

伴隨理論在四維變分及其在 氣象中之應用

曾建翰¹ 張忍成¹ 林 和²

¹中央氣象局科技中心

²國立臺灣大學大氣科學研究所

(中華民國八十五年十一月十八日收稿；中華民國八十六年九月五日定稿)

摘 要

伴隨理論(adjoint method)已被廣泛用來解決四維時空下氣象的最佳化問題，特別是在四維資料同化(4DDA)的研究領域上。其基本觀念在於，原控制方程組的微分算子和其伴隨算子，在事先所訂的一內積空間內完成交換。本研究的目的即在：利用一淺水模式和由其導出的伴隨模式，描繪出一簡單四維變分(4DVAR)的過程。一般說來有兩種方式可用以建構伴隨模式。其一，是利用解析方法從淺水方程中導出；另外一種，則是直接利用原淺水模式的正切線性模式(Tangent Linear Model)的數值差分格式中導出。

本文用一個 Rossby-Hauwitz 模波做控制組初始場；另外在此重力位高度場上，於任意地區疊加一相較於約原振幅百分之十的擾動，做為實驗組初始場，完成一所謂比對實驗(twin-experiment)。所求的成本函數(cost function)定義為這兩組初始場在積分 36 小時後差值的平方。如果成本函數中的組成變數沒有做尺度分析時，整個計算調整過程將受限於得出最終的重力位場，並指出造成此質量場的源頭是初始的風場變化。但是，當取用一個更合理的尺度分析時(如 $uscale = 1$, $vscale = 1$, $\Phi scale = 5 \times 10^{-5}$)，整個的調整過程將遍佈所有變數而得出一更適切的初始風場。另一值得討論的是：由解析法所得出的伴隨模式需要仔細小心地處理邊界條件，這在實際作業的運算中較為困難。另一方面來說，數值直接導出法不僅不用考慮邊界條件的設定，對於伴隨模式的建構和維護也較容易。

關鍵詞：伴隨模式、正切線性模式、比對實驗、成本函數

一、前 言

伴隨理論(adjoint method)在氣象上的應用最早可分為兩支(Daley,1991)，分別以 Marchuck(1974)和 Lions(1971)為代表。Marchuck 利用此法來探討氣候模式的敏感度(sensitivity)問題；而 Lions 則認為不但能研究敏感度問題外，對一個微分系統而言，應該可以透過伴隨方法求算出最佳初始值。現今，這兩支理論已合而為一，其在數學分析中的推導僅有先後之別；換言之，在求得敏感度後，接下來的挑戰就是如何修正初始場而得到一個最佳的結果，這在數學上是屬於變分學所探討的範圍。探討敏感度問題的代表文章有 Cacuci(1981)、Hall et al.(1982)、Errico and Vukicevic(1992)等；而探討初始值問題的代表則有 Farrel(1982)、Branstator(1985a,b)、Borges and Hartmann(1992)和 Borges and Sardeshmukh(1995)等。結合兩種探討問題的方式並利用變分法求解的有 LeDimet and Talagrand(1986)、Talagrand and Courtier(1987)和 Zou et al.(1992)。

早期的伴隨方程都要靠解析運算，而後再轉成數值模式求解。這個過程相當繁瑣，加上氣象變數較一般問題為多，方程組亦有相當的複雜性，故所遭遇的問題是如何應用至氣象中？往往需要簡化方程組來配合探討的問題。但隨著電腦硬體的進步，原先認為的龐大計算至今已不是問題；加上數值方法的進步，像早期的氣象學家(如 Kontarev(1980)、LeDimet and Talagrand(1986)、Talagrand and Courtier(1987)等)所使用的解析運算方法，如今可被 Talagrand(1991)所提出的方法取代，即由正切線性模式(tangent linear model)直接導出伴隨模式，這個方法我們姑且稱之為數值直接導出法，其對伴隨算子有新的認知，伴隨模式的發展遠較從前容易。隨後的 Errico and Vukicevic(1992)和 Zou et al.(1993)等所使用的方法，都已經是用此數值導出法求算伴隨模式。

在許多氣象問題中，常需要求出一個泛函的極值。所謂泛函即是由函數所組成的函數，例如： $\mathbf{x} \equiv (\mathbf{u}, \mathbf{v}, \Phi)^T$ ， $\hat{\mathbf{x}} \equiv (\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\Phi})^T$ ，上標 T 表矩陣向量的轉置，欲求 $J \equiv (\mathbf{x} - \hat{\mathbf{x}})^2 / 2$ 的極值，此處 J 即為泛函。數學上以變分法來求泛函的極值，而伴隨理論則為變分法中求泛函極值的標準計算步驟之一。

伴隨理論的精神在於利用算子交換的特性，而為了要方便完成算子的交換，伴隨理論在線性理論的架構下探討問題，由原線性方程有關的線性算子，經過算子交換過程後，得到作用在另一變數上的伴隨算子，其目的在規避對原方程變數做繁瑣的直接計算而求得解答的一種方法。回溯數學史，伴隨算子分別在微分方程和線性代數中有相似的定義和不同的應用。在微分方程中的伴隨算子，應用於滿足某特定邊界、初始條件而無法求得通解之情形下，找出特解的計算方法，常配合格林函數(Green's function)使用。在求泛函極值問題時，伴隨變數即為拉氏乘數(Lagrange multiplier)，其和極值的敏感度有關。而線性代數中的伴隨算子用於坐標轉換的討論，求得在非正交坐標上某一向量在各基底上的最佳投影。上述之

求解的過程將留在第二節中說明。

在變分最佳化分析中，經由伴隨算子計算所得之伴隨變數，即為四維空間變數(三維空間加上一維時間)的梯度，可用來搜尋泛函的極值。而在許多的物理問題中大都是有限制性條件的最佳化過程，也就是在所求的成本函數(cost function)—泛函中，必須滿足附帶的條件；像是考慮雨量的問題，是決不可能求當負雨量發生時某物理量的極值，故這一問題必需加上雨量不能為負之限制條件。但隨著數值方法的進步，求取無限制的變分最佳化問題(unconstrained optimization problems)較為容易，故限制性最佳化問題(constrained optimization problems)常轉換成無限制的變分問題求解，而實際上數值法求算限制性最佳化問題時是用一組無限制的變分問題替代，所以從無限制變分的問題出發是一個很好的嘗試，故本文將焦點放在用數值法求算無限制最佳化的問題上。

本文第三節中，我們嘗試用一淺水方程來推導傳統解析法所得出之伴隨模式和數值導出法所得出的伴隨模式。數值導出法伴隨模式的好處，在於填補線性代數和微分方程中伴隨算子恆等式的差異；此外，其所推導出之伴隨模式和原方程一階線性化後的正切線性方程有很好的相依存關係，於數值模式上的維護和發展較為容易。於第四節中，本文將討論解析和數值兩法的異同，並用一共軛梯度法(conjugate gradient method)來求取變分極值，進而應用至無限制最佳化的問題。第五節為結論和對運用伴隨理論時的建議。

二、伴隨理論和無限制最佳化

(一)線性代數中的伴隨算子

伴隨算子恆等式在線性代數中的定義如下：

$$\langle Lx, y \rangle = \langle x, L^*y \rangle \quad (1)$$

其中 L 是線性算子， x 和 y 分別為行向量， \langle, \rangle 表內積。(1)式的意義為 x 的行向量在經過 L 的運算後和 y 向量的內積，等於 x 向量和 L^* 對 y 運算後之內積(可參看圖 1)。換言之，設 $x, L^*y, \langle x, L^*y \rangle$ 在 A 空間內，而 x 經過 L 的運算映射到 B 空間；假設 B 空間存在一個 y ，可經過 L^* 的運算回到 A 空間，則在 A 空間上的 $\langle x, L^*y \rangle$ 內積值恆等於 B 空間之 $\langle Lx, y \rangle$ 內積值。此處， L^* 為 L 的伴隨算子，在複數空間中，上標 $*$ 表矩陣的取共軛轉置，對實數空間而言僅表示矩陣的轉置。上述恆等式可經由基本的線性代數運算證明成立，但其最重要的精神是：當在線性算子控制下的兩個變數，在做空間轉換時，在各自空間內所取之內積為一保守量，用以掌握所要關切之物理性質。此外， x 和 y 可形成一完整正交集，當任一向量分解成由 x 組成時，可經由 y 來決定在各 x 分量上之最佳投影(Branstator, 1985b、Zhang, 1988 和 Borges and Hartmann, 1992)。在這裏所探討的空間轉

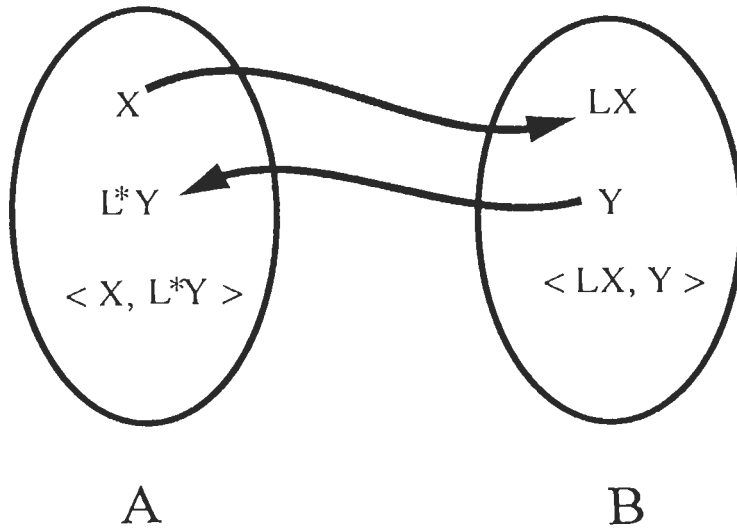


圖 1 伴隨算子運算示意圖，詳細說明請參看內文。

換的觀念好像很抽象，但任何一種求解過程都可視為由一運算子和其變數所組成的函數，經由運算子作用的過程皆可想成是空間的轉換，所以只要能將運算子線性化，都能得到伴隨算子而加以利用。

(二)格林函數

上小節所述之算子互相交換位置的手法亦見於求解微分方程時，而格林函數問題為其中之代表，其解法過程如下：

設有一微分方程如下，

$$Lx = f \tag{2}$$

其中 L 是微分算子，x 為變數所組成的行向量，f 是方程中的非均質項 (inhomogeneous term) 或是稱做強制項 (forcing term)。現設有一格林函數 G 存在，取其和 Lx 在運算空間的內積可寫成，

$$\int G Lx \, d\Omega \tag{3}$$

而運算空間的內積 $\int (...) \, d\Omega$ 最簡單的定義為微分方程所要處理問題之定義域或物理空間 (即 Ω) 的範圍，在氣象的問題中可為全球球面的範圍或一特定區域。現將 (3) 式做部分積分，得到

$$\int G Lx \, d\Omega = Gx|_{B.C} + \int L^*G x \, d\Omega \tag{4}$$

其中 B.C 表邊界條件。如果存在一 δ 函數，令 $L^*G = \delta$ (即 G 可由 $L^{-1}\delta$ 求得)，則

$$\int L^* G \mathbf{x} d\Omega = \int \delta \cdot \mathbf{x} d\Omega = \mathbf{x}$$

又假設 $G\mathbf{x}|_{BC}$ 為一常數值的話，為簡化問題起見可設為零，如此可得

$$\mathbf{x} = \int G L \mathbf{x} d\Omega = \int G f d\Omega$$

可發現上式為原微分方程的一個特解，也就是一微分方程的格林函數乘上其非均值項，對處理問題的區域做積分，即為原微分方程的一特解。這對處理當微分方程因某些特定邊界條件下，所產生的通解退化(trivial solution)現象很有幫助，處理當物理問題中，如無法改變邊界、初始條件的設定，而希望求得含有非均值項時滿足方程的特解。而微分方程中將

$$\int y L \mathbf{x} d\Omega = y \mathbf{x}|_{BC} + \int L^* y \mathbf{x} d\Omega \quad (5)$$

稱為伴隨算子恆等式，在(4)式中的格林函數 G 即為此處之 y ，為微分算子中的伴隨變數，和(1)式比較發現，其二者差異在 $y \mathbf{x}|_{BC}$ 這項。

(三)變分最佳化分析的伴隨方程

如果存在一非線性方程如下，

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{n}(\mathbf{x}) = \mathbf{F} \quad (6)$$

其中 \mathbf{n} 是非線性算子， \mathbf{x} 為變數所組成的行向量。為了要使交換算子的過程順利進行，常將所求之方程做線性化處理。設 $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{x}'$ ，其中 \mathbf{x}' 為擾動場，而 $\bar{\mathbf{x}}$ 為一參考變數：可為不穩定度分析時之基本場，或是四維資料同化時之觀測值。於此將 $\mathbf{n}(\mathbf{x})$ 做 Taylor 級數展開，即

$$\mathbf{n}(\mathbf{x}) = \mathbf{n}(\bar{\mathbf{x}}) + \frac{\partial \mathbf{n}(\bar{\mathbf{x}})}{\partial \mathbf{x}} \mathbf{x}' + \mathbf{O}(\mathbf{x}'^2)$$

