

MỘT VÀI MÔ HÌNH HỒI QUY DỰ BÁO NGẮN HẠN LẠM PHÁT

NGUYỄN HOÀI BẢO

Viện Công nghệ thông tin

Abstract. In recent time, inflation has once again become an urgent problem in the Vietnam economy. One of the challenges faced by economists is to forecast the inflation rate of the whole year as early as by mid-year, e.g. in June. The forecast of inflation is important for various reasons, especially for macro-economic analysis and management. In this article, I would like to present some models for short-term forecasting of inflation rate. These are regression models, which are built based on the rate of increase of food price and the consumer price index. Based on these data, we can calculate the inflation rate for the whole year using these models. The advantages of these models are that it is very easy to obtain the data and the calculation is fairly simple. In addition, there are many models to choose from, and one can compare results of these different models to come up with the best possible forecast.

Tóm tắt. Gần đây vấn đề lạm phát lại trở nên thời sự. Các chuyên gia thường đưa ra những con số dự báo lạm phát khác nhau, nhưng không nêu ra cơ sở lập luận, tính toán. Một trong những vấn đề đặt ra cho những người nghiên cứu và điều hành kinh tế là có thể dự báo chỉ số lạm phát cả năm ngay từ một thời điểm sớm giữa năm, như từ tháng 6 chẳng hạn. Việc dự báo sớm và tương đối chính xác chỉ số lạm phát có ý nghĩa quan trọng trong nghiên cứu và điều hành kinh tế vĩ mô. Trong bài báo này, chúng tôi đưa ra một kiểu mô hình hồi quy, dự báo ngắn hạn lạm phát chủ yếu chỉ dựa trên những thông tin về chỉ số lạm phát những năm trước, mức tăng giá lương thực 6 tháng đầu năm và mức tăng giá tiêu dùng 6 tháng đầu năm để đưa ra dự báo về chỉ số lạm phát của năm đó. Ưu điểm của việc sử dụng các mô hình kiểu này là dễ thu thập số liệu để xây dựng mô hình. Việc tính toán cũng có thể được thực hiện dễ dàng bằng sử dụng EXCEL. Ngoài ra, ta có thể lựa chọn từ nhiều mô hình dạng này mô hình thích hợp nhất hay so sánh các kết quả khác nhau để đưa ra dự báo tốt nhất như có thể.

1. ĐẶT VẤN ĐỀ

Gần đây, vấn đề về lạm phát lại trở nên thời sự. Các chuyên gia đưa ra những con số dự báo về lạm phát khác nhau, nhưng không nêu ra những cơ sở lập luận cho tính toán định lượng. Vậy chỉ số lạm phát năm 2004 này là bao nhiêu? là một con số hay hai con số? Việc xác định cơ sở cho tính toán định lượng và dự báo chỉ số lạm phát là việc cần thiết phải làm.

Trong một báo cáo tại Hội nghị khoa học Viện Công nghệ thông tin [1] chúng tôi có đề cập đến một mô hình dự báo lạm phát đơn giản, chủ yếu chỉ dựa trên thông tin về mức giá lương thực 6 tháng đầu của mỗi năm, để đưa ra dự báo về lạm phát của năm đó. Bài này phát triển ý tưởng trên, nêu ra một số mô hình khác nhau cũng như tính toán định lượng cụ thể, dựa trên những thông tin thống kê giai đoạn 1994 - 2003 và mức tăng giá tiêu dùng chung cũng như mức tăng giá lương thực 6 tháng đầu năm 2004 để đưa ra dự báo về chỉ số

lạm phát cho năm 2004 này. Các mô hình này có thể sử dụng riêng rẽ hay kết hợp với nhau trong dự báo tùy theo khả năng về số liệu cũng như mức độ chính xác của chúng trong từng giai đoạn kinh tế cụ thể. Nhưng nói chung, có thể xem đây là những mô hình dự báo cho một thời điểm tương lai gần, số liệu dễ thu thập, việc thiết lập và tính toán khá đơn giản, có ý nghĩa đối với những quyết định và điều chỉnh kinh tế vĩ mô.

