

# ÁP DỤNG PHƯƠNG PHÁP TÍCH PHÂN HỒI QUI BỘI VÀ MÔ HÌNH ARIMA TRONG VIỆC DỰ BÁO NĂNG SUẤT LÚA Ở HÀ NỘI

KS. Dương Văn Khảm  
Viện Khí tượng Thủy văn

## 1. Đặt vấn đề

Những năm gần đây đã có nhiều đề tài vận dụng các mô hình thống kê để dự báo năng suất và sản lượng cây trồng nói chung và cây lúa nói riêng ở các tỉnh, đã góp phần đắc lực phục vụ sự phát triển của ngành nông nghiệp và bảo đảm an ninh lương thực quốc gia.

Tuy nhiên các phương pháp dự báo thống kê luôn chịu sự ảnh hưởng của nhiều yếu tố làm giảm độ tin cậy của mô hình, ví dụ mức độ ổn định của phép ngoại suy năng suất xu thế, ý nghĩa vật lý và sinh vật cũng như phương pháp chọn các yếu tố trong mô hình dự báo, sự không đồng nhất của các giả thuyết thống kê v.v...

Để hạn chế những tồn tại của các phương pháp dự báo thống kê nhằm từng bước nâng cao độ tin cậy của mô hình, chúng tôi vận dụng phương pháp tích phân hồi qui bội và mô hình ARIMA trong việc dự báo năng suất lúa ở Hà nội.

## 2. Các thành phần của năng suất

Như đã biết năng suất thực tế ( $y_t$ ) (gọi tắt là năng suất) của cây trồng bao gồm hai thành phần. 1) Phần năng suất do trình độ thăm canh, sự thay đổi và gia tăng về giống, phân bón, cải tạo đất, tưới tiêu ... tạo nên gọi là năng suất xu thế ( $f_t$ ). 2) Phần năng suất do các điều kiện thời tiết khí tượng tạo nên gọi là năng suất thời tiết ( $p_t$ ). Về mặt toán học năng suất ( $y_t$ ) được thể hiện bằng công thức sau đây:

$$y_t = f_t + p_t \quad (1)$$

Trên thực tế, năng suất xu thế và năng suất thời tiết trong quá trình mô phỏng bao giờ cũng có độ lệch so với năng suất thực của nó, vì vậy tác giả định nghĩa độ lệch của năng suất với năng suất xu thế mô phỏng và năng suất thời tiết mô phỏng gọi là năng suất dư ( $\omega_t$ ).

$$\omega_t = y_t - \hat{f}_t - \hat{p}_t \quad (2)$$

Trong đó

$\hat{f}_t$ ,  $\hat{p}_t$  là năng suất xu thế và năng suất thời tiết mô phỏng.

Như vậy có thể cho rằng năng suất xã hội bao gồm 3 thành phần, năng suất xu thế  $f_t$ , năng suất thời tiết  $p_t$ , và năng suất dư  $\omega_t$ .

Dưới đây trình bày các phương pháp dự báo 3 thành phần năng suất trên của lúa ở Hà Nội với 2 vụ đông xuân và mùa.

## 3. Mô tả mô hình dự báo

### a. Mô hình dự báo năng suất xu thế

Mô hình dự báo năng suất xu thế có nhiều phương pháp ngoại suy, như trung bình trượt, trượt có trọng lượng điều hoà, các phương pháp ngoại suy hồi qui tuyến

tính, phi tuyến tính .... Để xây dựng mô hình cần căn cứ vào thực tế chuỗi số liệu để chọn phương pháp phù hợp.

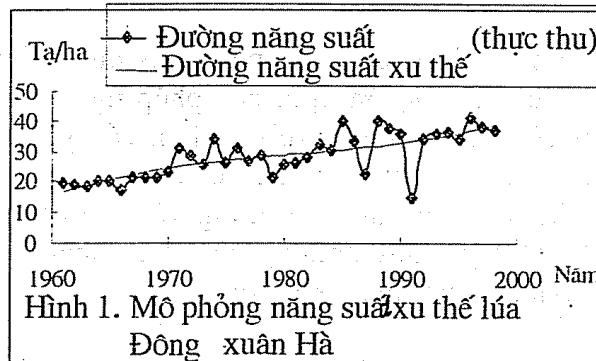
Với năng suất lúa ở Hà Nội áp dụng phương pháp hồi qui phi tuyến tính. Năng suất xu thế của vụ đông xuân và vụ mùa được mô phỏng theo các phương trình:

$$f_{\text{v}_0} = 0,0007t^3 - 0,0439t^2 + 1,27t + 15,616 + \varepsilon_t \quad (\text{Vụ đông xuân})$$

