

## II-8

## パターン認識手法による水文時系列の解析とシミュレーション

金沢大学工学部 正員 宇治橋康行

同 上 正員 高瀬 信忠

備 大 林 組 正員 家門 邦彦

1. まえがき、水文時系列中には明確なグループ（例えば、高水流量と低水流量）が存在することは古くから知られていたが、時系列解析の情報としては利用されていなかった。Panu らはこのようなグループを一つの変動パターンと見なし、パターン認識の概念を用いたデータシミュレーションと特徴予測手法を提案した。彼らの手法は簡単でしかも時系列の変動パターンが直観的で分かり易いという利点を持つ。しかし、実データ中の特殊なパターン（例えば、異常小雨や多雨）が必ずしもうまく分離・認識できない、旬・半旬のような短い時間単位のデータや降水データに対しては誤認が多い等の欠点がある。ここでは Isodata アルゴリズムによるクラスタリングを用いてこれらの点を改良したデータ解析とシミュレーション手法について述べる。

## 2. Isodata アルゴリズムによる認識システムとデータシミュレーション

詳細は文献にゆずり、ここでは図-1の流れ図に沿って概略を述べる。ステップ2までの時系列の対象への分割とパターンベクトルの形成は、コレログラム、ペリオドグラム解析の結果等を参考にシーズンの概念を用いて行われ一年をKシーズンに分割し、N年間のデータを $K \times N$ 個のパターンベクトル系列に変換する。ステップ3の標準ベクトルの形成と分類規則に関して Panu らは一シーズンを一パターンクラスと見て、各パターンクラスに属するパターンベクトルを正しく認識できるように特徴変換を行ない、特徴ベクトルに対し最小距離と誤認基準を用いた認識システムを構成している。しかし、一シーズンを一つのパターンクラスとし各シーズンごとに異なる特徴変換を行うため、異なるシーズンで類似したパターンを全く異なるものと分類認識したり、逆に類似していないパターンを同じパターンクラスのものとして誤認することがある。又、特徴変換を行うことによって原系列の変動パターンの各シーズンごとの比較ができなくなる。そこで本研究では、特徴変換は行わずパターンベクトルをそのまま特徴ベクトルと見なし、かつ一シーズンを一パターンクラスとはせず $N \times K$ 個のすべてのパターンベクトルを Isodata アルゴリズムを用いてクラスタリングを行うこととした。Isodata アルゴリズムはユークリッド距離を類似度とし、初期パターン中心ベクトルと適当なき値を与えるだけで発見的にクラスタリングを行うことのできる手法である。この結果パターンベクトルの分類が行われ標準ベクトルが形成されると同時に最小距離基準に基づく認識システムが構成される。定められた認識システムとパターン内およびパターン間の統計的構造に従ってデータシミュレーションを行うステップ4、5は Panu らの方法と同じである。

## 3. 適用結果と考察

解析に用いたデータは金沢の1886年～1985年までの99年間の月降水量データである。データのシーズンへの分割は3ヶ月を一シーズンとして4シーズンに分割した。第1シーズンは12、1、

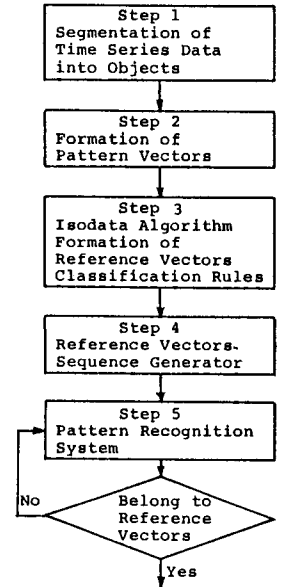


図-1 データシミュレーションの流れ図

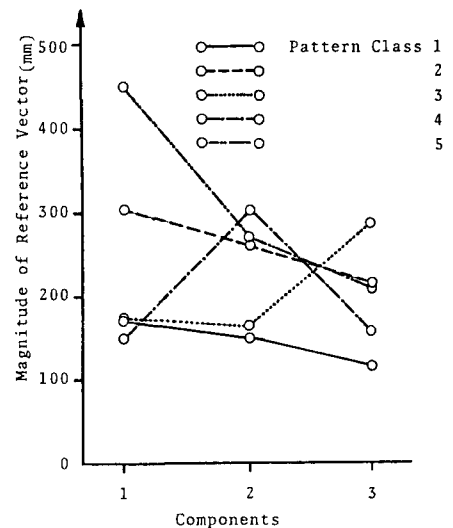


図-2 各パターンクラスの標準ベクトル

表-1 各シーズンに生起するパターンベクトル数

|                  |   | Season |    |    |    |  |
|------------------|---|--------|----|----|----|--|
|                  |   | 1      | 2  | 3  | 4  |  |
| Reference Vector | 1 | 3      | 40 | 13 | 5  |  |
|                  | 2 | 24     | 0  | 5  | 14 |  |
|                  | 3 | 0      | 6  | 11 | 23 |  |
|                  | 4 | 4      | 4  | 20 | 4  |  |
|                  | 5 | 19     | 0  | 1  | 4  |  |