2. MỐI LIÊN HỆ GIỮA GIÁ LƯƠNG THỰC VÀ LẠM PHÁT

Từ cách tính chỉ số giá cả hàng hóa tiêu dùng và dịch vụ chung PC (chỉ số biểu thị cho mức lạm phát), ta có sự phụ thuộc về hàm số sau:

$$PC(t) = f[P_1(t), P_2(t), \dots, P_n(t)], \quad t = \text{tháng hoặc năm} \quad (1)$$

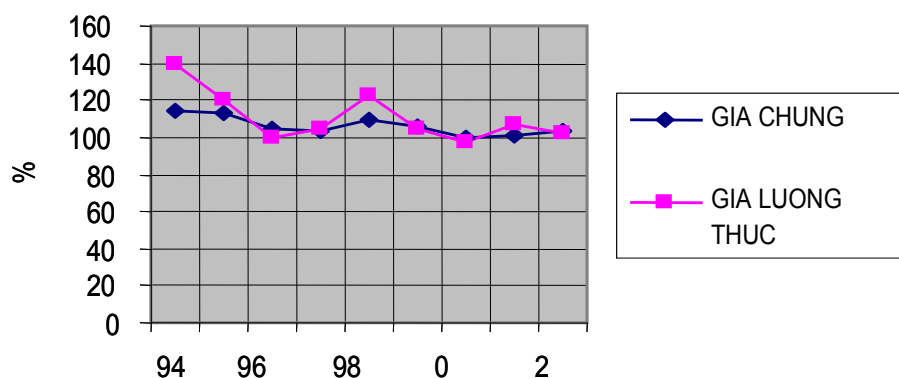
Trong đó $PC(t)$ là hàm biểu thị chỉ số giá hàng hóa tiêu dùng và dịch vụ, được gọi tắt là chỉ số giá tiêu dùng hay mức giá chung, $P_i(t), i = 1, \dots, n$ là chỉ số giá của từng nhóm mặt hàng hay dịch vụ chủ yếu được lựa chọn từ “giỏ hàng” tiêu dùng hay dịch vụ.

Ta cũng có thể chỉ sử dụng một biến duy nhất trong hàm số (1), đó là giá lương thực $PLT(t)$, khi đó ta có:

$$PC(t) = f[PLT(t)] \quad (2)$$

Trong đó PC là chỉ số giá hàng hóa và dịch vụ hàng năm, PLT là chỉ số giá lương thực hàng năm, P_6LT là mức tăng giá lương thực đến tháng 6 mỗi năm và Pc_6 là mức tăng đến tháng 6 hàng năm của giá tiêu dùng và dịch vụ.

Qua các số liệu thống kê từ 1994 đến 2002 (cũng tương tự như các giai đoạn khác trước đó) ta có thể thấy trên thực tế động thái của hai đại lượng này có diễn biến khá tương đồng. Điều này nói lên sự phụ thuộc tương quan khá chặt chẽ giữa mức giá chung và mức giá lương thực:



Hình 1. Mức tăng giá hàng năm (1994-2002)

Đồ thị trên cũng như các tính toán và mô hình trong bài chủ yếu đều được xây dựng trên cơ sở những chuỗi số liệu thống kê có được trong giai đoạn 1994-2003. Trong đó PC là chỉ số giá tiêu dùng và dịch vụ hàng năm, tức là mức tăng giá chung hàng năm; PLT là chỉ số giá lương thực hàng năm; P_6LT là chỉ số giá lương thực tính đến tháng 6 mỗi năm; và Pc_6

là chỉ số giá tiêu dùng và dịch vụ hàng năm chỉ được tính đến thời điểm tháng 6 mỗi năm (được tính bằng %, năm sau so với cùng kỳ năm trước).

Bảng 1. Diễn biến chỉ số giá cả các năm 1994 - 2003

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
PC	114,4	112,7	104,5	103,6	109,2	105,5	99,4	100,8	104	103
PLT	139	120,6	100,2	104,4	123,1	104,4	97,7	106,6	102,8	102,9
P6LT	114,8	101,3	96,2	92,5	115,1	97,1	92,4	94,3	100,6	98,8
Pc6	107,1	111,5	103,3	101,1	106	101,6	99,0	100,8	100,1	100,1

Từ hai dãy số liệu tương ứng giai đoạn 1994-2003, ta xây dựng được mối tương quan sau:

$$PC(t) = 0,337436 \times PLT(t) + 68,53471 \quad (2b)$$

Multiple R	0,902015	R Square	0,813632
Adjusted RSq	0,790336	Standard Erro	2,246,879

Phương trình cho thấy giá lương thực quả thực có mối ràng buộc khá chặt chẽ tới mức giá chung, hay nói một cách khác từ mức tăng giá lương thực có thể dự báo được chỉ số lạm phát hàng năm. (Trong những giai đoạn khác của kinh tế Việt Nam ta cũng thấy xuất hiện mối tương quan này, dù những giai đoạn đó có những đặc điểm khác hẳn, ngay như trong giai đoạn lạm phát cao giữa những năm 80 vẫn thấy xuất hiện quan hệ này).