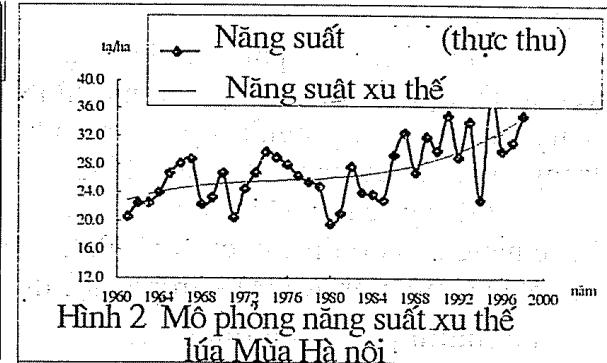
$$f_{\text{v}_1} = 0,0006t^3 - 0,0295t^2 + 0,513t + 22,415 + \varepsilon_t \quad (\text{Vụ mùa})$$

Trong đó  $t$  là thời gian (năm). Để các giá trị tham số của các phương trình trên không quá lớn qui ước  $t=1$  là năm 1961.

Năng suất xu thế có dạng theo các hình (1, 2).



Hình 1. Mô phỏng năng suất xu thế lúa  
Đông xuân Hà Nội



Hình 2 Mô phỏng năng suất xu thế  
lúa Mùa Hà nội

Các sai số ngẫu nhiên  $\varepsilon_t$ , sau khi thông qua kiểm nghiệm (dựa vào hàm tự hồi qui của chúng), thấy rằng chúng hoàn toàn là hàm ổn định ngẫu nhiên. Vì vậy các phương trình mô phỏng năng suất xu thế trên được chấp nhận.

### b. Mô hình dự báo năng suất thời tiết

Theo lý thuyết thống kê nếu sử dụng phương pháp hồi qui để phân tích tương quan thì độ dài của chuỗi phải lớn gấp từ 5 đến 12 lần số nhân tố, ví dụ chia cả quá trình phát triển của cây lúa từ khi gieo đến thu hoạch theo tuần 10 ngày sẽ có khoảng từ 12 đến 17 thời đoạn (hay 12 đến 17 nhân tố). Nếu mới chỉ xét 3 yếu tố nhiệt độ, lượng mưa và số giờ nắng với năng suất lúa đã có khoảng 36 nhân tố, như vậy độ dài chuỗi ít nhất cũng phải là 180 năm, trên thực tế hiện nay ở Việt Nam đối với số liệu năng suất cây trồng khó có chuỗi với một độ dài quan trắc như vậy. Vì vậy dùng phương pháp tích phân hồi qui bội có sử dụng hàm đa thức trực giao gốc sẽ hạn chế được vấn đề trên.

Nguyên lý cơ bản của phương pháp tích phân hồi qui bội như sau:

Giả sử thời gian  $t$  (từ 0 đến  $\tau$ ) là toàn bộ thời gian sinh trưởng và phát triển (từ khi gieo hạt đến khi chín) của cây trồng. Phân chia toàn bộ thời gian  $\tau$  thành các thời đoạn mỗi thời đoạn có độ dài là  $\Delta t$ . Chúng có thể có độ dài bằng nhau hoặc khác nhau. Năng suất thời tiết cây trồng dưới sự tác động của các yếu tố khí tượng được hình thành trong mỗi thời đoạn  $\Delta t$  là  $\Delta p$ . Như vậy năng suất thời tiết của cả quá trình sinh trưởng và phát triển ( $p$ ) của cây trồng là tổng của sự hình thành năng suất mỗi thời đoạn, tức là:

$$p = \Delta p_0 + \Delta p_1 + \dots + \Delta p_\tau = \sum_{t=0}^{\tau} \Delta p_t \quad (3)$$

Do đó đơn vị năng suất thời tiết ( $p$ ) tăng thêm trên một đơn vị thời gian là  $\Delta p/\Delta t$ . Mỗi một đơn vị năng suất thời tiết  $\Delta p/\Delta t$  đều chịu sự tác động tổng hợp của các

yếu tố khí tượng  $x_j(t)$ , tức là tồn tại quan hệ  $\Delta p/\Delta t \sim x_j(t)$ . Nếu quan hệ này là thực ta có

$$\Delta p/\Delta t = a_j(t) x_j(t) \quad (4)$$

Trong đó  $x_j(t)$  ( $j=1,2,\dots,m$ ;  $t=1,2,\dots,n$ ) là trị số của yếu tố khí tượng  $j$  tại thời điểm  $\Delta t$ ;  $a_j(t)$  là hàm tác động của yếu tố khí tượng  $x_j(t)$  tới năng suất khí tượng ( $p$ ). Gọi là hàm sinh trưởng. Khi  $\Delta t \rightarrow 0$  biểu thức (4) viết dưới dạng vi phân là

$$dp = a_j(t) x_j(t) dt \quad (4)$$

Tích phân 2 vế được

$$p = \int_0^t dp = \int_0^t a_j(t) x_j(t) dt + C_0 \quad (5)$$

Biểu thức trên gọi là tích phân hồi qui,  $C_0$  là hằng số tích phân.

