

BAN ĐẦU HOÁ BẰNG LỌC SỐ TRONG CÁC MÔ HÌNH KHÍ TƯỢNG

CN. Lê Đức, ThS. Đỗ Lê Thủy

CN. Nguyễn Thị Anh Đào, ThS. Võ Văn Hoà

Trung tâm Dự báo Khí tượng Thủy văn Trung ương

Nếu không có sự cân bằng giữa trường gió và trường khối lượng ban đầu trong mô hình dự báo (được xây dựng trên hệ phương trình nguyên thủy) thì sẽ xuất hiện những dao động giả có tần số cao. Những dao động này được tạo ra bởi những sóng trọng trường quán tính có biên độ rất lớn không tồn tại trong khí quyển. Mục đích chính của phương pháp ban đầu hóa là loại bỏ hoặc làm giảm những nhiễu động này bằng cách hiệu chỉnh các trường phân tích. Bài báo này trình bày việc áp dụng sơ đồ lọc số (Digital Filter Initialization-DFI) để ban đầu hóa số liệu cho mô hình HRM (đang chạy nghiệp vụ với sơ đồ các thành phần chính phi tuyến (Nonlinear Model Initialization - NMI) tại Trung tâm Dự báo KTTV Trung ương). Các thử nghiệm được tiến hành trong 6 tháng (từ tháng VII đến tháng XII năm 2004). Kết quả đánh giá với lượng mưa được dự báo từ mô hình cho thấy sơ đồ DFI cho kết quả tốt hơn NMI đối với các ngưỡng mưa từ 20 đến 50mm.

1. Một số kiến thức về xử lý tín hiệu số

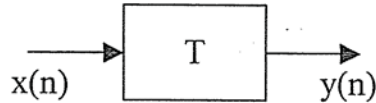
a. Tín hiệu và hệ thống xử lý tín hiệu số

Lọc số trong các mô hình khí tượng thực chất là thực hiện một bộ lọc thông thấp trên các biến dự báo đã được rời rạc hoá theo thời gian. Lý thuyết nền tảng trong việc thực hiện một bộ lọc bất kỳ liên quan đến lý thuyết về xử lý tín hiệu bao gồm cả tín hiệu tương tự và tín hiệu số. Các mô hình khí tượng đều là các mô hình sai phân rời rạc các yếu tố khí tượng trong không gian và theo thời gian cho nên ở đây chúng tôi chỉ trình bày về lý thuyết xử lý tín hiệu số.

Trước hết, tín hiệu được định nghĩa là dạng biểu diễn vật lý của thông tin. Thông tin là một khái niệm trừu tượng mô tả tri thức của con người. Trong cuộc sống hàng ngày, con người hành động và nhận thức theo những thông tin từ bên ngoài đưa lại. Lượng thông tin được thể hiện qua các tín hiệu vật lý.

Thông thường, tín hiệu đưa lại cho con người là các tín hiệu tương tự tức là các tín hiệu liên tục cả về không gian, thời gian và giá trị. Tuy nhiên, cùng với sự xuất hiện của máy tính, một dạng tín hiệu khác đang ngày càng trở nên quan trọng: tín hiệu số. Đó là tín hiệu rời rạc hoá trong không gian, thời gian cũng như về giá trị.

Để phù hợp với bài toán khí tượng, ta sẽ chỉ giả định tín hiệu rời rạc hoá theo thời gian. Xét tín hiệu $x(t)$ liên tục, ta sẽ rời rạc tín hiệu theo thời gian dưới dạng chuỗi $x(n)$ với n nguyên. Ví dụ: nếu coi x là nhiệt độ thì ta sẽ có một chuỗi số liệu nhiệt độ được rời rạc hoá theo thời gian (chính là nhiệt độ tại các bước thời gian liên tiếp trong mô hình). Một hệ thống xử lý tín hiệu tương tự sẽ biến đổi tín hiệu $x(t)$ thành $y(t)$. Khi $x(t)$ đã được rời rạc hoá ta có hệ thống xử lý tín hiệu số biến đổi $x(n)$ thành $y(n)$.



Hình 1. Hệ thống xử lý tín hiệu số

Đối với bài toán khí tượng, hệ thống xử lý tín hiệu số ở đây chính là bộ lọc. Thông thường, để đơn giản người ta thường chỉ xét các hệ thống tuyến tính.

$$T[ax_1(n)+bx_2(n)] = aT[x_1(n)] + bT[x_2(n)] \quad (1)$$

Biểu diễn chuỗi $x(n)$ qua hàm Dirac (chỉ bằng 1 khi $x=0$, còn lại là giá trị 0)

$$x(n) = \sum_{k=-\infty}^{+\infty} x(k)\delta(n-k) \quad (2)$$

rồi tác dụng toán tử T lên $x(n)$ ta được:

$$y(n) = T[x(n)] = \sum_{k=-\infty}^{+\infty} x(k)T[\delta(n-k)] = \sum_{k=-\infty}^{+\infty} x(k)h_k(n) \quad (3)$$

Từ $T[h(n)] = h_0(n)$ thay n bằng $n-k$ ta được $T[h(n-k)] = h_0(n-k)$.

Mà $T[h(n-k)] = h_k(n)$ do đó $h_0(n-k) = h_k(n)$.

Ký hiệu $h(n-k) = h_0(n-k)$ thay vào công thức (3) ta được:

$$y(n) = \sum_{k=-\infty}^{+\infty} x(k)h(n-k) = x(n) * h(n) \quad (4)$$

với $h(n)$ được gọi là hàm đáp ứng xung của hệ thống, còn $(*)$ là ký hiệu của toán tử tổng chập. Hàm $h(n)$ chỉ phụ thuộc vào hệ thống và là đặc trưng của mỗi hệ thống. Trên quan điểm vật lý có thể xem các hệ số $h(n)$ như là các trọng số có tác dụng làm trơn xác định lại chuỗi $x(n)$. Như vậy, một hệ thống xử lý tín hiệu số sẽ được xác định hoàn toàn bởi hàm đáp ứng xung của nó. Nhiệm vụ cơ bản khi thiết kế một hệ thống chính là việc xác định hàm đáp ứng xung của nó.

