土木学会第53回年次学術講演会（平成10年10月）

II -66

長期気候変動のパターン分類とその発生確率に関する研究

京都大学大学院　学生員　塩出貴保
京都大学防災研究所　正員　小尻利治
京都大学防災研究所　正員　友杉邦雄

1 はじめに 本研究では、時系列データを１つのベクトルパターンとして捉え、パターン認識概念を導入して降水量、気温時系列の特性把握を試み、長期気候変動によるパターンの発生特性の変化の検討を行う。さらに、降雨発生確率に対して、降水量系列の変動形状を考慮したパターンとしての発生確率評価手法を提案する。最後に、１つのパターンとして捉えるのに最適な期間分割を考慮したパターン分類を試みる。

2 パターン分類化手法 パターン分類化の１つであるISODATA手法[1]を用いて、水文事象の分類化を試みる。パターン分類を行うにあたって、最初に分類基準となる目的関数を設定しなければならない。全期間、かつ、全観測値を対象とし、時間持続性を考慮した分類を行うため、次式を目的関数として採用する。

$$OF(I,M)=\sqrt{\sum_J\{X(J,I)-CL(J,M)\}^2} \qquad (1)$$

ただし、$X(J,I)$をI年のJ月の観測値、$CL(J,M)$をパターンMのクラスターセンターのJ月の値とする。

ISODATA手法より、対象とするデータに対して任意のクラスター数への分類を行うことができる。しかし、クラスター数によって分類結果も異なってくるので、適切なクラスター数の判定を行うために、情報量規準AIC[2]を導入する。ここで、各クラスターで、サンプルデータの目的関数値が平均$\lambda(M)$、分散$\zeta^2(M)$の対数正規分布に従うとすると、クラスターS_Mでの対数尤度は、

$$LL(M)=-\sum_{X\in S_M}[\ln\{OF(I,M)\}+\tfrac{1}{2}\ln\{2\pi\zeta^2(M)\} + \frac{1}{2}\{\frac{\ln\{OF(I,M)\}-\lambda(M)}{\zeta(M)}\}^2] \qquad (2)$$

で表される。全情報量は$LL(M)$の和で求められるので、最適クラスター数は次の最小化問題として決定される。

$$AIC(N)=-2*\sum_{M=1}^{N}LL(M)+4*N\rightarrow min \qquad (3)$$

ただし、Nはクラスター数を表す。

また、１月から１２月を１つの期間、つまり１形状として取り扱うだけでなく、１年を何期間かに分割して、分割された期間を１つの形状として取り扱う方がより最適な分類が得られるはずである。そこで、どのような期間分割が最も適しているかについても考察を加える。すなわち、１年を適当な期間に分割し、それぞれの期間で最適クラスター数への分類を行う。それぞれの期間の各クラスターで、属するサンプルデータの目的関数値が平均$\lambda(M)$、分散$\zeta^2(M)$の対数正規分布していると仮定すると、各クラスターにおける目的関数値の対数尤度$LL(M)$は、

$$LL(M)=-\sum_I[\ln\{OF(I,M)\}+\tfrac{1}{2}\ln2\pi\zeta^2(M) + \frac{1}{2}\{\frac{\ln\{OF(I,M)\}-\lambda(M)}{\zeta(M)}\}^2] \qquad (4)$$

と表される。全情報量は対数尤度$LL(M)$の和で求められるので、情報量規準AICはNをクラスター数とすると、

$$AIC=-2*\sum_{M=1}^{N}LL(M)+4*N \qquad (5)$$

で与えられる。全ての期間分割に対する情報量規準AICのうち最小となるAICを与える期間分割を最適期間分割とする。

3 降雨の発生確率 従来は、年間総降水量分布を対象として、ある少雨の年の年間総降水量の非超過確率でもって少雨の発生確率としてきた。しかしこの方法では、分布の形状が全く考慮されていないので、本研究では次のような方法を提案する。すなわち、対象とする年間少雨のサンプルデータXがすべてのクラスターに属していると考えて、各クラスターでのクラスター自身の発生確率$P_c(M)$とそのクラスター内でのサンプルデータXの発生確率$P_D(X|M)$との積の和、

