

確率論的な流出予測に関する研究

——有色ノイズの導入——

高棹 琢馬・椎葉 充晴・宝 馨

A STOCHASTIC APPROACH TO PREDICTION OF FLOOD RUNOFF

——Incorporation of Colored Noise——

By *Takuma* TAKASAO, *Michiharu* SHIIBA and *Kaoru* TAKARA

Synopsis

The additive noise in stochastic state space model is regarded as the model error and usually assumed to be white Gaussian. But the assumption that the noise sequence is white is inadequate to describe a large and complicated system like as rainfall/runoff system.

In this paper, the authors design several runoff models with so-called colored noise. The numerical studies, where only statistical and mathematical models are dealt with, are presented to show the efficiency of incorporating the colored noise into the runoff models. Moreover they also give an outline of identifying the state space model in which both the physical factors in the runoff system and the colored noise sequence can be taken into consideration.

1. 序 論

1.1 はじめに

筆者らは、流出システムが stochastic かつ dynamic なシステムであるという認識のもとに、確率論的な短期流出予測方式を構成する理論を展開してきた^{1), 2), 3)}。それは、簡単に言うと、流出システムを状態空間法によって記述し、Kalman⁴⁾ にはじまるフィルタリング理論とその応用理論⁵⁾ を適用して流出予測計算を行う、といった内容であった。そこでのシステム記述から予測計算に至るまでの概要は Fig. 1 のようである。この流れの中において意図されることは、

- (a) 時々刻々得られる観測情報を最大限に利用し、逐次的な予測をすること。いわゆるオンライン予測を可能ならしめること。
- (b) 予測に際して、流量値のみならず、予測の精度までも分散（あるいは標準偏差）のような形で明らかにすること。

である。実際、電子計算機をはじめとする情報処理

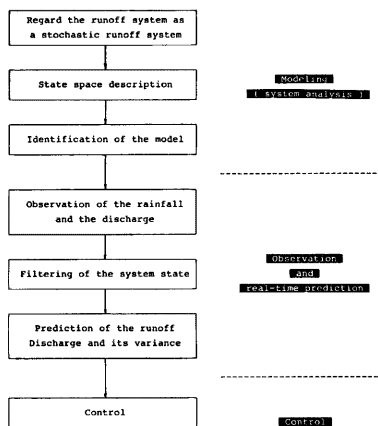


Fig. 1. Outline of the stochastic runoff prediction

技術の発達近年目覚ましいものがあり、(a)は時代の要請に則したものと言える。また、ダム・水門等の操作による洪水制御の面で、確率論的な対処法を要求される場面も多く、予測の精度が明らかにされているならば制御に対する有益な情報を与えることになり、(b)も洪水予報の不可欠な要素となる。これら2点を満足するような予測方式は短期流出予測の最適方式であると考えられる。そのような方式を構成するのに、前記のフィルタリング理論は有用なものである。

日野⁶⁾は、流出システムに Kalman フィルターを適用したが、それはモデル定数を状態量に読みかえたもので、いわゆるモデル同定のための議論であった。Todini⁷⁾は、このようなアプローチは、定数の最適推定は実行するが、出力としての物理量の推定値が最小分散推定とはならないことを指摘し、システムの状態を推定するフィルターと、モデル定数を逐次推定してゆくフィルターとを同時に稼動させる推定手法 (MISP, Mutually Interactive State-Parameter estimation) を提案している。

筆者らのこれまでの方法は、降水現象をうけもつ気象システムと流出現象をうけもつ流出システムとをそれぞれ状態空間モデルで表現し、それらを統合した形で議論を進めているところに1つの特徴がある。また、観測値の形では陽に値の得られない流域内貯留量のような物理量を状態量として最小分散推定値を求めようとする点で日野の方法とは全く趣を異にするものである。

1.2 本研究に至るまでの経過とその問題点

本研究に至るまでの経過とそこでの問題点を、ひとまずここにまとめておく。

筆者らによる気象・流出システムそれぞれの確率論的な状態空間モデルは、一般的な表現をすると次のようである。

[気象システム]

$$\left. \begin{aligned} S_{k+1} &= \phi(S_k, t_k) + \Gamma(t_k) \alpha_{k+1} \\ r_k &= h(S_k, t_k) + \beta_k, \quad k=0, 1, \dots \end{aligned} \right\} \dots\dots\dots (1-1)$$

ここに、 S_k は時刻 t_k における気象システムの状態量ベクトル、 r_k は時刻 t_{k-1} から時刻 t_k までの降雨強度、 ϕ , h は一般に非線形の関数、 $\Gamma(t_k)$ は非確率行列、 α_{k+1} , β_k は平均値 0 の白色正規ノイズである。

[流出システム]

$$\left. \begin{aligned} \frac{dx_t}{dt} &= f(x_t, r_t, t) + Gw(t), \quad t_{k-1} < t \leq t_k, \quad k=0, 1, \dots \\ y_k &= g(x_{t_k}, t_k) + v_k \end{aligned} \right\} \dots\dots\dots (1-2)$$

ここに、 x_t は時刻 t における流出システムの状態量ベクトル、 y_k は時刻 t_k の流出流量、 f , g は一般に非線形の関数、 G は非確率行列、 v_k は白色正規ノイズ、 $w(t)$ は連続白色正規ノイズ (Brown 運動過程の微分) である。

筆者らはまず、このようなシステムモデルについて、降雨・流量の模擬時系列を乱数発生させ、その系列を順次観測値としてフィルタリング計算することから手がけた¹⁾。そこでは、気象システムは台風性降雨を想定した仮想モデルを、流出システムは貯留関数モデルを用いているが、状態方程式が連続時間、観測方程式は離散時間であるところのいわゆる連続・離散形の非線形システムモデルを含んでいることもあって、単なる Kalman フィルターではなく、その拡張理論⁵⁾を取扱い、流出予測への適用性を検証した。

次に、実流域への適用を図った²⁾。上の数値シミュレーションでは、モデル定数、ノイズの統計量は既知としたが、それらには実際には何らかの方法で決定すべきものである。モデル定数の逐次決定手法として、決定すべき定数を状態量に組み込む方法 (次元拡大 (state augmentation) を行なう。たとえば、Wood らの文献⁸⁾を見よ) を用い、また、ノイズの統計量の推定のための Sage ら⁹⁾の方法に若干の修正を加えた。

これらの理論および手法を短期流出予測の実際の場面に持ち込もうとする際、その適用限界は明らかでない。また、2, 3の問題点も存在する。

第1の問題点は、気象システムモデルの設定にある。降雨予測は古来からの人類の課題であり、かなり研究が進んできてはいるが、まだ、完成された降雨予測モデルが存在するわけではない。熟練した気象予報官の経験や主観的判断に予測を委ねることも現実に行なわれている。また、降雨の観測値自体、流域内の個々の地点のものにすぎず、しかも単位時間内の積分値の形で得られるという特殊なデータであるため取扱いが難しい。

このようなことから、気象システムを状態空間モデルによって記述するのは現実性に乏しいのではないか、という疑問は今のところ否めないのである。しかし、たとえば「今後1時間のA流域の降水量は〇〇mmで、その標準偏差は△△mm」というふうな降雨予測がなされれば、具体的な気象システムモデルは考慮せずに（気象・流出の統合システムから気象システムを切離して）、それらの情報を入力として、流出システムのみに議論することが可能となる。

最近、降雨の確率予報¹⁰⁾や降雨観測のためのレーダー・サイトの設置等が行なわれてきているが、それらはまだ流出予測に直接役立つまでには至っていない。

第2の問題点は、流出システムのモデリングの問題である。流出モデルは現在まで種々提案されているが、流出システムが大規模かつ複雑であるため、単純なモデルではシステムの挙動を十分に表現しきれない。逆に、現象に忠実なモデルでは、次元が大きくなりすぎて膨大な計算量が要求されたり、偏微分方程式系の数値解法に困難があったりして処理上の問題が残る。これら双方のトレードオフの関係を解消する方向での最適なモデル構成は未だ確立されていない。実際にフィルタリング計算をするときには、モデルは単純である方が望ましいのである。