忽略掉高次項，則(6)式可變為，

$$\frac{\partial \bar{\mathbf{x}}}{\partial t} + \frac{\partial \mathbf{x}'}{\partial t} = \mathbf{n}(\bar{\mathbf{x}}) + \frac{\partial \mathbf{n}(\bar{\mathbf{x}})}{\partial \mathbf{x}} \mathbf{x}' = \bar{\mathbf{F}} + \mathbf{f}$$

又 $\frac{\partial \bar{\mathbf{x}}}{\partial t} = \mathbf{n}(\bar{\mathbf{x}}) = \bar{\mathbf{F}}$ ，則此時

$$\frac{\partial \mathbf{x}'}{\partial t} = \frac{\partial \mathbf{n}(\bar{\mathbf{x}})}{\partial \mathbf{x}} \mathbf{x}' = \mathbf{f} \quad (7a)$$

(7)式對 \mathbf{x}' 言為一線性方程。此方程又稱為正切線性方程(tangent linear model)和平常所認知的線性方程之不同處在於 $\frac{\partial \mathbf{n}(\bar{\mathbf{x}})}{\partial \mathbf{x}}$ 這項來自(6)式的運算，為一時間的函數。上式可用線

性算子 \mathbf{L} 來取代，簡寫成：

$$\frac{\partial \mathbf{x}'}{\partial t} \equiv \dot{\mathbf{x}}' = \mathbf{L}\mathbf{x}' = \mathbf{f} \quad (7b)$$

在變分最佳化分析中，所欲求極值的泛函數稱為成本函數，現今，假設所求的成本函數 J 定義成：

$$J \equiv \frac{1}{2} \int_{t_0}^{t_f} \mathbf{x}'^2 dt$$

其中 t_0 表初始時間， t_f 表積分終了時間。其義為求 \mathbf{x} 和 $\bar{\mathbf{x}}$ 在時間演化軌跡上的相異程度，而求算 J 之極小或極大值是一體兩面的事，解法過程類似，故本文下述的探討和實驗都以求極小值為例。

成本函數的選取和定義有很大的彈性。以圖 2 為例：對圖 2a 而言，所關心的問題僅在積分終點時之物理狀態，其中的粗黑點表某一物理量成本函數的最終狀態，而僅求這一特定時間的極值。圖 2b 則是對整個積分演化軌跡有興趣，圖中實線表一實驗組物理量隨時間的演化軌跡，而虛線表一控制組的演化軌跡，所欲求之成本函數定義為這兩條曲線所夾的斜線區面積，可計算當這個面積達到最小時，物理量隨時間的演化軌跡會是什麼狀況？圖 2c 則是對積分終點的現象取權重，圖中黑線表實驗的演化軌跡，而在積分終點時加入一強制因素

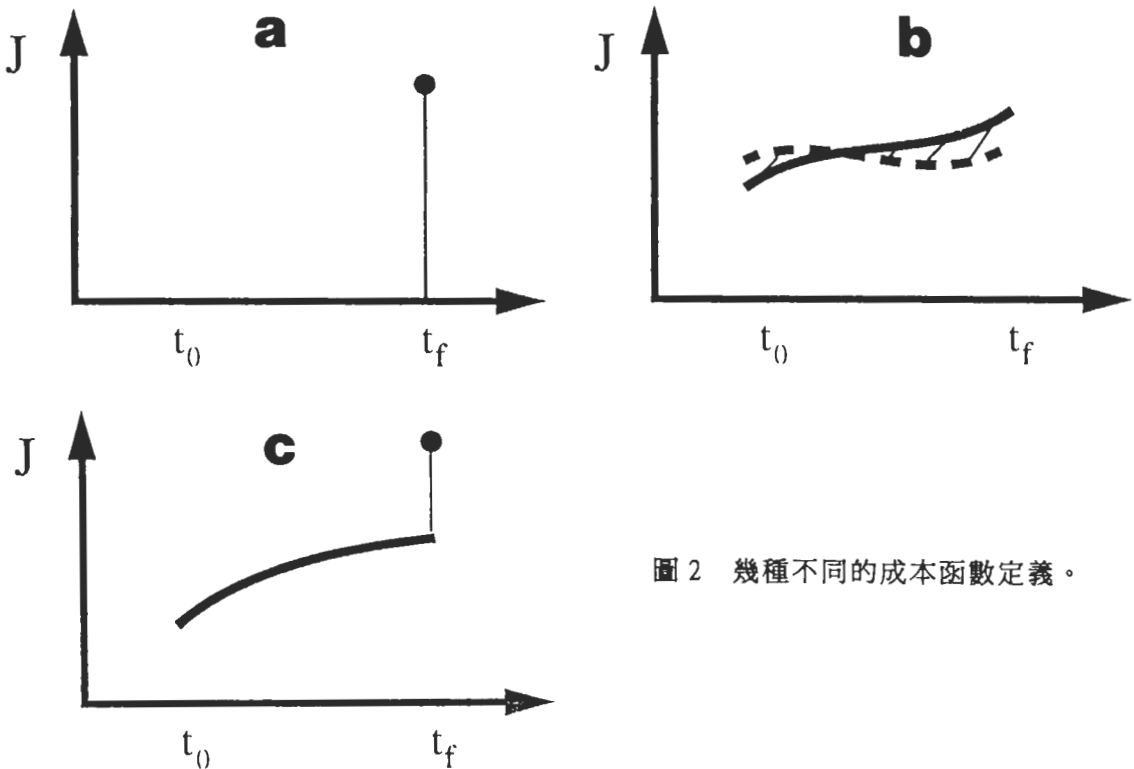


圖 2 幾種不同的成本函數定義。

(由圖中粗黑點所示)，所求之成本函數不單是考量終點的加權因素，同時也考量整個積分演化軌跡的變化。

考慮圖 2a 的簡單狀況，則成本函數可定義如下：

$$J = \frac{1}{2} \mathbf{x}'^2(t_f) + \int_{t_0}^{t_f} \mathbf{x}'_a{}^T (\mathbf{f} - \dot{\mathbf{x}}') dt \quad (8)$$

其中， $\int_{t_0}^{t_f} \mathbf{x}'_a{}^T (\mathbf{f} - \dot{\mathbf{x}}') dt$ 為此問題之限制條件，其義為在被控制方程組所規範之條件下求 J 之極值；而 \mathbf{x}'_a 為拉氏乘數，或稱為伴隨變數 (adjoint variables)，上標 T 表矩陣的轉置。此種型式，亦稱為有限制條件之最佳化問題 (Arora, 1989)。將上式右手邊積分項中之 $\mathbf{x}'_a{}^T \dot{\mathbf{x}}'$ 取部分積分可得

$$J = \frac{1}{2} \mathbf{x}'(t_f)^2 - \mathbf{x}'_a{}^T \mathbf{x}' \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} (\mathbf{x}'_a{}^T \mathbf{f} + \dot{\mathbf{x}}'_a{}^T \mathbf{x}') dt$$

而 J 的微量變化 δJ ，相對於 $\delta \mathbf{x}'$ 的改變，基本上可由 J 對 \mathbf{x}' 的一階導數來近似之，即

$$\delta J = \frac{\partial J}{\partial \mathbf{x}'} \delta \mathbf{x}'$$

代入(8)式，則成爲

$$\begin{aligned} \delta J &= \mathbf{x}'^T(t_f) \delta \mathbf{x}' - \mathbf{x}'_a{}^T \delta \mathbf{x}' \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left(\mathbf{x}'_a{}^T \frac{\partial \mathbf{f}}{\partial \mathbf{x}'} + \dot{\mathbf{x}}'_a{}^T \right) \delta \mathbf{x}' dt \\ &= [\mathbf{x}'^T(t_f) - \mathbf{x}'_a{}^T(t_f)] \delta \mathbf{x}' + \mathbf{x}'_a{}^T(t_0) \delta \mathbf{x}' + \int_{t_0}^{t_f} \left(\mathbf{x}'_a{}^T \frac{\partial \mathbf{f}}{\partial \mathbf{x}'} + \dot{\mathbf{x}}'_a{}^T \right) \delta \mathbf{x}' dt \end{aligned} \quad (9)$$

當有 J 之極值存在時， $\delta J = 0$ 。

假設 $\mathbf{x}'_a{}^T(t_0) \delta \mathbf{x}' \rightarrow 0$ ，則 $\delta J = 0$ 的必要條件爲

$$\mathbf{x}'_a(t_f) = \mathbf{x}'(t_f) \quad (10a)$$

$$\dot{\mathbf{x}}'_a - \mathbf{L}^* \mathbf{x}'_a = 0 \quad (10b)$$

此處， $\mathbf{L}^* = -\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}'}\right)^T$

(10b,c)式稱為伴隨方程組，或 Euler-Lagrange 方程組。由(9)式可看出 \mathbf{x}'_a 爲對 J 極值敏感的變數，於此，可將(10a,b)代入(9)中得到，

$$\frac{\delta J}{\delta \mathbf{x}'} = \mathbf{x}'_a(t_0)$$

，意即以(10a)式當做初始條件解(10b)式，由伴隨模式所解出的 $\mathbf{x}'_a(t_0)$ 可代表成本函數的敏感度，所解出之 \mathbf{x}'_a 代回(8)式可得 J 的極值，這也說明了為何伴隨理論可用來計算敏感度問

題。

應用數值直接導出法上的變分解法和上述者略有不同。首先，除保有算子交換的特性外，其求出之伴隨算子能滿足線性代數中的伴隨恆等式(即(1)式)，不會有邊界條件設定的限制。如此，可方便數值運算之進行，而不需直接求解方程組(9,10 式)來得到極值。數值法變分最佳化分析中，最簡單的是處理無限制條件之最佳化問題。換言之，即將(8)式中被控制方程組所限制的條件拿掉，將其隱含至求解所定義的內積中。而從解析法求有限制最佳化問題過渡到無限制最佳化的問題之理由是：

1. 變數或是限制條件太多，運用解析法計算不易。
2. 所考慮的問題屬高度非線性(highly nonlinear)者，於計算成本函數中控制方程組算子的部分積分過於複雜，解析方法難於處理。
3. 定義成本函數時將有困難，即變數和變數或變數和參數間之依存關係非常複雜，一階線性化後無法將其彼此分離，造成計算上的困難。
4. 限制條件最佳化問題可轉換成一組用數值法計算的無限制條件之最佳化問題。
5. 雖然(8)式在某些變分學的教科書中(Hildebrand,1965)視為一新的無限制條件方程，但在數值法求解變分極值的過程中和拉氏乘數相伴的這一項，仍是一個限制條件，須要再加以特別處理。這和本文以下所要探討的無限制主題有所不同。
6. 用解析法導出伴隨模式較數值直接導出法繁瑣，如以熟知的差分格式來近似伴隨微分方程，則在編輯程式時之伴隨模式和原控制模式間將無相依存關係，除編寫和維護程式不易外，其結果亦無法保證滿足伴隨算子恆等式。

上述理由中，1~4 點之詳細說明可參閱 Arora(1989)書中之討論，第 6 點則是近來氣象學者如 Talagrand(1991)，Errico(1992)，Navon et al.(1992)，Zou et al.(1993)等人所積極提倡。運用數值直接導出法求伴隨模式之推導過程簡述如下：

首先，定義廣義的成本函數為：

$$J[\mathbf{x}(t_0)] = \frac{1}{2} \sum_{t=t_0}^{t_f} [\mathbf{x}(t) - \bar{\mathbf{x}}(t)]^T \times \mathbf{W}(t) \times [\mathbf{x}(t) - \bar{\mathbf{x}}(t)] \quad (11)$$

其中 \mathbf{W} 為權重函數，在第四節中將詳細討論其特性。請特別注意此式中，沒有由控制方程組所構成的限制條件(和(8)式比較)；而用 Σ 取代 \int 之意為在數值積分中時間間隔是離散而非連續的。另外，我們亦認為成本函數可為初始變數 $\mathbf{x}(t_0)$ 所組成的函數，於是將(11)式左手邊項對 $\mathbf{x}'(t_0)$ 取一階導數，可得

$$\mathbf{J}'[\mathbf{x}(t_0)] = \{\nabla J[\mathbf{x}(t_0)]\}^T \mathbf{x}'(t_0) \quad (12)$$

而(11)式右手邊項對取 $\mathbf{x}'(t)$ 微分，

$$J'[\mathbf{x}(t_0)] = \sum_{t=t_0}^{t_f} \{ \mathbf{W}(t) [\mathbf{x}(t) - \bar{\mathbf{x}}(t)] \}^T \mathbf{x}'(t) \quad (13)$$