3. CÁC MÔ HÌNH DỰ BÁO LẠM PHÁT THÔNG QUA GIÁ LƯƠNG THỰC

Một nhu cầu khác thường nảy sinh trong dự báo kinh tế là câu hỏi: đến giữa mỗi năm ta có thể đưa ra dự báo gì cho kinh tế cả năm đó, khi mà thời điểm cần dự báo cách không quá xa thời điểm số liệu thu thập được? Một dự báo ngắn hạn quả thật có khá nhiều lợi ích cho những phân tích và điều hành có tính chất vĩ mô. Một loạt các hàm số dạng sau đây về mức giá lương thực có thể được xây dựng qua các dãy số liệu thống kê:

$$PLT(t) = f[P_m LT(t)]. \quad (3)$$

Trong đó PLT là chỉ số giá lương thực năm t , $P_m LT$ là chỉ số giá lương thực chỉ tính đến tháng m năm t . Ta có thể xây dựng các hàm này cho $m = 1, 2, 3, \dots, 12$. Trong số các hàm số này ta chọn $P_6 LT$, mức tăng giá lương thực đến tháng 6 để nghiên cứu. Đây là thời điểm phù hợp nhất cho những nghiên cứu đánh giá ngắn hạn kinh tế hàng năm. Và diễn biến về giá cả nói riêng cũng như về kinh tế nói chung có thể phản ánh và mô tả tương đối rõ nét xu thế phát triển kinh tế của năm đó. Nếu kết hợp giữa hai phương trình (2b) và (3), thì qua mức tăng giá lương thực đến tháng 6 hàng năm ta có thể có được dự báo khá tốt và khá sớm về tỷ lệ lạm phát hàng năm:

$$PC(t) = f[P_6 LT(t)]. \quad (4)$$

Từ những số liệu thống kê tương ứng cho giai đoạn 1994-2003, ta thiết lập được mô hình dự báo sau theo phương pháp hồi quy:

$$PC(t) = 0,467903 \times P_6 LT(t) + 58,77461 \quad (4b)$$

Multiple R	0,902015	R Square	0,813632
Adjusted RSq	0,790336	Standard Erro	2,246,879

Hàm số này phản ánh mối liên hệ tương quan khá tốt giữa giá lương thực 6 tháng đầu năm và chỉ số giá tiêu dùng chung cả năm, và có thể sử dụng để dự báo lạm phát. Tuy nhiên, ta vẫn có thể thiết lập trên cùng hai chuỗi số liệu cùng một mối quan hệ giữa hai đại lượng tương ứng thông qua một quan hệ hàm số khác, trong đó thay vì xây dựng một mô hình từ quan hệ trực tiếp giữa PC và P_6LT , ta xét tương quan hàm số giữa chúng và căn bậc hai của chúng $SQ(PC)$ và $SQ(P_6LT)$. Đó là:

$$PC(t) = f[SQ(P_6LT(t))], \quad (5)$$

$$SQ(PC(t)) = f[SQ(P_6LT(t))], \quad (6)$$

$$SQ(PC(t)) = f[P_6LT(t)]. \quad (7)$$

Như vậy để mô tả quan hệ giữa hai đại lượng này, ngoài các quan hệ tuyến tính ta có thể mô tả bằng các quan hệ phi tuyến. Với trong số hàng loạt các phương trình khác nhau này, sau khi tính toán xây dựng mô hình, ta có thể chọn ra mô hình mô tả tốt nhất mối quan hệ giữa hai đại lượng (tức là phương trình hồi quy có hệ số tương quan R lớn hơn cả) để tính toán dự báo cho mình.