さて、流出システムのような大規模で複雑なシステムを考えると、(1-2)式中の決定論的な部分である f や g がシステム内の本質的な動特性を十分に記述することはむずかしく、攪乱項として付加されたノイズが白色系列であるという仮定は必ずしも成立しない。よって、このノイズに何らかの動的構造を与えた確率論的モデルを構成することが、合理的な1つの対処法として考えられる。

1.3 本研究の目的と方針

本研究では、流出システムのモデリングの問題に焦点をおき、システムを十分に記述しえない場合の対応策を考える。前節での議論から、モデル誤差を補償する攪乱項が白色でないこと（有色性）を考慮し、決定論的な流出モデルに移動平均（moving average）型のノイズを付加した確率論的なモデルを構成して、流出予測計算を行うことにする。

すなわち、流量時系列 $y_1, y_2, \dots, y_k, \dots$ を表現する決定論的部分を f_k 、時点 k におけるモデル誤差を v_k とすると、

$$y_k = f_k + v_k, \quad k=1, 2, \dots \quad (1-3)$$

となり、 v_k の系列が白色でないとき、 v_k を

$$v_k = e_k + c_1 e_{k-1} + \dots + c_m e_{k-m} \quad (1-4)$$

(e_k は白色ノイズ、 $c_1, \dots, c_m$ は定係数)

のような移動平均過程と考え、

$$y_k = f_k + e_k + c_1 e_{k-1} + \dots + c_m e_{k-m} \quad (1-5)$$

という確率論的モデルを構成する。

本研究の目的は、確率論的な流出モデルとして従来からの一般的な記法である(1-3)式を、モデル誤差の有色性を考慮した(1-5)式のような形に書き換えて、さらに精度のよい流出予測値を得ようとするのである。(1-5)式は、流出モデルに限らず、システムを十分に記述できないような他のシステムモデルを確率論的に取扱う場合にも応用可能な形である。

本論文の構成は以下のようである。

まず, 2. において, 移動平均型ノイズを含むモデルとして, 時系列解析の分野で従来からよく用いられている, 自己回帰の項をもつ線形モデルを取扱う。次に, 3. においては, 変数組合せ法 (Group Method of Data Handling 以下 GMDH と略称する) によって同定された非線形部に移動平均型ノイズを付加したモデルについて考察する。

上記2つの異なるタイプのモデルは, それぞれ線形, 非線形に分類される。しかし, どちらも過去の降雨・流量のデータのみに依存したいわゆる統計的な解析によるものであって, 地形や流出現象を考慮した物理モデルとは異なる。

4. では, それらの物理的要素と, モデル誤差の有色性とを同時に考慮しうる状態空間モデルを構成する理論を展開する。本研究は, 「時々刻々得られる情報を最大限に利用し, 予測値のみならずその精度までも明らかにするオンライン予測方式を確立する」という最終目標をもつものである。ここでの議論により, そのような“最適”予測方式に一步接近することになるものとする。

最後に, 5. において本研究の総括を行なう。

2. ARMAX モデルによる流出予測

2.1 線形時系列モデル

時系列解析¹¹⁾の分野では線形定常過程として, 自己回帰過程 (autoregressive process, AR), 移動平均過程 (moving average process, MA), 両者をミックスした自己回帰・移動平均過程 (ARMA) などがよく用いられる。

流出過程については, たとえば, 流域を1つの線形貯水池 (linear reservoir) で表わすと, 降雨 r と流出量 Y との関係は

$$K \frac{dY}{dt} + Y = r \quad (\text{ただし } K: \text{貯留係数})$$

これを差分式に直すと

$$y_t = a_1 y_{t-1} + r'_t \quad (a_1: \text{定数}, r'_t: \text{差分化によりえられる項})$$

となって, 1次の自己回帰の形式に帰着できる¹²⁾。このようなことから, 近年, ARMA モデル等は, 流出解析にしばしば持ち込まれているのは周知のとおりである。

最近, ARMAX と呼ばれるモデルがよく用いられる^{7), 13)}。これは, ARMA モデルに外生的 (exogenous) な入力 (これが X-part) を加えたもので次式のような形式である。

$$y_t = a_1 y_{t-1} + \dots + a_n y_{t-n} + e_t + c_1 e_{t-1} + \dots + c_m e_{t-m} + b_1 u_{t-1} + \dots + b_n u_{t-n} \dots \dots \dots (2-1)$$

ここに, $a_1, \dots, a_n, c_1, \dots, c_m, b_1, \dots, b_n$ は定数係数で, y_t, u_t はそれぞれ, 時点 t における流量, 降雨, e_t は白色正規ノイズである。

ARMA モデルは, システムへの入力としての降雨が白色ノイズとして扱われることになるが, (2-1) 式のように X-part ($b_i u_{t-i}, i=1, \dots, n$) を付加することによって, ARMAX モデルにおけるノイズは, 降雨入力以外のノイズ部分を表わすことになる。(2-1) 式は, 結局, 時点 t の流量 y_t は, 過去の流量, 降雨, ノイズの線形結合で表わされるという構造となる。この ARMAX モデルの移動平均ノイズを白色ノイズに置き換えて

$$y_t = a_1 y_{t-1} + \dots + a_n y_{t-n} + b_1 u_{t-1} + \dots + b_n u_{t-n} + e_t \dots \dots \dots (2-2)$$

としたものを ARX モデルと呼ぶことにする。

本章では, ARX モデルと, 移動平均型の有色ノイズをもつ ARMAX モデルとによってそれぞれ流出予測計算を行ない, 有色ノイズを導入することの有効性を検討する。

以後の便宜のため、(2-2) の形式のものを ARX (l, n), (2-1) を ARMAX (l, m, n) というふうにしてそれぞれモデル次数をつけて表わすことにする。

2.2 同 定 手 法

前節で述べた ARX あるいは ARMAX 型の線形モデルの係数を決定するアルゴリズムを以下に示す。ここの理論展開はおおむね Katayama *et al.*¹⁴⁾ によった。

ARMAX (l, m, n) モデル (2-1) 式を書き直すと、

$$y(t) = \sum_{i=1}^l a_i y(t-i) + \sum_{i=1}^n b_i u(t-i) + e(t) + \sum_{i=1}^m c_i e(t-i) \quad \dots\dots\dots (2-3)$$

ここに、 $y_t \rightarrow y(t)$, $u_t \rightarrow u(t)$, $e_t \rightarrow e(t)$ と書き換えた。

簡単のため、

$$\begin{aligned} \mathbf{a} &= (a_1 a_2 \cdots a_l)' \\ \mathbf{b} &= (b_1 b_2 \cdots b_n)' \quad \dots\dots\dots (2-4) \\ \mathbf{c} &= (c_1 c_2 \cdots c_m)' \end{aligned}$$

とする (' は転置記号)。

問題は、利用できる入出力データ $U^N = \{u_1, \dots, u_N\}$ および $Y^N = \{y_1, \dots, y_N\}$ に基づいて、

i) (2-3) 式の最適次数 (l, m, n)_{opt} を決定すること

ii) パラメタ $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ を推定すること

である。

i) については後で議論することにして、まず、固定された次数に対してパラメタを推定する方法を述べる。

【1】最小二乗法による同定

一般に、パラメタ推定やシステム同定においては、先験的知識とみなされる3つの前提条件が存在する。すなわち、

(1) $p(e)$: ノイズ過程の確率密度関数。これにより、パラメタの値が θ であるとき、出力 $y(t)$ の条件付確率密度関数 $p(y | \theta)$ を得る。

(2) $p(\theta)$: パラメタ θ の先験的確率密度関数

(3) $l(\theta, \theta)$: θ の推定値として θ を用いたときの損失関数。

である。先験的な情報としてすべての条件(1), (2), (3)が使えればベイズの方法 (Bayesian approach) を適用することができる。もし(1)だけが使えれば最尤法が適当であろうし、もし先験的な知識が与えられていなければ最小二乗法が適用されることになる。

今、 $\mathbf{c} = 0$ として、 $\mathbf{a}$ と $\mathbf{b}$ の推定に最小二乗法を適用する。(2-3)式より、

$$y(t) = \sum_{i=1}^l a_i y(t-i) + \sum_{i=1}^n b_i u(t-i) + e(t) \quad \dots\dots\dots (2-5)$$