由(7b)，在 t 時的 \mathbf{x}' 向量變數可寫成線性傳遞算子形式，

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{P}_1(t - \Delta t) \mathbf{x}'(t - \Delta t) \\ &= \mathbf{P}_1(t - \Delta t) \mathbf{P}_2(t - 2\Delta t) \mathbf{x}'(t - 2\Delta t) \\ &\quad \cdot \\ &\quad \cdot \\ &\quad \cdot \\ &= \mathbf{P}(t) \mathbf{x}'(t_0) \end{aligned} \quad (14)$$

其中，

$$\begin{aligned} t &= t_0 + n\Delta t \\ \mathbf{P}_i(t - i\Delta t) &\equiv [\mathbf{I} + \Delta t \mathbf{L}(t - i\Delta t)], \quad i = 1, 2, \dots, n \end{aligned}$$

而

$$\mathbf{P}(t) \equiv \prod_{i=1}^n \mathbf{P}_i$$

將(14)式代回(13)式，並和(12)式比較，可得：

$$\{ \nabla J[\mathbf{x}(t_0)] \}^T \mathbf{x}'(t_0) = \sum_{t=t_0}^{t_f} \{ \mathbf{W}(t) [\mathbf{x}(t) - \bar{\mathbf{x}}(t)] \}^T \mathbf{P}(t) \mathbf{x}'(t_0)$$

其中暗示，

$$\nabla J[\mathbf{x}(t_0)] = \sum_{t=t_0}^{t_f} \mathbf{P}^T(t) \mathbf{W}(t) [\mathbf{x}(t) - \bar{\mathbf{x}}(t)] \quad (15)$$

如果定義伴隨方程組如下：

$$\mathbf{x}'_n(t) = \mathbf{P}^T(t) \mathbf{x}'_n(t) \quad (16)$$

且 $\mathbf{x}'_n(t) = \mathbf{W}(t) [\mathbf{x}(t) - \bar{\mathbf{x}}(t)]$ 則可得： $\mathbf{x}'_n(t_0) = \mathbf{P}^T(t) \mathbf{W}(t) [\mathbf{x}(t) - \bar{\mathbf{x}}(t)]$

代回(15)式，知

$$\nabla J[\mathbf{x}(t_0)] = \sum_{t=t_0}^{t_f} \mathbf{x}'_n(t_0) = \sum_{t=t_0}^{t_f} \mathbf{P}^T(t) \mathbf{W}(t) [\mathbf{x}(t) - \bar{\mathbf{x}}(t)]$$

由此可看出成本函數在四維空間上用初始變數場所表現的梯度，可用一組伴隨變數值求得。

由(14)，(16)式可輕易驗證出：

$$\begin{aligned} \langle \mathbf{x}'_s(t), \mathbf{x}'(t) \rangle &= \langle \mathbf{x}'_s(t), \mathbf{P}(t) \mathbf{x}'(t_0) \rangle \\ &= \langle \mathbf{P}^T(t) \mathbf{x}'_s(t), \mathbf{x}'(t_0) \rangle = \langle \mathbf{x}'_s(t_0), \mathbf{x}'(t_0) \rangle \end{aligned}$$

滿足伴隨算子恆等式。

舉圖 2a 的情況為例，則成本函數(即(11)式)可簡化成

$$\begin{aligned} J[\mathbf{x}(t_0)] &= \frac{1}{2} \sum_{t=t_f}^{t_f} [\mathbf{x}(t) - \bar{\mathbf{x}}(t)]^T \times \mathbf{W}(t) \times [\mathbf{x}(t) - \bar{\mathbf{x}}(t)] \\ &= \frac{1}{2} [\mathbf{x}(t_f) - \bar{\mathbf{x}}(t_f)]^T \times \mathbf{W}(t_f) \times [\mathbf{x}(t_f) - \bar{\mathbf{x}}(t_f)] \end{aligned} \quad (17)$$

且其初始敏感度(即(15)式)為

$$\nabla J[\mathbf{x}(t_0)] = \sum_{t=t_f}^{t_f} \mathbf{x}'_s(t) = \mathbf{P}^T(t_f) \times \mathbf{W}(t_f) \times [\mathbf{x}(t_f) - \bar{\mathbf{x}}(t_f)] = \mathbf{x}'_s(t_0)$$

在下一節中，將採用一個淺水方程為例子，分別以解析法和數值直接導出法求得伴隨模式。總結來說，採用數值直接導出法求得之伴隨算子，也就是式中的 \mathbf{P}^T ，是按照數值方法排出。其精確的表達出 \mathbf{P} 和 \mathbf{P}^T 間互為轉置的關係，這將保證(1)式之伴隨算子恆等式成立，將可利用線性代數中計算的便利性來探索物理現象，不會有微分算子中之伴隨算子恆等式尚需考量邊界條件的困擾。此外，將限制條件隱含至成本函數中，使變分的問題成為無限制條件的問題，這在求變分極值時有很大的便利性，同時也為往後以數值方法求限制條件最佳化問題的基礎。

三、淺水伴隨模式

本節嘗試用一個淺水方程來討論伴隨理論，分別用解析和數值直接導出兩種方法來建立淺水伴隨模式。所利用之淺水方程如下：

$$\begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - f v + \frac{\partial \Phi}{\partial x} &= 0 \\ \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + f u + \frac{\partial \Phi}{\partial y} &= 0 \\ \frac{\partial \Phi}{\partial t} + u \frac{\partial \Phi}{\partial x} + v \frac{\partial \Phi}{\partial y} + \Phi \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) &= 0 \end{aligned}$$

模式採用 C 網格式差分方程，時間上為調整性的跳蛙積分格式(filtered leap frog time

integration scheme), 水平東西向範為 45000 公里, 南北向 20000 公里, 水平網格點數 96×96 點, 時間間距 600 秒, 特徵高度 2000 公尺 (按照 Zou et al., 1992)。其淺水模式的正切線性方程推導如下,

若將原淺水方程表成如(6)式,

$$\text{則 } \mathbf{x} = \begin{pmatrix} u \\ v \\ \Phi \end{pmatrix}, \quad \mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} + fv - \frac{\partial \Phi}{\partial x} \\ -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - fu - \frac{\partial \Phi}{\partial y} \\ -u \frac{\partial \Phi}{\partial x} - v \frac{\partial \Phi}{\partial y} - \Phi \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \end{pmatrix}$$

設 $u = \bar{u} + u'$, $v = \bar{v} + v'$, $\Phi = \bar{\Phi} + \Phi'$, 利用(7)式, 則得到淺水方程之正切線性方程為:

$$\frac{\partial}{\partial t} \begin{pmatrix} u' \\ v' \\ \Phi' \end{pmatrix} - \begin{pmatrix} \frac{\partial n_1}{\partial u} & \frac{\partial n_1}{\partial v} & \frac{\partial n_1}{\partial \Phi} \\ \frac{\partial n_2}{\partial u} & \frac{\partial n_2}{\partial v} & \frac{\partial n_2}{\partial \Phi} \\ \frac{\partial n_3}{\partial u} & \frac{\partial n_3}{\partial v} & \frac{\partial n_3}{\partial \Phi} \end{pmatrix} \begin{pmatrix} u' \\ v' \\ \Phi' \end{pmatrix} = 0$$

或寫為

$$\frac{\partial}{\partial t} \begin{pmatrix} u' \\ v' \\ \Phi' \end{pmatrix} + \begin{pmatrix} \bar{u} \frac{\partial}{\partial x} + \frac{\partial \bar{u}}{\partial x} + \bar{v} \frac{\partial}{\partial y}, & \frac{\partial \bar{u}}{\partial y} - f, & \frac{\partial}{\partial x} \\ \frac{\partial \bar{v}}{\partial x} + f, & \bar{u} \frac{\partial}{\partial x} + \bar{v} \frac{\partial}{\partial y} + \frac{\partial \bar{v}}{\partial y}, & \frac{\partial}{\partial y} \\ \frac{\partial \bar{\Phi}}{\partial x} + \bar{\Phi} \frac{\partial}{\partial x}, & \frac{\partial \bar{\Phi}}{\partial y} + \bar{\Phi} \frac{\partial}{\partial y}, & \bar{u} \frac{\partial}{\partial x} + \frac{\partial \bar{u}}{\partial x} + \bar{v} \frac{\partial}{\partial y} + \frac{\partial \bar{v}}{\partial y} \end{pmatrix} \begin{pmatrix} u' \\ v' \\ \Phi' \end{pmatrix} = 0$$

(一) 解析淺水伴隨模式

將上式左手邊第二項微分算子寫成矩陣 A, 則

$$\mathbf{A} \equiv \begin{pmatrix} \bar{u} \frac{\partial}{\partial x} + \frac{\partial \bar{u}}{\partial x} + \bar{v} \frac{\partial}{\partial y}, & \frac{\partial \bar{u}}{\partial y} - f, & \frac{\partial}{\partial x} \\ \frac{\partial \bar{v}}{\partial x} + f, & \bar{u} \frac{\partial}{\partial x} + \bar{v} \frac{\partial}{\partial y} + \frac{\partial \bar{v}}{\partial y}, & \frac{\partial}{\partial y} \\ \frac{\partial \bar{\Phi}}{\partial x} + \bar{\Phi} \frac{\partial}{\partial x}, & \frac{\partial \bar{\Phi}}{\partial y} + \bar{\Phi} \frac{\partial}{\partial y}, & \bar{u} \frac{\partial}{\partial x} + \frac{\partial \bar{u}}{\partial x} + \bar{v} \frac{\partial}{\partial y} + \frac{\partial \bar{v}}{\partial y} \end{pmatrix}$$

利用第二節中所討論之伴隨算子的解析推導過程, 即定義一內積範圍做部分積分和一些邊界

條件的設定(本例中定義一半球為內積空間，東西向取周期性邊界條件，南北向為固定邊界)，再將 A 取轉置即得其伴隨算子，並得出解析之淺水伴隨模式如下：

$$-\frac{\partial}{\partial t} \begin{pmatrix} u'_a \\ v'_a \\ \Phi'_a \end{pmatrix} + \begin{pmatrix} -\bar{u} \frac{\partial}{\partial x} - \bar{v} \frac{\partial}{\partial y} - \frac{\partial \bar{v}}{\partial y}, & \frac{\partial \bar{v}}{\partial x} + f, & -\bar{\Phi} \frac{\partial}{\partial x} \\ \frac{\partial \bar{u}}{\partial y} - f, & \frac{\partial}{\partial x} - \frac{\partial \bar{u}}{\partial x} - \bar{v} \frac{\partial}{\partial y}, & -\bar{\Phi} \frac{\partial}{\partial y} \\ -\frac{\partial}{\partial x}, & -\frac{\partial}{\partial y}, & -\bar{u} \frac{\partial}{\partial x} - \bar{v} \frac{\partial}{\partial y} \end{pmatrix} \begin{pmatrix} u'_a \\ v'_a \\ \Phi'_a \end{pmatrix} = 0$$

詳細推導過程請見如 Kontarev(1980)，LeDimet and Talagrand(1986)中所述。

(二) 數值淺水伴隨模式

如果能得出一組方程式線性化後的運算子和其變數所組成的行向量，理論上來說，將此微分算子取轉置即得伴隨算子(對實數空間而言)。此法由 Talagrand(1991)提出，今僅用線性化後之淺水方程中的一項舉例說明如下：

考慮 $\dot{u}' = \frac{\partial u'}{\partial t} = -\bar{u} \frac{\partial u'}{\partial x} - u' \frac{\partial \bar{u}}{\partial x}$ 這項，其寫成差分方程時成為，

$$\begin{pmatrix} \cdot \\ \cdot \\ \dot{u}'(i, j) \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \cdot \cdot \bar{u}(i, j) \frac{1}{2\Delta x}, & -\frac{\bar{u}(i+1, j) - \bar{u}(i-1, j)}{2\Delta x}, & -\bar{u}(i, j) \frac{1}{2\Delta x} \cdot \cdot \end{pmatrix} \begin{pmatrix} \cdot \\ \cdot \\ u'(i-1, j) \\ u'(i, j) \\ u'(i+1, j) \\ \cdot \\ \cdot \end{pmatrix}$$

$\mathbf{L} \qquad \mathbf{x}'$

此時 L 為一線性算子，而 \mathbf{x}' 為一行向量變數，當對 L 取轉置時，可得其伴隨算子 L^T ，和其對應之伴隨方程

$$\begin{pmatrix} \vdots \\ \dot{u}'_a(i-1, j) \\ \dot{u}'_a(i, j) \\ \dot{u}'_a(i+1, j) \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \vdots \\ \bar{u}(i, j) \frac{1}{2\Delta x}, \\ -\frac{\bar{u}(i+1, j) - \bar{u}(i-1, j)}{2\Delta x}, \\ -\bar{u}(i, j) \frac{1}{2\Delta x}, \\ \vdots \\ \vdots \end{pmatrix} \begin{pmatrix} \vdots \\ \vdots \\ u'_a(i-1, j) \\ u'_a(i, j) \\ u'_a(i+1, j) \\ \vdots \\ \vdots \end{pmatrix}$$

