

110學年度高等水文分析作業四

時間序列分析

個人作業，請於12月15日(星期三) 3am以前繳交電子檔至CEIBA
請三、四組準備簡報12/16星期四8:10抽籤由一位同學報告作業成果

HW4_Temp.txt檔案內容為中央氣象局19個測站2003到2009年無缺漏的小時溫度紀錄，第一欄為年、月、日、時，2-20欄為19站的溫度記錄觀測值，其順序和StnList.txt檔案中的測站順序相同。請去除2004和2008年的2月29日的24筆溫度資料，且假設去除2月29日的24筆溫度資料後，對於計算時間序列的週期項、自迴歸關係的差異影響可以忽略。

1. 假設資料的長期趨勢可以忽略，且時間序列的模型為週期項與殘差項兩者的和，請估計所有測站的共同週期項。(1+1+2+2分)

A. 參考作業2-2，首先計算各測站的平均溫度 $\bar{T}_i = \sum_i T_i(t) / (7 \times 365 \times 24)$ ，
 $i = 1 \sim 19$ ；

B. 計算所有測站去均溫的平均值時間序列：

$$\tilde{T}(t) = \sum_{i=1}^{19} [T_i(t) - \bar{T}_i] / 19 ;$$

C. 將年、月、日、時改為 $7 \times 365 \times 24$ 的編號，再利用調和分析法計算包含年週期(8760小時)和日週期(24小時)的所有364+24個傅利葉係數值，保留超過 $\alpha = 0.01$ 顯著水準的所有傅立葉階項（即使部分低階項不顯著、高階項顯著時，仍保留高階項）。（註：因為資料完整、沒有缺漏，所以不同階的傅立葉項彼此正交，只用各階係數的平方和便可評估是否顯著，但先後去除有自由度的差異，建議按照係數平方和大小，依序評估顯著性）

$$\begin{aligned} \tilde{T}(t) = & \mu + \sum_{j=1}^{12} \left\{ \alpha_j \sin \left[\frac{2j\pi t}{24} \right] + \beta_j \cos \left[\frac{2j\pi t}{24} \right] \right\} \\ & + \sum_{k=1}^{182} \left\{ a_k \sin \left[\frac{2k\pi t}{8760} \right] + b_k \cos \left[\frac{2k\pi t}{8760} \right] \right\} + \bar{Z}(t) \end{aligned}$$

D. 繪製顯著日週期項的和 $\tilde{P}_{d,s}(t) = \sum_{j,s} \left\{ \alpha_{j,s} \sin \left[\frac{2j\pi t}{24} \right] + \beta_{j,s} \cos \left[\frac{2j\pi t}{24} \right] \right\}$ ；繪

製顯著的年週期項 $\tilde{P}_{y,s}(t) = \sum_{k,s} \left\{ a_{k,s} \sin \left[\frac{2k\pi t}{8760} \right] + b_{k,s} \cos \left[\frac{2k\pi t}{8760} \right] \right\}$ ；計算各

顯著傅利葉項解釋變異數百分比，繪顯著階序—解釋變異數百分比圖。

（註1： j, s 、 k, s 都未必連續，繪圖時請依 $\alpha_{j,s}^2 + \beta_{j,s}^2$ 、 $a_{k,s}^2 + b_{k,s}^2$ 的大小順

序混合排列，圖座標依大小序標示顯著階數，另列表表示圖中橫座標序的傅立葉階數。註2： $j=1$ 可標示為 $k=365$ ， $j=2$ 可標示為 $k=730$ 。）

2. 計算各測站溫度的序率項 $z_i(t) = T_i(t) - \bar{T}_i - \tilde{T}(t)$ ；假設19站序率項的時間序列模型為AR(1)、AR(2)或AR(3)，請建立模型、估計參數。辨識假設三階（亦可刪除）序率模型如下（本作業不考慮四階和高階序率模型）：

$$\hat{\mathbf{T}}_3(t) = \bar{\mathbf{T}} + \tilde{\mathbf{P}}_{d,s}(t) + \tilde{\mathbf{P}}_{y,s}(t) + \mathbf{A}_1 \mathbf{Z}(t-1) + \mathbf{A}_2 \mathbf{Z}(t-2) + \mathbf{A}_3 \mathbf{Z}(t-3)$$

模型估計的誤差向量為：

$$\boldsymbol{\varepsilon}(t) = \mathbf{T}(t) - \hat{\mathbf{T}}(t)$$

最佳化的目標，是使估計誤差變異數的總和為最小：

$$\min \sum_t \boldsymbol{\varepsilon}^T(t) \boldsymbol{\varepsilon}(t)$$

- A. 請導出 \mathbf{A}_1 、 \mathbf{A}_2 、 \mathbf{A}_3 矩陣中各元素的數值，判斷1至3項是否顯著？（3分）
- B. 若擬利用上述判定顯著的模型，產生更多數量的19站聯合數據，請寫出產生19站聯合數據的表示式。（1分）