

GIẢI PHÁP DỰ BÁO NGẮN HẠN VỀ TĂNG TRƯỞNG KINH TẾ VIỆT NAM

ĐỖ VĂN THÀNH

Bộ Kế hoạch và Đầu tư

Abstract. In order to regulate the Vietnam market economy in an efficient manner, it is important to have forecasting short-term growth of the economy.

A problem is arisen as follows: With respect to current macroeconomic data of Viet Nam which forecast models are suitable with forecasting short-term growth of the economy and which main macroeconomic issues can be also forecasted by that models? The problem is a challenge for institutions that are in charge of providing socio - economic forecasting information for the Government and the Premier Minister. The goal of this paper is to resolve a part of the arisen problem.

Tóm tắt. Để điều hành hiệu quả nền kinh tế thị trường ở nước ta hiện nay điều quan trọng phải dự báo ngắn hạn về tăng trưởng kinh tế.

Một vấn đề đặt ra là: trong trường hợp dữ liệu kinh tế vĩ mô ở nước ta hiện nay thì mô hình dự báo nào là phù hợp với dự báo ngắn hạn về tăng trưởng kinh tế và những vấn đề kinh tế vĩ mô chủ yếu nào có thể được dự báo bởi mô hình đó? Vấn đề này hiện đang là thách thức đối với các đơn vị, tổ chức có nhiệm vụ cung cấp thông tin dự báo kinh tế-xã hội cho Chính phủ và Thủ tướng Chính phủ. Mục đích của bài báo này là giải quyết một phần của câu hỏi được đặt ra.

ĐẶT VẤN ĐỀ

Một trong 3 nhiệm vụ chủ yếu của Chính phủ nước ta năm 2007 là duy trì được tốc độ tăng trưởng kinh tế cao ở mức 8,5%/ năm. Để đạt được mức tăng trưởng này, cần có nhiều giải pháp trong đó sự chỉ đạo điều hành của Chính phủ và Thủ tướng chính phủ được xác định là một trong những giải pháp quan trọng nhất. Để thực hiện tốt giải pháp này cần xây dựng được báo cáo tổng quan kinh tế quốc gia theo quý và thông tin dự báo tương đối chính xác khả năng phát triển kinh tế ở một vài quý tiếp theo. Đây là nhiệm vụ và cũng là thách thức lớn đối với những đơn vị có nhiệm vụ nghiên cứu, cung cấp cho Chính phủ thông tin, nhất là thông tin dự báo về tình hình phát triển kinh tế của đất nước.

Mục đích của bài báo này nhằm tham gia giải quyết vấn đề đặt ra ở trên bằng cách đi tìm những mô hình dự báo phù hợp với thực tiễn số liệu kinh tế Việt Nam. Cụ thể bài báo sẽ đề xuất giải pháp dự báo ngắn hạn tăng trưởng kinh tế Việt Nam theo cách tiếp cận của khai phá dữ liệu đó là dựa vào số liệu thực tế của nền kinh tế ở thời điểm hiện tại và quá khứ để dự báo giá trị của chúng trong tương lai gần.

Bài báo gồm 5 mục. Mục 1 của bài báo nhằm xác định và giới hạn những vấn đề cần phân tích và dự báo ngắn hạn về tăng trưởng kinh tế. Mục 2 trình bày cách tiếp cận lựa chọn mô hình và phần mềm công cụ phục vụ dự báo. Mục 3 trình bày nhanh quy trình ứng dụng mô hình được lựa chọn để dự báo ngắn hạn về tăng trưởng kinh tế. Mục 4 trình bày một cách tóm tắt quá trình xây dựng mô hình dự báo và kết quả dự báo về tốc độ tăng

trường GDP theo quý và lạm phát chỉ số giá tiêu dùng CPI theo tháng, quý. Mục 5 sẽ đề xuất những vấn đề kinh tế - xã hội chủ yếu của Việt Nam có thể được dự báo tốt theo mô hình được đề xuất.

