制御理論と多変量解析を組み合せた環境汚染の統計的予測法

東京工業大学　　日野 幹雄

要旨：著者は先に制御理論におけるカルマン・フィルターによる大気汚染の予測の方法[1]について発表した。本報告では，これを更にすゝめて回帰分析法と因子分析法の考え方を加えて，環境汚染その他の確率的な現象のより信頼度の高い予測法を提案する。

I. 序　論

a) 物理学的自然感とその限界：自然現象というものは，たとえそれがどんなに複雑であっても，小数の単純普遍的な基礎的法則の上に成り立っているというのが，われわれの物理学的自然観である。われわれの生活圏である大気や水の環境汚染は，乱流拡散という物理現象に支配されていることは疑のない事実である。それにもかゝわらず，現象を微分方程式等で記述される基礎法則のみで解明しえないのは，現象のスケールおよび自由度が極めて大きく，かつ重ね合せの原理の成立しない非線型性の強い乱流状態が対象であり，確率性が高いためであると考えられる。

さらに，大気汚染や水質汚濁の場合のように化学的プロセスが入り込んだり，また汚染質の排出源の状況が適確に把握しえない点も問題を複雑にしている。

こゝに，環境予測において確率過程論的方法の役割の重要性があると考えられる。

b) 大気汚染予測法の分類：現在行われている大気汚染の予測方法を分類したのが表－1である。水質汚濁についてもほゞ同様である。これらの方法は物理法則を基本とする方法と物理法則にとらわれない非物理的な方法とに大別される。物理的方法はさらに次のように分類される。

① 基礎方程式を厳密に解く方法[2,3,4,5]：この場合には，地形や気象條件などの複雑な條件を考慮しうる反面，拡散係数や気象的境界條件の変化はそれほど厳密に理論的に取扱えず，また計算時間も尨大でオンラインの予測に適していない。

② 単純化された物理的方法：Taylor の拡散理論は均一定常な場に対して立てられているが，大気は均一でも定常でもない。しかし，局所的にまた短時間で考える場合にはこの仮定がほゞ成立するから汚染質の変動を時間や拡りを限って追跡する方法である。この方法は実用上の簡易さと精確さが程良く実用の面で有利である。

③ 制御理論による修正を考慮する物理的方法：汚染質の輸送連続式を制御理論による修正を行いつゝ解く方法で，混合層とか多重ボックスなど設定しその内部では濃度は一様と考えている。[6,7,8]

非物理的な方法には，次の二つがある。

④ 統計的方法：これには回帰分析法と因子分析法がある。これらは過去の資料の statistical な処理結果を利用するもので，静的予測法であり，現在いくつかの地方自治体で採用されている方法である。[9]

⑤ 適応制御理論：予測と時々刻々の実測とのズレを学習的に予測に用いる方法[1]で，特にカルマン・フィルター法が有効である。この場合には，物理モデルにおける輸送方程式は用いずに，環境システム方程式を未知と考え，これを実測データーにより同定しようとする点に特徴がある。

II. 因子分析法を組み込んだ回帰分析法による予測

a) 回帰分析法による予測

良く知られているように，一つの要素にはそれに影響を及ぼし，あるいは関連するいくつかの要素を見出すこ

表－1　大気汚染予測モデルの分類

分類	方法	モデル名
物理モデル	Physical model	数値シミュレーション (日野, APCS) (大西) (島貫)
	simplified physical model	プルーム　(横山) プルーム　(林) 二層プルーム(公害気象) 二層パフ　(日立)
	simplified physical model/control theory	混合層　(高松－least sq.) 多重ボックス(日立－Kalman filter)
非物理モデル	Control theory	集中定数系(日野－Kalman filter)
	statistical method	回帰分析　(東京都) 因子分析　(藤田)

とができる。物理学的方法はこれらのいくつかの要素の間を支配する単純な普遍的な法則を発見し，それをもとにして諸要素の間の複雑な関係を解明しようとするものである。これに反して，心理学とか医学では諸要素の間の関係を支配する単純な定量的法則を発見することはほとんど困難であるので，多くの実験データを客観的に処理してそれらの間の因果関係の一般則を見出そうとする，このような立場から，今世紀以後約3/4世紀の時間をかけて体系化されつゝある方法が多変量解析である。これには大別して回帰分析法と因子分析法がある。