Hoặc có thể viết dưới dạng tích phân hồi qui bội như sau:

$$p = \sum_{j=1}^m \int_0^t a_j(t) x_j(t) dt + C_0 \quad (6)$$

Trong đó dấu  $\Sigma$  chỉ sự ảnh hưởng tổng hợp của các yếu tố ( $j = 1,2,\dots,m$ ) đến năng suất cây trồng.

Để giải phương trình tích phân trên trước tiên ta phân giải hàm ảnh hưởng  $a_j(t)$  thành 2 thành phần: một thành phần thay đổi theo thời gian  $\phi_k(t)$  và thành phần không đổi  $\alpha_{jk}$ ; tức:

$$a_j(t) = \alpha_{j0} + \alpha_{j1}\phi_1(t) + \dots + \alpha_{jq}\phi_q(t) = \sum_{k=0}^q \alpha_{jk}\phi_k(t) \quad (7)$$

Như vậy ta có thể dùng hàm đa thức trực giao gốc bội để thay thế vào thức (7), với  $\phi_k(t)$  là trị số của đa thức trực giao;  $k$  là cấp số của đa thức trực giao,  $\alpha_k$  là hệ số của đa thức.

Do đó (6) được viết lại là

$$p = \sum_{j=1}^m \int_0^t x_j(t) \sum_{k=0}^q \alpha_{jk} \phi_k(t) dt + C_0 = C_0 + \sum_{j=1}^m \sum_{k=0}^q \alpha_{jk} \int_0^t x_j(t) \phi_k(t) dt \quad (8)$$

Ta đặt phần tích phân

$$\int_0^t x_j(t) \phi_k(t) dt = H_{jk}(t). \quad (9)$$

Thay vào (8) được

$$p = \sum_{j=1}^m \int_0^t dp = C_0 + \sum_{j=1}^m \sum_{k=0}^q \alpha_{jk} H_{jk}(t). \quad (10)$$

$$H_{jk}(t) = \sum_{t=0}^{\tau} x_{jt} \Phi_{jk}(t). \quad (11)$$

$x_j(t)$  là một biến lượng rời rạc  $x_{jt}$  do vây tích phân biến lượng  $H_{jk}$  ( $j=1, n$  độ dài của chuỗi) có thể viết dưới dạng:

Với biểu thức (10) có thể hoàn toàn áp dụng phương pháp hồi qui bội để tìm các giá trị  $C_0$  và  $\alpha_{jk}$ .

Việc sử dụng phương pháp tích phân hồi qui bội ngoài việc xây dựng được mô hình dự báo năng suất, thông qua giá trị hàm  $a_j(t)$  còn cho ta biết được mỗi yếu tố khí tượng ở mỗi thời đoạn tăng hoặc giảm 1 đơn vị sẽ làm tăng hoặc giảm bao nhiêu đơn vị năng suất.

#### Xây dựng mô hình dự báo năng suất lúa thời tiết ở Hà Nội

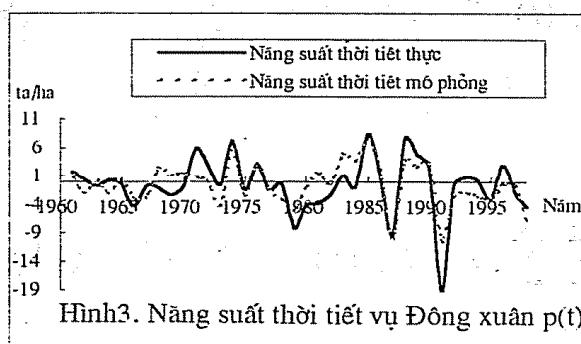
Đối với lúa đông xuân, chia cả quá trình sinh trưởng và phát triển (lấy trị số trung bình ở thời vụ chính) thành 17 thời đoạn mỗi thời đoạn là 1 tuần (10 ngày), thời đoạn 1 là tuần thứ 2 của tháng 12 (thời kỳ gieo), thời đoạn 17 là tuần cuối của tháng 5 (thời kỳ lúa chín). Chọn các phần tử đưa vào mô hình là năng suất lúa thời tiết, nhiệt độ trung bình tuần, nhiệt độ tối cao trung bình tuần, nhiệt độ tối thấp trung bình tuần, tổng lượng mưa tuần và tổng số giờ nắng tuần với độ dài quan trắc từ năm 1961 đến 1998. Như vậy tổng số nhân tố là 85 và độ dài chuỗi là 38.