2月, 第2シーズンは3, 4, 5月, 第3シーズンは6, 7, 8月, 第4シーズンは9, 10, 11月とした。前半50年間のデータに対する Isodataによるクラスタリングの結果, 全データは5つのパターンクラスに分割された。各クラスの標準ベクトルを図-2に, 各シーズンに生起する各パターンクラスに属するパターンベクトルの数を表-1に示す。この結果から各シーズンは一つのパターンクラスと見なすことができ, 第1シーズンでは2個(第2, 4パターン), 第2シーズンでは1個(第1パターン), 第3シーズンでは3個(第1, 3, 4パターン), 第4シーズンでは2個(第2, 3パターン)のサブパターンクラスが存在することが分かる。また各シーズンには, Panu らの方法では誤認となるような生起度数の少ないサブパターンクラスに属さないパターンが生起することや, 2つのシーズンのサブパターンとなるパターンが存在することが示され実際の良い分類認識結果となっている。各パターンクラスに属するパターンベクトルは有意水準5%で多変数正規性を満たし, シーズン内のパターンの遷移を示す一次マルコフ遷移行列は表-2のようであった。ここで構成された認識システムと統計的構造に基づくデータシミュレーションの結果を表-3, 図-3, 4に示すが, 結果の詳細については講演時に述べる。

## &lt;参考文献&gt;

- 1) U. S. Panu, T. E. Unny : Stochastic Synthesis of Hydrologic Data Based on Concepts of Pattern Recognition, Jour. Hydrology, 46, pp.5-34, 1980.
- 2) 家門邦彦 : パターン認識手法による水文データの解析に関する研究, 金沢大学修士論文, 1987年.
- 3) Tou, J. T., Gonzalez, R. C. : Pattern Recognition Principles; Addison-wesley Reading, 1974.

表-2 各シーズンの一次マルコフ遷移行列

|                  |    | Reference Vector |    |       |       |    |
|------------------|----|------------------|----|-------|-------|----|
|                  |    | 21               | 22 | 23    | 24    | 25 |
| Reference Vector | 11 | 1.0              |    |       |       |    |
|                  | 12 | 0.75             |    | 0.083 | 0.167 |    |
|                  | 13 |                  |    |       |       |    |
|                  | 14 | 1.0              |    |       |       |    |
|                  | 15 | 0.789            |    | 0.211 |       |    |

|                  |    | Reference Vector |       |       |       |       |
|------------------|----|------------------|-------|-------|-------|-------|
|                  |    | 31               | 32    | 33    | 34    | 35    |
| Reference Vector | 21 | 0.25             | 0.1   | 0.2   | 0.425 | 0.025 |
|                  | 22 |                  |       |       |       |       |
|                  | 23 | 0.333            | 0.167 | 0.167 | 0.333 |       |
|                  | 24 | 0.25             | 0.5   | 0.25  |       |       |
|                  | 25 |                  |       |       |       |       |

|                  |    | Reference Vector |       |       |       |       |
|------------------|----|------------------|-------|-------|-------|-------|
|                  |    | 41               | 42    | 43    | 44    | 45    |
| Reference Vector | 31 |                  | 0.231 | 0.538 | 0.077 | 0.154 |
|                  | 32 |                  | 0.4   | 0.6   |       |       |
|                  | 33 | 0.273            | 0.273 | 0.364 | 0.091 |       |
|                  | 34 | 0.1              | 0.3   | 0.4   | 0.1   | 0.1   |
|                  | 35 |                  |       | 1.0   |       |       |

|                  |    | Reference Vector |       |    |       |       |
|------------------|----|------------------|-------|----|-------|-------|
|                  |    | 11               | 12    | 13 | 14    | 15    |
| Reference Vector | 41 |                  | 0.6   |    | 0.2   | 0.2   |
|                  | 42 | 0.071            | 0.5   |    |       | 0.429 |
|                  | 43 | 0.067            | 0.478 |    | 0.044 | 0.391 |
|                  | 44 |                  | 0.75  |    |       | 0.25  |
|                  | 45 |                  |       |    | 0.5   | 0.5   |

表-3 時系列レベルでの統計量の比較

|       | mean     | s.d.     | lag 1 AC | skew.  | kurt.  |
|-------|----------|----------|----------|--------|--------|
| X     | 214.4283 | 101.5970 | 0.1406   | 0.7293 | 3.4138 |
| S     | 19.7550  | 14.1906  | 0.2421   | 0.5332 | 1.6067 |
| Hist. | 211.4940 | 98.6359  | 0.1978   | 0.9848 | 4.2828 |

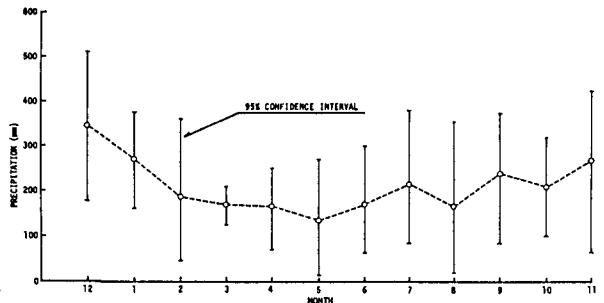


図-3 月レベルでの平均値の比較

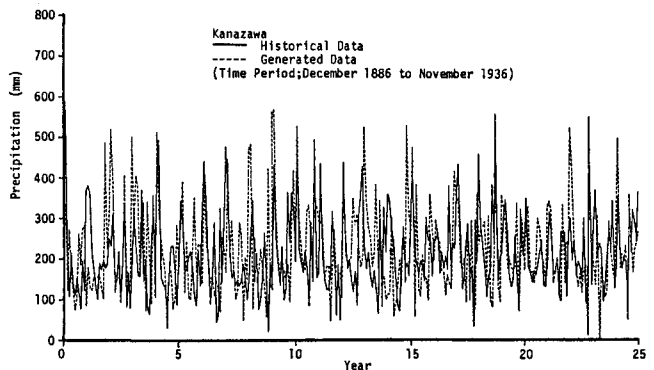


図-4 シミュレーション結果の一例