Tín hiệu thường được một hệ thống số xử lý gồm có tín hiệu âm thanh và hình ảnh. Trong thực tế, các tín hiệu này thường mang trong nó những dạng nhiễu ký sinh khác nhau. Với tín hiệu khí tượng thì đây chính là những dao động phát sinh từ các sóng chuyển động nhanh trong khí quyển bao gồm sóng âm và sóng trọng trường. Do đó,

một yêu cầu quan trọng thường được đặt ra cho một hệ thống số là khả năng khử nhiễu trong tín hiệu. Khảo sát trong thực tế cho thấy các thông tin hữu ích nằm trong vùng tần số thấp của tín hiệu trong khi nhiễu nằm trong vùng tần số cao. Bởi vậy, thay vì khảo sát hệ thống số trên miền thời gian người ta thường khảo sát hệ thống số trên miền tần số. Trên quan điểm toán học thì điều này tương đương với việc biểu diễn tín hiệu chuyển từ hệ hàm cơ sở Dirac sang hệ hàm cơ sở Fourier. Tuy nhiên, trước hết ta sẽ khảo sát tín hiệu trên miền phức bởi biểu diễn tín hiệu trên miền tần số chỉ là một trường hợp riêng của tín hiệu trên miền phức.

b. Hệ thống số trên miền phức

Khi khảo sát trên miền phức ta có thể dễ dàng thu được nhiều tính chất quan trọng của hệ thống mà sẽ rất khó tìm ra khi khảo sát trên miền thời gian. Phép biến đổi qua lại giữa miền thời gian và miền phức được định nghĩa như sau. Biến đổi từ miền thời gian sang miền phức được xác định bằng phép biến đổi Z .

$$X(z) = \sum_{n=-\infty}^{+\infty} x(n)z^{-n} \quad (5)$$

với z là một số phức. Phép biến đổi Z ngược có dạng:

$$x(n) = \frac{1}{2\pi i} \oint_C X(z) z^{n-1} dz \quad (6)$$

với i là đơn vị ảo, C là chu trình kín trên mặt phẳng phức.

Phép biến đổi Z có một số tính chất sau:

- Tuyến tính: $X(z) = aX_1(z) + bX_2(z)$
với $x(n) = ax_1(n) + bx_2(n)$

- Trễ: $Y(z) = z^{-k}X(z)$ với $y(n) = x(n-k)$

- Đảo: $Y(z) = X(1/z)$ với $y(n) = x(-n)$

- Tổng chập: $Y(z) = X(z)H(z)$ với $y(n) = x(n)*h(n)$

Bây giờ ta xét hai dạng hiện của hệ thống xử lý tuyến tính:

Hệ không đệ quy:

$$y(n) = b_0x(n) + \dots + b_Mx(n-M) \quad (7)$$

Áp dụng phép biến đổi Z sử dụng tính chất tuyến tính và trễ ta được

$$Y(z) = b_0X(z) + \dots + b_Mz^{-M}X(z) \quad (8)$$

Từ tính chất của tổng chập $Y(z) = X(z)H(z)$ ta thu được:

$$H(z) = b_0 + \dots + b_Mz^{-M} \quad (9)$$

Hệ thống có hàm đáp ứng xung như trên được gọi là hệ thống có đáp ứng xung hữu hạn (Finite Impulse Response - FIR).

Để thu được hàm đáp ứng xung $h(n)$ ta chỉ cần thực hiện phép biến đổi Z ngược.

Hệ đệ quy:

$$a_0y(n) + \dots + a_Ny(n-N) = b_0x(n) + \dots + b_Mx(n-M) \quad (10)$$

Cũng áp dụng phép biến đổi Z sử dụng tính chất tuyến tính và trễ ta được:

$$a_0Y(z) + \dots + a_Nz^{-N}Y(z) = b_0X(z) + \dots + b_Mz^{-M}X(z) \quad (11)$$

Từ tính chất của tổng chập $Y(z) = X(z)H(z)$ ta thu được:

$$H(z) = \frac{b_0 + \dots + b_Mz^{-M}}{a_0 + \dots + a_Nz^{-N}} = \sum_{k=-\infty}^{+\infty} c_k z^{-k} \quad (12)$$

trong đó, ta đã chia tử số cho mẫu số và thu được một chuỗi vô hạn. Hệ thống có hàm đáp ứng xung như trên được gọi là hệ thống có đáp ứng xung vô hạn (Infinite Impulse Response - IIR). Tương tự như trên để thu được hàm đáp ứng xung $h(n)$ ta chỉ cần thực hiện phép biến đổi Z ngược. Bây giờ ta sẽ chuyển sang khảo sát tín hiệu trên miền tần số.

c. Hệ thống số trên miền tần số

Biến đổi tín hiệu giữa miền thời gian và miền tần số được thực hiện thông qua phép biến đổi Fourier. Dưới đây, ta sẽ thấy phép biến đổi Fourier thực chất là một trường hợp riêng của phép biến đổi Z . Tuy nhiên, phép biến đổi Fourier riêng lại cho phép ta khảo sát tín hiệu trên miền tần số và ta tách ra xét riêng trong một phần không trùng với phép biến đổi Z .

Phép biến đổi Fourier:

$$X(e^{j\theta}) = \sum_{n=-\infty}^{+\infty} x(n)e^{-j\theta n} \quad (13)$$

với i là đơn vị ảo, θ tần số số không có thứ nguyên. Ta có phép biến đổi Fourier ngược:

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\theta}) e^{j\theta n} d\theta \quad (14)$$

Tần số số liên hệ với tần số góc thực qua công thức $\theta = \omega\Delta t$ tương tự như n trong các công thức trên là một đại lượng không thứ nguyên cần được hiểu là nt trong đó t là khoảng thời gian giữa hai tín hiệu liên tiếp.

Từ (5) và (13) dễ dàng nhận thấy phép biến đổi Fourier là trường hợp

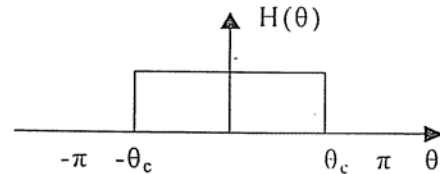
riêng của phép biến đổi Z với $z = e^{j\theta}$. Bởi thế, mọi tính chất của phép biến đổi Z đều áp dụng được với phép biến đổi Fourier. Dưới đây ta chỉ nêu lên các tính chất riêng của phép biến đổi Fourier.

Một số tính chất của phép biến đổi Fourier:

- Tuần hoàn: tín hiệu trên miền tần số tuần hoàn với chu kỳ 2π . Sau này, khi khảo sát tín hiệu trên miền tần số ta chỉ cần xét tín hiệu trong khoảng $[0, 2\pi]$ hoặc $[-\pi, \pi]$. Khi khảo sát trên khoảng $[-\pi, \pi]$ dễ thấy tần số thực lớn nhất có thể biểu diễn là $f = 1/2\Delta t$ tương ứng với chu kỳ nhỏ nhất $T = 2t$. Điều này hoàn toàn dễ hiểu do tín hiệu của ta là tín hiệu số, không liên tục.