となるが、これは ARX (l, n) にほかならない。簡単のため、 $l = n$ とし、 $t = (-n+1), \dots, -1, 0, \dots, N$ を考え、

$$\mathbf{A} = \begin{bmatrix} y(0) & y(-1) & \cdots y(-n+1) & u(0) & u(-1) & \cdots u(-n+1) \\ y(1) & y(0) & \cdots y(-n+2) & u(1) & u(0) & \cdots u(-n+2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y(N-1) & y(N-2) & \cdots y(-n+N) & u(N-1) & u(N-2) & \cdots u(-n+N) \end{bmatrix}$$

$$\mathbf{y} = [y(1) \ y(2) \ \cdots \ y(N)]'$$

$$\mathbf{e} = [e(1) \ e(2) \ \cdots \ e(N)]' \quad \mathbf{a} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$$

と定義すれば (2-5) は

$$\mathbf{y} = \mathbf{A}\boldsymbol{\alpha} + \mathbf{e} \dots\dots\dots(2-6)$$

となる。ノイズ項 $\mathbf{e}$ が係数行列と独立ならば、 $\boldsymbol{\alpha}$ の最小二乗推定値 $\hat{\boldsymbol{\alpha}}$ は

$$\hat{\boldsymbol{\alpha}} = [\mathbf{A}'\mathbf{A}]^{-1}\mathbf{A}'\mathbf{y}$$

で与えられる。ここに、 $\mathbf{A}'\mathbf{A}$ は正則であるとする。

(2-3) で系列相関をもつノイズ $\{e(t) + \sum_{i=1}^N c_i e(t-i), t=1, 2, \dots, N\}$ が、(2-5) において無相関なノイズに置き換えられたわけで、よって最小二乗推定値 $\hat{\boldsymbol{\alpha}}$ は偏りをもつであろう。しかしながら、最小二乗法は、他の推定法に比べて少ない仮定しか必要とせず、また計算時間も少なくすむ。そのうえ、実際的な見地から合理的な結果を与えることも少なくない。ARX モデルはこの方法により容易に同定される。

【2】最尤法による同定

パラメタ $\boldsymbol{\theta}$ に関する出力 $y(t)$ の条件付確率密度関数 $p(y|\boldsymbol{\theta})$ が与えられると、尤度関数は

$$L(\boldsymbol{\theta}) = p(y|\boldsymbol{\theta}) \dots\dots\dots(2-7)$$

のように定義される、与えられた y の標本値に対して $L(\boldsymbol{\theta})$ は単にパラメタの関数である。最尤推定値は $L(\boldsymbol{\theta})$ を最大にする $\boldsymbol{\theta}$ の値で $\hat{\boldsymbol{\theta}}_{ML}$ で表される。換言すれば、最尤推定値は y の標本値を最も“尤もらしく”なるように決められる。

以下においては、 $\boldsymbol{\theta}$ は未知パラメタ $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ の集合を表わす。(2-3) において $\{e(t), t=1, \dots, N\}$ は平均値 0、分散 σ^2 をもつ正規白色ノイズであるとする。そのとき、 U^{N-1} , $\boldsymbol{\theta}$, σ^2 , 初期条件が与えられたときの Y^N の条件付確率密度関数は次のように書ける。

$$\begin{aligned} p(Y^N|U^{N-1}, \boldsymbol{\theta}, \sigma^2, ICs) &= \prod_{i=1}^N p(y(i)|u(i), y(i), i \leq t-1, \boldsymbol{\theta}, \sigma^2) \\ &= \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} e^2(t) \right\} \dots\dots\dots(2-8) \end{aligned}$$

ここに、 ICs は初期条件 (2-9) である。

$$\begin{aligned} y(0), y(-1), \dots, y(-l+1) \\ u(0), u(-1), \dots, u(-n+1) \end{aligned} \dots\dots\dots(2-9)$$

$e(t)$ は、

$$e(t) = y(t) - y(t|t-1) \dots\dots\dots(2-10)$$

ここに

$$y(t|t-1) = \sum_{i=1}^l a_i y(t-i) + \sum_{i=1}^n b_i u(t-i) + \sum_{i=1}^N c_i e(t-i)$$

である。よって、条件付対数尤度関数は

$$\begin{aligned} L_N(\boldsymbol{\theta}, \sigma^2) &= \log p(Y^N|U^{N-1}, \boldsymbol{\theta}, \sigma^2, ICs) \\ &= -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N e^2(t) \dots\dots\dots(2-11) \end{aligned}$$

となり、これを最大にするような $\boldsymbol{\theta}$ や σ^2 を求める。(2-11) を σ^2 に関して微分し、0 とおくことによって

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N e^2(t) \dots\dots\dots(2-12)$$

を得る。(2-12) を (2-11) に代入して

$$-L_N(\theta, \sigma^2) = \frac{N}{2}(1 + \log 2\pi) + \frac{N}{2} \log \sigma^2 \dots\dots\dots (2-13)$$

したがって、 L_N を最大にするような最尤推定値 $\hat{\theta}_{ML}$ は、(2-3) の制約条件のもとで $\log \sigma^2$ を最小にすると、すなわち、

$$L_N^*(\theta) = \frac{1}{N} \sum_{t=1}^N e^2(t) \dots\dots\dots (2-14)$$

を最小にすることによって得られる。しかして、最尤法は等号制約をもつパラメトリック最適化問題へと導かれたことになり、それは(2-10)式のイノベーション過程の標本分散を最小化する。上のすべての議論は、 $e(t)$ が正規白色ノイズであるという前提に基づいている。

上のように最尤法は、(2-3) の等号制約のもとで (2-14) の L_N^* の最小点を求める最適化問題となる。本研究では、Katayama¹⁴⁾と同様に Davidon のアルゴリズムを用いて最適化を図る。具体的な事項は付録にまとめておいた。

【3】モデル次数の決定

(2-1) や (2-2) のモデルを構成するとき、各 part (AR, MA, X) の次数 l, m, n を決定しなければならない。

Akaike¹⁵⁾ は、統計的なモデル同定において、最適なモデル次数を決定するための情報理論的基準を提案している。それは AIC (an information criterion) と呼ばれるもので、

$$AIC(k) = -2\log(\text{最大尤度}) + 2k \dots\dots\dots (2-15)$$

で定義される。ここに、 k はモデルパラメタの個数である。すなわち、尤度が大きくなればなるほど AIC の値は小さくなるが、パラメタの個数を増やすと AIC は大きくなる。このトレードオフ関係の中で AIC を最小にするようなパラメタの個数 k_0 をもつモデルが最良のものである、というのが赤池の考えである。

Katayama¹⁶⁾ は、群馬県烏川の日流量予測に ARMAX 型のモデルを用い、最尤法を用いてモデル同定を行っている。そこでは AIC を計算しているが、最小 AIC を与える次数のモデルが予測モデルとして最適であるという結果を得ていない。モデル次数の決定のための簡単ではあるが効果的な方法として、Katayama は同定用のデータを training 用と checking 用とに分け、checking データに対して予測誤差の分散が最小になるようなモデルの次数を最適次数としている。本研究では、日流量ではなく、時間単位の流量予測計算を行なう。

2.3 有色ノイズ導入効果の検討

本節において、ARX 型のモデルと移動平均型の有色ノイズをもつ ARMAX 型のモデルとを比較し、予測モデルとしての性能を検討する。検討のために用いたデータは群馬県神流川流域 (面積 373.6 km²) の時雨量・時流量データである¹⁷⁾。観測された出水のうち、同定用データとして No. 2-9 の出水を、予測計算用のデータとして No. 10, 11 の出水を対象とした。

時雨量は算術平均によって流域平均のものとし、流量は流出高 (mm/hr) に換算し直した。流域図を Fig. 2 に示す。

【1】同定・予測の結果

本研究では、簡単のため ARX (l, n), ARMAX (l, m, n) の次数は、 $l=m=n=k$ とした。

Table 1 は、 $k=1, \dots, 10$ の ARX と ARMAX の training データに対する再現誤差の二乗平均、AIC, checking データに対する予測誤差の二乗平均をそれぞれ求めて一覧表にしたものである。