1. XÁC ĐỊNH VÀ GIỚI HẠN VẤN ĐỀ DỰ BÁO

1.1. Những vấn đề kinh tế vĩ mô chủ yếu của Việt Nam cần được phân tích và dự báo ngắn hạn

Dựa vào phân tích nội dung Pha 1 dự án: “Hỗ trợ xây dựng Hệ thống thông tin Quốc gia phục vụ lập kế hoạch kinh tế” do EU tài trợ [10-12]; dựa vào thực tiễn của công cuộc công nghiệp hoá, hiện đại hoá đất nước và xây dựng nền kinh tế thị trường của nước ta hiện nay [7-8], việc phân tích và dự báo ngắn hạn về kinh tế vĩ mô cần tập trung vào những vấn đề sau: Dựa vào phân tích nội dung Pha 1 dự án: “Hỗ trợ xây dựng Hệ thống thông tin Quốc gia phục vụ lập kế hoạch kinh tế” do EU tài trợ [10-12]; dựa vào thực tiễn của công cuộc công nghiệp hoá, hiện đại hoá đất nước và xây dựng nền kinh tế thị trường của nước ta hiện nay [7-8], việc phân tích và dự báo ngắn hạn về kinh tế vĩ mô cần tập trung vào những vấn đề sau:

a. Phân tích và dự báo nhóm các chỉ tiêu kinh tế vĩ mô quan trọng nhất về GDP, về xuất nhập khẩu, đầu tư, lạm phát, tiêu dùng, lao động, tiền tệ,... như: phân tích và dự báo tăng trưởng GDP, GDP của các ngành công nghiệp, nông nghiệp và dịch vụ và cơ cấu đóng góp của các ngành này trong GDP theo quý và theo năm; phân tích và dự báo về tình hình xuất nhập khẩu, giá trị xuất nhập khẩu, chỉ số giá xuất nhập khẩu theo tháng, quý, năm,...; phân tích và dự báo tình hình thu hút đầu tư trực tiếp nước ngoài, tình hình giải ngân nguồn vốn ODA theo quý, năm; phân tích và dự báo về tỷ giá hối đoái giữa VNĐ và một số ngoại tệ mạnh cũng như giữa các ngoại tệ mạnh; phân tích và dự báo tình hình tiêu dùng trong nước, dự báo chỉ số giá tiêu dùng, tình hình sản xuất và giá cả chính theo các ngành kinh tế, theo các loại hình kinh tế,...

b. Phân tích và dự báo một số cân đối vĩ mô quan trọng: cân đối ngân sách nhà nước, cân đối vốn đầu tư phát triển, ngoại thương và cân đối xuất nhập khẩu, nợ nước ngoài và cán cân thanh toán quốc tế,... phân tích và dự báo các mối quan hệ giữa: vốn đầu tư trực tiếp nước ngoài với tăng trưởng xuất, nhập khẩu và GDP, giữa tỉ lệ đầu tư trên GDP với kim ngạch nhập khẩu máy móc thiết bị, tăng tiêu dùng tư nhân và nhập khẩu hàng hoá tiêu dùng,...

Bài báo này chỉ tập trung đề xuất giải pháp dự báo ngắn hạn cho những vấn đề thuộc hạng mục nội dung a) của mục này.

1.2. Đặc điểm số liệu kinh tế - xã hội

Nói chung dữ liệu phản ánh tình hình phát triển kinh tế - xã hội đều được đo theo những kỳ thời gian nhất định như ngày, tuần, tháng, quý, năm,... Do đặc điểm của hệ thống kinh tế - chính trị và trình độ phát triển kinh tế - xã hội ở nước ta, dữ liệu phản ánh tổng thể nền kinh tế chỉ có thể thu thập được tương đối đầy đủ theo quý [2, 5, 12], do vậy việc xây dựng báo cáo tổng quan kinh tế quý và phân tích, dự báo những vấn đề chủ yếu của kinh tế - xã hội Việt Nam theo quý là hết sức quan trọng [3, 11 – 12]. Bài báo này tập trung giới thiệu giải pháp dự báo ngắn hạn về tăng trưởng kinh tế Việt Nam theo quý. Tuy nhiên hoàn toàn có thể ứng dụng giải pháp được đề xuất để dự báo ngắn hạn về tăng trưởng kinh tế Việt Nam theo tháng, năm cho những vấn đề thuộc hạng mục nội dung a) Mục 1.1 nêu trên.