- Chẵn, lẻ: biên độ tín hiệu là hàm chẵn (đối xứng qua Oy), pha tín hiệu là hàm lẻ (đối xứng qua gốc tọa độ). Biểu

diễn tín hiệu trên miền tần số dưới dạng $X(e^{j\theta}) = A(\theta)e^{j\varphi(\theta)}$ ta sẽ có $A(\theta) = A(-\theta)$ và $\varphi(\theta) = -\varphi(-\theta)$. Bây giờ ta sẽ xây dựng bộ lọc thông thấp lý tưởng. Bộ lọc sẽ cắt bỏ hoàn toàn các tín hiệu có tần số lớn hơn một tần số cắt c nào đó. Hàm đáp ứng tần số $H(e^{j\theta})$ của bộ lọc này có dạng sau:



Hình 2. Đáp ứng tần số của bộ lọc thông thấp

Chú ý rằng ở đây ta đã áp dụng tính chẵn của hàm biên độ với giả định hàm pha bằng không cho nên hàm đáp ứng tần số đúng bằng hàm biên độ và là hàm chẵn, áp dụng phép biến đổi Fourier ngược ta có hàm đáp ứng xung:

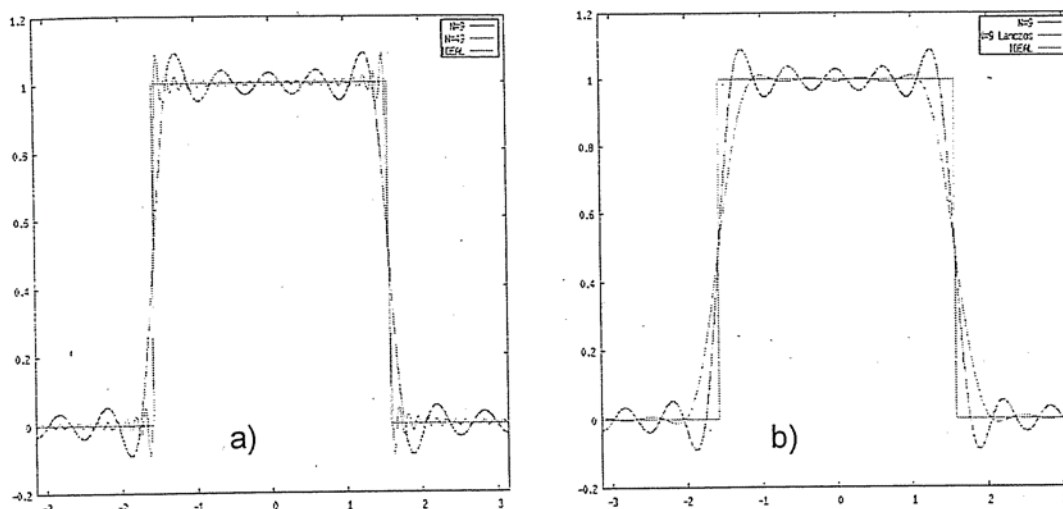
$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\theta}) e^{j\theta n} d\theta = \frac{1}{2\pi} \int_{-\theta_c}^{\theta_c} e^{j\theta n} d\theta = \begin{cases} \frac{1}{n\pi} \sin \theta_c n & (n \neq 0) \\ \frac{\theta_c}{\pi} & (n = 0) \end{cases} \quad (15)$$

Thực tế, khi sử dụng bộ lọc thông thấp lý tưởng ta chỉ có thể nhân và lấy tổng với một số hữu hạn $h(n)$ do đó công thức với tổng vô hạn (4) chuyển thành:

$$y(n) = \sum_{k=-N}^N x(k)h(n-k) \quad (16)$$

trong đó: N là bước thời gian cắt. Dạng gần đúng này của tổng chập so với (7)

mô tả một hệ xử lý số không đệ quy. Với dạng hàm đáp ứng xung ở công thức (15) đây là bộ lọc số không đệ quy thông thấp. Tuy nhiên, ở dạng gần đúng này tại lân cận tần số cắt xuất hiện những dao động nhỏ lệch so với hàm đáp ứng tần số lý tưởng. Dao động này được gọi là dao động Gibbs (hình 3, a).



Hình 3. Đáp ứng tần số lý tưởng so với đáp ứng tần số trong thực tế khi cắt chuỗi tại N bằng 9 và 49 (a). Đáp ứng tần số trước và sau khi lọc bằng cửa sổ Lanczos N bằng 9) so với đáp ứng tần số lý tưởng (b).

Để loại bỏ dao động Gibbs, Lanczos đã đề xuất việc làm trơn hàm đáp ứng xung thay cho dạng gãy khúc không liên tục.

$$\hat{H}(\theta) = \frac{N}{2\pi} \int_{-\pi/N}^{\pi/N} H(\theta) d\theta \quad (17)$$

Việc này tương đương với nhân hàm đáp ứng xung với hệ số sau (cửa sổ Lanczos)

$$\omega(n) = \frac{\sin(n\pi/N)}{n\pi/N} \quad (18)$$

và ta có

$$\hat{h}(n) = \omega(n) * h(n) \quad (19)$$

Một kết quả minh họa được biểu diễn trên hình 3 (b).

2. Ban đầu hoá trường khí tượng bằng lọc số

a. Tạo chuỗi thời gian

Khi áp dụng lọc số cho các trường khí tượng, mỗi trường khí tượng tại một

điểm lưới không gian được xem như một tín hiệu x. Để thu được chuỗi tín hiệu theo thời gian người ta làm như sau:

- Tích phân tiến mô hình N bước sẽ có $x_1, x_2, \dots, x_N$.

- Tích phân lùi mô hình M bước sẽ có $x_{-1}, x_{-2}, \dots, x_{-M}$.

Từ đó ta thu được chuỗi thời gian với $N+M+1$ phần tử (tính cả phần tử x_0 ban đầu). Sau đó, lọc với bộ lọc có hàm đáp ứng xung tùy chọn xác định lại x_0 . Thông thường, x_0 nằm giữa nên người ta thường lấy $N=M$. Chú ý rằng bài toán ban đầu hoá chỉ yêu cầu xác định trường khí tượng tại thời điểm ban đầu, vì vậy, ta không lọc toàn bộ chuỗi mà chỉ xác định duy nhất giá trị chuỗi sau khi lọc tại thời điểm $t=0$.