Đối với vụ mùa do thời gian sinh trưởng ngắn hơn vụ đông xuân vì vậy chia cả quá trình sinh trưởng và phát triển thành 12 thời đoạn, thời đoạn 1 là tuần 2 của tháng 6 (thời kỳ gieo), thời đoạn 12 là tuần 1 của tháng 10 (thời kỳ chín). Các phần tử của mô hình và độ dài chuỗi cũng tương tự như vụ đông xuân. Sử dụng phương pháp trên, thông qua F kiểm nghiệm với mức tin cậy  $\alpha=0,1$  đã nhận được mô hình dự báo năng suất thời tiết của vụ đông xuân và mùa lần lượt là:

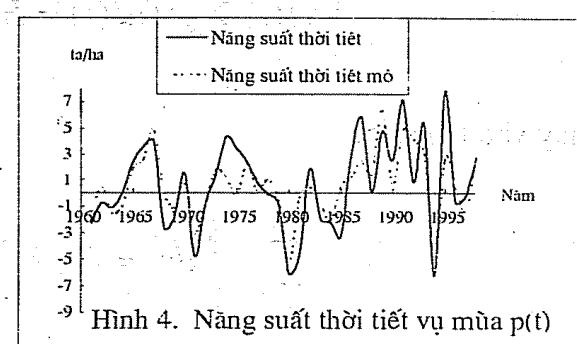
$$\begin{aligned} P_t &= 60,60477 - 0,19266H_{u,0}(t) - 0,000735H_{r,3}(t) - 0,002378H_{s,3}(t) - 0,000677H_{s,4}(t) \\ P_t &= 6,8947 + 0,00877H_{t,5}(t) - 0,0112217H_{tc,3}(t) + 0,0052319H_{tc,5}(t) - 0,0054774H_{r,0}(t) + \\ &0,0001368H_{r,2}(t) + 0,0004655H_{s,2}(t) + 0,0007445H_{s,3}(t) \end{aligned}$$

Trong đó  $H$  là nhân tố thứ cấp đã được biến đổi theo công thức (11) với các chỉ số bằng chữ dưới  $H_{(u, tc, t, r, s)}$  lần lượt là các yếu tố khí tượng nhiệt độ trung bình tuần, nhiệt độ tối cao trung bình tuần, nhiệt độ tối thấp trung bình tuần, tổng lượng mưa tuần và tổng số giờ nắng tuần; các chỉ số bằng số dưới  $H_{(0, 1, 2, 3, 4, 5)}$  là các thành phần trực giao đầu tiên.

Kết quả của mô hình được biểu hiện ở các hình 3, 4



Hình 3. Năng suất thời tiết vụ Đông xuân  $p(t)$



Hình 4. Năng suất thời tiết vụ mùa  $p(t)$

### c. Mô hình dự báo năng suất dư

Năng suất dư được dự báo bằng mô hình ARIMA.

Sau khi phân tích số liệu, chọn các trị số p, d, q thông qua các bước kiểm nghiệm cần thiết đã xác định được mô hình dự báo năng suất dư của vụ lúa đông xuân và mùa tại Hà Nội như sau:

Đối với vụ đông xuân mô hình có dạng:

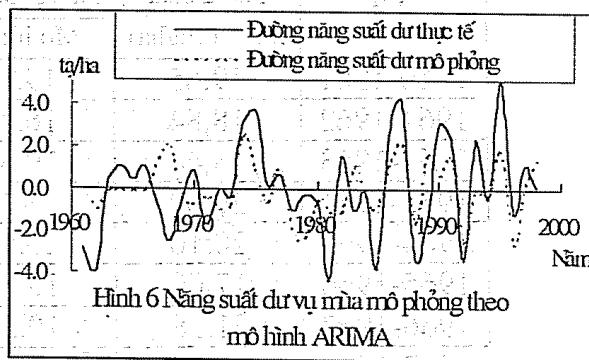
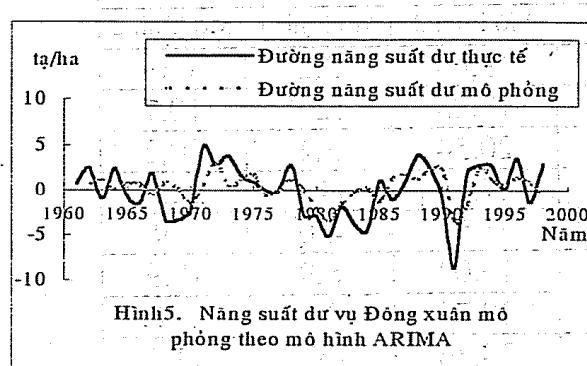
$$(1-B)(1-0,29046B)(1-0,76674B^3)\omega_t = (1-B)(1-0,97993B^3)a_t - 0,0250044$$