1.3. Vấn đề dự báo ngắn hạn về tăng trưởng kinh tế

Theo cách tiếp cận được các nhà tin - kinh tế sử dụng là dựa vào chu kỳ dữ liệu để phân loại dự báo ngắn hạn, trung hạn và dài hạn, cụ thể là những dự báo cho 1 hoặc 2 kỳ dữ liệu sau đó được gọi là dự báo ngắn hạn, những dự báo cho 3 đến 5 kỳ dữ liệu tiếp theo được gọi là dự báo trung hạn và trên 5 kỳ dữ liệu sẽ được gọi là dự báo dài hạn [9].

Mặt khác, khi nói đến tăng trưởng kinh tế của một quốc gia, thông thường phải xét đồng thời ít nhất 2 vấn đề là tốc độ tăng trưởng kinh tế và mức độ lạm phát của nền kinh tế [7-8].

Tốc độ tăng trưởng của nền kinh tế được thể hiện tập trung, rõ nét qua tốc độ tăng trưởng GDP chung của quốc gia, GDP của các ngành kinh tế như công nghiệp, nông nghiệp, dịch vụ,... và cũng được thể hiện một phần qua tốc độ tăng trưởng kim ngạch xuất, nhập khẩu và đầu tư trực tiếp nước ngoài,... Thực tiễn của nền kinh tế Việt Nam cũng như các nước khác đều cho thấy GDP thường được tăng trưởng mạnh vào những tháng cuối năm, nói cách khác tốc độ tăng trưởng kinh tế mang yếu tố mùa vụ cao. Hiện tại ở Việt Nam, số liệu về tốc độ tăng trưởng GDP chưa được thống kê đầy đủ theo tháng, mà chỉ được thống kê theo quý, năm. Trong bài báo này sẽ tập trung nghiên cứu và dự báo tăng trưởng GDP chung theo quý. Việc nghiên cứu và dự báo tăng trưởng GDP chung theo năm và GDP của các ngành công nghiệp, nông nghiệp và dịch vụ theo quý, năm hoặc tốc độ tăng trưởng của xuất, nhập khẩu, của đầu tư trực tiếp nước ngoài theo tháng, quý, năm được thực hiện theo cách tương tự.

Đi liền với tăng trưởng kinh tế phải tính đến mức độ lạm phát. Khi nói đến lạm phát, thường phải xét đầy đủ cả 3 khía cạnh cơ bản nhất của nó là [7-8]: chỉ số giá tiêu dùng, tỷ giá hối đoái và cung tiền cho nền kinh tế. Hiện tại số liệu về cung tiền cho nền kinh tế chưa được thống kê thường xuyên, đầy đủ và được phổ biến chính thức bởi một cơ quan quản lý Nhà nước Việt Nam, trong khi số liệu về chỉ số giá tiêu dùng (CPI) và tỷ giá hối đoái đã được Tổng cục Thống kê tính toán và công bố thường xuyên hàng tháng. Thực tiễn cho thấy chỉ số giá tiêu dùng và tỷ giá hối đoái thường biến động rất mạnh vào những tháng thuộc quý đầu và quý cuối trong năm, nói cách khác chúng cũng mang yếu tố mùa vụ cao. Xét về ý nghĩa kinh tế, chỉ số giá tiêu dùng phản ánh rõ nét tình trạng lạm phát của nền kinh tế hơn là tỷ giá hối đoái, do đó ở bài báo này cũng giới hạn chỉ tập trung vào nghiên cứu dự báo mức độ lạm phát chỉ số giá tiêu dùng theo tháng, và từ đó theo quý. Việc dự báo tỷ giá hối đoái theo tháng hoàn toàn được thực hiện theo cách tương tự.

2. LỰA CHỌN MÔ HÌNH DỰ BÁO VÀ PHẦN MỀM CÔNG CỤ

2.1. Lựa chọn mô hình

Dữ liệu phụ thuộc thời gian được chia làm 2 loại: dữ liệu chuỗi thời gian (series time data) và dữ liệu chuỗi (sequence data). Nếu dữ liệu chuỗi thời gian bao gồm chuỗi các sự kiện được đo tại các khoảng thời gian như nhau thì dữ liệu chuỗi bao gồm các sự kiện được sắp thứ tự có thể không cần đòi hỏi khái niệm cụ thể về thời gian. Các dữ liệu kinh tế - xã hội thuộc loại dữ liệu chuỗi thời gian.

