

行政院國家科學委員會補助專題研究計畫成果報告

級值序列應用於時間序列之模式判定

計畫類別： 個別型計畫

計畫編號： NSC89 - 2211 - E - 032 - 039

執行期間： 89 年 8 月 1 日 至 90 年 7 月 31 日

計畫主持人： 虞國興

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：淡江大學水資源及環境工程研究所

中 華 民 國 九 十 年 七 月 三 十 一 日

行政院國家科學委員會專題研究計畫成果報告

國科會專題研究計畫成果報告撰寫格式說明

Preparation of NSC Project Reports

計畫編號：NSC 89-2211-E-032-039

執行期限：89年8月1日至90年7月31日

主持人：虞國興(淡江大學水資源及環境工程研究所)

共同主持人：xxxxxx 執行機構及單位名稱

計畫、參與人員： 雍 賓

一、中文摘要

在河川治理與水資源工程規劃設計上，常因水文資料記錄長度不足或資料短缺，使得資料在分析上難有一致之結果。為解決水資源規劃設計因資料短缺造成之模式誤判以及參數推估之精確度不足，本研究根據莊(1999)提出之級值序列理論，針對小樣本數所造成之不確定性作一系統之研究，並提出級序繁衍法以降低參數推估之不確定性。本研究比較分析級序繁衍法所繁衍出來之合成資料與原始資料其 ACF、PACF 統計特性之保存能力，以及在不同樣本數之情況下，其模式判定與參數推估之表現。另外，亦將級序繁衍法應用於實測資料之模式判定上。研究結果顯示，級序繁衍法可有效保存原始資料 ACF 及 PACF 統計特性。在樣本數小於 100 且參數不顯著之情況下，級序繁衍法於參數推估及模式滿足 AR(1) 模式下模式判定之結果均優於原始序列。而在樣本數大於 100 或參數顯著之情況下，級序繁衍法於參數推估及模式判定之表現與原始序列結果一致。然，當模式滿足 AR(2) 時，級序繁衍法與傳統模式判定方法之精確度一致，特別在 AR(2) 模式參數較不顯著且趨近 AR(1) 模式時，無法有效解決模式不確定性之問題。

關鍵詞：時間序列、級值序列、模式判定

Abstract

In river management and water resources planning, the hydrologic data analyzed was usually not enough to obtain consistent results. To solve this problem, the order series theorem proposed by Chuang(1999) is used in this study to analyze uncertainty caused by small sample size. In this study, order series generation(OSG) method is used to degrade uncertainty of parameter estimation and improve model uncertainty at small sample size. Properties of the synthetic data generated by OSG method is compared with original data including ACF and PACF, model identification, and parameter estimation. The results indicated that the synthetic data generated by OSG method preserved the ACF and PACF of original one. In parameter estimation and model identification for AR(1) model, OSG method is

better than the conventional one, especially when sample size under 100 and the AR parameters are weak. When sample size over 100 or the AR parameters are strong, OSG method is as good as the conventional method. But for AR(2) model, OSG method and the traditional one make not much difference.

Keywords: Time Series, Order Series, Model Identification

二、緣由與目的

在水資源工程規劃設計上，就實務工程而言，一般之水文分析均屬定率分析(Deterministic Analysis)，並無考慮到在分析過程中所存在之各種不確定性(Uncertainties)；但是在實際水文分析與模式應用上，因自然現象之不確定性問題確實存在，故任何水文模式分析所獲得之結果亦須考慮到不確定性。其中，造成不確定之存在主要包括下列幾項：

- (1) 自然現象之隨機性：自然界水文發生過程之隨機性。
- (2) 模式本身之不確定性：模式本身不能完全解釋與描述該現象。
- (3) 參數推估之不確定性：參數是利用有限之資料做一推估而得，其真值並無從得知。
- (4) 其他因素：如記錄短缺、數據誤差及分析者本身之主觀判定等人為、非人為之因素。

時間序列分析中，常使用之推估法及模式判定準則均會受到參數推估之準確與否及模式選取是否正確等因素影響，進而造成參數不確定性(Uncertainty of Parameters)以及模式不確定性(Model Uncertainty)。造成參數推估或模式判定不確定性之主要原因之一在於資料個數之長短，即樣本數大小；在樣本數大時，無論使用何種分析方法其結果均差異不大；但是在樣本數小時，分析結果會因分析方法不同而有不同之結論。

實務工程規劃上，如河川治理與水資源規劃，常發生水文資料記錄長度不足或資料短缺等困擾。惟今，為解決水資源工程規劃上模式不確定性問題以改善規劃結果之適用性，本研究藉由級值序列理論(莊，1999)於小樣本數下所造成之不確定性

作一系統之研究，並提出一降低不確定性之參數推估方法，以改善不確定性於資料長度短缺之問題。

三、結果與討論

本章節主要分成合成資料及實測資料來探討。合成資料部份主要分為三部份：(1)級序繁衍法於資料結構之保存能力。(2)級序繁衍法於參數推估準確性之探討。(3)級序繁衍法於模式選取能力優劣之探討。實測資料部份則探討級序繁衍法應用於實測資料之模式選取能力。