Đối với vụ mùa mô hình có dạng:

$$\omega_t = (1-0,5336B^6+0,4664B^{11})a_t - 0,09735$$

Trong đó  $\omega_t$  là năng suất dư; B là phép tính chuyển lùi  $B\omega_t = \omega_{t-1}$ ;  $a_t$  là sai số ngẫu nhiên;

Mô hình dự báo năng suất dư vụ đông xuân và vụ mùa được thể hiện trên các hình 5, 6



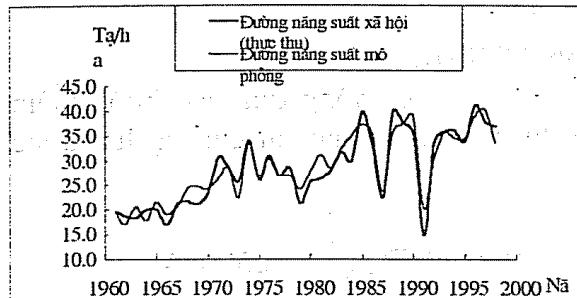
Theo kết quả tính toán trên mô hình (bảng 1, 2) tỷ lệ sai số phần trăm so với năng suất thực ở mức dưới 5% của 38 năm quan trắc của cả 2 mô hình, khi chưa có mô phỏng năng suất dư (gọi là mô hình 1) và có mô phỏng năng suất dư (gọi là mô hình 2) khoảng 40% ở vụ đông xuân và trên 50% ở vụ mùa.

Đối với vụ đông xuân với các mức sai số thấp dưới 5% thì mô hình 2 cao hơn mô hình 1, nhưng với các mức sai số cao trên 15% thì tỷ lệ sai số của mô hình 1 cao hơn mô hình 2. Tuy cả 2 mô hình đều có 1 năm mô phỏng với sai số ở mức trên 20% (năm mất mùa lớn 1990-1991) nhưng sai số của mô hình 1 so với năng suất thực là 57.7% còn của mô hình 2 đã giảm xuống còn sai số 25.9%.

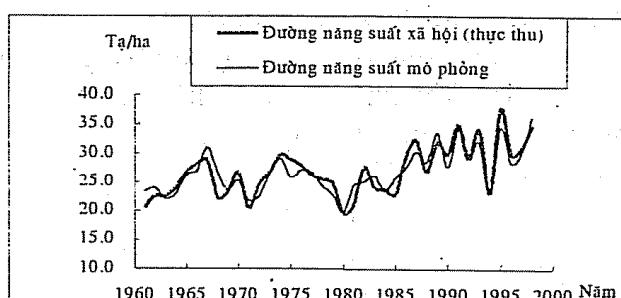
Đặc biệt với vụ mùa khi áp dụng mô hình 2 không còn sai số ở mức trên 15%, còn sai số ở mức từ 11%-15% chỉ còn chiếm 11% tổng số năm mô phỏng.

Tính độ tán xạ của cả 2 mô hình so với năng suất thực thu thấy rằng: vụ đông xuân mô hình 1 tán xạ là 2.9 tạ/ha trong khi đó của mô hình 2 là 2.5 tạ/ha. Vụ mùa mô hình 1 tán xạ là 2.3 tạ/ha trong khi đó của mô hình 2 là 1.7 tạ/ha. Như vậy, cả hai vụ tán xạ của mô hình 2 đều nhỏ hơn mô hình 1 từ 0,4 đến 0,6 tạ/ha.

Có thể nhận xét với sự áp dụng mô hình ARIMA việc mô phỏng thêm phần năng suất dư đã nâng cao đáng kể đến độ tin cậy của mô hình. Đặc biệt đã giảm được nhiều những sai số ở mức cao, đối với vụ đông xuân khoảng 3%, đối với vụ mùa khoảng 5%. Có nghĩa là đã nâng cao được độ tin cậy của mô hình từ 3% đến 5%



Hình 7. Mô phỏng dự báo năng suất lúa Đông xuân



Hình 8. Mô phỏng dự báo năng suất lúa Mùa tại Hà Nội

Bảng 1. So sánh kết quả mô phỏng của mô hình 1 và mô hình 2 với năng suất xã hội của vụ đông xuân ở Hà Nội