Qua chuỗi số liệu được tính toán tương ứng, tương tự như trên ta lần lượt có:

$$PC(t) = 12,4204 \times SQ(P_6LT(t)) - 18,0332 \quad (5b)$$

Multiple R	0,852147	R Square	0,726154
Adjusted RSq	0,691293	Standard Error	2,723622

$$SQ(PC(t)) = 0,599652 \times SQ(P_6LT(t)) + 4,304931 \quad (6b)$$

Multiple R	0,851496	R Square	0,725045
Adjusted RSq	0,690676	Standard Error	0,131862

$$SQ(PC(t)) = 0,022622 \times P_6LT(t) + 8,00996 \quad (6b)$$

Multiple R	0,791266	R Square	0,626103
Adjusted RSq	0,579365	Standard Error	0,153768

Rõ ràng so với (4b), hàm số (5b), (6b) và (7b) mô tả tốt hơn quan hệ giữa hai đại lượng là: mức tăng giá lương thực 6 tháng đầu năm và tỷ lệ lạm phát hàng năm, trong đó (5b) là tốt hơn cả. Vì thế, trong trường hợp này ta có thể chọn (5b) làm hàm dự báo chỉ số lạm phát.

Với mức tăng giá lương thực 6 tháng đầu năm 2004 được xác định là 11,5% (trong mô hình là 111,5% so với chỉ số giá tháng 12 năm 2003), từ phương trình (5b) ta có thể tính được mức tăng chung của chỉ số giá tiêu dùng cho năm 2004 (hay chỉ số lạm phát) là 13,12% (trong mô hình cũng tính theo mức tăng so với tháng 12 năm trước).

4. CÁC MÔ HÌNH DỰ BÁO LẠM PHÁT THÔNG QUA MỨC GIÁ CHUNG 6 THÁNG ĐẦU NĂM

Trong tự cách thức trên, ta có thể xác định mối quan hệ giữa chỉ số giá chung tháng 6 hàng năm P_{c6} với mức độ lạm phát cả năm PC , và có được mối liên hệ hàm số có hệ số tương quan khá cao sau:

$$PC(t) = 1,123562 \times P_{c6}(t) - 10,0742 \quad (8)$$

Multiple R	0,903937	R Square	0,817102
Adjusted RSq	0,794239	Standard Error	2,225863

Nhưng nếu xét đến các mối liên hệ phi tuyến giữa hai đại lượng này, tương tự như những hàm được nói đến ở phần trên, ta có thể tính toán và chọn ra được hàm số mô tả chính xác hơn cả quan hệ giữa chúng là:

$$PC(t) = 23,07473 \times SQ[P_{c6}(t)] - 128,494 \quad (8b)$$

Multiple R	0,905409	R Square	0,819765
Adjusted RSq	0,77236	Standard Error	22,09596

Với mức tăng giá chung 6 tháng đầu năm nay là 7,2% (trong phương trình là 107,2%), theo phương trình (8b) ta tính toán được PC dự báo cho cả năm 2004 sẽ là 10,42 %. Ta thấy, mức lạm phát dự báo tính theo mức tăng giá lương thực đến tháng 6 ở mô hình (5b) cao hơn khá rõ rệt con số lạm phát được tính theo mức tăng giá chung đến tháng 6 tại mô hình (8b). Có thể lý giải điều này không mấy khó khăn: mức độ tăng giá chung phụ thuộc khá mạnh vào mức độ tăng giá lương thực, và trong 6 tháng đầu năm vừa qua giá lương thực lại tăng rất cao, nên xu thế tăng mạnh của mức giá chung được biểu lộ rõ qua (5b). Nếu xu thế này vẫn tiếp diễn trên thực tế, thì chỉ số lạm phát có khả năng khá cao. Ngược lại, nếu giá lương thực bình ổn hoặc thông qua các biện pháp hạn chế được mức tăng giá của mặt hàng này trong 6 tháng cuối năm, thì có thể hạn chế được đáng kể mức độ tăng của lạm phát cả năm.

5. CÁC MÔ HÌNH KẾT HỢP

Bây giờ ta kết hợp giữa mức tăng giá chung đến tháng 6 hàng năm và mức tăng giá lương thực đến tháng 6 hàng năm để dự báo về chỉ số lạm phát của năm đó. Tức là ta xét đến hàm tương quan bội sau đây:

$$PC(t) = f[SQ(P_{c6}(t)); SQ(P_6LT(t))]. \quad (9)$$

Với các số liệu đã cho, ta thiết lập được tương quan bội cụ thể sau:

$$PC(t) = 15,59554 \times SQ(P_{c6}(t)) + 6,854334 \times SQ(P_6LT(t)) - 120,871 \quad (9b)$$

Multiple R	0,977134	R Square	0,954792
Adjusted RSq	0,941875	Standard Error	1,183038