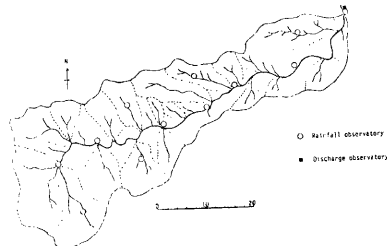


Fig. 2. Kanna River basin (373.6km)

Table 1. Comparison of the Performances of Linear Models

model order	number of data	reproduced M. S. E.			predicted M. S. E.			
		flood No. 2-9			No. 10		No. 11	
<i>k</i>	N	ARX	ARMAX	AIC(3 <i>k</i>)	ARX	ARMAX	ARX	ARMAX
		$\times 10^{-2}$	$\times 10^{-2}$		$\times 10^{-2}$	$\times 10^{-2}$	$\times 10^{-1}$	$\times 10^{-1}$
1	1080	7.882	6.083	-2.794	1.945	1.081	4.454	3.181
2	1072	4.390	4.362	-3.121	0.9453	0.9561	2.299	2.294
3	1064	4.312	4.291	-3.132	0.8938*	0.9273	2.295	2.232*
4	1056	4.183	4.173	-3.153	0.9445	0.9499	2.241*	2.284
5	1048	4.205	4.177	-3.147	1.124	0.9539	2.295	2.330
6	1040	3.220	3.138	-3.427	0.9674	0.5751	2.326	2.659
7	1032	2.082	1.893	-3.927	0.9443	0.5731*	2.351	2.872
8	1024	2.426	2.221	-3.761	0.9574	0.6422	2.396	3.041
9	1016	0.06518*	0.03392*	-7.936*	0.9607	1.669	2.457	3.750
10	1008	0.1698	0.08121	-7.057	1.042	4.663	2.432	5.589

* denotes minimum values

$$AIC(3k) = \log \sigma_k^2 + \frac{2(3k)}{N}$$

M. S. E. : mean square error

$\hat{\sigma}_k^2$: reproduced M. S. E.

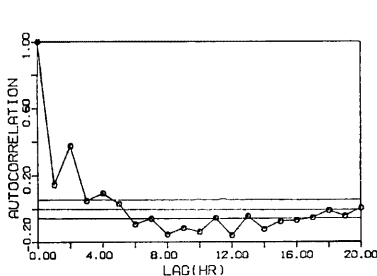


Fig. 3. Autocorrelation function of the reproduction error in ARMAX (1,1,1). 95% confidence level of whiteness is 0.06.

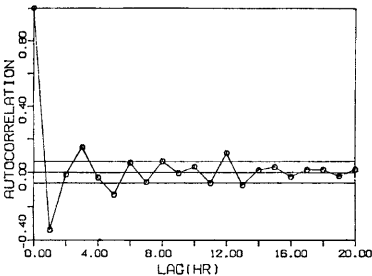


Fig. 5. Autocorrelation function of the reproduction error in ARMAX (7,7,7). 95% confidence level of whiteness is 0.06.

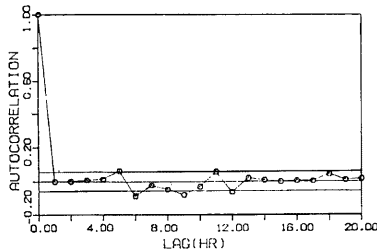


Fig. 4. Autocorrelation function of the reproduction error in ARMAX (3,3,3). 95% confidence level of whiteness is 0.06.

再現誤差の二乗平均や AIC を最小にする次数のモデル (ここでは ARX (9, 9), ARMAX (9, 9, 9)) は, 予測誤差を必ずしも最小としない。すわち, 再現誤差の二乗平均や AIC は, 最良予測モデルを判定する基準とはならない。Fig. 3~5 は, ARMAX (1, 1, 1), ARMAX (3, 3, 3), ARMAX (7, 7, 7) の再現誤差の自己相関のコログラムである。白色と仮定された系列 $\{e(t)\}$ の自己相関

$$\rho_{ee}(\tau) = \frac{1}{N \hat{\sigma}_N^2} \sum_{t=1}^N e(t)e(t+\tau), \quad \tau=0, 1, 2, \dots$$

.....(2-16)

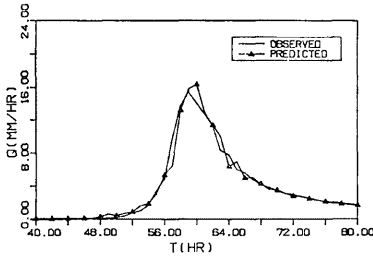


Fig. 6. One-hour ahead prediction of No. 11 flood by ARX (4,4)

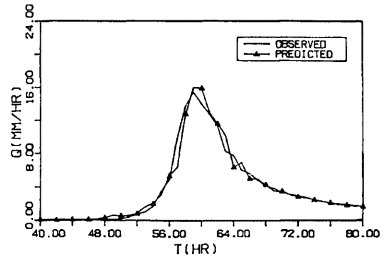


Fig. 8. One-hour ahead prediction of No. 11 flood by ARMAX (3,3,3)

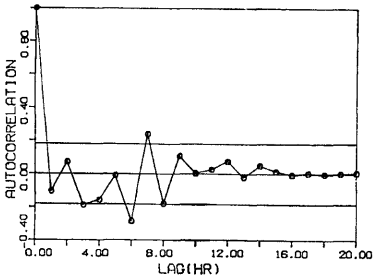


Fig. 7. Autocorrelation function of the prediction error of No. 11 flood by ARM (4,4). 95% confidence level of whiteness is 0.18

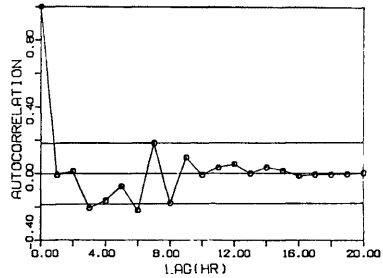


Fig. 9. Autocorrelation function of the prediction error of No. 11 flood by ARMAX (3,3,3). 95% confidence level of whiteness is 0.18

は、 $\tau \approx 0$ で 0 となるはずである。モデル同定が不十分であれば必ずしもそうはならない。Fig. 3~5 のうちでは ARMAX (3, 3, 3) のものが良好な結果を示している。実際、Table 1 によれば、ARMAX (3, 3, 3) は大きい方の出水 No. 11 の予測誤差を最小にするモデルとなっている。モデル同定の良否を判定するのに、このような白色性のチェックはかなり有効であると思われる。

Fig. 6 は ARX (4, 4) の予測ハイドログラフ、Fig. 7 は予測誤差のコログラムである。Fig. 8, 9 は ARMAX (3, 3, 3) のそれらを示している。ARX (4, 4) は ARX 型のうち最良のもの、ARMAX (3, 3, 3) は ARMAX 型のうち最良のものである。1 時間先の予測モデルとしてはどちらも実用に供しうる程度のハイドログラフと考えられる。

【2】有色ノイズ導入の効果について

Fig. 7 と Fig. 9 を比べてみると、明らかに有色ノイズを導入した ARMAX の方が良好な結果を示している。これは、有色ノイズ導入の有効性を示すものであり、実際、Table 1 によっても予測誤差の二乗平均は ARX (4, 4) よりも ARMAX (3, 3, 3) の方がすぐれている。

本章では、時系列解析で用いられるタイプの線形モデルについて、移動平均型の有色ノイズを導入することの有効性を確認した。ここで取扱ったような線形モデルは適当なベクトル、行列の操作を行えば状態空間法で記述できる。したがって、ARMAX モデルは、予測の精度までをも明らかにしうるオンライン流出予測方式を構成するのに、実用的観点から見ると、適当なモデルの 1 つである。

3. GMDH を用いた非線形モデルによる流出予測

3.1 概 説

前章では、過去のデータのみに依存する線形モデルを取扱い有色ノイズ導入の効果をみた。本章では、非線形モデルについて同様の考察を行なう。すなわち、(1-5) 式の f_s の部分が非線形である場合を考えるのであるが、その部分の同定には変数組合せ法 (GMDH) を用いる。