\mathbf{L}^T
 \mathbf{x}'_a

此時，可發現原先線性方程中僅對 $\dot{u}'(i, j)$ 有影響的右手邊項，在經過轉置求伴隨方程時變為和 $\dot{u}'_a(i-1, j)$ ， $\dot{u}'_a(i, j)$ ， $\dot{u}'_a(i+1, j)$ 項有關，故可根據此種規律計算得出空間上的伴隨算子。本文中的淺水伴隨模式的空間算子即照此法得出，因為要寫成 C 網格式且邊界需小心安排，其過程僅具技術性，於此不贅述太多。

當處理完空間的算子部分，於時間上的積分格式也要處理，Talagrand and Courtier(1987)曾舉蛙跳積分格式為例，於此，利用相同的原理，導出調整性的蛙跳時間積分格式，其結果如下，當線性模式的積分格式為：

$$\begin{cases} \mathbf{x}'(1) = \mathbf{x}'(0) + \Delta t \mathbf{F}(\mathbf{x}'(0)) \\ \tilde{\mathbf{x}}'(0) = \mathbf{x}'(0) \\ \mathbf{x}'(p+1) = \mathbf{x}'(p-1) + 2\Delta t \mathbf{F}(\mathbf{x}'(0)), \quad p \geq 1 \\ \tilde{\mathbf{x}}'(p) = \mathbf{x}'(p) + \alpha (\mathbf{x}'(p+1) - 2\mathbf{x}'(p) + \tilde{\mathbf{x}}'(p-1)) \end{cases}$$

其中“(~)”表調整過後之值， α 為時間過濾參數， p 是時間指標。其伴隨模式的積分格式可寫成設 $\tilde{\mathbf{x}}'_a(p-1) = 0$

$$\begin{cases} \mathbf{x}'_a(k) = (\mathbf{I} - 2\alpha + 2\alpha\Delta t \mathbf{F}_k^T) \tilde{\mathbf{x}}'_a(k) + 2\Delta t \mathbf{F}_k^T \mathbf{x}'_a(k+1) \\ \tilde{\mathbf{x}}'_a(k-1) = 2\alpha \tilde{\mathbf{x}}'_a(k) + \mathbf{x}'_a(k+1), \quad k = p-1, \dots, 1 \\ \mathbf{x}'_a(0) = \tilde{\mathbf{x}}'_a(0) + \mathbf{x}'_a(1) + \Delta t \mathbf{F}_0^T \mathbf{x}'_a(1) \end{cases}$$

其精神在於原線性模式是如何將物理量在時間上積分出去，如今伴隨模式也要如何積分回來，這和空間上處理方式的精神完全一致。經過空間和時間上如此處理的伴隨模式，將滿足(1)式中線性代數的伴隨恆等式，將邊界的訊息隱含至所定義的內積中，如此一來，線性代數算子運算的便利，便可以存用，於處理數值最佳化的問題時也較方便。現今氣象上運用的伴隨模式多以此數值直接導出法計算，僅有當波譜伴隨模式之傅立業(Fourier)算子不若本節的探討外，其空間非線性算子的處理和時間積分格式之求算過程，基本上是相似的。

四、敏感度測試之討論和變分極小化

(一) 敏感度測試

模式的敏感度測試不僅用來檢測不同參數、不同的強制因素(forcing)對模式的反應，以校驗模式的表現，亦用來推估真實世界中，在類似環境下可能有之變化。根據第二節的討論可知，初始時間的伴隨變數組成成本函數的梯度，標示出所求問題的敏感度區域((9),(15)式)。故本節將討論由解析法和數值導出法伴隨模式，比較所得到敏感度結果之異同處。在討論問題之初；設定一控制初始場對原非線性淺水模式積分，當做本節探討問題的控制組，以變數 \bar{x} 表示。在此，於初始質量場中另加入一個數量級為此場 10% 的擾動，且風場與質量場呈地轉平衡，以 x' 表示。控制組變數和擾動場相加得到 x ，代入原非線性淺水模式積分當做實驗組。用數學型式重新表示，則上述之情形即為：

$$x = \bar{x} + x'$$

模式中用一個 Rossby-Hauwitz 模波且風場與質量場呈地轉平衡的控制基本場 \bar{x} (圖 3) 做控制組的初始場，在沒有基本流之情形下，Rossby-Hauwitz 模波使得淺水方程的非線性平流項失去作用，僅受 β 效應影響，以正比於水平尺度大小的相速向西移動，接近南邊界赤道地區的相速較快，整個槽脊相位線將成西南-東北走向，圖 4 為積分 36 小時後之情形。而 x' 為選擇任一區域的擾動場，和 \bar{x} 相加所得的 x ，當做實驗組變數的初始場，其分佈如圖 5 所示，而所加之擾動場在圖 5 的右下方。 x' 變數的意義相當於實驗組模式和控制組變數在初始狀態的差值。在此，如所欲求之成本函數為，

$$J = \frac{1}{2} (x(t_f) - \bar{x}(t_f))^T W (x(t_f) - \bar{x}(t_f)) \quad (18)$$

意謂加入一相較於控制組外的擾動量，應修正若干以得出在積分終點時成本函數的極小值。

當使用 36 小時後實驗組變數減去控制組變數的值 x' (圖 6) 來定義成伴隨模式的初始值，可利用第三節中所探討之解析和數值兩種伴隨模式，得出解析(圖 7)和數值(圖 8)兩種做法之伴隨變數值(於第 0 時)，顯示出模式對此 x' 擾動所產生的敏感度，表現在各變數間應佔有多少的數量級，所在的位置應在何處。此即為成本函數的梯度，和成本函數的極值有關。比較圖 7 和圖 8，可發現兩種方法最大的差異在 v 變數場上(圖 7c、圖 8c)。圖 8c 接近赤道區處有一圈極值區域，而圖 7c 則無，這表示用數值法運算時，邊界條件的作用已被隱含進來，另外，這亦指示出本文使用的簡單模式，必須注意南邊界條件的處理，因其將會影響模式的積分結果。若忽略圖中變數值之大小和 v 場的南邊界之差異，其二者的型態極為類似。

使用我們熟悉的差分格式製做出解析導法的伴隨模式，將無法滿足伴隨恆等式((1)式)，故需要另行推導能保守伴隨恆等式之差分格式。如此一來，將增加發展模式的時間。

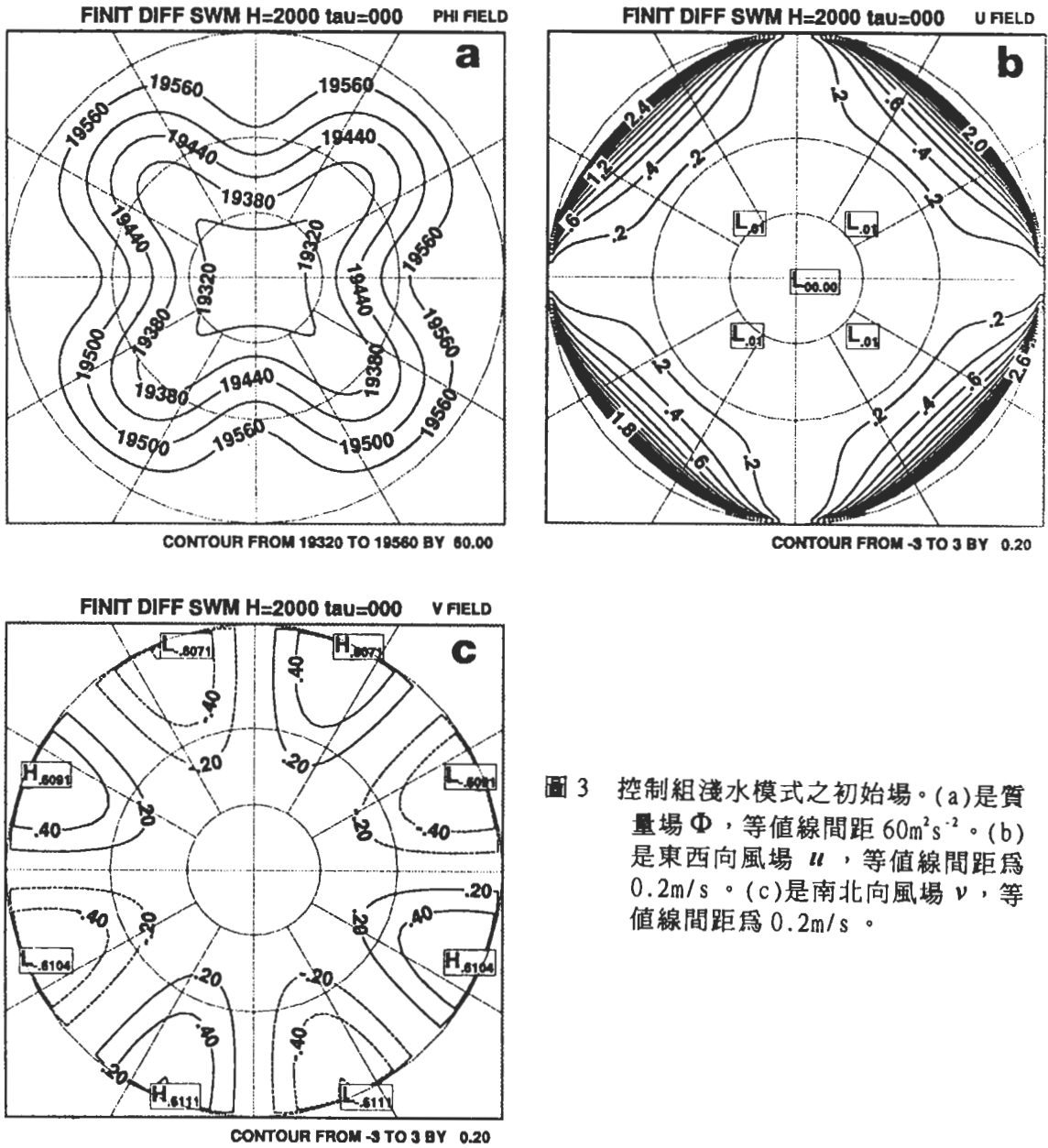


圖 3 控制組淺水模式之初始場。(a)是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u ，等值線間距為 0.2m/s 。(c)是南北向風場 v ，等值線間距為 0.2m/s 。

再者，於運用解析推導的過程中，從部分積分分離出的邊界條件項必須給定，如果要能滿足(1)式，則要設成爲零。本文中用解析法推導伴隨模式時，爲簡化問題起見，是設定邊界的變量爲零。至於運用數值直接導出法時，其邊界條件的處理是配合正切線性模式而來的，並不需要人爲給定，其數值差分格式亦不需另行推導，所以，現今的伴隨模式多採用數值導出法計算。