Hàm số này mô tả khá sát thực và chính xác động thái diễn biến theo thời gian của chỉ số giá tiêu dùng và dịch vụ. Mặt khác, có thể thấy mức lạm phát của năm trước cũng có tác động phần nào đến chỉ số lạm phát của năm sau, nên bây giờ ta thử bổ sung vào (9) thêm một quan hệ nữa là chỉ số giá tiêu dùng và dịch vụ chung của năm trước $PC(t-1)$ xem có nhận được một mô hình khả quan hơn hay không. Như vậy ta sẽ có:

$$PC(t) = f[SQ(PC(t-1)); SQ(P_{c6}(t)); SQ(P_6LT(t))]. \quad (10)$$

Trên cơ sở các số liệu thống kê thực tế, ta xây dựng được hàm số sau đây:

$$PC(t) = 2,477069 \times SQ(PC(t-1)) + 12,82786 \times SQ(PC_6(t)) \\ + 7,819743 \times SQ(P_6LT(t)) - 127,886 \quad (10b)$$

Multiple R	0,980518	R Square	0,961416
Adjusted RSq	0,942124	Standard Error	1,180497

Trong số các mô hình thiết lập được đến nay, đây là hàm số có hệ số tương quan cao nhất, nó phản ánh khá chính xác quan hệ giữa chỉ số lạm phát cả năm với diễn biến lạm phát năm trước đó, với mức tăng giá tiêu dùng và dịch vụ đến tháng 6 mỗi năm cũng như mức tăng giá lương thực đến tháng 6 mỗi năm. Do vậy, dùng (10b) để dự báo sẽ đưa ra kết quả chính xác hơn cả về mặt toán học. Kết quả tính toán từ (10b) theo số liệu tương ứng năm 2003 (PC là 103% so với năm trước) và 2004 (PC_6 là 107,2% và P_6LT là 111,5% so với tháng 12 năm trước) cho ta dự báo chỉ số lạm phát năm 2004 là 12,64%.

Về định lượng, theo các tính toán qua các mô hình trên, thì chỉ số lạm phát năm 2004 dự báo sẽ có dao động trong khoảng từ 10,42% đến 13,12%. Và nếu mọi việc vẫn diễn biến theo xu thế sáu tháng đầu năm và nếu không có sự can thiệp vĩ mô mạnh mẽ của nhà nước để bình ổn giá của những mặt hàng cơ bản hay điều chỉnh chính sách tiền tệ, thì mức lạm phát của 2004 sẽ là khoảng 12,64%. Trên thực tế, con số này sẽ thấp hơn do có tác động nhất định của các biện pháp điều khiển vĩ mô kiên quyết của nhà nước. Ngoài ra, các phương trình dự báo (4) và (8) cũng có thể được khái quát hóa thành một loạt các phương trình sau đây:

$$PC(t) = f[P_mLT(t)]. \quad (11)$$

Trong đó PC là mức lạm phát hàng năm, còn P_mLT là mức tăng giá lương thực cho đến tháng m mỗi năm. Và

$$PC(t) = f[PC_m] \quad (12)$$

với PC_m là mức tăng giá chung cho đến tháng thứ m mỗi năm.

Loạt các mô hình (11) và (12) có thể sử dụng cho việc dự báo chỉ số lạm phát cả năm từ sau tháng thứ m mỗi năm trở đi (tất nhiên với mức độ chính xác khác nhau, theo m). Ta cũng có thể xét các quan hệ phi tuyến giữa hai đại lượng tương tự như trường hợp (5), (6), (7) hoặc tương quan bội như (9), (10) hoặc kết hợp cả hai loại này như ở trên để chọn ra mô

hình dự báo tốt nhất và kết hợp các kết quả tính toán từ đây cùng với kết quả các dự báo đã nêu trên vào chùm kết quả dự báo chung để có thể phân tích và rút ra dự báo chỉ số lạm phát sát thực và phù hợp nhất.