GMDH は、システム構造が複雑でモデル化が困難なシステムに対するブラックボックス的アプローチとして、最近研究が進められている非線形同定手法である。大気汚染予測や台風性降雨の予測など¹⁸⁾にも用いられている。流出予測の方面では、池田ら¹⁹⁾が日流量の予測に導入し、さらに、岩崎ら²⁰⁾や市川ら²¹⁾が検討を加えた。橋本²²⁾は低水流出予測に GMDH を適用した。時間単位の流量予測への適用例は、筆者らの知るかぎりでは、今のところ見当たらない。その適用性については、後で述べる。

さて、今、GMDH によって同定した流出モデルを GMDH モデルと呼ぶことにし、それに移動平均型の有色ノイズを付加したものを GMDHMA (m) モデルと呼ぶことにする。GMDH モデルが

$$y(k) = f(Y^{k-1}, U^{k-1}) \dots\dots\dots (3-1)$$

で、 $y(k)$ は時点 k の流量、 Y^{k-1} 、 U^{k-1} は時点 $k-1$ 以前の流量、降雨の系列であるとき、GMDHMA (m) モデルは、

$$y(k) = f(Y^{k-1}, U^{k-1}) + e(k) + \sum_{i=1}^m c_i e(k-i) \dots\dots\dots (3-2)$$

で表わされる。ここに、 $\{e(j)\}$ 、 $j=0, 1, \dots$ は平均値 0 の白色正規ノイズ、 c_i 、 $i=1, \dots, m$ は定係数で、2. 章の最尤法を用いて同定する。(3-1)、(3-2) 式の f は、修正 (modified) GMDH アルゴリズム¹⁹⁾によるものとする。これは、流量の急激な増加に対する追従性を改善する目的で、池田らの提案したものである。

本研究は、GMDH の構造およびアルゴリズムそのものに興味をもつものではない。したがって、それに関する詳細な記述は他に譲ることとする。

3.2 GMDH による時流量予測モデルの同定について

GMDH は少ない入力データで複雑な多変数・非線形系の同定・予測が可能であると言われている。しかしながら、同定に用いるデータに偏りがあれば (たとえば、既往のデータとして小出水のみしか存在しないなどの)、次に比較的大きな出水が生じた場合の流量予測値は、非常に精度が悪いが、全くでたらめな値を算出するなど必ずしも良好な結果が得られない。時流量予測に GMDH を適用する際には、既往の大小出水をとりまぜ、なるべくデータを数多く用いて同定すべきであるとの結論を筆者らは得ている²³⁾。

本章でも、前章と同様、神流川流域のデータを用い、出水 No. 2-9 を同定用、出水 No. 10, 11 を予測用とした。No. 2-9 の中には、既往の大出水をも含んでいる。Fig. 10 は、このデータを用いて同定した GMDH モデルによって得た、出水 No. 11 の 1 時間先の予測ハイドログラフであって、かなり良好な結果を示している。Fig. 11 は予測誤差のコログラムであるが、誤差の系列の白色性は保証されない。同定時の再現誤差のコログラムは Fig. 12 である。これを見ると、GMDH による同定が十分でないことを示している。

3.3 有色ノイズ導入効果の検討

そこで、GMDH モデルと、それに有色ノイズを付加した GMDHMA (m) モデルとを比較し、有色ノイズを導入効果を検討する。

【1】同定・予測の結果

Table 2 は、GMDH モデルと $m=1, \dots, 10$ の GMDHMA (m) モデルの、training データに対する再現誤差の二乗平均、AIC、checking データに対する予測誤差の二乗平均をそれぞれ求めてまとめたものである。

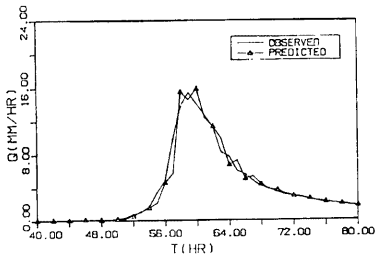


Fig. 10. One-hour ahead prediction of No.11 flood by GMDH

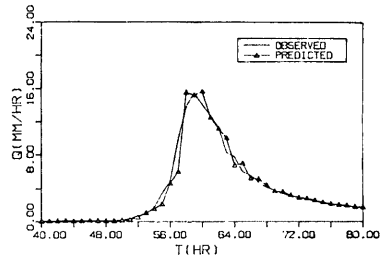


Fig. 13. One-hour ahead prediction of No.11 flood by GMDHMA(1)

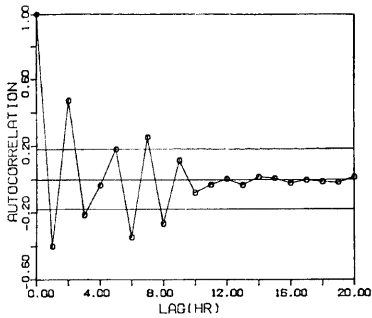


Fig. 11. Autocorrelation function of the prediction error of the No.11 flood by GMDH, 95% confidence level of whiteness is 0.18.

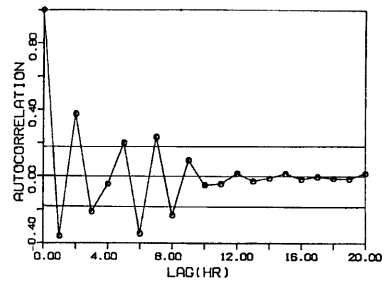


Fig. 14. Autocorrelation function of the prediction error of the No.11 flood by GMDHMA(1). 95% confidence level of whiteness is 0.18.

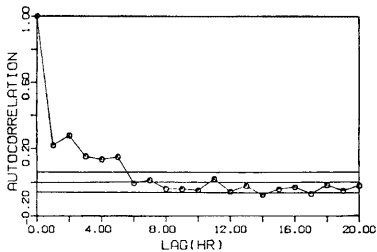


Fig. 12. Autocorrelation function of the reproduction error in GMDH, 95% confidence level of whiteness is 0.06.

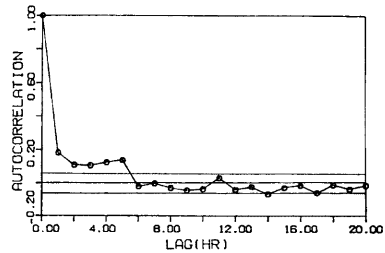


Fig. 15. Autocorrelation function of the reproduction error in GMDHMA(1). 95% confidence level of whiteness is 0.06.

Table 2. Comparison of the Performances among GMDH model and GMDHMA(m) models

model order	number of data	reproduced		predicted		
		flood No. 2-9		No. 10	No. 11	
m	N	M. S. E.	AIC(m)	M. S. E.	M. S. E.	
0	1024	6.211×10^{-2}	-2.779	0.8692×10^{-2}	0.2831	GMDH
1	1024	4.967×10^{-2}	-3.000	0.8081×10^{-2}	0.2537*	GMDHMA
2	1024	4.673	-3.059	0.7735*	0.2965	
3	1024	4.756	-3.040	0.8700	0.3024	
4	1024	4.656	-3.059	0.8181	0.2921	
5	1024	4.514	-3.088	0.8719	0.3123	
6	1024	4.530	-3.083	0.9409	0.3252	
7	1024	4.494*	-3.089*	0.9320	0.3406	
8	1024	4.539	-3.077	0.9779	0.3558	
9	1016	4.562	-3.069	1.011	0.3683	
10	1008	4.618	-3.065	1.016	0.3778	

* denotes minimum values

$$AIC(m) = \log \hat{\sigma}_m^2 + \frac{2m}{N}$$

M. S. E. : mean square error.

$\hat{\sigma}_m^2$: reproduced M. S. E.