Để phát hiện và phân tích các mối quan hệ nhân quả giữa các dữ liệu chuỗi thời gian về kinh tế có thể sử dụng phương pháp đa nhân tố trong kinh tế lượng [7, 13] hoặc kỹ thuật học mạng Nơron [9]; cũng có thể ứng dụng kỹ thuật phát hiện luật kết hợp nhằm phát hiện các mối quan hệ kết hợp và khuynh hướng biến đổi của dữ liệu [4, 20]...

Để dự báo dữ liệu kinh tế có thể sử dụng một trong 3 mô hình là mô hình làm trơn hàm mũ Holt-Winter, mô hình tích hợp trung bình trượt tự hồi quy ARIMA hoặc mô hình tự hồi quy véc tơ VAR. Dưới đây chỉ giới thiệu một cách rất tóm tắt về những mô hình được

bài báo đề xuất lựa chọn là mô hình ARIMA và mô hình được phát triển từ nó được gọi là SARIMA. Chi tiết của các mô hình Holt-Winter, ARIMA và SARIMA, VAR (hay ARMA nhiều biến); mối liên hệ, ưu nhược điểm và khả năng ứng dụng của chúng đã được khảo cứu và thử nghiệm ban đầu trong [6, 14, 16].

a. Mô hình ARIMA

ARIMA là mô hình được tổ hợp từ các mô hình của 3 quá trình:

- *Quá trình tự hồi quy*: Mô hình tự hồi quy bậc $p - AR(p)$. Giả sử $\{z_t\}$ là một chuỗi dừng và giá trị của biến z_t được xác định bởi giá trị của p quan sát ngay trước đó, khi đó mô hình $AR(p)$ có dạng:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (1)$$

trong đó, a_t là nhiễu trắng và không tương quan với $z_t, \phi_1, \phi_2, \dots, \phi_p$ là p hệ số tự hồi quy được lựa chọn để sinh ra giá trị dự báo phù hợp nhất cho z_t qua các giá trị trong quá khứ $z_{t-1}, z_{t-2}, \dots, z_{t-p}$.

$$(1) \Leftrightarrow \phi(B)z_t = a_t \text{ với } \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

B là toán tử trễ, tức là $Bz_t = z_{t-1}$.

- *Quá trình trung bình trượt*: Mô hình trung bình trượt bậc $q - MA(q)$. Giả sử chuỗi $\{z_t\}$ là dừng và giá trị của biến z_t được xác định bởi nhiễu ngẫu nhiên hiện tại cùng với q nhiễu ngay trước đó, khi đó mô hình $MA(q)$ có dạng:

$$z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2)$$

ở đây a_t là các nhiễu ngẫu nhiên, q là bậc trung bình trượt, $\theta_1, \theta_2, \dots, \theta_q$ là các hệ số trung bình trượt.

(2) $\Leftrightarrow z_t = \theta(B)a_t$ trong đó $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$, B là toán tử trễ.

- *Quá trình tích hợp*: Phép sai phân bậc $d - I(d)$. Các mô hình trên chỉ được ứng dụng khi dữ liệu chuỗi thời gian là dừng [6, 14, 16]. Tuy nhiên, dữ liệu chuỗi thời gian trong thực tế rất hiếm khi dừng. Vì thế cần thiết phải biến đổi dữ liệu chuỗi thời gian trở thành chuỗi dừng trước khi xây dựng mô hình trên nó. Phương pháp chủ yếu được sử dụng là phép sai phân, chuỗi sau khi được biến đổi thành chuỗi dừng được gọi là chuỗi tích hợp (Integration).

Sai phân bậc d có dạng: $\Delta^d z_t = \Delta(\Delta^{d-1} z_t) \equiv (1 - B)^d z_t \quad (3)$

Ở đây sai phân bậc 1 được xác định như sau: $\Delta z_t = z_t - z_{t-1} \equiv (1 - B)z_t$.