キーワード：パターン分類、気候変動、AIC、降雨発生確率
連絡先：宇治市五ケ庄防災研究所水資源センター、Tel:0774-38-4269、Fax:0774-32-3093

$$P_X = \sum_{M=1}^{N} P_c(M) * P_D(X|M) \tag{6}$$

をもってサンプルデータXの発生確率P_Xとする。ここで、$P_c(M)$は各クラスターに属するサンプルデータの個数を全サンプルデータ数で割ったもので与えられる。また、$P_D(X|M)$は各クラスターで、属するサンプルデータの目的関数値が対数正規分布と仮定した場合のサンプルデータXの目的関数値における超過確率で与えられるとする。

4 適用と考察 名古屋市における1891年から1996年までの106年間の月単位の降水量、気温時系列を適用した結果、降水量は12パターン、気温は10パターンに分類された。（ただし、ここでは１月から１２月までの時系列を１つの形状としている。）この分類結果から、各年のサンプルデータが経年的にどの降水量、気温パターンに属しているかを示したものが図１、図２である。ただし、縦軸は上から年総降水量の多い、年平均気温の高いパターンの順に離散的に並べている。同図から、降水量に関しては、気候変動による発生特性の変化はそれほど顕著に現れなかった。しかし、過去２０年に１回の割合でしか現れなかったパターン７のような少雨がここ１０年で２回も現れているのは気候変動の影響と言えよう。また、気温に関しては、１９５４年頃まではパターン３からパターン２の間で各年のパターンが分布していたが、それ以後、それまで全く現れなかったパターン１、６、８が近年多発していることが分かる。以上のことから、気温のパターンの発生特性が気候変動により、かなり変化してきていると言える。

また、降水量、気温パターン間の相関性を的中率を用いて調べると、気温のみでは確定的に降水量を推測することは難しいが、ある気温パターンに対応する降水量パターングループが存在することが分かる。

次に、降雨の発生確率について検討する。記録的な少雨となった1994年の降水量データを対象として降雨発生の確率的評価を試みる。まず、解析対象サンプルデータを年間総降水量分布のみを対象とした従来の方法で評価する。年間総降水量分布を対数正規分布と仮定すると、1994年の年間総降水量の非超過確率、つまり1994年の少雨の発生確率は0.0104となり、Return Periodは96.2年であった。次に、パターン認識概念を用いて形状としての降水量の異常性を基準とした少雨の評価すると、1944年の少雨の発生確率は0.0136，となり、Return Periodは73.5年であった。

最後に、期間分割を考慮したパターン分類を行った結果、前述の結果では現れなかった形状をしたパターンが得られるなど、より形状の特性に応じた分類が得られた。

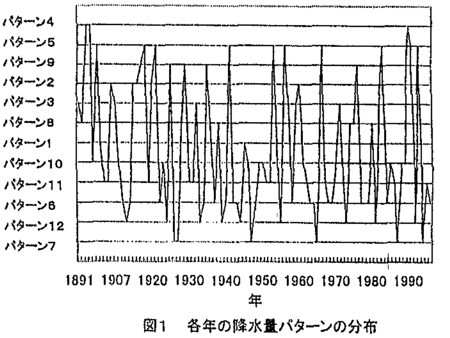

図1　各年の降水量パターンの分布

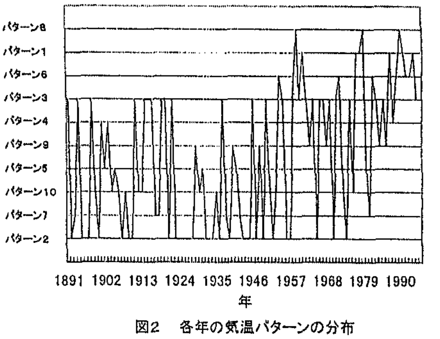

図2　各年の気温パターンの分布

5 結論 ISODATA手法とAICを組み合わせることによって、普遍的なパターン分類化手順を確立することができた。また、地球温暖化の影響とみられる近年のパターン特性を明らかにすることができた。さらに降雨の発生確率については、従来の方法では評価できなかった形状の相違を考慮した頻度を得ることができた。

参考文献

[1]Tou.J.T. and Gonzalez.R.C.:"Pattern Recognition-Principle." Applied Mathematics and Computation, No.7,Addison Wesley Publishing Company, pp.75-109, 1974.

[2]鈴木義一朗：情報量規準による統計解析入門，講談社出版，pp.66-151, 1995.