Vụ	Năng suất xã hội (tạ/ha)	Năng suất mô phỏng (tạ/ha)		Sai số (%)	
		Mô hình 1	Mô hình 2	Mô hình 1	Mô hình 2
1960-1961	19,65	19,01	19,65	3,2	0,0
1961-1962	18,84	16,41	17,03	12,9	10,7
1962-1963	18,44	19,43	20,54	5,4	10,2
1963-1964	20,09	17,72	17,82	11,8	12,7
1964-1965	20,10	20,91	21,59	4,0	6,9
1965-1966	17,00	18,51	19,08	8,9	10,9
1966-1967	21,28	19,37	20,79	9,0	2,4
1967-1968	21,65	25,06	24,36	15,8	11,1
1968-1969	21,22	24,60	24,70	15,9	14,1
1969-1970	23,47	25,83	24,23	10,0	3,1
1970-1971	30,71	25,84	26,36	15,9	16,5
1971-1972	28,59	25,71	28,72	10,1	0,4
1972-1973	25,73	21,93	22,47	14,8	14,5
1973-1974	34,02	32,42	33,53	4,7	1,5
1974-1975	26,05	25,14	26,86	3,5	3,0
1975-1976	31,00	31,03	30,37	0,1	2,1
1976-1977	26,96	27,06	26,94	0,4	0,1
1977-1978	28,55	25,84	26,93	9,5	6,0
1978-1979	21,37	24,26	24,19	13,5	11,7
1979-1980	25,56	28,48	27,54	11,4	7,2
1980-1981	26,36	31,52	31,17	19,6	15,4
1981-1982	28,20	30,11	28,53	6,8	1,2
1982-1983	31,90	35,73	32,45	12,0	1,7
1983-1984	30,10	34,69	34,96	15,3	13,9
1984-1985	39,80	38,78	37,42	2,6	6,4
1985-1986	33,10	34,19	35,45	3,3	6,6
1986-1987	22,60	21,88	23,61	3,2	4,3
1987-1988	39,90	36,07	35,70	9,6	11,8
1988-1989	37,68	35,22	37,53	6,5	0,4

Tiếp theo bảng 1.

Vụ	Năng suất xã hội (tạ/ha)	Năng suất mô phỏng (tạ/ha)		Sai số (%)	
		Mô hình 1	Mô hình 2	Mô hình 1	Mô hình 2
1989-1990	35,79	36,56	39,00	2,1	8,2
1990-1991	15,00	23,66	20,25	57,7	25,9
1991-1992	33,70	31,87	30,18	5,4	11,7
1992-1993	36,00	33,23	35,70	7,7	0,8
1993-1994	36,20	33,52	34,68	7,4	4,4
1994-1995	34,00	33,95	34,53	0,1	1,5
1995-1996	41,30	37,79	39,14	8,5	5,5
1996-1997	37,90	39,34	40,22	3,8	5,8
1997-1998	37,10	34,26	33,71	7,7	10,1

Bảng 2 . So sánh kết quả mô phỏng của mô hình 1 và mô hình 2 với  
năng suất xã hội của vụ mùa ở Hà Nội

Vụ	Năng suất xã hội (tạ/ha)	Năng suất mô phỏng (tạ/ha)		Sai số (%)	
		Mô hình 1	Mô hình 2	Mô hình 1	Mô hình 2
1961	20,61	23,50	23,40	14,0	11,9
1962	22,59	26,46	25,02	17,1	10,1
1963	22,59	22,19	22,09	1,8	2,3
1964	24,02	23,03	22,93	4,1	4,7
1965	26,72	26,28	26,18	1,6	2,1
1966	28,13	27,12	27,03	3,6	4,1
1967	28,80	29,63	31,02	2,9	7,2
1968	22,35	24,88	23,63	11,3	5,4
1969	23,26	23,91	23,55	2,8	1,2
1970	26,75	25,94	25,26	3,0	5,9
1971	20,54	22,25	21,87	8,3	6,1
1972	24,55	24,57	23,58	0,1	3,7
1973	26,78	27,16	26,49	1,4	1,1
1974	29,76	26,64	29,15	10,5	2,1
1975	28,99	25,33	25,90	12,6	11,9
1976	27,96	27,77	27,13	0,7	3,0
1977	26,35	25,64	26,77	2,7	1,6
1978	25,51	26,58	24,40	4,2	4,6
1979	24,86	25,10	22,77	1,0	9,2
1980	19,63	20,29	19,73	3,4	0,5
1981	21,05	25,37	24,32	20,5	13,4
1982	27,70	26,23	25,17	5,3	10,0
1983	24,00	25,07	26,11	4,5	8,1
1984	23,80	23,93	23,37	0,5	1,8
1985	22,90	26,62	25,70	16,3	10,9

Tiếp theo bảng 2

Vụ	Năng suất xã hội (tạ/ha)	Năng suất mô phỏng (tạ/ha)		Sai số (%)	
		Mô hình 1	Mô hình 2	Mô hình 1	Mô hình 2
1986	29,30	26,14	27,54	10,8	6,4
1987	32,40	28,28	30,31	12,7	6,9
1988	26,90	30,19	28,49	12,2	5,6
1989	31,90	33,50	33,45	5,1	4,6
1990	30,00	27,00	27,65	10,0	8,5
1991	35,00	32,72	34,07	6,5	2,7
1992	29,10	32,46	29,90	11,5	2,7
1993	34,10	31,83	31,84	6,7	7,1
1994	23,10	23,53	23,30	1,9	0,9
1995	37,74	32,66	34,44	13,5	9,6
1996	30,00	31,14	28,51	3,8	5,0
1997	31,20	30,07	30,30	3,6	3,0
1998	35,00	34,91	36,21	0,3	3,4