ここでも、再現誤差の二乗平均や AIC を最小にする次数のモデルは、予測モデルとして最良ではない。この表からは、GMDHMA (1) が最良予測モデルであると結論できる。GMDHMA (1) による予測ハイドログラフ、予測誤差のコレログラム、再現誤差のコレログラムをそれぞれ、Fig. 13, 14, 15 に示した。

【2】有色ノイズ導入の効果について

Fig. 10 と Fig. 13, Fig. 11 と Fig. 14, Fig. 12 と Fig. 15 をおのおの比べてみると、GMDH モデルより、GMDHMA(1) の方がすぐれたモデルであることがわかる。このことより、有色ノイズを導入することの有効性が確認された。

2., 3. 章の検討の結果、適当な次数の移動平均型ノイズを付加することによって、線形モデルであれ、非線形モデルであれ、流量の予測値を改善することは可能であった。しかし、これまで取扱った有色ノイズは、移動平均型の線形構造をもつにすぎず、モデル誤差の有色性がこれのみによって代表されるものでもないであろうから、他の形式の有色ノイズも考えられよう。

2., 3. 章で取扱ったモデルは、過去の降雨・流量データのみに基づく、いわゆる統計的なモデルであって、流出システム内の物理的要素を考慮したものではない。しかし、得られた結果は、物理的基礎をもつ状態空間モデルに有色ノイズを導入することによって、最適な確率論的流出予測モデルを構成することができるのではないかと期待を抱かせるものである。

4. モデル誤差の有色性を考慮する状態空間モデルの同定・フィルタリング・予測

4.1 概 説

2. および 3. で検討した線形および非線形システムモデルは、現在までの既知の降雨・流量データのみ依存する、いわゆる統計的な数理モデルであった。それは、流域内の地形や流出現象などの物理的な要素をモデル構造の中に全くとり入れていないので、予測モデルとしての適合度はともかくも、言わばその場しのぎのシステムモデルとみなしうる。そのような過去のデータのみ依存したモデルは、対象流域内の様相に何

らかの変化をきたしたとき、まるで使いものにならない、といったトラブルも生じよう。一方、物理的な基礎をもつモデルは、水理学・水文学・地質学・地形学等に基づく幅広い考察によって、完全ではないにしても、システム内に生じた何らかの変化に対してある程度対応することができる。

計算の簡便さやとつき易さ等の実用的見地から、統計的な数理モデルによる予測・制御の有用性は否定できないが、上のような理由から、物理的システム構造を解明していくことの意義もおのずと明らかである。

筆者らはすでに、攪乱項の系列が白色系列をなすと仮定した場合（すなわち、モデルが現象の動特性を完全に記述することのできる場合）の状態空間法による一般的な流出予測理論をひととおりまとめた³⁾。その理論をもとに、本研究のここまでの議論により、物理的要素を考慮すると同時に、モデル誤差（攪乱項）の有色性をも考慮する状態空間型の一般的なシステムモデルについて、4.2ではその同定方法を、さらに4.3ではフィルタリング・予測の方法を述べてゆくことにする。

4.2 有色ノイズをもつ状態空間モデルの同定

【1】モデル構造についての考察

状態空間法による確率論的な流出システムモデルは次のような形式であった。

$$\begin{cases} dx_t/dt = f(x_t, r_t) + Gw_t, & t_{k-1} < t \leq t_k, \quad k=0, 1, \dots \quad \dots\dots\dots (4-1) \\ y_{t_k} = g(x_{t_k}) + v_{t_k} \quad \dots\dots\dots (4-2) \end{cases}$$

記号の意味はすべて1.の(1-2)式に準ずるので詳しい説明は省略する。

ノイズ w_t, v_{t_k} が白色であるという仮定は必ずしも保証されない。また、連続時間型の状態方程式(4-1)に付加されているノイズ w_t の有色性を考慮するのは非常にむずかしい。したがって、(4-1)式からノイズを除去し、(4-2)式の v_{t_k} の系列に動的構造を与えて、次のようなモデルを構成する。

$$\begin{cases} dx_t/dt = f(x_t, r_t), & t_{k-1} < t \leq t_k, \quad k=0, 1, 2, \dots \quad \dots\dots\dots (4-3) \\ y_{t_k} = g(x_{t_k}) + v_{t_k} \quad \dots\dots\dots (4-4) \\ v_{t_k} = a_1 v_{t_{k-1}} + \dots + a_l v_{t_{k-l}} + c_1 e_{t_{k-1}} + \dots + c_m e_{t_{k-m}} + e_{t_k} \quad \dots\dots\dots (4-5) \end{cases}$$

ここに、 x_t は n 次元ベクトル、 e_t は白色正規ノイズ、 $a_1, \dots, a_l, c_1, \dots, c_m$ は定係数である。

連続時間の状態方程式(4-3)はノイズ項が除去されて決定論的となるが、観測(出力)方程式(4-2)の v_{t_k} がそもそも白色性が保証されないのであるから、(4-5)式のように v_{t_k} を有色ノイズとして積極的に定式化することによって、(4-1)式からノイズを除去したことと補償と、(4-2)式自体のもつモデル誤差の補償とを兼ねることになる。

【2】同定のアルゴリズム

同定のために用いるデータとして、降雨データの系列 $r_1, \dots, r_N$ 、流量データの系列 $y_{t_1}, \dots, y_{t_N}$ が与えられているとする。

Step 1: 初期状態量 x_0 と f, g のパラメタ θ_f, θ_g を

$$\frac{1}{N} \sum_{i=1}^N \{y_{t_i} - g(x_{t_i})\}^2 \longrightarrow \min.$$

となるように決定する。 $r_i, y_{t_i}, i=1, \dots, N$ が既知であるので、これは、(4-3)式のような常微分方程式形の等号制約と、 x_0, θ_f, θ_g の物理的制約とをもつ非線形最適化問題である。

Step 2: Step 1 で求められた x_0, θ_f を用いて、状態量の推定値の系列 $x_{t_1}, x_{t_2}, \dots, x_{t_N}$ を(4-3)式にしたがって順次求める。

Step 3: $\theta_g, x_{t_i}, i=1, \dots, N$ を用いて $y_{t_i} = g(x_{t_i}), i=1, \dots, N$ を計算する。

Step 4: $v_{t_i} = y_{t_i} - \hat{y}_{t_i}, i=1, \dots, N$ の系列のコログラムを解析する。もし、 $\{v_{t_i}\}$ が白色と認められればシステムは(4-3)、(4-4)のみで記述しえたことになり、(4-5)は不要である。そうでなければ次へ進む。

Step 5: $\{v_{t_k}\}$ は既知であるので2. で用いた最尤法によって、パラメタ $a_1, \dots, a_l, c_1, \dots, c_m$ が決定できる。

Step 1 における多変数の非線形最適化問題は一般に難解で、種々のアプローチが提案されている。 f, g の関数形により、適宜手法を選択する必要がある。

4.3 有色ノイズをもつ状態空間モデルのフィルタリング・予測

前節の同定アルゴリズムにより同定されたシステムモデルによって確率論的な流出予測を行なう方法について考察する。ここでの議論は1. で述べた筆者らの確率論的な流出予測理論を、モデル誤差の有色性を考慮できるように応用したものと考えてよい。

以下、フィルタリング→予測の手順を概説する。

【1】フィルタリング

流量観測値 y_{t_k} が時々刻々得られてゆき、各時点 t_k においてフィルタリング計算を行なうことになる。

(4-3)～(4-5) は連続時間の状態方程式を含むので、その扱いはかなり複雑である。

状態ベクトルを次元拡張し、

$$X_{t_k} = \begin{bmatrix} x_{t_k} \\ v_{t_k} \\ e_{t_k} \end{bmatrix}, \text{ここに } v_{t_k} = \begin{bmatrix} v_{t_k} \\ \vdots \\ v_{t_k-l+1} \end{bmatrix}, e_{t_k} = \begin{bmatrix} e_{t_k} \\ \vdots \\ e_{t_k-m+1} \end{bmatrix} \dots\dots\dots (4-6)$$

とする。 x, v, e の次元はそれぞれ $(n \times 1), (l \times 1), (m \times 1)$ であるので、 X_{t_k} は $(n+l+m)$ 次列ベクトルである。

(4-3)式は

$$\frac{dX_t}{dt} = \begin{bmatrix} f(x_t, r_t) \\ 0 \\ 0 \end{bmatrix}, t_{k-1} < t \leq t_k \dots\dots\dots (4-7)$$

と表現される。(4-3), (4-5)式は統合された形で状態ベクトル推移式を構成することができる。すなわち、(4-3)式は参考経路のまわりに線形化されて離散形の状態推移式が得られ次のようになる。

$$X_{t_k} = \phi_{k,k-1} X_{t_{k-1}} + B_k + \Gamma e_{t_k} \dots\dots\dots (4-8)$$