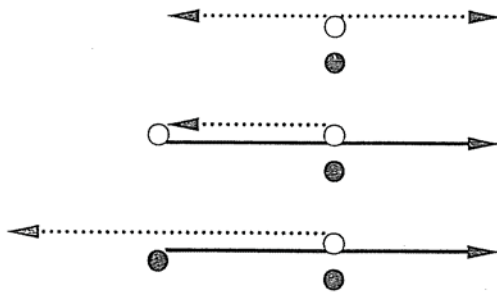
Thực hiện tích phân tiến rất đơn giản do mô hình cũng thực hiện tích phân tiến. Thực hiện tích phân lùi khó hơn bởi một số quá trình vật lý là bất thuận

ngịch tức là chỉ thực hiện theo chiều tiến của thời gian. Bởi vậy, khi tích phân lùi người ta chỉ thực hiện tích phân đoạn nhiệt, các quá trình phi đoạn nhiệt được bỏ qua. Giá trị của những biến khí tượng trên biên được cho bằng hằng số trong quá trình tích phân. Tích phân tiến có thể thực hiện đoạn nhiệt hoặc phi đoạn nhiệt. Hiện tại có ba cách thực hiện như sau:

- Tích phân tiến đoạn nhiệt N bước, lùi đoạn nhiệt N bước xác định $2N+1$ phần tử chuỗi rồi lọc xác định x_0 .

- Tích phân lùi đoạn nhiệt N bước. Từ x_N , tích phân tiến phi đoạn nhiệt $2N$ bước, xác định $2N+1$ phần tử chuỗi rồi lọc xác định x_0 .

- Tích phân lùi đoạn nhiệt $2N$ bước xác định $2N+1$ phần tử chuỗi rồi lọc xác định x_N . Từ x_N tích phân tiến phi đoạn nhiệt $2N$ bước, xác định $2N+1$ phần tử chuỗi rồi lọc xác định x_0 .



Hình 4. Ba phương pháp lọc xác định trường ban đầu (hình tròn trắng: chưa lọc, hình tròn đen: lọc, mũi tên nét đứt - tích phân đoạn nhiệt, mũi tên nét liền - tích phân phi đoạn nhiệt).

b. Các bộ lọc tạo trường ban đầu

Bộ lọc thông dụng nhất là bộ lọc không đệ quy (15) như đã trình bày ở phần 1. Đây là bộ lọc lý tưởng đã cắt bỏ các giá trị của hàm đáp ứng xung lớn hơn N . Khi thực hiện, có thể dùng cửa sổ Lanczos để hàm đáp ứng tần số liên

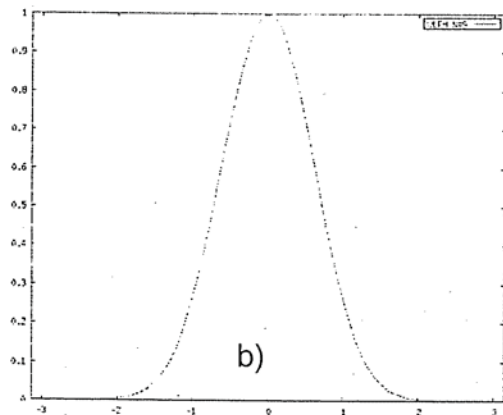
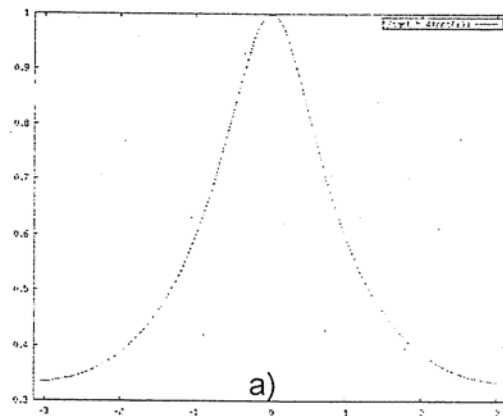
tục hơn so với đáp ứng tần số lý tưởng. Bộ lọc này (19) cũng đã được nhắc đến ở phần 1.

Các bộ lọc đệ quy đa dạng hơn, do nhiều tác giả khác nhau đưa ra. Như đã nói ở phần 1 đáp ứng tần số của bộ lọc đệ quy có dạng phân thức. Để đưa ra một bộ lọc thông thấp, nguyên lý chung vẫn là lựa chọn dạng hàm phân thức này gần với dạng của bộ lọc lý tưởng. Dưới đây là một bộ lọc đệ quy đơn giản

$$y(n) - ay(n-1) = x(n) \quad (20)$$

trong đó: a là tham số với đáp ứng tần số

$$H(\theta) = 1 / \sqrt{a^2 - 2a \cos \theta + 1} \quad (21)$$



Hình 5. Đáp ứng tần số của bộ lọc (20) với $a=2$ (a) và bộ lọc Dolph với $N=9$ (b).

Do hàm đáp ứng xung giảm dần (hình 5, a) nên khác với bộ lọc lý tưởng, bộ lọc đệ quy không cắt bỏ hoàn toàn các tần số cao mà chỉ giảm bớt tác động của những thành phần tín hiệu có tần số cao bằng cách giảm dần biên độ của các thành phần này.

Hiện tại, các trung tâm dự báo lớn trên thế giới đang hướng đến sử dụng bộ lọc tối ưu. Bộ lọc này được xây dựng dựa trên lý thuyết luân phiên Chebyshev cực tiểu hoá sai số cực đại trong dải thông và dải chắn. Lý thuyết toán học cho bộ lọc này khá phức tạp (sử dụng hệ phương trình phi tuyến phức) nên chúng tôi sẽ không nêu ở đây. Giải thuật được đưa ra để xác định hàm đáp ứng xung cũng rất phức tạp. Vì vậy, ở đây chúng tôi chỉ đưa ra bộ lọc đơn giản nhất trong các bộ lọc tối ưu là bộ lọc Dolph. Bộ lọc này sử dụng đa thức Chebyshev xác định đáp ứng tần số.

$$T_n(x) = \begin{cases} \cos(n \cos^{-1} x) & |x| \leq 1 \\ \cosh(n \cosh^{-1} x) & |x| > 1 \end{cases} \quad (22)$$

Hàm đáp ứng tần số có dạng:

$$H(\theta) = \frac{T_{2N}(x_0)}{T_{2N}(x_0 \cos(\theta/2))} \quad (23)$$

với x_0 liên hệ với tần số cắt c qua công thức:

$$x_0 \cos(\theta_c/2) = 1 \quad (24)$$

và N là số bước tích phân tiến hoặc lùi đã nói ở trên. Dạng đáp ứng tần số của bộ lọc Dolph với $N = 9$ được cho trên hình 5 (b).