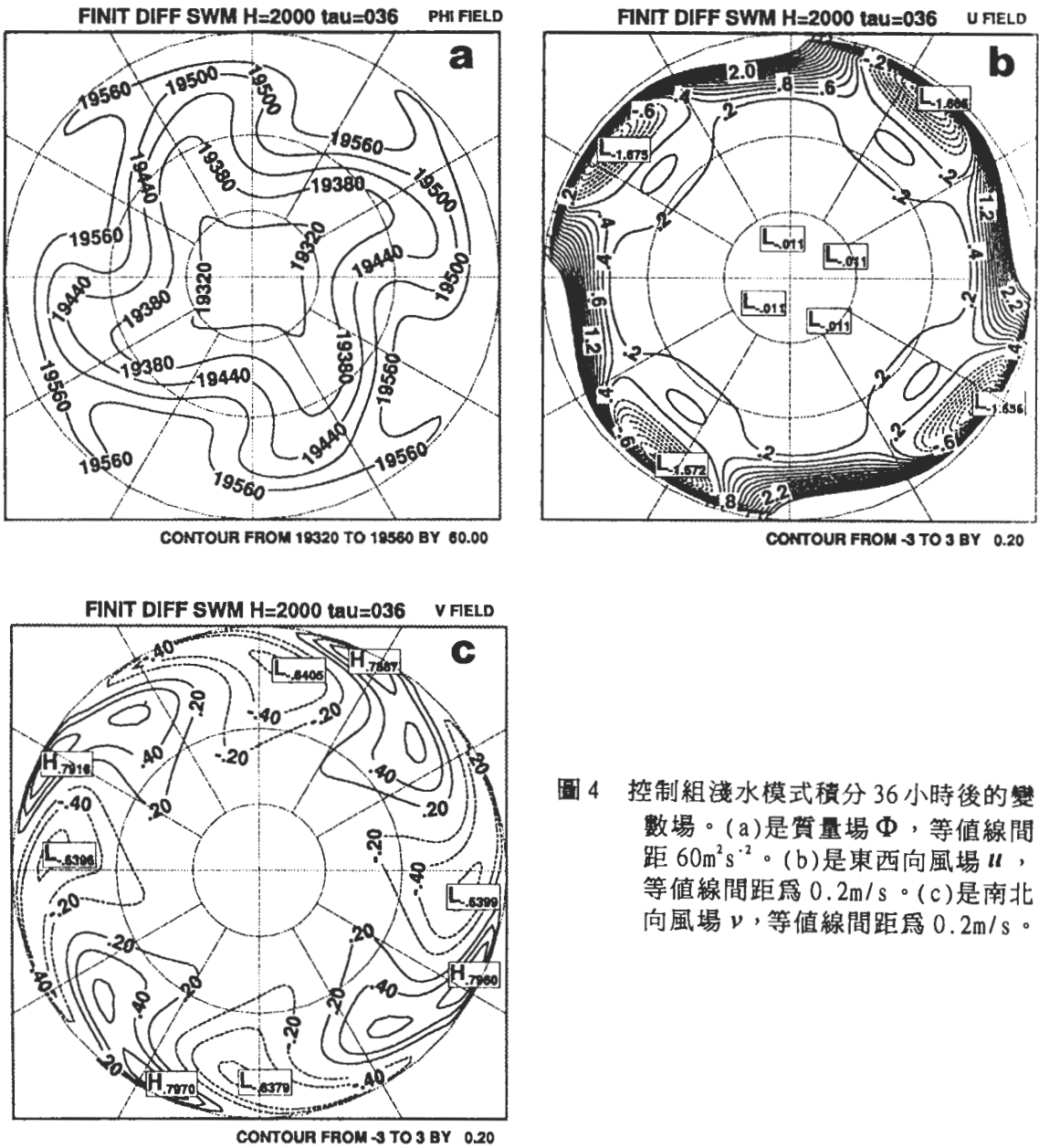


圖 4 控制組淺水模式積分 36 小時後的變數場。(a)是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u ，等值線間距為 0.2m/s 。(c)是南北向風場 v ，等值線間距為 0.2m/s 。

此外，本例中成本函數之權重函數矩陣為一單位矩陣，並沒有設定特別的權重關係。於此間發現：風場的伴隨變數場比伴隨質量場的值大上許多，也就是說，造成積分終點時擾動質量場比風場來的大之原因是源自初始的擾動風場；而調整初始風場的大小，是得到最後結果的最佳途徑，這點符合淺水方程的動力特徵，即質量場向風場調整。值得注意的是，伴隨變數所提供的僅是一種滿足求算結果之最佳化方式，並不保證物理現象的發展過程一定要

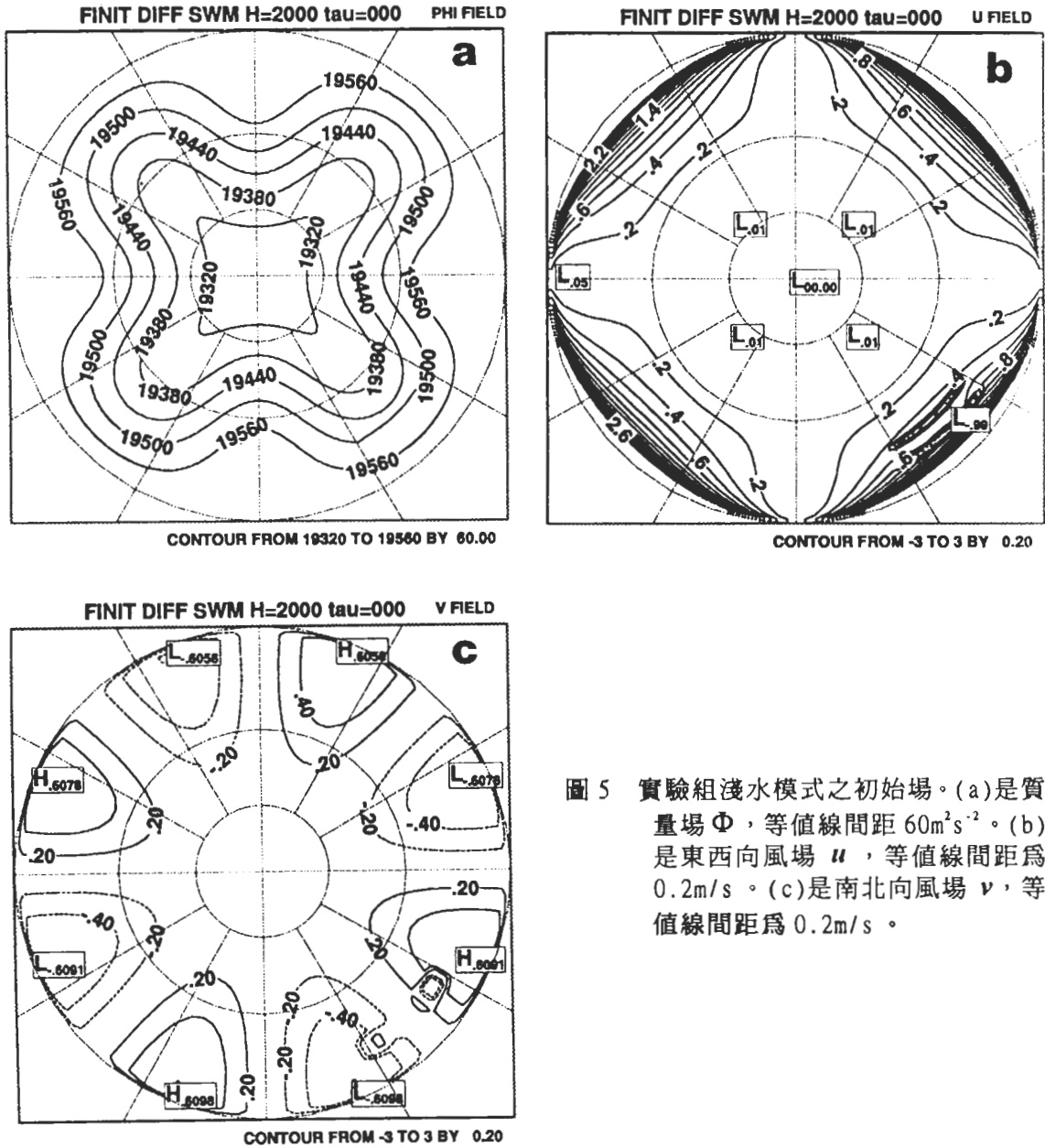


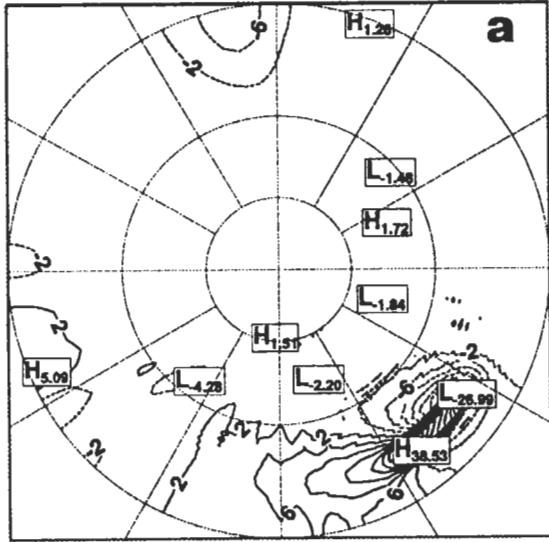
圖 5 實驗組淺水模式之初始場。(a)是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u ，等值線間距為 0.2m/s 。(c)是南北向風場 v ，等值線間距為 0.2m/s 。

遵照這個程序。

(二)變分極小化

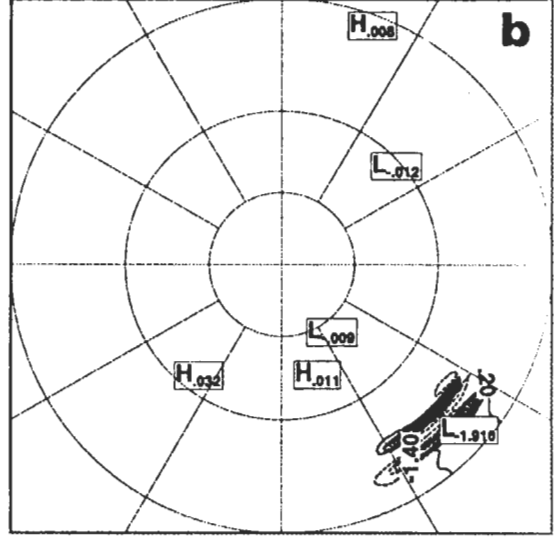
由伴隨變數所決定之四維空間的成本函數梯度，可配合數值方法求極值。意即當沿著梯度方向，選取適當的疊代間距，將可找出成本函數的最大值；反之，沿著負梯度的方向將可求到極小值。在數值變分法中，將初始狀態的變數 x_0 修正成爲 $x_0 + \alpha_i g_i$ ，將可經由疊

FINIT DIFF SW-ADJ.M H=2000 tau=-000 PHI ADJ FIELD



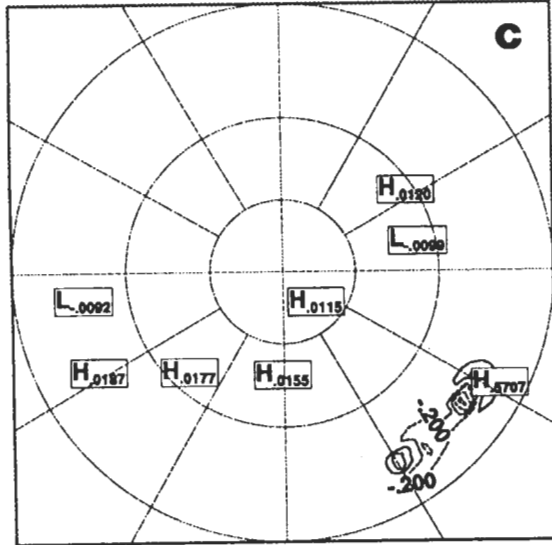
CONTOUR FROM -30 TO 30 BY 4.00

FINIT DIFF SW-ADJ.M H=2000 tau=-000 U ADJ FIELD



CONTOUR FROM -10 TO 10 BY 0.20

FINIT DIFF SW-ADJ.M H=2000 tau=-000 V ADJ FIELD

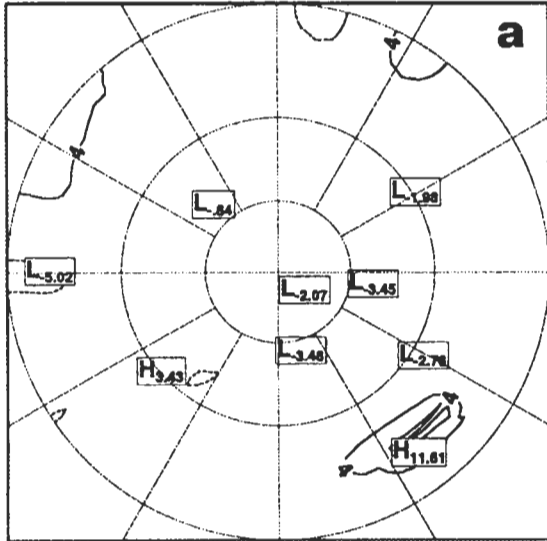


CONTOUR FROM -10 TO 10 BY 0.20

圖 6 實驗組積分 36 小時後之變數場減去控制組積分 36 小時後之變數場，用來當做伴隨模式的初始值 ($\tau=0$)。(a)是質量場 Φ' ，等值線間距 $4\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u' ，等值線間距為 0.2m/s 。(c)是南北向風場 v' ，等值線間距為 0.2m/s 。

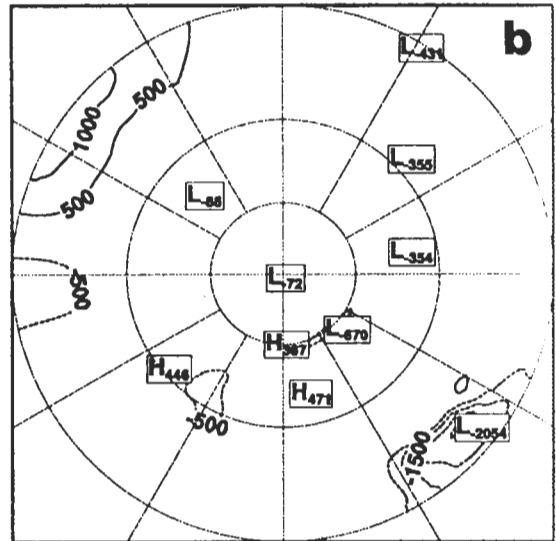
代運算而求到成本函數 $J(x)$ 極小值，其中 g 表負梯度向量， α 表取多少的梯度單位， i 表疊代次數。將 $x_0 + \alpha_1 g_1$ 再代入 $J(x)$ 中，可求出此時之伴隨變數，並再決定出一負梯度 g_2 的方向，將此值再代入變成 $x_0 + \alpha_1 g_1 + \alpha_2 g_2$ ，可決定出另一個初始值，再代入 $J(x)$ 中，重覆上述的過程直到求出可以接受的極值為止。於計算中，當 x_0 和 g_1 都已知時，成本函數 $J(x_0 + \alpha_1 g_1)$ 的極值問題變為單變數 α_1 的極值問題，一般常用線性疊代法去決定 α_1 ，每次疊