ここに $\phi_{k,k-1}, B_k$ はそれぞれ(4-7)式の参考経路のまわりで線形化したことによって得られる推移行列 $((n+l+m) \times (n+l+m))$ および定数ベクトル $((n+l+m) \times 1)$ である。

$$\phi_{k,k-1} = \begin{bmatrix} F_{k,k-1} & 0 & 0 \\ & a_1 \dots a_l & c_1 \dots c_m \\ & 1 & 0 & 0 \\ & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 \\ & & & 0 \dots 0 \\ & & & 1 & 0 & 0 \\ & & & \vdots & \vdots & \vdots \\ 0 & 0 & & 0 & 1 & 0 \end{bmatrix}, l = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \left. \begin{array}{l} n \text{ 次} \\ l \text{ 次} \end{array} \right\} \dots\dots\dots (4-9)$$

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \left. \begin{array}{l} m \text{ 次} \end{array} \right\}$$

で、 $F_{k,k-1}$ は線形化による x_t 部の推移行列 $(n \times n)$ である。

観測（出力）方程式は

$$y_{t_k} = G(X_{t_k}) \dots\dots\dots (4-10)$$

となる。ここに、 G は次元拡張したことによって θ を変換したものである。

結局、(4-3)～(4-5)のシステム方程式は、(4-8)、(4-10)の2式によって表わされた。フィルタリング計算を実行して、 $\hat{X}_k$ の推定値 $\hat{X}(t_k|t_k)$ とその共分散行列 $P_X(t_k|t_k)$ を得ることができる。

【2】予 測

時点 t_{k-1} において、フィルタリングによって、 $\hat{X}(t_{k-1}|t_{k-1})$ 、 $P_X(t_{k-1}|t_{k-1})$ を得ている。また、 $t_{k-1} < t \leq t_k$ の降雨 r_k については、気象予報官によって予測値 $\bar{r}_k$ が与えられているとする。

$\hat{X}(t_{k-1}|t_{k-1})$ を初期値として、(4-7)式にしたがって $t_{k-1} < t \leq t_k$ における参考経路を求める。参考経路のまわりで(4-7)式を線形化し、フィルタリングの場合と同様に $\phi_{k,k-1}$ 、 $\bar{F}_{k,k-1}$ 、 B_k を求め次式によって状態量を予測する。

$$\hat{X}(t_k|t_{k-1}) = \bar{\phi}_{k,k-1} \hat{X}(t_{k-1}|t_{k-1}) + B_k \quad \dots\dots\dots (4-11)$$

ここに、

$$\bar{\phi}_{k,k-1} = \left[\begin{array}{c|cc} \bar{F}_{k,k-1} & 0 & 0 \\ \hline & a_1 \cdots a_l & c_1 \cdots c_m \\ & 1 & 0 & 0 \\ & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 \\ \hline & & & 0 \cdots 0 \\ & & & 1 & 0 & 0 \\ & 0 & 0 & \vdots & \vdots & \vdots \\ & & & 0 & 1 & 0 \end{array} \right] \left. \begin{array}{l} n \text{ 次} \\ \\ l \text{ 次} \\ \\ m \text{ 次} \end{array} \right\} \quad \dots\dots\dots (4-12)$$

で、 $\bar{\phi}_{k,k-1}$ 、 $\bar{F}_{k,k-1}$ 、 B_k はそれぞれ $(n+l+m) \times (n+l+m)$ 、 $n \times n$ 、 $(n+l+m) \times 1$ の次元をもつ。

よって、予測値は

$$\hat{y}(t_k|t_{k-1}) = G(\hat{X}(t_k|t_{k-1})) \quad \dots\dots\dots (4-13)$$

で得られる。

次に、 $\hat{y}(t_k|t_{k-1})$ の分散を求めることを考える。このとき、 $t_{k-1} < t \leq t_k$ における降雨 r_k の予測誤差 ($r_k - \bar{r}_k$) も考慮に入れなければならない。ここで、状態量ベクトル X_k の次元をさらに拡張することにし、 $(1+n+l+m)$ 次列ベクトル

$$X_k = \begin{bmatrix} r_k \\ X_k \end{bmatrix} \quad \dots\dots\dots (4-14)$$

とすれば、 X_k の状態の推移式は(4-8)式のようになる。ただし、次元が1次増大し、(4-9)式において、 $\phi_{k,k-1}$ は $(1+n+l+m) \times (1+n+l+m)$ 行列、 $F_{k,k-1}$ は $(n+1) \times (n+1)$ 行列、 B_k は $(1+n+l+m)$ 次列ベクトルとなる。

(4-8)式より、 $\hat{X}(t_k|t_{k-1})$ の誤差の共分散行列が求められることになり、さらに、(4-10)式を線形化して、予測流量 $\hat{y}(t_k|t_{k-1})$ の分散が求められる。

5. 結 論

より精度のよい流出予測を行なうためには、流出システムを合理的に解析し、すぐれたシステムモデルを構成することが肝要であるのは言うまでもない。しかしながら、流出システムは大規模かつ複雑であるがゆえに、その解析は一筋縄ではいかないところがある。つまるところ、モデル同定が十分でなく流出予測がう

まくゆかないということになる。大規模で複雑なシステムを単純なモデルで記述しようとするとき、モデル同定の不十分さの問題は避けて通れない。したがって、場合によっては、モデルパラメータを逐次更新 (update) してゆく手法や、確率論的な手法が流出システムの解析・予測・制御にしばしば持ち込まれるのである。

本研究では、モデルパラメータの更新は行なわず、モデル同定の不十分さを系列相関をもつノイズ (有色ノイズ) を導入することによって補おうと試みた。

本論文は、2つのパートからなっている。1つは、有色ノイズを導入した確率論的モデルについて、その導入効果を検証するための部分 (2, 3.) である。そこでは、実流域のデータを用いて同定・予測計算を行ない、有色ノイズを導入したモデルとそうでないモデルとを比較検討した。2. では、線形モデルとして、近年動的システムの記述によく用いられる、自己回帰の項を含むモデルのうち、ARMAX 型のものと ARX 型のものとを取扱った。3. では、非線形モデルとして、GMDH によって同定されたモデル (GMDH モデルと称した) とそれに移動平均型ノイズを付加したモデル (GMDHMA モデルと称した) とを比較検討した。このパートでは、適当な次数の有色ノイズを付加することにより、流出予測が改善されうること明らかにした。

しかし、ARMAX モデルも GMDHMA モデルも過去の降雨・流量データのみ依存する統計的な数理モデルであること、また、予測の精度までも与えうる状態空間型のモデル構成で議論していないことから、ここでは、有色ノイズ導入の有効性を確認したにとどまっている。

もう一つのパートは、前半の検証に力を得て、物理的基礎をもつ状態空間モデルの攪乱項に有色ノイズを導入したモデルを構成する部分 (4. 章) である。そこでは、そのような確率論的モデルの同定法と、フィルタリング理論を応用した予測手法を概説した。短期流出予測において、筆者らが目標とするのは、「時々刻々得られる観測情報を最大限に利用し、予測値のみならずその精度までも明らかにするオンライン予測方式を確立する」ことである。4. での議論はそのための1つの方策を展開したことになるが、具体的な適用については、今後の検討課題である。

参 考 文 献

- 1) 高棹琢馬・椎葉充晴：流出システムのフィルタリングと予測，第16回自然災害科学総合シンポジウム講演論文集，1979，pp. 133-136.
- 2) 高棹琢馬・椎葉充晴・宝 馨：流出系のフィルタリングと予測に関する基礎的研究 (第2報)，土木学会第35回年次学術講演会講演概要集第2部，1980，pp. 185-186.
- 3) 高棹琢馬・椎葉充晴：状態空間法による流出予測，京都大学防災研究所年報，第23号B-2，1980，pp. 211-226.
- 4) Kalman, R. E. : "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME, J. Basic Eng.*, Vol. 82, 1960, pp. 35-45.
- 5) Jazwinski, A. H. : *Stochastic Process and Filtering Theory*, Academic Press, 1968.
- 6) 日野幹雄：水文流出系予測へのカルマン・フィルター理論の適用，土木学会論文報告集，第221号，1974，pp. 39-47.
- 7) Todini, E. : "Mutually Interactive State-Parameter Estimation", *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources*, ed. Chao-Lin Chiu, Proc. of AGU Chapman Conf., Pittsburg, Penn., 1978, pp. 135-151.
- 8) Wood, E. F. and Szöllösi-Nagy, A. : "An Adaptive Algorithm for Analysing Short-Term Structural and Parameter Changes in Hydrologic Prediction Models," *Water Resources Research*, Vol. 14, No. 4, 1978, pp. 577-581.
- 9) Sage, A. P. and Husa, G. W. : "Adaptive Filtering with Unknown Prior Statistics," *10th Joint Automatic Control Conf. of AACC*, Boulder, Colo., August, 1969, pp. 760-769.