Mô hình gộp $ARIMA(p, d, q)$ chứa đựng mô hình tự hồi quy $AR(p)$, mô hình trung bình trượt $MA(q)$, phép sai phân $I(d)$ có dạng như sau:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d z_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)a_t \quad (4)$$

hay $\phi(B)\Delta^d z_t = \theta(B)a_t$ trong đó a_t là nhiễu trắng $\sim WN(0, \sigma_a^2)$.

Mô hình $ARIMA(p, d, q)$ chỉ được ứng dụng để dự báo dữ liệu chuỗi thời gian không có yếu tố mùa vụ, trong khi có rất nhiều dữ liệu chuỗi thời gian thể hiện các mẫu lặp lại theo mùa vụ. Yếu tố mùa vụ chỉ ra rằng các mẫu lặp lại đều đặn trong khoảng thời gian cố định như hàng ngày, hàng tuần, hàng tháng, hàng năm, hàng thập kỷ. Ký hiệu khoảng thời gian cố định biểu thị cho kỳ dữ liệu là s và nếu dữ liệu được tính theo tháng thì s là 11 hoặc 12 tháng trong năm, nếu theo quý thì s là 4, theo ngày trong tuần thì s là 7 hoặc theo ngày làm việc thì s lại là 5,...

Trước đây, để dự báo dữ liệu chuỗi thời gian có yếu tố mùa vụ người ta thường sử dụng các mô hình $X_{11}ARIMA$ với khoảng thời gian là 11 tháng, hoặc $X_{12}ARIMA$ với khoảng thời gian là 12 tháng [9, 18]. Mô hình $SARIMA$ ra đời muộn hơn, vào nửa cuối những năm

1990, và cũng được phát triển từ mô hình $ARIMA$ [6, 14, 17] đồng thời cho cả 3 quá trình: tự hồi quy (AR), tích hợp hay sai phân (I) và trung bình trượt (MA) vì thế $SARIMA$ có thể được viết dưới dạng là $SARIMA(p, d, q) \times (P, D, Q)_s$ và nó phức tạp hơn $X_{11}ARIMA$ và $X_{12}ARIMA$, là những mô hình được phát triển từ $ARIMA$ chỉ ở quá trình trung bình trượt (MA).

b. Mô hình $SARIMA$

Mô hình $SARIMA(p, d, q) \times (P, D, Q)_s$ tổng quát có dạng:

$$\phi(B)\phi_S(B)\Delta^d\Delta_S^D(z_t) = \theta(B)\theta_S(B)a_t \quad (5)$$

trong đó, $\phi_S(B)$ và $\theta_S(B)$ lần lượt là đa thức tự hồi quy theo mùa bậc $P - SAR(P)$ và đa thức trung bình trượt theo mùa bậc $Q - SMA(Q)$.

$$\phi_S(B) = 1 - \phi_{S,1}B^S - \phi_{S,2}B^{2S} - \dots - \phi_{S,P}B^{SP}$$

$$\theta_S(B) = 1 - \theta_{S,1}B^S - \theta_{S,2}B^{2S} - \dots - \theta_{S,P}B^{PS}$$

$\phi(B)$ và $\theta(B)$ là đa thức tự hồi quy thường bậc $p - AR(p)$ và đa thức trung bình trượt bậc $q - MA(q)$. a_t là nhiễu trắng $\sim WN(0, \delta^2)$ và sai phân theo trễ mùa vụ bậc $D : \Delta_S^D z_t = (1 - B^S)^D z_t$.

Khác với $X_{11}ARIMA$ và $X_{12}ARIMA$, mô hình $SARIMA$ phù hợp với bất kỳ dữ liệu chuỗi thời gian mùa vụ nào, tức là nó đáp ứng mọi độ dài về mùa vụ. Mùa vụ có thể là 7 ngày trong tuần; 4 quý trong năm; 11, 12 tháng trong một năm,... Mô hình này đang được ứng dụng để dự báo đầu tư nước ngoài vào Mỹ, dự báo chỉ số chứng khoán S&P500 [18-19],...