Một bộ lọc khác cũng đang được nghiên cứu triển khai là bộ lọc Padé. Điểm đặc biệt của bộ lọc này ở chỗ tác giả không xét hàm đáp ứng xung trên

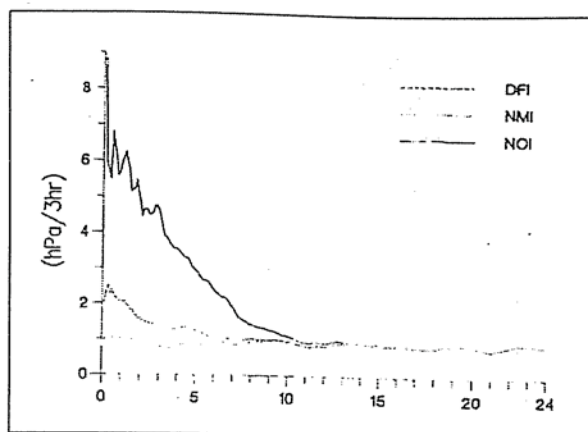
miền tần số mà xét trên miền phức Z . Các nghiên cứu bộ lọc này hứa hẹn sẽ đem lại nhiều kết quả tốt trong tương lai.

c. Hiệu năng của lọc số

Dưới đây là tổng kết các ưu điểm của lọc số so với các phương pháp ban đầu hoá khác đặc biệt là phương pháp ban đầu hoá các thành phần chính phi tuyến (NMI):

- Không cần tính toán hay lưu trữ các thành phần chính.
- Không cần tách biệt các thành phần thẳng đứng.
- Hoàn toàn tương thích với tính rời rạc của mô hình.
- Áp dụng cho mọi loại lưới trên mọi miền.
- Không có các thuật toán lặp có thể không hội tụ.
- Dễ thực hiện và bảo trì.
- Áp dụng cho mọi biến dự báo của mô hình.
- Áp dụng được cho cả các mô hình phi tuyến tính.
- Kinh tế và hiệu quả khi đưa vào ràng buộc trong 4DVAR.

So sánh giữa kết quả dự báo độ biến thiên áp suất mực biển theo thời gian dự báo trong 3 trường hợp không ban đầu hoá (NOI), ban đầu hoá theo phương pháp NMI và ban đầu hoá bằng lọc số DFI được cho trên hình 8. Kết quả này được lấy lại theo Lynch và Huang, 1992 [6] khi ứng dụng lọc số vào ban đầu hoá cho mô hình HIRLAM. Để thấy rằng, lọc số đã lọc rất tốt nhiễu trong trường ban đầu. Kết quả do lọc số đem lại tốt hơn so với phương pháp NMI.



Thời gian (giờ)

Hình 6. Độ biến thiên áp suất mực biển theo thời gian dự báo trong ba trường hợp: không ban đầu hoá NOI, ban đầu hoá theo phương pháp NMI và ban đầu hoá bằng lọc số DFI

3. Ban đầu hóa bằng lọc số cho mô hình HRM

Trong phần dưới đây, các kết quả dự báo từ mô hình HRM với trường ban đầu được ban đầu hóa theo phương pháp lọc số DFI sẽ được trình bày. So sánh các kết quả này với các kết quả dự báo khi thực hiện ban đầu hóa theo phương pháp cũ NMI sẽ cho ta một đánh giá về hiệu năng của lọc số khi sử dụng cho mô hình HRM. Giá trị được chọn đối với tần số cắt tương ứng với chu kỳ $\tau_c = 6h$. Giá trị này được chọn trên cơ sở thử nghiệm với mô hình chính áp (Lynch 1990, [5]). Nếu sử dụng chu kỳ là 3h, ta cũng nhận được kết quả tương tự. Song nếu sử dụng chu kỳ là 12h lại dẫn đến sự tắt dần của những sóng khí tượng thông thường. Thời gian tích phân tiến cũng như lùi là 3h.

a. Mô hình HRM

HRM là mô hình gồm các phương trình nguyên thủy cho khu vực hạn chế, sử dụng sơ đồ sai phân hữu hạn bậc hai

trên lưới xen C và hệ tọa độ lai chạy nghiệp vụ trên hệ máy tính song song hiệu năng cao tại Trung tâm Dự báo KTTV Trung ương. Mô hình được mô tả đầy đủ trong các tài liệu [1], [2], [3]... Sơ đồ ban đầu hoá các thành phần chính phi tuyến (NMI) hiện đang sử dụng trong nghiệp vụ được mô tả chi tiết trong chương III của công trình [1]. Độc giả quan tâm có thể tham khảo trong các tài liệu này.

Mô hình toàn cầu GME của CHLB Đức (sử dụng phương pháp lọc số trong ban đầu hoá trường phân tích) cung cấp trường ban đầu và biên cho mô hình HRM [7]. Để thống nhất trong việc ban đầu hoá của cả hai mô hình GME và HRM, các tác giả đã thử nghiệm áp dụng sơ đồ lọc số DFI cho mô hình HRM. Sơ đồ được thử nghiệm cho số liệu 00Z trong 6 tháng (từ tháng VII đến tháng XII/2004). Ở đây, tác giả chỉ trích dẫn đặc điểm tình hình thời tiết trong tháng VII (tháng có đợt mưa từ 19 đến 24/VI với tổng lượng mưa và cường độ lớn nhất trong năm tại Bắc Bộ) và tháng XI (đợt mưa từ 23 đến 28/XI với tổng lượng mưa và cường độ lớn nhất trong năm tại Trung Bộ).

b. Diễn biến thời tiết trong tháng VII/2004

Trong tháng VII có ba đợt mưa, tập trung chủ yếu vào nửa tháng sau và tại các khu vực Bắc Bộ, Bắc Trung Bộ. Lượng mưa phổ biến từ 100 đến 300mm. Có một số nơi mưa to như Hải Dương (461mm), Đông Xuyên (Hải Phòng): 532mm. Hình thể synopt của giai đoạn từ 19 đến 24 đặc trưng bởi hoạt động của rãnh thấp có trục theo

hướng Tây Bắc - Đông Nam ở tầng thấp, kết hợp với rãnh thấp trên cao nên vùng núi và đông bắc Bắc Bộ có mưa rào và dông. Kể từ sau ngày 19, rãnh áp thấp mạnh dần lên, áp cao cận nhiệt đới có xu hướng lẩn về phía tây. Ngày 20 đến 21: rãnh áp thấp có trục tây bắc đông nam vắt qua Bắc Bộ kết hợp với hội tụ mạnh của đới gió đông nam tồn tại từ mặt đất lên 5.000m, áp cao cận nhiệt đới vẫn tiếp tục lẩn dần về phía tây. Tới ngày 23: dải hội tụ nhiệt đới (ITCZ) dịch dần về Bắc Trung Bộ.