3-1 合成資料

3-1.1 級序繁衍法於 ACF、PACF 統計特性之探討

(1) AR(1) :

在參數不顯著的情況下，級序繁衍法平均 ACF 之 MSE 值均比原始資料小，但在參數顯著的情況下，級序繁衍法平均 ACF 之 MSE 值則比原始資料大。隨著參數顯著的情形愈增，級序繁衍法之 $Bias^2$ 與原始資料之 $Bias^2$ 差距愈大。

(2) AR(2) :

由這三部份結果顯示：級序繁衍法只有在 AR(2)兩模式參數均不顯著的情況下，其平均 ACF 之 MSE 值會小於原始資料，其他模式參數下，級序繁衍法之 MSE 值則會比原始資料大，其原因在於級序繁衍法之 $Bias^2$ 會隨著參數顯著情形增加而變大。隨著參數顯著性減弱兩者 $Bias^2$ 之差距亦隨著變小。

級序繁衍法之 $Bias^2$ 在參數顯著的情況下較為顯著，此為級值序列之一主要特性，但是其數值相對來說並不大，故級序繁衍法的確可以保留原始資料 ACF、PACF 之特性，尤其在參數不顯著的情形下。隨著樣本數增大，級序繁衍法之特性符合前述參數不確定性之理論，其平均 ACF 之變異數亦隨著減小。

3-1.2 級序繁衍法於參數推估之探討

(1) AR(1) :

參數不顯著之情況下，級序繁衍法推估參數之 MSE 值均較原始資料為小，尤其在樣本數小的時候，更為明顯；隨著參數之顯著情形增加，原始序列之參數推估 MSE 值會比級序繁衍法為小。

(2) AR(2) :

級序繁衍法推估參數之 MSE 值會隨著模式參數顯著性減小而由大於原始序列推估參數之 MSE 值變為小於原始序列之 MSE 值。值得注意的是在樣本數小的情況下，參數顯著性均微弱時，級序繁衍法推估參數之 MSE 值小於原始序列推估參數之 MSE 值。

級序繁衍法於參數推估之適用性於此又可在一次得到驗證。在參數不顯著的情況下，級序繁衍法於參數推估較傳統方法來的好，尤其在樣本數較小的情況下。

3-1.3 級序繁衍法於模式選取能力之探討

(1) AR(1) :

在樣本數為 50，參數不顯著之情況下，原始資料在 $W_1 = \pm 0.1 \sim 0.2$ 時會判定出 AR(0) 模式，而級序繁衍法只發生在 $W_1 = 0.1$ ，在 BIC2 之準則下並不會有此情形發生；隨著參數之顯著情形增加，各準則均可選取出正確模式，其中級序繁衍法所判定出之正確組數比原始資料直接判定之正確組數為多。在樣本數 100、200 及 500 時，級序繁衍法判定出之模式與原始資料模式判定結果一致，且其選取能力均在九成五以上，尤以 BIC2 準則為最。

(2) AR(2) :

在樣本數為 50，模式參數顯著情形極為明顯(特徵方程式根的絕對值倒數均在 0.65 以上)的情況下，級序繁衍法方可判定出正確的模式且判定出正確模式之組數比原始資料為多。反之，級序繁衍法及原始資料判定出之模式均會發生模式低估的情況。在樣本數為 100、200 及 500 時，均可以看到級序繁衍法之表現與原始資料判定結果一致。

級序繁衍法在模式判定上之表現能力可說優於原始序列，尤其在樣本數小、參數不顯著之情況下；由級序繁衍法選取出之模式組數比原始序列多之結果可知級序繁衍法彰顯出原始序列之特性；而在樣本數大時，其結果與原始序列之結果一致更說明了級序繁衍法於模式判定之適用性。

3-2 實測資料

級序繁衍法應用於實測資料時，因為所繁衍出之合成資料多數均通過適合度檢定，造成無法選取正確模式。故，將該步驟予以省略。由表 1 可以知道級序繁衍法亦適用於實測資料，但仍須更進一步的探討。

四、計劃成果自評

本研究應用級值序列於模式判定之分析，其主要目的在避免傳統分析方法上直接以該筆資料判定之，易產生模式判定不正確，即發生模式之不確定性。其概念在於利用級值序列能保有資料之結構特性，藉由資料之繁衍並配合資料本身之級值序列，繁衍保有原始序列模式結構之合成資料數組(本研究採用 100 組)，由該合成資料判定模式所得最多組數之模式即為該原始序列應滿足之模式。如此，避免了傳統模式判定方法所造成之不確定性。研究中，亦針對級序繁衍法所得序列之統計特性、參數推估作一完整之探討。其結果可歸納得下列幾點結論：

- (1) 在統計特性方面，級序繁衍法合成資料之 ACF 與 PACF 統計特性與原始序列相同。
- (2) 在參數推估方面，樣本數 100 以下且參數不顯著時，級序繁衍法之參數推估優於原始序列之參數推估。在樣本數 100 以上