- 10) 立平良三・保科正男：降水の確率予報，天気，Vol. 27, No. 6, 1980, pp. 433-439.
- 11) Box, G. E. P. and Jenkins, G. M. : *Time Series Analysis—forecasting and control*, Holden-Day, 1970.
- 12) 岸 力：線形確率過程の解析と予測，第11回水工学に関する夏期研修会講義集，A-6, 1975.
- 13) Bolzern, P., Ferrario, M. and Fronza, G. : “Adaptive Real-Time Forecast of River Flow-Rates from Rainfall Data,” *Journal of Hydrology*, 47, 1980, pp. 251-267.
- 14) Katayama, T., Akimoto, K. and Sawaragi, Y. : “Application of Maximum Likelihood Method to Boiler System Identification,” *the Memoirs of the Faculty of Engineering*, Kyoto Univ., Vol. XXXIX, Part 1, Jan., 1977, pp. 148-167.
- 15) Akaike, H. : “A New Look at the Statistical Model Identification,” *IEEE Trans. on A. C.*, Vol. AC-19, No. 6, pp. 716-723.
- 16) Katayama, T. : “Application of Maximum Likelihood Identification to River Flow Prediction,” *Real-Time Forecasting/Control of Water Resource Systems*, ed. Wood, E. F., IASIA proc. series, PERGAMON PRESS, 1976, pp. 63-82.
- 17) 建設省土木研究所：神流川流域水文観測資料
- 18) たとえば，計測と制御，Vol. 19, No. 7, 1980, pp. 725-728 および pp. 733-737.
- 19) Ikeda, S., Ochiai, M. and Sawaragi, Y. : “Sequential GMDH Algorithm and Its Application to River Flow Prediction,” *IEEE Trans. on System, Man and Cybernetics*, vol. SMC-6, No. 7, pp. 473-479.
- 20) 岩崎敏夫・西田吉男：カルマンフィルター及び GMDH による流出計算の実際，土木学会第31回年次学術講演会講演概要集第2部，1976, pp. 168-169.
- 21) 市川良輔・砂田憲吾・三沢宏治：GMDH の流出解析への適用について，土木学会第34回年次学術講演会講演概要集第2部，1979, pp. 45-46.
- 22) 橋本 健：流出成分を考慮した GMDH による低水流出予測，第21回水理講演会論文集，1977, pp. 165-170.
- 23) 宝 馨：確率論的な流出予測に関する研究，京都大学大学院工学研究科修士論文，1981, pp. 26-35.

付 録

【最尤法における最適化アルゴリズム¹⁴⁾】

2. 章の(2-3)の等号制約のもとで，(2-14)の $L^*(\theta)$ の最小点を求めるアルゴリズムを展開する。(2-3)より， $e(t)$ は $\mathbf{a}$ ， $\mathbf{b}$ に関しては線形であるから， $L^*(\theta)$ は $\mathbf{a}$ ， $\mathbf{b}$ に関して2次式となるが， $\mathbf{c}$ に関しては高度に非線形であることに着目する。よって，数値解を求めるために Davidon の方法を用いる。

まず，(2-3)の制約のもとに $L^*(\theta)$ のグラディエント (gradient) を計算する。以下の方法は補助変数を用いたものである。ハミルトン演算子 (Hamiltonian)

$$H(\mathbf{e}, \lambda, \theta) = \frac{1}{N} \sum_{t=1}^N e^2(t) + \frac{1}{N} \sum_{t=1}^N \lambda(t) [e(t) + \sum_{i=1}^m c_i e(t-i) + \sum_{i=1}^n b_i u(t-i) - y(t) + \sum_{i=1}^l a_i y(t-i)] \quad (1)$$

を定義する。ここに， $\{\lambda(t), t=1, \dots, N\}$ は補助変数である。

$$\frac{\partial}{\partial e(t)} H(\mathbf{e}, \lambda, \theta) = 0, \quad t=1, \dots, N \quad (2)$$

より，補助変数が満足すべき方程式は

$$\lambda(t) + \sum_{i=1}^m c_i \lambda(t+i) + 2e(t) = 0 \quad (3)$$

となる。終端条件は

$$\lambda(t)=0, \quad t=N+1, \dots, N+m \dots\dots\dots (4)$$

である。このとき、 $L_N^*(\theta)$ の傾きは次の偏導関数の形で得られる。ここに、 i はそれぞれの次数 l, m, n までの値をとる。

$$\begin{aligned} \frac{\partial L_N^*(\theta)}{\partial a_i} &= \frac{\partial H}{\partial a_i} = \frac{1}{N} \sum_{t=1}^N \lambda(t) y(t-i) \\ \frac{\partial L_N^*(\theta)}{\partial b_i} &= \frac{\partial H}{\partial b_i} = \frac{1}{N} \sum_{t=1}^N \lambda(t) u(t-i) \dots\dots\dots (5) \\ \frac{\partial L_N^*(\theta)}{\partial c_i} &= \frac{\partial H}{\partial c_i} = \frac{1}{N} \sum_{t=1}^N \lambda(t) e(t-i) \end{aligned}$$

よって、グラディエントを計算する手法は以下のように要約される。

(i) パラメタ θ の値が仮定されたとき

$$e(t) + \sum_{i=1}^m c_i e(t-i) = y(t) - \sum_{i=1}^l a_i y(t-i) - \sum_{i=1}^n b_i u(t-i) \dots\dots\dots (6)$$

なる関係にしたがって、 $\{e(t), t=1, \dots, N\}$ を計算する。 $e(t)$ に関する初期条件は

$$e(t)=0, \quad t=0, -1, \dots, -m+1$$

によって与えられるものとする。

(ii) 終端条件(4)を用いて(3)を逆上って加算してゆき補助変数 $\lambda(t)$ を得る。

(iii) (5)によってグラディエントを計算する。

さて、Davidon のアルゴリズムについて述べる。

$(l+m+n) \times 1$ ベクトル

$$g_v = - \left. \frac{\partial L_N^*(\theta)}{\partial \theta} \right|_{\theta = \theta_v}$$

を定義する。添字 v は反復回数を示す。このとき、アルゴリズムは以下のようにまとめることができる。

Step 1 : $\theta_1 = \begin{bmatrix} d \\ 0 \end{bmatrix}$, $d_1 = -g_1$ とする。

Step 2 : $v=1, 2, \dots$ について、 $\phi_v(\lambda) = L_N^*(\theta_v + \lambda d_v)$ を最小とするような $\lambda = \lambda_v$ の値を見つける。

ここで、 λ_v の決定には、黄金分割法やフィボナッチ法や多項式近似法などの線形探索法を用いる。本研究では Davies らによる低次多項式近似法^{*}によった。

Step 3 : $\theta_{v+1} = \theta_v + \lambda_v d_v$, $p_v = g_{v+1} - g_v$ と定義する。

Step 4 : $(l+m+n) \times (l+m+n)$ 行列

$$H_v = H_{v-1} + \lambda_v \frac{d_v d_v'}{g_v' H_{v-1} g_v} - \frac{H_{v-1} p_v p_v' H_{v-1}'}{p_v' H_{v-1} p_v}$$

を計算する。ただし、 H_v は単位行列である。 $d_{v+1} = -H_v g_v$ とする。

Step 5 : Step 2-4 をくり返し、適当な収束条件が満足されるまでつづける。

^{*} Davies らによる低次多項式近似法については、関根智明訳：非線形最適化問題の反復解法，培風館，1976，pp. 64-65.