FINIT DIFF SW-ADJ.M H=2000 tau=-036 PHI ADJ FIELD



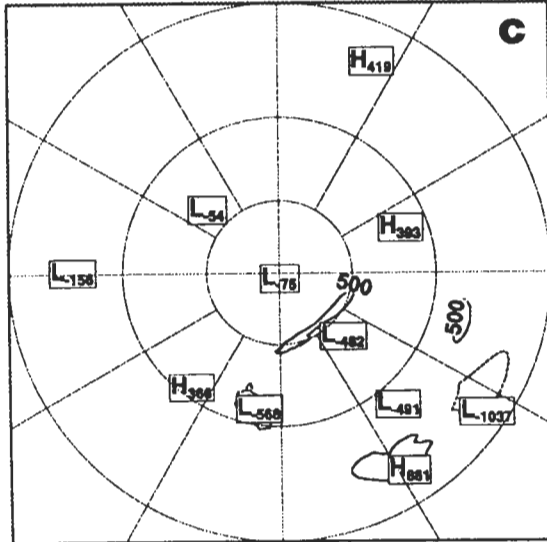
CONTOUR FROM -12 TO 12 BY 4.00

FINIT DIFF SW-ADJ.M H=2000 tau=-036 U ADJ FIELD



CONTOUR FROM -3000 TO 3000 BY 500.00

FINIT DIFF SW-ADJ.M H=2000 tau=-036 V ADJ FIELD

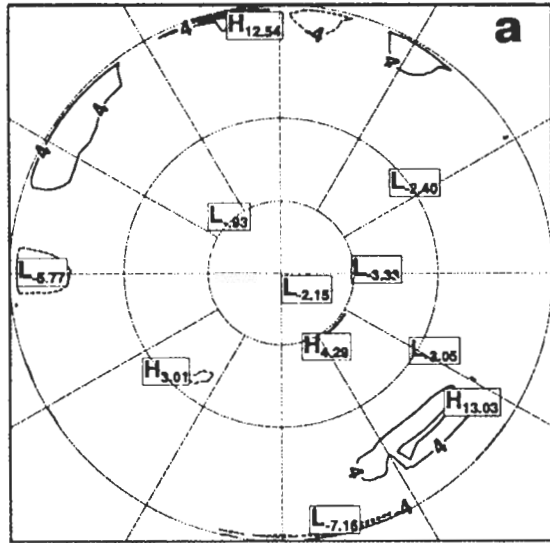


CONTOUR FROM -3000 TO 3000 BY 500.00

圖 7 解析法伴隨模式時間逆向積分 36 小時後的變數場，即回到淺水模式的初始時間 (tau=-36)。(a)是 Φ'_a 場，等值線間距 $4m^2s^{-2}$ 。(b)是 u'_a 場，等值線間距 $500m/s$ 。(c)是 v'_a 場，等值線間距 $500m/s$ 。

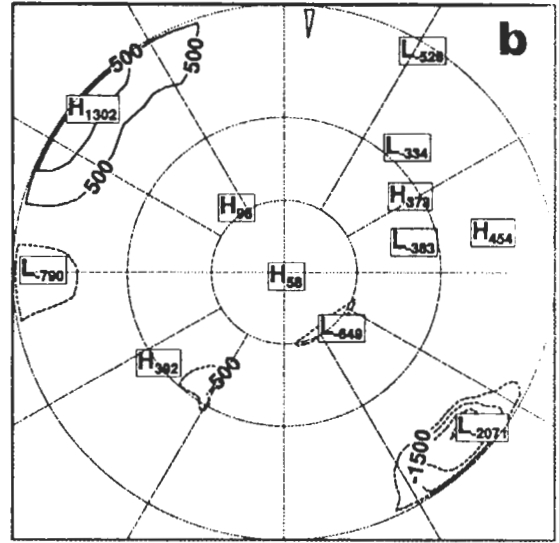
代運算時可照此決定其 α_i 值之大小。上述之過程在 Polak (1971)、Navon and Legler(1987)或是 Arora(1989)的書中皆有詳細的探討，本文採用其中最常用的共軛梯度法做成本函數極小化的實驗。本小節中，我們將以數值導出法所得之伴隨變數當成梯度來求取變分極值，主要原因是其滿足(11)至(16)式的推導。那麼，圖 8c 中南北風向場 v 接近赤道地區的一圈值區扮演著何種角色呢？由於本文所用的淺水模式南北邊界的 v 風場被強迫為零，因此該區

FINIT DIFF SW-ADJ.M H=2000 tau=-036 PHI ADJ FIELD



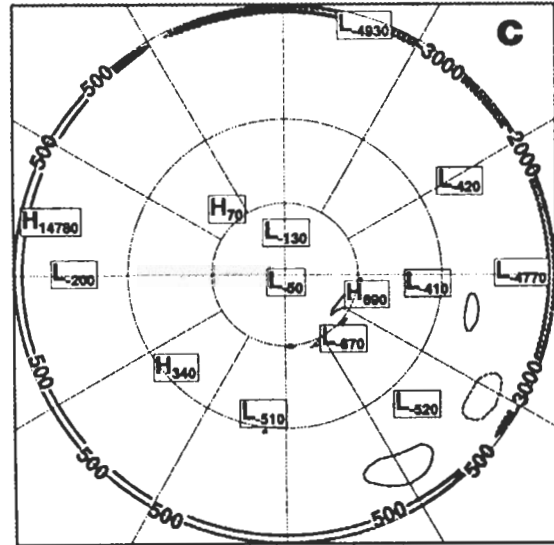
CONTOUR FROM -12 TO 12 BY 4.00

FINIT DIFF SW-ADJ.M H=2000 tau=-036 U ADJ FIELD



CONTOUR FROM -3000 TO 3000 BY 500.00

FINIT DIFF SW-ADJ.M H=2000 tau=-036 V ADJ FIELD



CONTOUR FROM -3000 TO 3000 BY 500.00

圖 8 數值直接導出法伴隨模式時間逆向積分 36 小時後的變數場，即回到淺水模式的初始時間 ($\tau = -36$)。
 (a) 是 Φ'_a 場，等值線間距 $4\text{m}^2\text{s}^{-2}$ 。
 (b) 是 u'_a 場，等值線間距 500m/s 。
 (c) 是 v'_a 場，等值線間距 500m/s 。

在做變分求算極值的疊代分析時，不起任何作用。換言之，即使此處利用解析法伴隨模式所求得之梯度來做運算，預料其結果會是相類似的。另一方面，若回顧上節的敏感度測試討論，這表示如採用不同的邊界條件給淺水模式，對所求之成本函數將有影響。這對了解模式本身的特性有所幫助，但其並不影響此處變分疊代分析之運作過程。

在此，成本函數的定義如(17)式之定義時，將淺水模式積分 36 小時後，實驗組變數減

去控制組變數的值 \mathbf{x}' (圖 6) 來定義成伴隨模式的初始值。首先，取權重函數為單位矩陣，並利用共軛梯度法疊代 60 次求算成本函數的極小值，其結果如圖 9a 所示，縱軸代表每次疊代的成本函數 $J(\mathbf{x})$ 與未經調整時成本函數 $J(\mathbf{x}_0)$ 的比。另外，成本函數可解釋成實驗組變數對控制組變數的差值，像統計學上的變異數。該圖中，可發現 J 隨疊代次數增加而減小，表示伴隨模式很有效的決定此四維函數的梯度，協尋所求之極小值。當疊代 60 次後，成本函數的大小僅有原先之 20%。由於上述成本函數中的權重皆設為 1，而質量場的單位要比風場大上 10^4 倍，故本變分問題本質上可說成是解決當積分終了時，怎樣的初始場可造成最後兩組模式質量場的差異為最小？於是，在 Navon et al. (1992) 中建議，欲求成本函數的最小值應做變數間的正則化 (normalization)，亦即各變數間的權重應有所不同，根據 Navon et al. (1992) 的建議，另外設定一組權重關係為 $\mathbf{W}_\phi = 5 \times 10^{-5}$ ， $\mathbf{W}_{u,v} = 1$ 來重新定義成本函數，所做出之疊代結果如圖 9b，經過相同次數的疊代，成本函數的大小僅有原先之 2%。另外除了因上述權重函數差異造成不同結果的現象外，由於所定義之成本函數不同，變分極小化的結果亦將不同，所得出的修正初始狀態亦將各異。故如何針對所要解決的問題定義成本函數非常重要。舉例來說：四維資料同化的問題在對所有積分時間的物理變數做調整 (如圖 2b)；而討論動力最不穩定模時所關心的是物理變量的最終狀態 (圖 2a)。

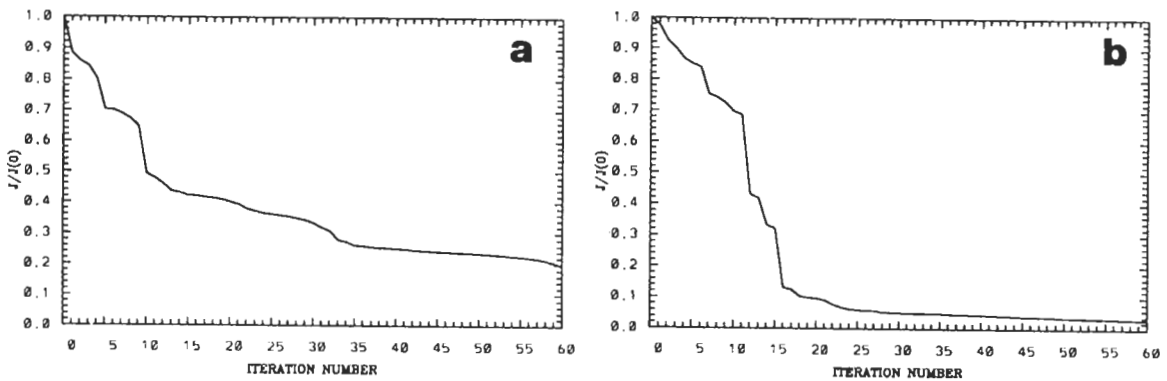


圖 9 (a) 當 $\mathbf{W}_\phi = 1, \mathbf{W}_{u,v} = 1$ 時，縱軸代表每次疊代之價值函數與初始價值函數之比，橫軸表是疊代次數。

(b) 同 (a)，但 $\mathbf{W}_\phi = 5 \times 10^{-5}, \mathbf{W}_{u,v} = 1$

圖 10 為未經權重分析所得之初始場 (即圖 9a 的情形)，於圖 5 相較，可發現質量場幾乎未做調整，但風場則略做修正。其積分 36 小時後的情況如圖 11，如果我們認為圖 4 為所要求的標準答案，可發現圖 11a 中質量場部份已做了修正，但圖 11b, c 中的風場部份則不盡理想。造成這樣的結果並不難想像，因為未經權重分析下所定義的成本函數，質量場的數量級