6. KIỂM ĐỊNH ĐỘ CHÍNH XÁC CỦA CÁC MÔ HÌNH DỰ BÁO

Trên cơ sở các số liệu của quá khứ, ta có thể dễ dàng kiểm tra lại mức độ chính xác và khả năng dự báo của các mô hình bằng cách so sánh kết quả dự báo được tính toán thông qua mô hình và con số thực tế đã diễn ra. Chẳng hạn, trên cơ sở dữ liệu thống kê giai đoạn 1994 - 2002, ta dự báo chỉ số lạm phát cho năm 2003. Ta sẽ chọn kiểu mô hình tương đối đơn giản sau đây:

$$PC(t) = f[PC(t-1); Pc_6(t); P_6LT(t)] = a + b \times PC(t-1) + c \times Pc_6(t) + d \times P_6LT(t). \quad (13)$$

Tức là ta dự báo lạm phát hàng năm thông qua chỉ số giá tiêu dùng và dịch vụ năm trước, mức tăng giá chung cho đến tháng 6 và mức tăng giá lương thực tính đến tháng 6 của năm cần dự báo. Tính toán hồi quy cụ thể cho ta hàm số dự báo sau (cho các chuỗi số liệu 9 năm):

$$PC(t) = -5,68625 + 0,115578 \times PC(t-1) + 0,698715 \times Pc_6(t) + 0,270671 \times P_6LT(t) \quad (13b)$$

Multiple R	0,9558809	R Square	0,913571
Adjusted RSq	0,861714	Standard Error	1,898661

Hệ số tương quan khá cao ở đây cho phép ta hy vọng có được kết quả dự báo khả quan. Các số liệu thống kê thực tế tương ứng là:

$$PC(2002) = 104; Pc_6(2003) = 100,1; P_6LT(t) = 98,8.$$

Thay thế số liệu vào (13b), ta có $PC_{\text{dự báo}}(2003) = 103,0175$. Trong khi đó trên thực tế chỉ số lạm phát năm 2003 là: $PC_{\text{thực tế}}(2003) = 103(\%)$. Như vậy kết quả dự báo theo mô hình đã chọn và con số thực tế là tương đối khớp nhau.

Tương tự, nếu ta chọn sử dụng mô hình

$$PC(t) = f[SQ(PC(t-1)); SQ(Pc_6(t)); SQ(P_6LT(t))] \quad (10)$$

thì từ số liệu thống kê cho giai đoạn 1994 - 2002 ta có mô hình hồi quy cụ thể:

$$PC(t) = 2,624407 \times SQ(PC(t-1)) + 12,28523 \times SQ(Pc_6(t)) + 7,999423 \times SQ(P_6LT(t)) - 125,628 \quad (10c)$$

Multiple R	0,980518	R Square	0,961416
Adjusted RSq	0,942124	Standard Error	1,180497

Thay thế số liệu tương ứng vào, với mô hình này ta có $PC_{\text{dự báo}}(2003) = 103,563$, trong khi đó $PC_{\text{thực tế}}(2003) = 103(\%)$. Tuy hệ số tương quan trong mô hình này cao hơn mô hình trên, nhưng kết quả dự báo lại có độ chênh lệch lớn hơn so với con số thực tế.

7. KẾT LUẬN

Ưu điểm có thể thấy được của loạt các mô hình dự báo lạm phát ngắn hạn trên đây là sự dễ dàng và thuận lợi trong thu thập số liệu tính toán (trong trường hợp đơn giản nhất có thể sử dụng giá gạo hàng tháng mỗi năm). Với một loạt các mô hình khác nhau có thể thiết lập được, ta có điều kiện chọn ra mô hình có kết quả phù hợp và sát thực nhất với từng thời kỳ, giai đoạn phát triển kinh tế khác nhau. Tại đây có thể kết hợp những kết quả dự báo khác nhau từ những mô hình khác nhau để đưa ra những phân tích và đánh giá định lượng về kinh tế. Ta cũng có thể lựa chọn mô hình dự báo tùy theo điều kiện và khả năng thu thập thông tin có thể có. Sử dụng giá lương thực làm tham số dự báo là rất thuận lợi về phương diện thực tế và thông tin. Nhưng mặt khác, việc thử nghiệm xây dựng mô hình dự báo trên cơ sở mức tăng giá của một mặt hàng nào đó khác giá lương thực cũng rất có ích trong một hệ thống mô hình dự báo và phân tích tổng hợp. Việc thiết lập mô hình và tính toán qua các số liệu thống kê ở đây tương đối đơn giản: chỉ cần sử dụng EXCEL.

TÀI LIỆU THAM KHẢO

- [1] Nguyễn Hoài Bảo, Phân tích định lượng trong kinh tế nông nghiệp, *Kỷ yếu HNKH Viện Công nghệ thông tin 2001*.
- [2] *Niên giám Thống kê các năm 1994-2003*, NXB Thống kê, Hà Nội.

Nhận bài ngày 9-8-2004

Nhận lại sau sửa ngày 20-12-2004