或參數顯著時，級序繁衍法與原始序列之推估結果一致。

在模式判定方面，當模式滿足 AR(1)，樣本數在 100 以下且參數不顯著時，級序繁衍法之判定結果優於原始資料之判定結果。在樣本數 100 以上或參數顯著時，級序繁衍法與原始資料判定之結果一致。然，當模式滿足 AR(2)時，結果顯示，本研究分析方法與傳統模式判定方法之精確度相同，尤以 AR(2)模式參數較不顯著且趨近 AR(1)模式時，並無法有效解決模式不確定性之問題。

五、參考文獻

1. 莊明德(1999)，虞國興博士指導，「級值序列法應用於非常態水文時間序列之分析」博士論文。
2. 虞國興、劉治均(1995)，「時間序列模式選取準則之比較」，農業工程學報，第 41 卷，第一期，pp.15-26。
3. 虞國興、金士凱(1999)，「水文短序列模式判定與預測之研究」，農業工程學報，第 45 卷，第一期，pp.13-23。
4. Akaike, H.(1974), “A New Look at the Statistical Model Identification”, IEEE Trans. Autom. Control. 19, 716-723.
5. Akaike, H.(1974), “A Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting”, Biometrika 66, 237-242.
6. Box, G.E.P., Jenkins, G.M. and Reinsel, G.C.(1994), Time Series Analysis Forecasting and Control, 3rd Edition, Printice Hall, New Jersey.
7. Brockwell, P.J. and Davis, R.A.(1991), Time Series: Theory and Methods, 2nd Edition, Springer-Verlag, New York.
8. Chen, C., Davis, R.A., Brockwell, P.J. and Bai, Z.D.(1993), “Order Determination for Autoregressive Processes Using Resampling Methods”, Statist. Sinica 3, 481-500.
9. Hurvich, C.M.(1987), “Automatic Selection of a Linear Predictor Through Frequency Domain Cross-Validation”, Comm. Statist. Theory Methods 16, 3199-3234.
10. Hurvich, C.M. and Tsai, C.-L.(1989), “Regression and Time Series Model Selection in Small Samples”, Biometrika 76, 297-307.
11. Hurvich, C.M. and Tsai, C.-L.(1991), “Bias of the Corrected AIC Criterion for Underfitted Regression and Time Series Models”, Biometrika 78, 499-509.
12. Rissanen, J.(1978), “Modelling by Shortest Data Description”, Automatica 14, 465-471.
13. Salas, J.D., Delleur, J.W., Yevjevich, V. and Lane, W.L.(1985), Applied Modeling of Hydrologic Time Series, 2nd Edition, Water Resources Publications, Michigan.
14. Shaman, P. and Stine, R.A.(1988), “The Bias of Autoregressive Coefficient Estimator”, J. Amer. Statist. Assoc. 83, 842-848.
15. Shibata, R.(1976). “Selection of the Order of an

Autoregressive Model by Akaike’s Information Criterion”, Biometrika 63, 117-126.

16. Shanmugan, K.S., and Breipohl, A.M.(1988), “Random Signals: Detection, Estimation and Data Analysis”, John Wiley & Sons, Canada.

17. Wei, W.S.(1990), Time Series Analysis Univariate and Multivariate Methods, Addison-Wesley, New York.

表 1 實測資料之模式判定結果

代號	已知模式	AIC		BIC1		BIC2	
		原始	合成	原始	合成	原始	合成
W1	AR(1)	AR(1)	AR(1)	AR(0)	AR(0)	AR(1)	AR(1)
W2	AR(2), AR(9)	AR(9)	AR(9)	AR(9)	AR(9)	AR(9)	AR(9)
W3	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)
W5	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)
W7	AR(3)	AR(4)	AR(4)	AR(4)	AR(4)	AR(4)	AR(4)
BJ-D	AR(1)	AR(1)	AR(1)	AR(0)	AR(0)	AR(1)	AR(1)
BJ-E	AR(2), AR(3)	AR(2)	AR(2)	AR(2)	AR(2)	AR(2)	AR(2)
BJ-F	AR(2)	AR(2)	AR(2)	AR(1)	AR(1)	AR(1)	AR(1)
C1	AR(2)	AR(2)	AR(2)	AR(2)	AR(2)	AR(2)	AR(2)
C2	AR(1)	AR(8)	AR(7)	AR(7)	AR(6)	AR(7)	AR(6)
LIN1	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)
HW1	AR(2)	AR(2)	AR(2)	AR(1)	AR(2)	AR(2)	AR(2)
HW3	AR(1)	AR(1)	AR(1)	AR(0)	AR(0)	AR(1)	AR(1)
L1	AR(2)	AR(3)	AR(2)	AR(0)	AR(0)	AR(2)	AR(1)
L2	AR(2)	AR(2)	AR(1)	AR(0)	AR(0)	AR(2)	AR(1)
選取率	-	12/15	12/15	6/15	7/15	12/15	10/15