第二章 研究資料與研究方法

2.1 研究資料：

本論文所採用的研究資料為法文縮寫之 AVISO（Archivage，Validation et Interpretation des donnees des Satellites Oceanographiques)的衛星測高海表面高度距平(Sea Surface Height Anomaly，簡稱 SSHA)，以及做過衛星測高資料同化處理的三維海洋環流數值模式(Wu et al.，1999)所輸出的海表面高度距平，兩種資料的空間範圍皆選定為 0˚N～25˚N 以及 99˚E～124˚E，時間的範圍則是由 1993 年 1 月 1 日到 2002 年 12 月 31 日。兩種資料不同的地方在於，AVISO 的資料由於受到衛星軌道繞行地球的影響，所以七天才有一筆資料，而本研究在時間的範圍內所選取的第一筆資料為 1993 年 1 月 6 日，最後一筆資料則是 2002 年 12 月 25 日；至於模式所輸出的資料則沒有此一限制，所以是採用一天輸出一筆的資料，因此，由 1993 年 1 月 1 日至 2002 年 12 月 31 日止，每天都有一筆資料。

2.1.1 AVISO衛星遙測資料

AVISO 整合的衛星測高資料主要是來自於 Topex/Poseidon、Jason-1、ERS-1、ERS-2、EnviSat 以及 Doris 等衛星的實測資料，因此AVISO就相當於一個衛星的整合資料庫，其海表面測高資料可以說是將上述的每個衛星的資料整合在一起，並且互相比對進行修正，所以在資料的精確度上，會比其他單一衛星所觀測到的海表面測高值來得準確些。另外，這些海表面高度值的資料會再經過減去長時間海表面高度平均值的處理，變成海表面高度距平。這些經過AVISO所整合處理的資料現在也廣泛地被使用在海表面高度的監測、全球暖化研究、聖嬰現象(El Niño)和反聖嬰現象(La Niña)事件的追蹤探討、氣候預測、海洋環流、風、潮汐、波浪、海洋氣象模式等範疇之中。

2.1.2 三維海洋環流數值模式資料

本論文所使用的三維海洋環流數值模式為做過衛星測高資料同化處理的一個模式(Wu et al.，1999)，這樣所輸出的模式結果會比沒有做過資料同化的三維海洋數值模式更接近於實際現場所觀測到的情況，我們利用這個模式輸出的一天一筆海表面高度值，減去十年

（1993～2002）的海表面高度平均值後所得到的海表面高度距平，進一步研究南海海表面高度及流場的變化。

2.2 研究方法：

本論文所使用的研究方法為經驗正交函數(Empirical Orthogonal Function，簡稱 EOF）分析法，在大氣的應用上又稱為主要成份分析法（Principle Component Analysis，簡稱 PCA)。使用經驗正交函數分析法的目的，即在於找出同一組資料在「空間」以及「時間」上具有代表性的統計量，來解釋這組資料的組合。因此，本論文主要是以經驗正交函數分析法來分析 AVISO 的衛星測高海表面高度距平和模式的海表面高度距平，藉以找出南海海表面高度場在空間與時間上的某種合成，以突顯出支配南海海盆海表面高度升降的主要因素。經驗正交函數（EOF）分析的方法說明如下：

首先，令 Z 為一個大小為 n × m 的矩陣：

$$\mathbf{Z} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{X}_{13} & \Lambda & \mathbf{X}_{1m} \\ \mathbf{X}_{21} & \mathrm{O} & & & \mathrm{M} \\ \mathbf{X}_{31} & & \mathrm{O} & & \mathrm{M} \\ \mathrm{M} & & & \mathrm{O} & \mathrm{M} \\ \mathbf{X}_{n1} & \Lambda & \Lambda & \Lambda & \mathbf{X}_{nm} \end{bmatrix},$$

其中，n 為每一個時間序列的長度，而 m 則是時間序列的總數，因此 X_{nm} 即表示在第 m 個測站的第 n 個時間點的資料(在本論文中即為海表面高度)，也就是說，Z 為在 m 個測站裡，每個測站都有 n 個時間點的資料所構成的行向量（column vector）所組成的一個矩陣。而主要成份的獲得，則是從方陣（square matrix）$\mathbf{S}=\mathbf{ZZ}'$ 裡找出特徵向量（eigenvectors），此特徵向量主要是以複數共軛轉置（complex conjugate transpose）的方式呈現。接著，再令 Λ 為一個對角線由方陣 S 的特徵值(eigenvalues)，對角線之外的每個元素皆為 0 所共同組成的矩陣。

即

$$\mathbf{\Lambda} = \begin{bmatrix} \boldsymbol{\lambda}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda}_2 & & & \mathbf{0} \\ \mathbf{0} & & \ddots & & \vdots \\ \vdots & & & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\lambda}_n \end{bmatrix},$$

其中，λ_1～λ_n 即為矩陣 S 的特徵值，且 $\lambda_1 > \lambda_2 > ... > \lambda_n$。

而這些特徵值是由 $Se=e\lambda$ 的方程式所解出，這裡的 e 包含了特徵向量。

即

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_{11} & \mathbf{e}_{12} & \mathbf{e}_{13} & \Lambda & \mathbf{e}_{1n} \\ \mathbf{e}_{21} & \mathrm{O} & & & \mathrm{M} \\ \mathbf{e}_{31} & & \mathrm{O} & & \mathrm{M} \\ \mathrm{M} & & & \mathrm{O} & \mathrm{M} \\ \mathbf{e}_{n1} & \Lambda & \Lambda & \Lambda & \mathbf{e}_{nn} \end{bmatrix},$$

其中，e_{nn} ，左邊的 n 代表測站的時間點，右邊的 n 代表模式的編號，因此每一個直行(column)代表一個模式的特徵向量，每一個特徵向量解釋了每個模式在時間尺度上的變化，彼此互相獨立而且已經標準化為單位長度，並且是以「解釋變異度」(explained variance) 由大到小來排列。由於 S 是一個 n 階的方陣，所以我們可以找出 n 組的 e 和 λ 的解，使其滿足 Se=eλ 的方程式。至於在空間獨立的分佈上，是以 $\mathbf{X} = \mathbf{e}'\mathbf{z}$ 裡的列向量 (row vectors) 來表示，這裡的 X 是一個大小為 n × m 的矩陣。既然 e 裡的每一行向量都已經過標準化，那不同模式的相對重要性，也就反應在空間分佈的圖上。