いま$\mathbb{X}=[x_1, x_2, \cdots, x_n]^T$を予測すべき確率変量を成分とするベクトル，$\mathbb{Y}$を$\mathbb{X}$に影響をする諸要素を成分とするベクトルとする。時刻tにおいてk時間先の$\mathbb{X}$の値を推定するのに，回帰分析では$\mathbb{X}$と$\mathbb{Y}$との間に線型の関係を仮定し，[10]

(単純模型)　$\mathbb{X}(t+k)=\mathbb{A}\mathbb{Y}(t)+\mathbb{N}(t)$　(1)

とする。こゝに，$\mathbb{A}$は予測行列，$\mathbb{N}$は誤差ベクトルである。われわれの対象とする自然現象や社会現象は一般に日周期や年周期など明確な同期性を伴うから，$\mathbb{A}$は位相Tの関数と考えれば，

(周期模型)　$\mathbb{X}(t+k)=\mathbb{A}(T)\mathbb{Y}(t)+\mathbb{N}(t)$　(2)

こゝに，$\mathbb{Y}$は一般に$\mathbb{X}$の過去の値とそれ以外の入力$\mathbb{Z}$より成る。

$$\mathbb{Y}(t)=\begin{bmatrix} \mathbb{X}(t) \\ \mathbb{X}(t-1) \\ \vdots \\ \mathbb{X}(t-l) \\ \mathbb{Z}(t) \end{bmatrix} \tag{3}$$

予測行列$\mathbb{A}$の推定値は式(1)又は(2)より直ちに

$$\overline{\mathbb{A}} = (\overline{\mathbb{X}\mathbb{Y}^{T}})(\overline{\mathbb{Y}\mathbb{Y}^{T}})^{-1} \tag{4}$$

と求まる。こゝに，$\overline{\mathbb{X}\mathbb{Y}^{T}}$は$\mathbb{X}$と$\mathbb{Y}$との母集団についての共分散行列，$\overline{\mathbb{Y}\mathbb{Y}^{T}}$は$\mathbb{Y}$の母集団の分散行列である。実際計算においては，これらを標本平均で置き換えるが，標本数が少ない場合には不合理な結果を生まないよう統計的前処理が必要である。

b) 因子分析法による次元数の減少：

さて，上述の回帰分析の方法により予測を行おうとする場合には，変量ベクトルの次数が大きくなりすぎて実際計算には適当でなくなる場合が多い。そこで，取扱う変数の次元を減すために，因子分析法を導入する。[10,11]

一般に不規則な一つの現象に付随するいくつかの要素は互に独立とは限らずこれより少ない数の独立な成分より成り立つと考えられる。この独立な成分は因子(factor)と呼ばれる。因子は回帰分析における要素(Component)と異り，具体的な意味は必ずしもなく抽象的で非実在的なものである。一つの標本$\mathbb{X}$において各成分が変化するのは，この因子ベクトル$\mathbb{f}$の変化によるパターンの組合せの変化によると考えられる。これを次のように表わす。

$$\mathbb{X}(t+T) = \mathbb{F}(T)\,\mathbb{f}(t, T) \tag{5a}$$

$$\mathbb{Y}(t+T) = \mathbb{G}(T)\,\mathbb{g}(t, T) \tag{5b}$$

こゝに，$\mathbb{F}$・$\mathbb{G}$は行列で因子負荷行列(Factor loading matrix)と呼ばれるが，わかり易くいえば$\mathbb{X}$あるいは$\mathbb{Y}$を構成する基本パターンの集りである。また，$\mathbb{f}$・$\mathbb{g}$はそれぞれ$\mathbb{X}$・$\mathbb{Y}$より次元数の少ないベクトルで因子評点(factor Score)と呼ばれる。

式(5a)(5b)の関係を式(2)に代入するとき，

$$\mathbb{f}(t+k, T') = \mathbb{B}(T, k)\,\mathbb{g}(t, T) \tag{6}$$

の関係を得る。こゝに，

$$\mathbb{B} = [\mathbb{F}^{T}\mathbb{F}]^{-1}\mathbb{F}^{T}\mathbb{A}\,\mathbb{G} \tag{7}$$

であるが，わざわざ上式により計算しなくても，式(6)の回帰分析より求めうる。

Ⅲ. 適応制御理論との組合せ(KAREA法の提案)