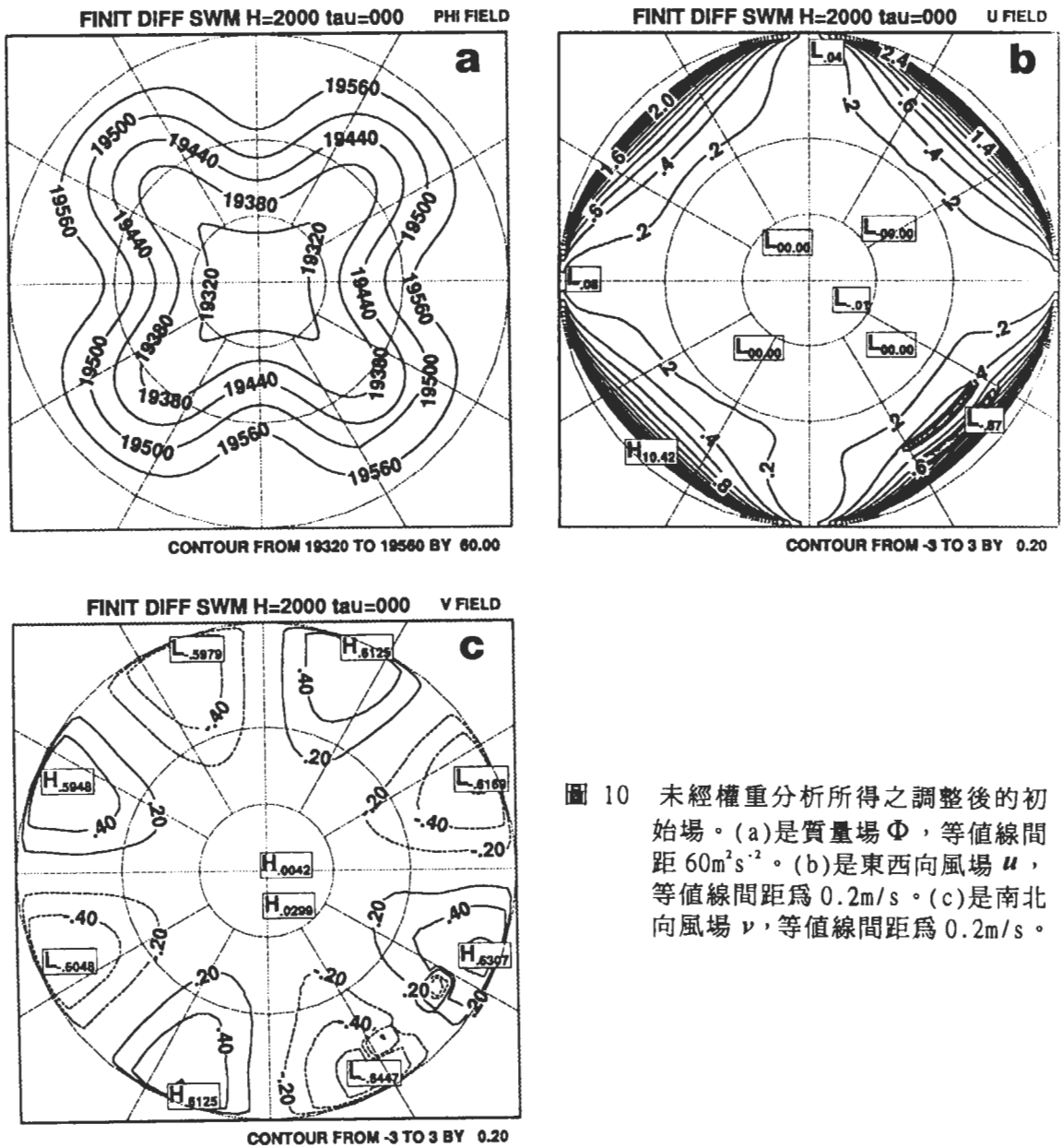


圖 10 未經權重分析所得之調整後的初始場。(a)是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u ，等值線間距為 0.2m/s 。(c)是南北向風場 v ，等值線間距為 0.2m/s 。

比風場大上許多，造成只對質量場做修正的後果。另一方面，經權重分析所得之初始場如圖 12。同樣地，初始質量場之調整有限，但風場則修改了很多。圖 13 為其所得積分 36 小時後的情況，和圖 11 比較，發現圖 13 中質量場和風場都做了修正，而和圖 4 更加接近。我們曾設定權重關係為 $W_{\Phi}=10^{-4}$ ， $W_u, v=1$ ，得到和圖 13 類似的結果(未圖示)，但其成本函數收斂速度較圖 9a 和圖 9b 中為慢，為什麼有這樣的結果，尚待進一步分析。但無論如何，在

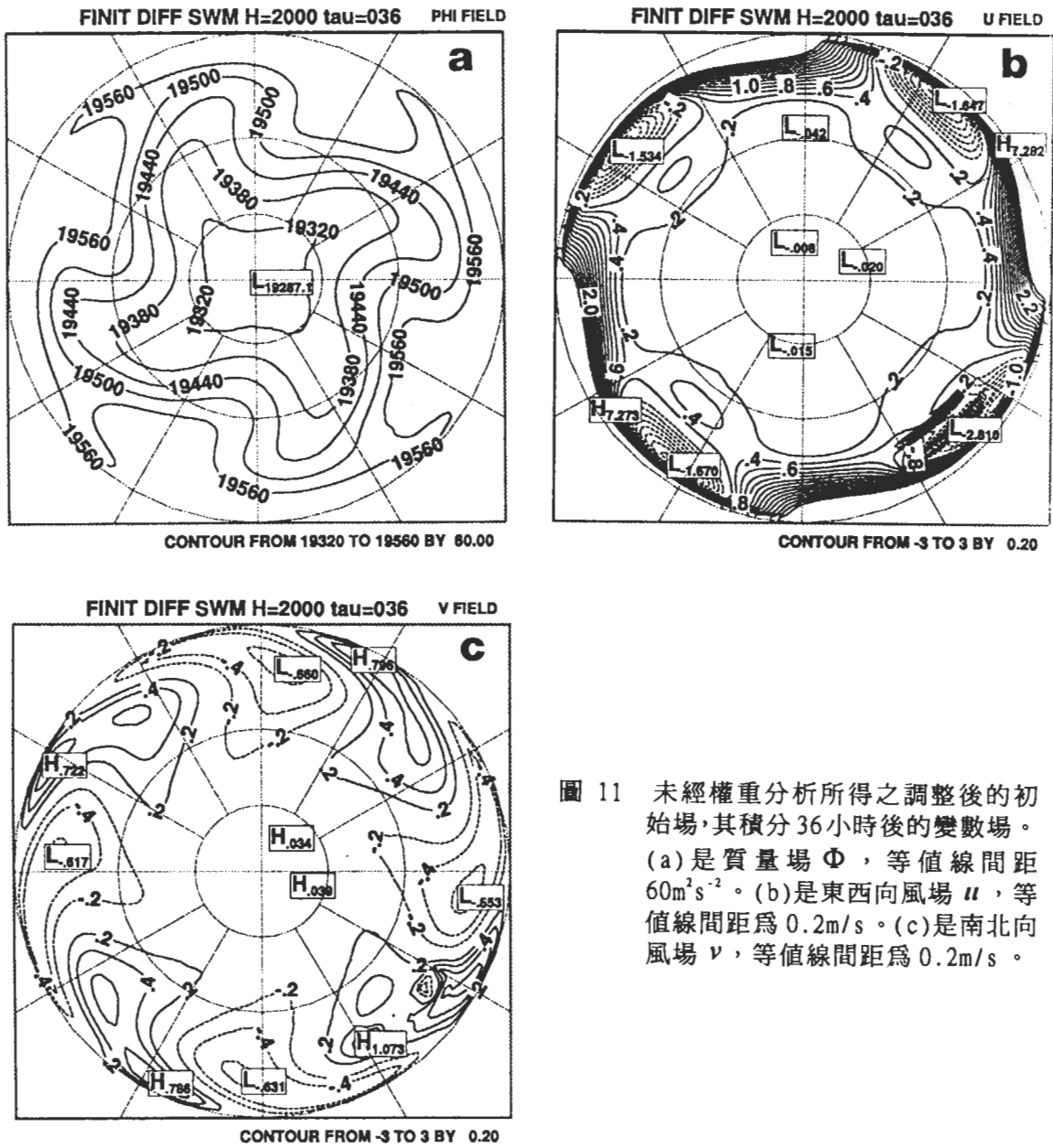


圖 11 未經權重分析所得之調整後的初始場，其積分 36 小時後的變數場。(a) 是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b) 是東西向風場 u ，等值線間距為 0.2m/s 。(c) 是南北向風場 v ，等值線間距為 0.2m/s 。

四維變分分析中，如 Navon et al.(1992)所建議，權重分析扮演重要的角色！這關連到如何適當的定義成本函數，在計算效率和所求問題上取得一平衡點。

五、結論與建議

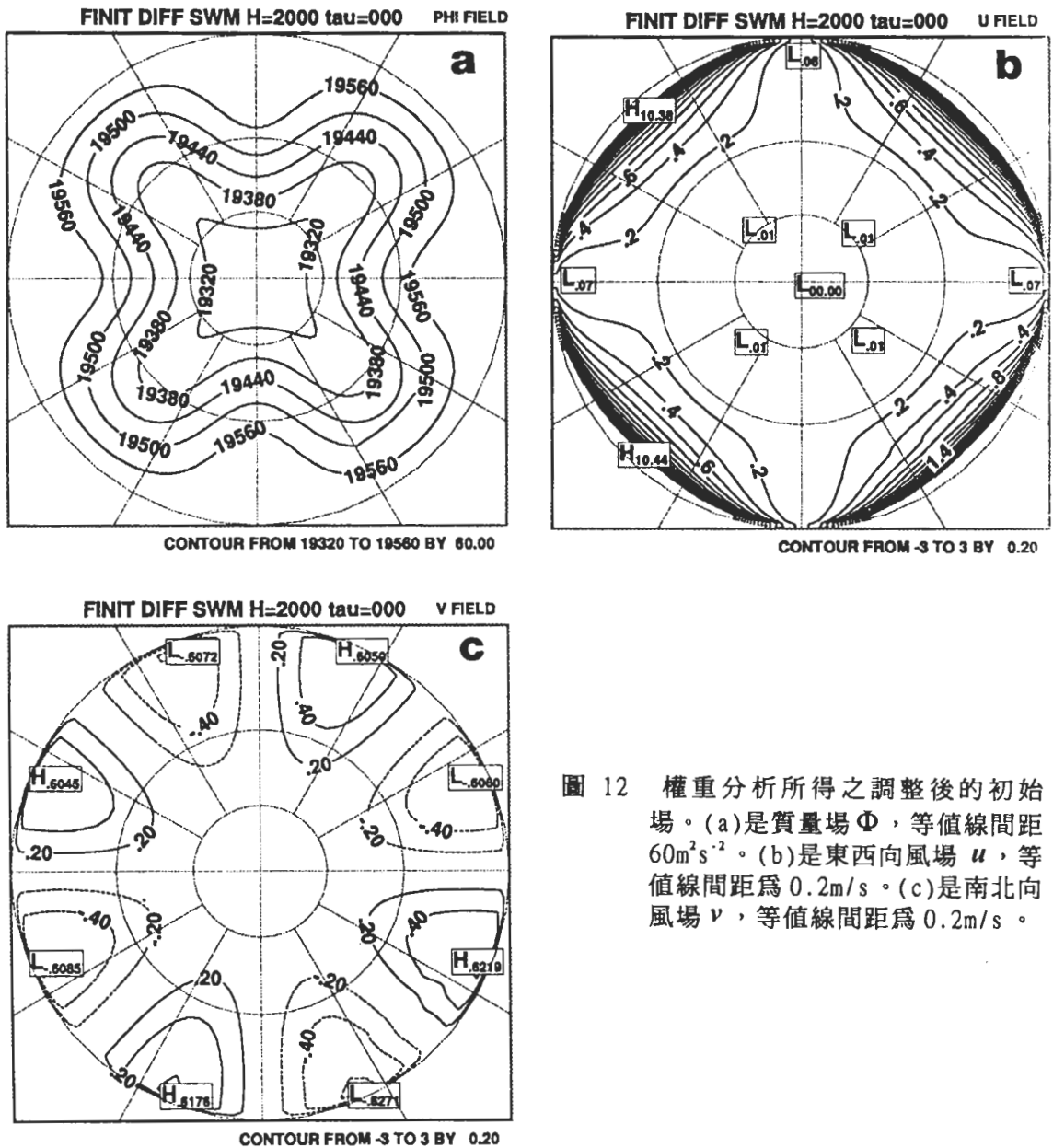


圖 12 權重分析所得之調整後的初始場。(a)是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u ，等值線間距為 0.2m/s 。(c)是南北向風場 v ，等值線間距為 0.2m/s 。

在本文中，我們首先介紹了伴隨理論的淵源和基本原理。伴隨理論源自於線性代數和微分方程算子，其後應用於變分最佳化分析。由於求算伴隨模式數值直接導出法的提出 (Talagrand, 1991)，可消弭線性代數和微分方程中伴隨算子恆等式的差異，而毋須事先預設所求問題之邊界條件來推導伴隨方程。所求出之伴隨變數，正是所定義成本函數以初始場變數表示之梯度。根據此梯度值可用數值疊代法求解當初始場為若干時，其成本函數的極值。