Bảng 3. Sai số % của 2 mô hình khi chưa có mô phỏng năng suất dư  $\omega_t$  (mô hình 1) và khi có mô phỏng năng suất dư  $\omega_t$  (mô hình 2) so với năng suất xã hội

Sai số so với năng suất xã hội (%)		Vụ đông xuân				Vụ mùa			
		Mô hình 1		Mô hình 2		Mô hình 1		Mô hình 2	
		Số năm	Tỷ lệ %	Số năm	Tỷ lệ %	Số năm	Tỷ lệ %	Số năm	Tỷ lệ %
0 - 5	12	32	15	40	20	52	19	50	
6 - 10	13	34	8	21	6	16	15	39	
11 - 15	7	18	12	31	9	24	4	11	
16 - 20	5	13	2	5	2	5	0	0	
>20	1	3	1	3	1	3	0	0	

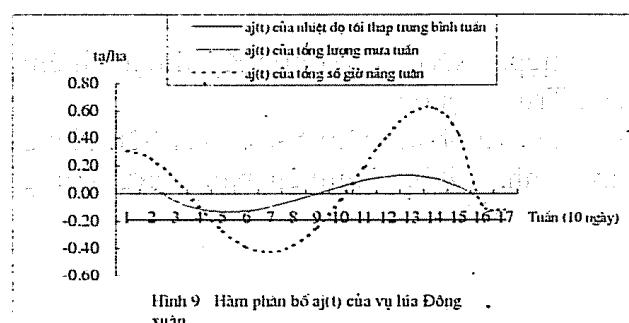
d. Một vài đánh giá sự tác động của các yếu tố khí tượng đến năng suất lúa

Như đã trình bày ở trên, dùng phương pháp tích phân hồi qui bội ngoài việc xây dựng được mô hình dự báo năng suất, thông qua giá trị hàm sinh trưởng  $a_j(t)$  còn cho ta biết được mỗi yếu tố khí tượng ở mỗi thời điểm tăng hoặc giảm 1 đơn vị sẽ làm tăng hoặc giảm bao nhiêu đơn vị năng suất.

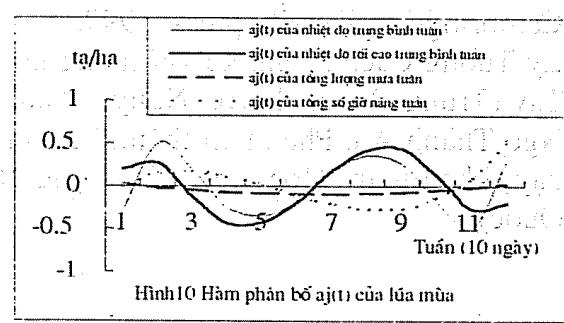
Sau khi tính toán, các giá trị  $a_j(t)$  của vụ đông xuân và vụ mùa được miêu tả tại các hình 9, 10 trong đó nhiệt độ đơn vị là  $1^{\circ}\text{C}$ , lượng mưa là 10 mm, số giờ nắng là 10 giờ. Căn cứ vào các hàm phân bố  $a_j(t)$  ta thấy được sự ảnh hưởng định tính và định lượng của các yếu tố khí tượng tới năng suất cây trồng trong từng thời đoạn phát triển của nó. Điều này rất có ý nghĩa trong việc xác định cơ cấu thời vụ, tận dụng tối đa những thời đoạn có lợi, né tránh những thời đoạn bất lợi nhằm tăng năng suất và sản lượng cây trồng. Đặc biệt là khi đánh giá tác động của biến đổi khí hậu nó làm rõ được vai trò của nhiệt độ tăng, lượng mưa, số giờ nắng giảm và sự thay đổi phân bố

của chúng trong năm có ảnh hưởng như thế nào đến năng suất cây lúa nói riêng và cây trồng nói chung.