Trong [1, 6, 14] nhận xét rằng các mô hình $ARIMA$ và $SARIMA$ tương ứng là những mô hình 'tốt nhất' để dự báo ngắn hạn dữ liệu chuỗi thời gian không có hoặc có yếu tố mùa vụ. Bài báo này sẽ sử dụng mô hình $SARIMA$ để giải quyết 2 bài toán được đặt ra ở trên.

2.2. Phần mềm hỗ trợ xây dựng mô hình dự báo EVIEW

EView được sử dụng rộng rãi trong các lĩnh vực kinh tế. Nó rất hiệu quả để xử lý các dữ liệu chuỗi thời gian về kinh tế. Nó được sử dụng để phân tích, hồi quy, đánh giá, ước lượng, mô phỏng và dự báo dữ liệu chuỗi thời gian [16]. Với EView người ta có thể phát hiện nhanh những quan hệ thống kê từ dữ liệu và sử dụng quan hệ đó để dự báo giá trị tương lai của dữ liệu. EView được sử dụng trong các lĩnh vực phân tích, dự báo các dữ liệu khoa học, tài chính; mô phỏng và dự báo kinh tế vĩ mô, mô phỏng và dự báo nhu cầu tiêu dùng, giá cả thị trường [7, 16]...

3. QUY TRÌNH ỨNG DỤNG MÔ HÌNH SARIMA

Các giai đoạn xây dựng mô hình $SARIMA$ cũng tương tự như với mô hình $ARIMA$, nhưng phải thực hiện thêm một số nội dung khác nữa trong từng giai đoạn [17-19], cụ thể:
Bước 1 - Giai đoạn nhận dạng mô hình. Xác định giá trị của các tham số (D, d, p, P, q, Q) và kỳ dữ liệu s , trong đó, trước hết phải xác định bậc sai phân theo mùa vụ D và sai phân thường d nhằm biến đổi dữ liệu chuỗi thời gian thành chuỗi dừng; sau đó kiểm tra ACF và $PACF$ để xác định s và xác định bậc tự hồi quy p và tự hồi quy theo mùa vụ P , bậc trung bình trượt q và trung bình trượt theo mùa vụ Q .

Bước 2 - Giai đoạn ước lượng mô hình. Ước lượng các hệ số ϕ, ϕ_S và θ, θ_S tương ứng là các hệ số tự hồi quy và tự hồi quy theo mùa vụ, các hệ số trung bình trượt và trung bình trượt theo mùa vụ. Sử dụng phương pháp ước lượng cực đại hợp lý để ước lượng giá trị của các hệ số này.

Bước 3 - Giai đoạn kiểm định. Kiểm định tính hợp lý của mô hình SARIMA được lựa chọn, bao gồm kiểm định các tham số và kiểm định phần dư. Nếu kiểm định mô hình được lựa chọn không thỏa mãn thì bắt đầu lại từ giai đoạn nhận dạng để lựa chọn mô hình khác hợp lý hơn.

Bước 4 - Dự báo. Dựa trên mô hình được lựa chọn thực hiện dự báo giá trị tương lai của dữ liệu chuỗi mùa vụ cũng như khoảng tin cậy của dự báo. Giá trị tương lai có thể được sử dụng, cập nhật vào mô hình để dự báo cho kỳ dữ liệu hoặc mùa vụ kế tiếp.

Dự báo dữ liệu chuỗi thời gian có thể được thực hiện theo một trong hai phương pháp: dự báo trước một bước ngoài mẫu hoặc dự báo trước nhiều bước ngoài mẫu. Thực tế cho thấy rằng dự báo trước một bước ngoài mẫu là chính xác hơn dự báo trước nhiều bước ngoài mẫu [6, 14, 17]. Bài báo này chỉ ứng dụng phương pháp dự báo trước một bước ngoài mẫu để dự báo dữ liệu chuỗi thời gian cho một số kỳ tiếp theo.