上述の因子分析を併用した回帰分析法により，可成りの精度で長周期の変動を予測することができる。しかし，不規則変量の間には正確に線型関係が成立するものでもないし，また考慮されていない要素の影響もあるので，予測誤差はまぬがれ得ない。予測誤差の中には全くの雑音もあろうが，短周期の変動成分もかなりあると思われる。こうした平均的挙動からの徐々のズレは制御理論における学習的制御いわゆる適応制御理論により予測することが可能である。前論文ではカルマン・フィルター理論の書き換えによる方法について述べた。すなわち，カルマン・フィルター[12]は本来システム方程式の既知の系からの出力の状態推定法であるが，これを未知の系のシステム方程式の同定に応用した。[1,13]

時刻$t+k$での$\mathbb{X}$の因子ベクトルを回帰分析法による予測分$\overline{\mathbb{f}}(t+k)$とそれに対する誤差分$\widehat{\mathbb{f}}(t+k)$の和と考える。

$$\mathbb{f}(t+k) = \overline{\mathbb{f}}(t+k) + \widehat{\mathbb{f}}(t+k) \tag{8}$$

こゝに，

$$\bar{f}(t+k) = \mathbb{B}(T,k)\, g(t) \tag{9}$$

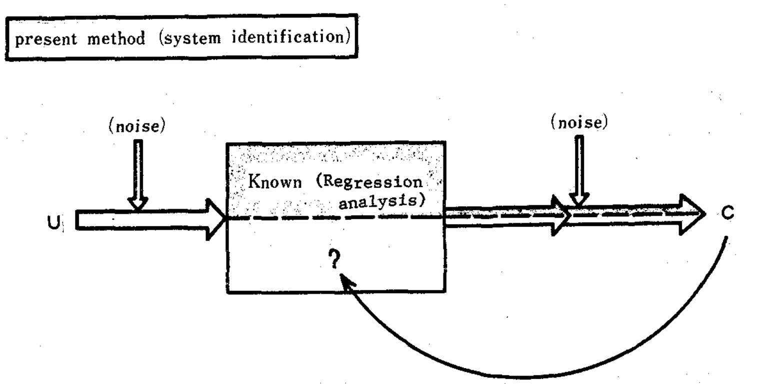
予測誤差因子$\tilde{f}$は過去のfの予測誤差により

$$\tilde{f}(t+k) = \tilde{B}_0 \tilde{f}(t) + \tilde{B}_1 \tilde{f}(t-1) + \cdots + \tilde{B}_m \tilde{f}(t-m) \tag{10}$$

の形式で推定しうると考え，係数行列$\tilde{B}_i$をカルマン・フィルターにより推定する。最後に

$$\hat{X}(t+k) = \mathbb{F}(T)\left[\bar{f}(t+k) + \hat{f}(t+k)\right] \tag{11}$$

によりXを予測する。



図－1： 本論文の方法の基本的な考え方 ,（システムの安定分はstatisticalに，変動分はstochastic processとして予測する）

本論文で提案する方法（これを"stochastic prediction through Kalman filtering combined with regression and factor analysis"を縮めてKARFA法と名付けたい）の基本的な考え方は，したがって図－1のようである。変動のうち比較的安定した部分は，過去のデータの統計的処理（statistical analysis）により取り除く。これには，具体的には回帰分析法を用いる。前論文ではもっと単純に日週期分を除いている。残りの変動分はシステムの非定常変動分などであり，この部分は適応制御理論とくにKalman filterにより予測する。

IV. 例　題

本論文の方法の妥当性を調べるために，電子計算機（HITAC 8700）によりシミュレーション実験を行った。変数$X(t)$は一次マルコフ過程として$X(t-1)$に依存し，さらに別の要素$Y(t)$の影響を受ける。

$$X(t) = \alpha(t)\, X(t-1) + \beta Y(t-1)$$

こゝに，αは$\alpha_0 + \alpha(t)$と考える。$\|\alpha\|/\alpha_0$の値を順次0より1まで変えて，単純な回帰分析による予測（+印：f_0）とKARFA法による予測（○印：fest）を実際の値（＊印：f）と比較したのが次頁の図である。

参考文献

1) 日野 幹雄・森 義一・吉川 信二郎： カルマン・フィルターによる大気汚染の予測法の開発， 東工大・土木工学科研究報告，No. 14 (1973)

2) 日野 幹雄： 実地形における排煙拡散の数値解析， 電力中央研究所技術研究所報告，No. 66100 (1967)