c. Diễn biến thời tiết trong tháng XI/2004

Theo quy luật hàng năm, khu vực từ Thừa Thiên - Huế đến Khánh Hoà thường có những đợt mưa có cường độ và tổng lượng mưa lớn nhất so với các vùng khác trên cả nước, càng về cuối mùa mưa, trên khu vực này các đợt mưa càng lớn và dữ dội hơn, điển hình là đợt mưa từ 23 đến 28/XI trên khu vực từ Hà Tĩnh đến Khánh Hòa, do ảnh hưởng kết hợp của không khí lạnh và hội tụ gió của rìa tây bắc cơn bão số 4, gây ra đợt mưa lớn diện rộng với tổng lượng mưa phổ biến từ 300 đến 500mm, riêng Thừa Thiên - Huế từ 900 đến 1.200mm, một số nơi lớn hơn như: Huế (trạm khí tượng) 1.257mm; Nam Đông (trạm thủy văn) 1.265mm, Nam Đông (trạm khí tượng) 1.415mm.

d. Kết quả dự báo

Dự báo chính xác mưa phụ thuộc chủ yếu vào dự báo chính xác chuyển động khí quyển và lượng ẩm. Như vậy, dự báo tốt mưa trên một vùng lớn có nghĩ khác. Vì thế mà nhiều trung tâm dự báo lớn sử dụng kỹ năng dự báo định lượng mưa (QDF) làm số đo tiêu biểu của khả năng mô hình [4]. Báo cáo cũng sử

dụng các chỉ số thống kê thường được dùng để đánh giá chất lượng của mô hình [8]. Dưới đây là một số định nghĩa của các chỉ số dự báo thông dụng:

Các kí hiệu:

- A: số trường hợp dự báo đúng (cho các hiện tượng có xảy ra).

- B: số trường hợp dự báo không.

- C: số trường hợp dự báo sót.

- D: số trường hợp dự báo đúng (cho các hiện tượng không xảy ra).

- E: dự báo đúng ngẫu nhiên, $E=(A+B)(A+C)/N$.

- N: tổng số trường hợp (bằng tổng của A, B, C, D).

FB (Frequency Bias Index): chỉ số đánh giá tần số dự báo, $0 \leq FB \leq \infty$.

$$FB = \frac{A+B}{A+C} \quad (25)$$

TS (Threat Score): chỉ số đánh giá tổng thể dự báo xuất hiện hiện tượng có tính đến dự báo không và dự báo sót, $0 \leq TS \leq 1$.

$$TS = \frac{A}{A+B+C} \quad (26)$$

ETS (Equitable Threat Score): tương tự như chỉ số TS nhưng loại bỏ dự báo đúng do ngẫu nhiên: $1/3 \leq ETS \leq 1$.

$$ETS = \frac{A-E}{A+B+C-E} \quad (27)$$

POD (Probability of Detection): Xác suất phát hiện hiện tượng: $0 \leq POD \leq 1$.

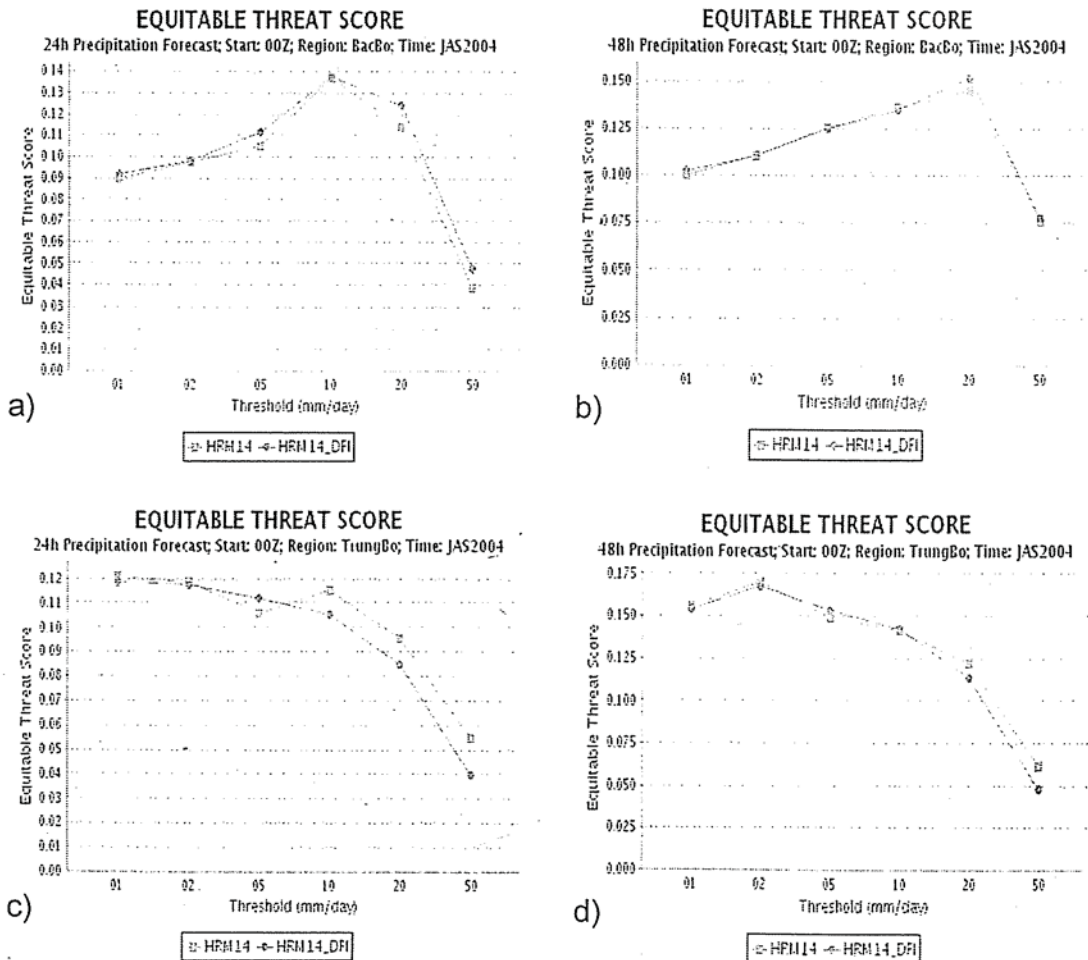
$$POD = \frac{A}{A+C} \quad (28)$$

FAR (False Alarm Ratio): Tỷ lệ cảnh báo sai: $0 \leq FAR \leq 1$.