4. DỰ BÁO NGẮN HẠN TĂNG TRƯỞNG KINH TẾ VÀ LẠM PHÁT BẰNG MÔ HÌNH SARIMA

4.1. Dự báo tăng trưởng GDP Việt Nam theo quý

a. Đặc điểm số liệu

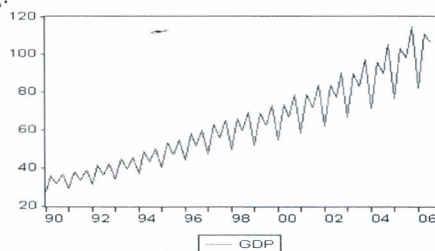
Số liệu về GDP theo quý do Tổng cục Thống kê công bố được tính theo giá so sánh 1994 nhằm loại trừ biến động của yếu tố giá cả qua các năm bao gồm GDP của các quý từ quý 1 năm 1990 (Q1/1990) đến quý 4 năm 2006 (Q4/2006). Tập số liệu này gồm 68 quan sát, có giá trị trung bình là 63,28 nghìn tỷ đồng; trung vị: 59,77 nghìn tỷ; giá trị lớn nhất là 125,32 nghìn tỷ và giá trị nhỏ nhất là 27,64 nghìn tỷ; độ lệch chuẩn là 23,8 nghìn tỷ.

Để xây dựng mô hình dự báo GDP quý, ta chia chuỗi số liệu đã có thành hai tập dữ liệu. Tập số liệu thứ nhất dùng để phát hiện và xây dựng mô hình dự báo gồm dữ liệu GDP quý từ Q1/1990 đến Q3/2006; tập thứ 2, dữ liệu GDP của Q4/2006 dùng để kiểm thử mô hình mà nội dung thực chất là đánh giá sai số của kết quả dự báo so với thực tế xảy ra;

b. Xây dựng mô hình dự báo và kiểm nghiệm

Bước 1. Nhận dạng mô hình

- Phân tích đồ thị dữ liệu chuỗi GDP quý (Hình 1) cho thấy GDP có xu thế tăng và biến động có tính chu kỳ: tăng ở quý 2 và quý 4; giảm vào quý 1 và quý 3 hàng năm. Điều đó gợi ý rằng cần phải khử tính xu thế và tính mùa vụ của dữ liệu chuỗi GDP quý và biến đổi dữ liệu chuỗi thành chuỗi dừng.



Hình 1. Đồ thị chuỗi GDP

- Biến đổi dữ liệu chuỗi thành chuỗi dừng: Biến đổi chuỗi GDP qua hàm logarit thành chuỗi LNGDP nhằm làm giảm mạnh tính xu thế của chuỗi ban đầu, mặc dù chuỗi kết quả sau biến đổi vẫn còn tính mùa vụ, sau đó thực hiện một số phép sai phân trên chuỗi LNGDP nhằm chuyển nó thành chuỗi dừng.

Date: 12/07/06 Time: 15:16
 Sample: 1990Q1 2006Q3
 Included observations: 67

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.732	0.732	37.514	0.000
		2	0.792	0.552	82.110	0.000
		3	0.606	-0.179	108.61	0.000
		4	0.701	0.288	144.66	0.000
		5	0.476	-0.329	161.54	0.000
		6	0.538	0.059	183.50	0.000
		7	0.380	0.020	194.61	0.000
		8	0.470	0.086	211.94	0.000
		9	0.284	-0.123	218.39	0.000
		10	0.347	-0.023	228.16	0.000
		11	0.215	0.033	231.99	0.000
		12	0.299	0.029	239.48	0.000
		13	0.148	-0.043	241.36	0.000
		14	0.207	-0.042	245.11	0.000
		15	0.100	0.025	246.00	0.000
		16	0.173	0.001	248.70	0.000
		17	0.054	-0.008	248.97	0.000
		18	0.105	-0.041	250.01	0.000

Hình 2. Lược đồ ACF-PACF của chuỗi LNGDP

Phân tích lược đồ tương quan của chuỗi LNGDP (Hình 2). Trong thực hành, vì dữ liệu chuỗi thời gian được coi là dừng khi có trên 95% quan sát có giá trị ACF xấp xỉ bằng 0 một cách có ý nghĩa (tức là nằm giữa 2 đường kẻ vạch) và do trong 67 quan sát của chuỗi LNGDP, có đến trên 10 quan sát có giá trị ACF khác 0 (Hình 2) một cách có ý nghĩa, vượt quá giới hạn 5% số quan sát, nên chuỗi LNGDP là không dừng.