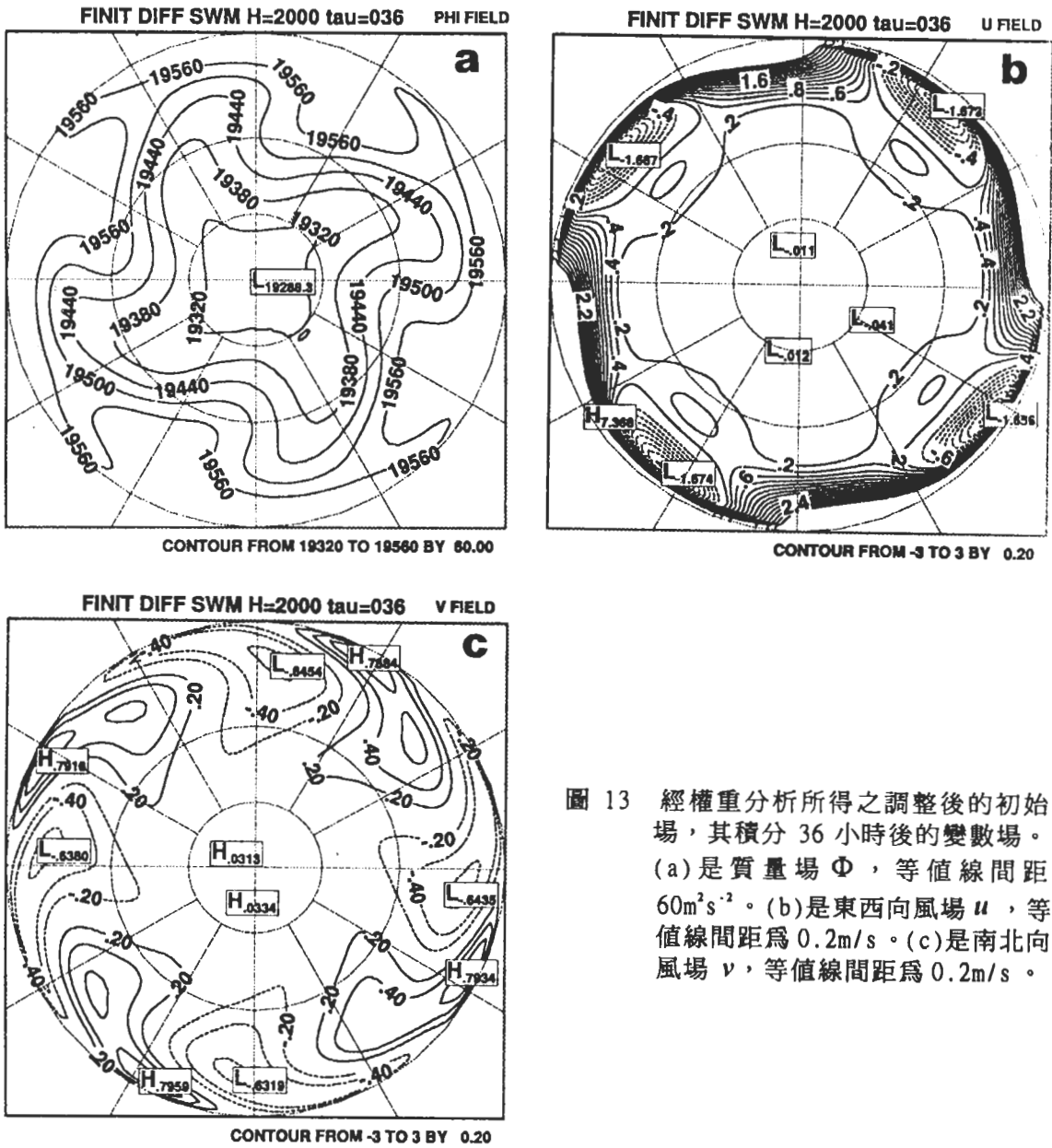


圖 13 經權重分析所得之調整後的初始場，其積分 36 小時後的變數場。(a)是質量場 Φ ，等值線間距 $60\text{m}^2\text{s}^{-2}$ 。(b)是東西向風場 u ，等值線間距為 0.2m/s 。(c)是南北向風場 v ，等值線間距為 0.2m/s 。

由於所興趣的對象不同，所定義的成本函數當亦不同，將決定出不同的伴隨變數，其代表對所求問題之敏感性。伴隨變數值愈大的區域，表示這塊地區對所求問題的影響愈大，任何一個存於此區的小擾動將影響後來的結果甚鉅，需要比其它地區多注意其正確性。但由本文利用淺水模式的實驗中發現：真實世界的變化，並不一定要照著伴隨變數所指示的數量級做調整才能得到所要的結果，即微量修改初始風場可得到積分終點時符合定義的質量場。換

言之，伴隨理論乃提供一最佳化的方法，亦即是調整最少量的初始場，以達到最接近所設想狀態之目的。然而，另一方面，假如成本函數中各變數間之權重給定得宜，由本文的例子中，伴隨變數又能指出近乎正確無誤之修改方向。所以，運用伴隨模式確可得到接近真實的結果，而適當權重之給予則為成敗之關鍵。正切線性模式在本文中只扮演用來推導和驗證伴隨模式的角色，於做敏感度測試和變分最佳化分析中並不需要利用到它，可謂功成而身退，但其在解決奇異向量分析方面(即最佳擾動之激發問題，如 Hartmann et al., 1995)，則又扮演吃重的角色，取代原非線性模式的地位。

舉凡在氣象中牽涉到變分求極值的問題，都可藉助伴隨理論求解。因為任何一種求解過程，都可視為由一運算子和其變數所組成的函數，只要能將運算子線性化，都能得到伴隨算子而求解極值。但實際作業時方程組的複雜性、電腦計算資源的大小，都可能影響到伴隨模式的實用性。儘管如此，由於其可明確計算出四維空間變數的梯度值，在處理變分極值問題時仍有很大潛力。此外，其對不容易直接診斷的模式敏感度測試亦有很大幫助；而在不穩定度和可預測度的問題(Molteni and Palmer, 1993)亦是不錯的工具。

本文中討論了利用解析法和數值直接導出法發展伴隨模式的異同，運用後者推導伴隨模式似乎是較好的選擇，因其可以藉由計算伴隨算子的恆等式是否相等，來驗證伴隨模式之正確性，在 64bit 的 Cray 機器上，其精確度可達 10 位數以上。另一方面，模式的發展維護亦上較為容易。必須特別注意的是，伴隨模式並非是一個時間逆向的模式而已，它所得出的值是一個梯度，一個敏感度區，決非當定義伴隨變數初始值為 $\mathbf{x}'_n(t_f) \equiv \mathbf{x}(t_f) - \bar{\mathbf{x}}(t_f)$ ，其積分終了時所得到的變數是 $\mathbf{x}'_n(t_0) = \mathbf{x}(t_0) - \bar{\mathbf{x}}(t_0)$ ，這是一個嚴重的認知錯誤。

文中以簡單的淺水方程做例子，並沒有對南北邊界做完善的處理，只強迫在其上的 v 值為零。而經由伴隨模式的反求，它指出模式的南北邊界處理不當的問題，足見伴隨模式在對原非線性模式參數做微調之功能。另外，亦藉由伴隨模式的幫助，了解到淺水方程中，正確初始風場的重要性。作者正利用這套系統，繼續從事更多的敏感度測試問題，所得結果將陸續發表於此或其他刊物上。

致 謝

非常感謝兩位審查者之意見，指出本文許多缺失。研究期間，感謝陳建河先生對本文模式部分提供意見。本文在國科會研究計畫 NSC85-2111-M-052-006-AP4 和 NSC86-2111-M-052-001-AP4 的部份經費補助下完成，特致謝意。

參考文獻

Arora, J. S., 1989: *Optimum design*. McGraw-Hill, New York, 625pp.

- Borges, M. D., and D. L. Hartmann, 1992: Barotropic instability and optimal perturbations of observed nonzonal flows. *J. Atmos. Sci.*, 49, 335-354.
- , and P. D. Sardeshmukh, 1995: Barotropic Rossby wave dynamics of zonally varying upper-level flows during northern winter. *J. Atmos. Sci.*, 52, 3779-3796.
- Branstator, G., 1985: Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. Part I: Forced Solutions. *J. Atmos. Sci.*, 42, 2225-2241.
- , 1985: Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. Part II: Eigenanalysis. *J. Atmos. Sci.*, 42, 2242-2254.
- Cacuci, D. G., 1981: Sensitivity theory for nonlinear system. I: Nonlinear functional analysis approach. *J. Math. Phys.*, 22, 2794-2802.
- Courtier, P., J. Derber, R. M. Errico, J.-F. Louis and T. Vukicevic, 1993: Important literature on the use of adjoint, variational methods and the Kalman filter in Meteorology. *Tellus*, 45A, 257-342.
- Daley, R., 1991: *Atmospheric data analysis*. Cambridge, New York, 457pp.
- Errico, R. M. and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of PSU-NCAR mesoscale model. *Mon. Wea. Rev.*, 120, 1644-1660.
- Farrel, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, 39, 1633-1686.
- Hall, M. C. G., D. G. Cacuci, M. E. Schlesinger, 1982: Sensitivity analysis of a radiative model by adjoint method. *J. Atmos. Sci.*, 39, 2083-2050.
- Hartmann, D. L., R. Buizza, and T. N. Palmer, 1995: Singular vectors: The effect of spatial scale on linear growth of disturbances. *J. Atmos. Sci.*, 52, 3885-3894.
- Hildebrand, F. B., 1965: *Methods of applied Mathematics*. Prentice-Hall, New Jersey, 362pp.
- Kontarev, G., 1980: The adjoint equation technique applied to meteorological problems. *ECMWF Tech. Rep., No. 21*, Shinfield Park, Reading, RG29AX, U.K., 21pp.
- LeDimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, 38A, 97-110.
- Lions, J., 1971: *Optimal control of systems governed by partial differential equations*. Springer-verlag, Berlin, 396pp.
- Marchuck, G. I., 1974: *Numerical methods in weather prediction*. Academic Press, New York, 277pp.
- Molteni, F. and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, 119, 269-298.
- Navon, I. M., X. Zou, J. Derber, and J. Sela, 1992: Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.*, 120, 1433-1466.
- , and D. M. Legler, 1987: Conjugated-gradient methods for large scale minimization in meteorology. *Mon. Wea. Rev.*, 115, 1479-1502.
- Polak, E., 1971: *Computational methods in optimization*. Academic Press, New York.
- Stengel, E., 1994: *Optimal control and estimation*. Dover, New York, 639pp.

- Talagrand, O., 1991: The use of adjoint equations in numerical modeling of the atmospheric circulation. In: (A. Griewank, G. Corlisse, editor) Automatic differentiation of algorithms: theory, implementation and application, pp. 169-180.
- , and P. Courtier, 1987: Variational assimilation of Meteorological observations with adjoint vorticity equation. Part I. Theory. *Quart. J. Roy. Meteor. Soc.*, 113, 1311-1328.
- Zhang, Z., 1988: The linear study of zonally asymmetric barotropic flows. *Ph. D. thesis*, University of Reading.
- Zou, X., I. M. Navon, and F. X. Le Dimet, 1992: Incomplete observations and control of gravity waves in variational data assimilation. *Tellus*, 44A, 272-296.
- , -----, and J. G. Sela, 1993: Variational data assimilation with moist threshold processes using the NMC spectral model. *Tellus*, 45A, 370-387.

Application of 4 Dimensional Variations Using Adjoint Method in Meteorology

Chien-Han John Tseng¹ J.-C. Joseph Chang¹ LinHo²

¹Research and Development Center, Central Weather Bureau

²Department of Atmospheric Sciences, National Taiwan University

ABSTRACT

Adjoint method has been widely applied to solve four dimensional (space-time) meteorological optimization problems, especially on researches of the so-called 4-dimensional variational (4DDA) data assimilations. Its basic concept on the exchanges between the original operator and its adjoint of a given set of governing equations, both of which are bounded by a predefined inner-product norm. In this study, a shallow water model (SWM) was employed to illustrate the derivation of its adjoint model, and the process to tackle a simple 4DVAR problem. Two approaches have been adopted in constructing the adjoint model: One is analytically derived from the SW equations, and the other is directly formulated from the numerical schemes of the SWM's Tangent Linear Model(TLM).

We performed a so-called "twin-experiment" with the evolution of a single Rossby-Hauwitz wave as the controlled fields, and a small patch of random fluctuation (10% of total amplitude) in

geopotential height field as the perturbation. The cost function to be minimized is defined as the square of model variable difference in the final time of 36-hour integration. Without any scaling on different variables of the cost function, the adjustment mainly emphasized on the correction of the final geopotential height field, which pointed to a sensitivity in the initial wind field rather than the height field itself. With a more proper set of scaling factors (e.g. $uscale = 1$, $vscale = 1$, $\Phi scale = 5 \times 10^{-5}$), a whole adjustment on the final total fields can be achieved, which requires more corrections in the initial velocity field. It is noted that the analytical approach in deriving an adjoint model has to deal with the boundary conditions carefully, and is much more difficult in a real-world operation. On the other hand, the numerical approach not only has no such a concern on boundary conditions, but it is also easy to construct and therefore to maintain.

Key words: Adjoint model, Tangent linear model, Twin experiment, Cost function.