$$FAR = \frac{B}{A + B} \quad (29)$$

Các tác giả đã thử nghiệm áp dụng lọc số DFI cho sơ đồ ban đầu hoá của mô hình HRM độ phân giải ngang 14km, phân giải thẳng đứng 31 mực (kí hiệu H14_DFI) cho 6 tháng cuối năm 2004, chia làm hai giai đoạn: mùa mưa cho khu vực Bắc Bộ: tháng VII, VIII và IX (JAS); mùa mưa cho khu vực Trung và Nam Bộ: tháng X, XI và XII (OND). Sau đó, tiến hành đánh giá kỹ năng dự

báo của từng sơ đồ (HRM14_DFI và mô hình nghiệp vụ với độ phân giải ngang 14km, phân giải thẳng đứng 31 mực, sử dụng sơ đồ ban đầu hoá các thành phần chính phi tuyến NMI, kí hiệu HRM14) theo các chỉ số thống kê FB, TS, ETS, POD và FAR. Vì kết quả đánh giá rất nhiều nên ở đây chỉ đưa ra minh hoạ một chỉ số, đó là ETS cho Bắc Bộ và Trung Bộ. Kết quả được trình bày trong các hình 7 (cho JAS) và hình 8 (cho OND). ETS càng lớn thì kỹ năng dự báo càng tốt và ngược lại.

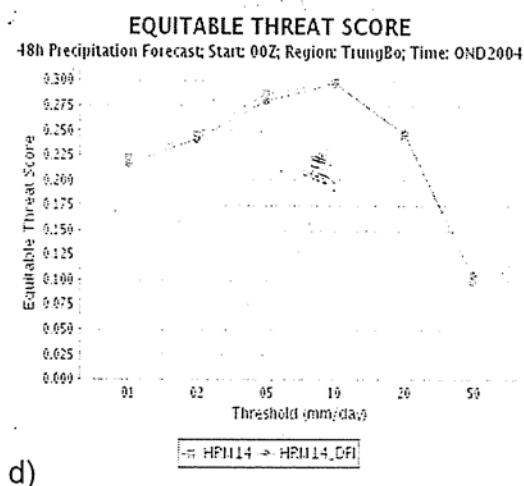
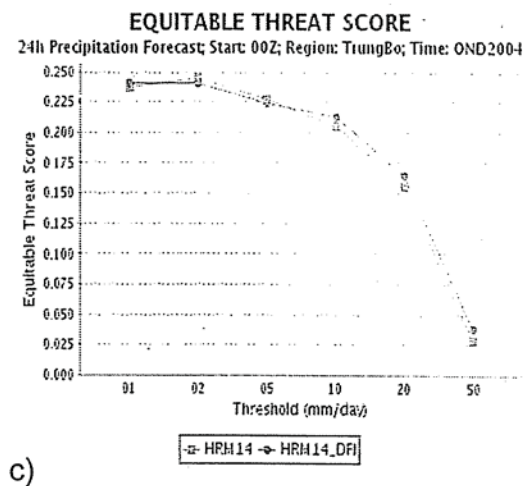
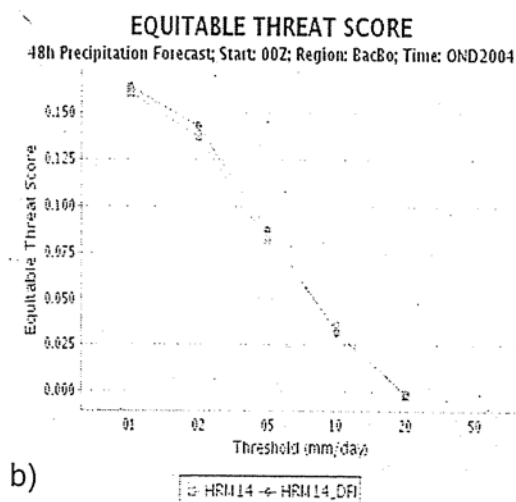
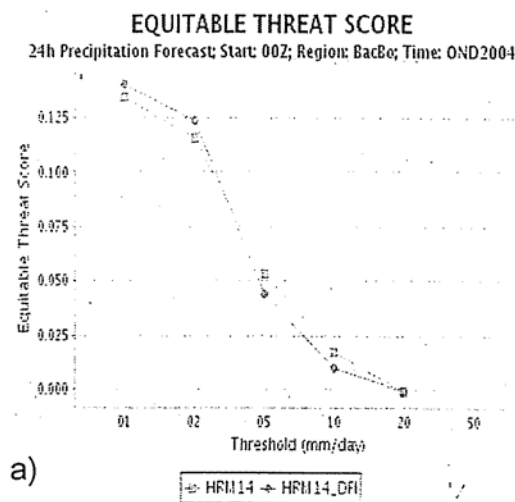


Hình 7. Diễn biến theo ngưỡng của chỉ số thống kê ETS của dự báo mưa 24h và 48h của HRM14, HRM14_DFI, lấy trung bình cho khu vực Bắc Bộ (a,b), Trung Bộ (c,d) cho 3 tháng VII, VIII và IX/2004.

Phân tích kết quả cho thấy, trong ba tháng mùa mưa ở Bắc Bộ, sơ đồ lọc số DFI tỏ ra có kỹ năng dự báo tốt hơn sơ đồ NMI ở các ngưỡng mưa lớn (trên 20mm/24 giờ), đặc biệt cho hạn dự báo 24 giờ (hình 7 a). Ngược lại, tại Trung Bộ, sơ đồ này tỏ ra chiếm ưu thế hơn cho ba tháng X, XI và XII (hình 8, c). Để củng cố thêm cho nhận xét này, tác giả đã tiến hành đánh giá riêng tháng VII cho Bắc Bộ (có ba đợt mưa

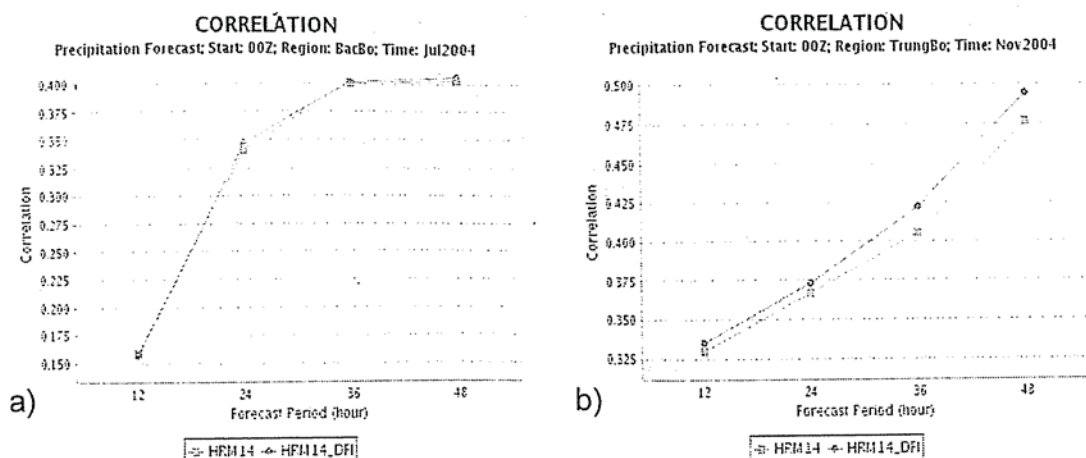
lớn) và tháng XI cho Trung Bộ (có hai đợt mưa lớn).