Nếu ta giả sử những yếu tố khác ở cùng thời điểm có giá trị không đổi thì sự tăng hoặc giảm 1 đơn vị giá trị của yếu tố này sẽ dẫn đến sự tăng hoặc giảm năng suất bao nhiêu tạ/ha. Ví dụ, vụ lúa đông xuân (hình 9) nhiệt độ trung bình tối thấp là hàm ảnh hưởng tuyến tính trong cả giai đoạn sinh trưởng, giả sử các yếu tố khác không đổi nhưng nhiệt độ trung bình tối thấp tăng  $1^{\circ}\text{C}$  thì sẽ làm giảm năng suất khoảng 0,3 tạ/ha (kết quả này trùng với nhiều nghiên cứu và nhận xét trước đây: năm mùa đông lạnh năng suất lúa thường cao hơn năm mùa đông ấm). Đổi với lượng mưa và số giờ nắng là hàm ảnh hưởng phi tuyến tính đối với năng suất như tuần thứ 13, 14 (trung tuần và cuối tháng 4 kỳ lúa đồng, trổ bông nở hoa) nếu tổng lượng mưa tăng 10 mm/tuần hoặc số giờ nắng tăng 10 giờ/tuần năng suất sẽ tăng khoảng từ 0,15 đến 0,6 tạ/ha



Hình 9 Hàm phân bố aj(t) của vụ lúa Đông xuân



Hình 10 Hàm phân bố aj(t) của lúa mùa

Hoặc của vụ mùa (hình 10) ở tuần 4, 5 (thời kỳ lúa bén rễ hồi xanh) nếu nhiệt độ trung bình tuần hoặc nhiệt độ tối cao trung bình tuần giảm  $1^{\circ}\text{C}$  năng suất sẽ tăng khoảng 0,3-0,4 tạ/ha...

### 3. Kết luận

a) Với việc áp dụng phương pháp tích phân hồi qui bội và ARIMA trong việc dự báo năng suất thời tiết và năng suất dư của vụ lúa đông xuân (ngoại trừ một năm mất mùa lớn) đạt độ chính xác ở mức trên 80% mà tuyệt đại đa số là ở mức trên 90%, chiếm 70% số năm mô phỏng. Đổi với vụ lúa mùa đạt độ chính xác ở mức trên 85% mà tuyệt đại đa số ở mức trên 90%, chiếm 90% số năm mô phỏng, trong đó tới hơn 1/2 số năm mô phỏng đạt mức chính xác trên 95%.

b) Với việc áp dụng mô hình ARIMA để dự báo năng suất dư đã nâng đáng kể độ chính xác của mô hình, đặc biệt là những năm có năng suất dư lớn. Với lúa đông xuân đã nâng độ chính xác khoảng 3% và lúa mùa khoảng 5%.

c) Sử dụng phương pháp tích phân hồi qui bội đã hạn chế được sự đòi hỏi độ dài của chuỗi so với các biến. Với phương pháp này ngoài việc xây dựng được mô hình dự báo năng suất thời tiết, thông qua hàm phân bố aj(t) còn cho biết một cách định hình và định lượng sự ảnh hưởng của các yếu tố khí tượng đến năng suất cây trồng ở mỗi thời đoạn phát triển của nó.

d) Toàn bộ việc tính toán, xử lý, kiểm nghiệm số liệu và xây dựng mô hình được thực hiện trên phần mềm SAS (Statistical analysis system) vì vậy đảm bảo độ tin cậy cao. Nếu có các số liệu thích ứng cho mô hình sẽ dễ dàng dự báo năng suất lúa của 2 vụ đông xuân và mùa trong những năm sau.

Cuối cùng tác giả xin chân thành cảm ơn Trung tâm Nghiên cứu Khí tượng Nông nghiệp đã cung cấp các số liệu cần thiết. Cảm ơn Giáo sư Lý Tương Các,

nguyên chủ nhiệm khoa khí tượng ứng dụng, Học viện khí tượng Nam Kinh, Trung Quốc đã giúp đỡ thực hiện bài viết này.

### Tài liệu tham khảo

1. Nguyễn Văn Việt, Đăng Quang Vinh. Cơ cấu thời vụ gieo trồng và công thức luân canh lúa ở vùng Hà Nội - KTTV số 10 - 1996.
2. Ngô Sỹ Giai. Xây dựng các phương pháp dự báo khí tượng nông nghiệp các thời kỳ phát triển chủ yếu, năng suất và sản lượng lúa ở Việt Nam (Đề tài nghiên cứu Tổng cục Khí tượng Thủy văn). - Hà Nội, 1989.
3. Box, G.E.P and G.M. Jenkins. Time series analysis for forecasting and control. Holden-Dap, San. Fran., revised edition, 1976.
4. Hans Von storck and Francis W. Zwiers. Statistical analysis in climate research. Cambridge University press, 1999.
5. Lý Tương Các. Thống kê khí tượng nông nghiệp.- NXB Khoa học Kỹ thuật Thiểm Tây (Trung Quốc). 1996. (Nguyên bản tiếng Trung Quốc).
6. Ngô Thành Âu. Phân tích thống kê hàm đa biến và phần mềm SAS.- Tài liệu giảng dạy, khoa toán, Học viện Khí tượng Nam Kinh., 1989. (Nguyên bản tiếng Trung Quốc).