
(2,2)	0.304 0.130	0.476 0.300	0.703 0.508	0.360 0.161	0.559 0.365	0.810 0.604	2.396 0.646	3.317 1.208	4.273 2.116

表11. 予測精度 (近畿1地点)

(d,q)	A			B			C		
	1	2	3	1	2	3	1	2	3
(1,1)	0.408 0.134	0.517 0.227	0.566 0.287	0.442 0.166	0.548 0.276	0.590 0.350	3.792 0.647	4.507 1.183	6.775 2.062
(1,2)	0.408 0.134	0.517 0.227	0.566 0.287	0.442 0.165	0.548 0.275	0.589 0.350	3.784 0.640	4.494 1.179	6.756 2.054
(2,1)	0.410 0.189	0.640 0.385	0.826 0.576	0.487 0.235	0.470 0.694	0.912 0.694	4.448 0.901	7.600 2.318	13.729 4.800
(2,2)	0.409 0.189	0.639 0.383	0.824 0.575	0.467 0.236	0.709 0.469	0.910 0.692	4.443 0.899	7.582 2.311	13.686 4.783

表12. 予測精度 (中国1地点)

(d,q)	A			B			C		
	1	2	3	1	2	3	1	2	3
(1,1)	-								
(1,2)	0.506 0.137	0.733 0.289	0.958 0.443	0.541 0.169	0.777 0.338	0.986 0.507	5.956 0.579	10.102 1.226	10.956 2.241
(2,1)	0.431 0.189	0.715 0.433	1.198 0.737	0.513 0.228	0.854 0.487	1.286 0.801	5.224 0.777	12.315 1.880	14.468 3.740
(2,2)	0.436 0.195	0.803 0.499	1.363 0.860	0.520 0.235	0.945 0.560	1.460 0.942	5.303 0.792	13.669 2.235	17.251 4.728

表13. 予測精度 (九州2地点)

(d,q)	A			B			C		
	1	2	3	1	2	3	1	2	3
(1,1)	0.468 0.131	0.615 0.248	0.784 0.355	0.526 0.179	0.646 0.300	0.745 0.412	8.832 0.624	9.491 1.659	12.07 1.838
(1,2)	0.466 0.130	0.610 0.244	0.744 0.351	0.524 0.176	0.643 0.297	0.742 0.409	8.761 0.612	9.393 1.634	11.82 1.811
(2,1)	0.519 0.243	0.780 0.482	1.012 0.722	0.649 0.330	0.878 0.618	1.172 0.927	8.945 1.205	8.905 2.166	8.747 3.927
(2,2)	-								

データの誤差率であり、下段の値は対平滑化データの誤差率である。対平滑値の単純平均誤差率で評価すると、1時間先の予測では(d,q)の組み合わせの選び方次第では(この適用例では(1,2)が最もよい)、10数%の誤差率が得られ、実用に耐える可能性を示している。しかしながら、最大誤差率は50%以上であり、かなり大きいし、2時先、3時先の予測では誤差率は急激に増えてゆくという問題点を残している。

図6および図7は、関東A市と中国C市の予測結果を4日分グラフにプロットしたものであるが、平滑値を用

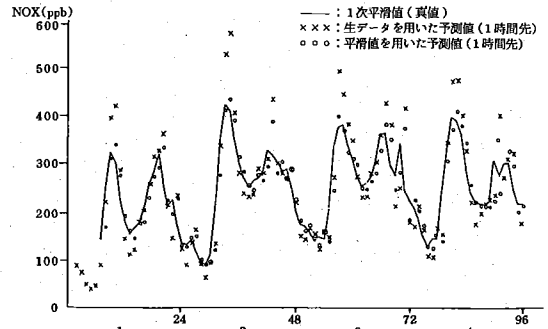


図5. 関東A市の予測値と真値

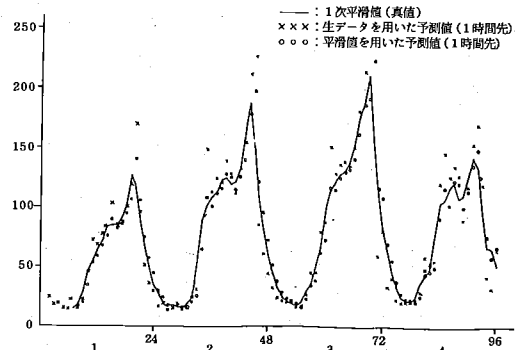


図6. 中国C市の予測値と真値

いた予測値は真値の平滑化データにある程度近い形で変化していることがわかる。とくに、中国の場合はよい予測値を与えているようである。

こうした結果から判断すると、汚染濃度に影響を与えらると思われる多くの状態量を予測モデル式に持ち込んで複雑な予測過程を作成するよりも、少なくとも1時先程度の短期間予測ならば、単なる汚染濃度の時系列のみで予測をする方がかえって精度が高くなるといえそうである。

参考文献

芝祐順：相関分析法 東大出版会

R. K. Kalman: A New Approach to Linear Filtering and Prediction Problem, Trans. Soc. Mech. Engrs., 1960, 82, D, pp. 35-45

Box & Jenkins: Timeseries Analysis, Holden-Day