3) Hino, M. ; Computer experiment on smoke diffusion over a complicated topography, Atmospheric Environment, Vol.2, p541 (1968)

4) Estoque, M. & Bhumralkar, C. M. ; A method for solving the planetary boundary layer equations, Boundary-Layer Meteorology, Vol.1, No.2, p169 (1970)

5) Delage, Y. & Taylor, P. A. ; Numerical studies of heat island circulations, Boundary-Layer Meteorology, Vol.1, No.2, p201 (1970)

6) 高松 武一郎・内藤 正明： 大気汚染の計算機制御，計測と制御，Vol. 8，No. 12 (1969)

7) 平岡 正勝・池田 有光： 大都市内の局地大気汚染濃度予測に関する研究， 土木学会論文報告集，No. 209，P. 63 (1973)

8) 船橋 誠寿： 多排出源による広域汚染シミュレーション・モデル，APM8資料

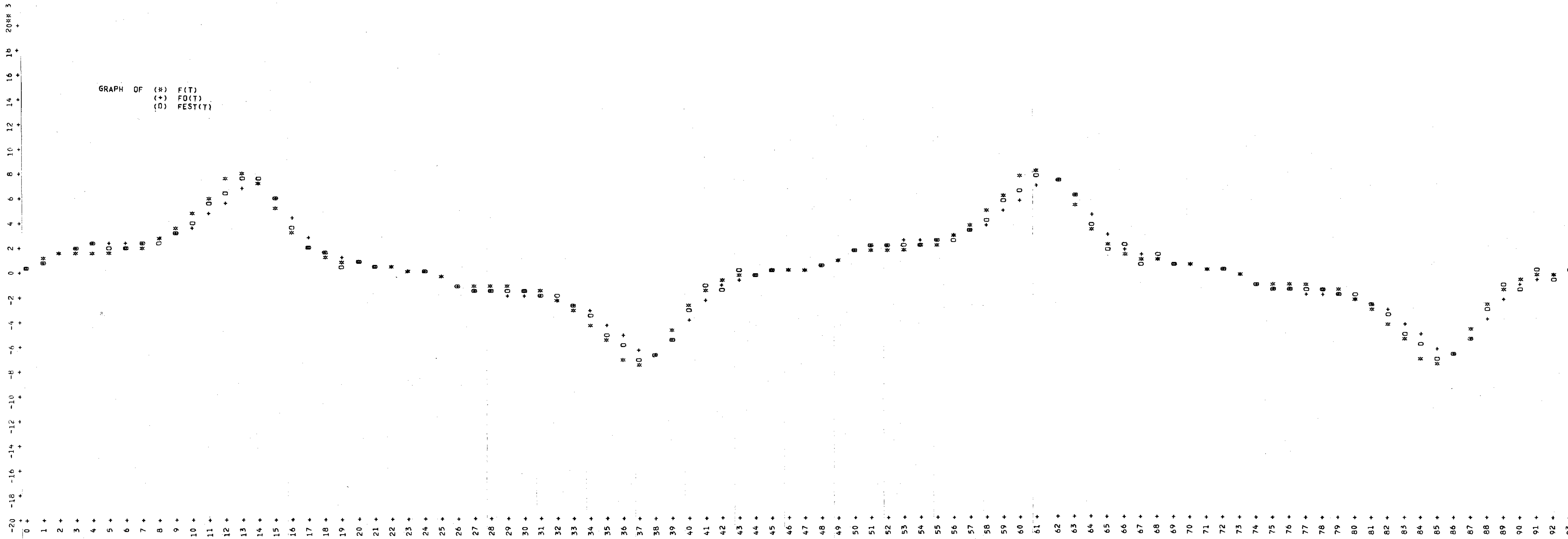
9) 藤田 威雄・古谷忠義： 公害監視網データによる大気汚染パターンの解析，「環境汚染制御」研究経過報告書，昭和47年度文部省科学研究費 特定研究(1)，P. 244，(1973)

10) 塩谷 実・浅野 長一郎： 多変量解析論，共立出版，(1967)

11) 浅野 長一郎： 因子分析法通論，共立出版，(1971)

12) Kalman, R.E. : A new approach to linear filtering and prediction problems, J. of Basic Eng., Trans.ASME Ser.D, Vol.82, p35 (1960)

13) 相良 節夫： 同定問題，計測と制御，第8巻，第4号 (1969)


GRAPH OF (*) F(T)
(+) FO(T)
(O) FEST(T)