Kết quả được minh họa trên hình 9. Ta thấy rất rõ là hệ số tương quan đều tăng khi hạn dự báo tăng, nhất là đối với khu vực Trung Bộ, tại thời điểm dự báo 48 giờ, hệ số tương quan là 0,475 đối với HRM14 và 0,5 đối với HRM14_DFI (hình 9 b).



Hình 8. Diễn biến theo ngưỡng của chỉ số thống kê ETS của dự báo mưa 24h và 48h của HRM14, HRM14_DFI, lấy trung bình cho khu vực Bắc Bộ (a,b), Trung Bộ (c,d) cho 3 tháng X, XI và XII/2004

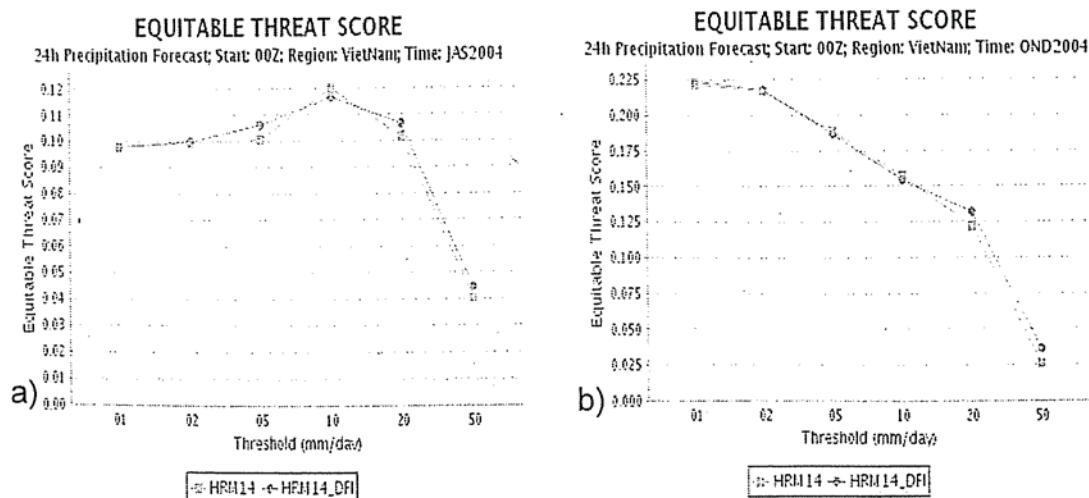
NGHIÊN CỨU & TRAO ĐỔI



Hình 9. Hệ số tương quan giữa lượng mưa quan trắc và lượng mưa dự báo theo hạn dự báo 12, 24, 36 và 48 giờ của HRM14 và HRM14_DFI cho Bắc Bộ (VII/2004, hình a) và Trung Bộ (XI/2004, hình b).

Nhìn chung cho toàn bộ Việt Nam (hình 10) đối với các ngưỡng mưa nhỏ (nhỏ hơn 5mm/24 giờ), sơ đồ DFI và NMI gần như không có sự khác biệt. Đối với ngưỡng mưa lớn, đặc biệt là từ 20-50mm/24 giờ, sơ đồ DFI thể hiện rõ tính ưu việt của mình, các nhiễu động có tần số cao giả đã bị loại bỏ, cho ta xu

thế dự báo chính xác hơn, phù hợp với những khu vực có địa hình phức tạp. Trong thời gian tới, các tác giả sẽ tiến hành thử nghiệm ứng dụng sơ đồ DFI với nhiều trường hợp, đồng thời áp dụng các hướng cải tiến khác với mục đích nâng cao hơn nữa chất lượng dự báo của mô hình HRM.



Hình 10. Diễn biến theo ngưỡng của chỉ số thống kê ETS của dự báo mưa 24 giờ của HRM14 và HRM14_DFI, lấy trung bình cho Việt Nam cho 3 tháng VII, VIII, và IX (hình a), và 3 tháng X, XI, XII/2004 (hình b).

Tài liệu tham khảo

1. Lê Công Thành và cctv. Xây dựng hệ thống nghiệp vụ HRM trên hệ máy tính song song hiệu năng cao. Báo cáo kết quả thực hiện Dự án ứng dụng tiến bộ kỹ thuật cấp Bộ, năm 2003, 146 tr.
2. Kiều Thị Xin, Lê Công Thành, Phan Văn Tân, Đỗ Lệ Thủy và Nguyễn Văn Sáng. Mô hình dự báo số phân giải cao HRM và thử nghiệm áp dụng dự báo thời tiết khu vực Đông Nam Á - Việt Nam. *Tạp chí Khí tượng Thủy văn*, số 8(488)/2001, Tổng cục KTTV, tr. 36-44.
3. Kiều Thị Xin và cctv. Nghiên cứu ứng dụng bộ mô hình số trị khu vực cho dự báo chuyển động của bão ở Việt Nam. Báo cáo kết quả thực hiện đề tài KHCN độc lập cấp Nhà nước, năm 2002, 184 tr.
4. Kiều Thị Xin, Lê Đức và Vũ Thanh Hằng. Cải tiến mô hình dự báo thời tiết phân giải cao HRM cho dự báo mưa lớn gây lũ lụt ở Việt Nam. Báo cáo tại Hội nghị Khoa học công nghệ dự báo và Phục vụ dự báo lần thứ VI, Tập I, Trung tâm Dự báo KTTV Trung ương, năm 2005, trang 1-14.
5. Lynch, Peter, Initialization using a digital filter. In research Activities in Atmospheric and Ocean Modeling, CAS/JSC Working Group on Numerical Experimentation. Report No 14, WMO Secretariat, Geneva, 1.5-1.6, 1990.
6. Lynch, P. and X. Y, Huang, Initialization of the HIRLAM Model using a Digital Filter, *Amer. Met. Soc.*, 120, 1019-1034, 1992.
7. Quarterly Report of the German NWP-System. Part 2: Description of the German NWP-System, DWD, Business Area Research and Development, December 2002.
8. COST17-Working group 2 Working Area2: Approaches to verification: Clive Wilson, Meteorological Office, UK.