Kiểm định nghiệm đơn vị (hay kiểm định Dickey-Fuller: DF) cũng cho thấy LNGDP không phải là chuỗi dừng. Giá trị tuyệt đối được tính toán qua kiểm định DF bằng 1,85, nhỏ hơn giá trị tuyệt đối tại các mức ý nghĩa 1% (2,6) và 5% (1,94) như hình dưới.

Null Hypothesis: LNGDP has a unit root
 Exogenous: None
 Lag Length: 8 (Automatic based on AIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	1.855603	0.9838
Test critical values:		
1% level	-2.605442	
5% level	-1.946549	
10% level	-1.613181	

*MacKinnon (1996) one-sided p-values.

Phân tích lược đồ tương quan của chuỗi LNGDP (Hình 2) ta thấy ACF khác 0 tại các trễ thứ 4, 8, 12 đã gợi ý rằng trong chuỗi LNGDP có sự tự tương quan theo mùa vụ với trễ mùa vụ $s = 4$, vì thế cần phải loại bỏ quan hệ tự tương quan theo mùa vụ bằng cách sai phân bậc 1 theo trễ mùa vụ.

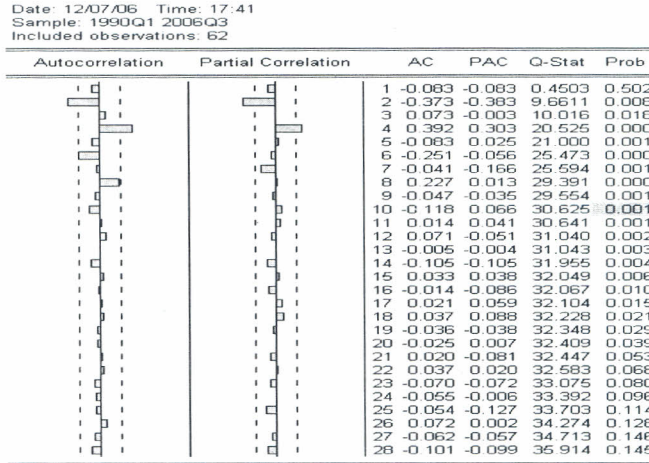
Chuỗi thu được sau khi lấy sai phân mùa vụ là GDP2. Kiểm định nghiệm đơn vị cho thấy chuỗi GDP2 không dừng do giá trị tuyệt đối được tính toán qua tiêu chuẩn DF nhỏ hơn giá trị tuyệt đối tại các điểm tới hạn 1%, 5% và 10% như được mô tả ở dưới.

Null Hypothesis: GDP2 has a unit root
 Exogenous: None
 Lag Length: 4 (Automatic based on AIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.442027	0.5188
Test critical values:		
1% level	-2.605442	
5% level	-1.946549	
10% level	-1.613181	

*MacKinnon (1996) one-sided p-values.

Tiếp tục lấy sai phân thường bậc 1 của chuỗi GDP2, được chuỗi GDP3. Lược đồ tương quan của chuỗi GDP3 (Hình 3) cho thấy ACF chỉ khác 0 một cách có ý nghĩa tại các trễ 2, 4, còn lại giá trị ACF tại các trễ khác là bằng 0 có ý nghĩa. Điều này gợi ý rằng GDP3 có thể là chuỗi dừng.



Hình 3. Lược đồ ACF-PACF của chuỗi GDP3

Thực hiện kiểm định nghiệm đơn vị cho thấy chuỗi GDP3 là dừng.

Null Hypothesis: GDP3 has a unit root
 Exogenous: None
 Lag Length: 3 (Automatic based on AIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.401773	0.0010
Test critical values:		
1% level	-2.605442	
5% level	-1.946549	
10% level	-1.613181	

*MacKinnon (1996) one-sided p-values.

Bước 2. Ước lượng mô hình

- Mô hình $SARIMA(p, d, q) \times (P, D, Q)_s$ đối với chuỗi GDP3 được xác định như sau:
- Mùa vụ $s = 4$, bậc sai phân thường $d = 1$, sai phân mùa vụ $D = 1$;
- Phân tích giá trị ACF trong lược đồ tương quan của GDP3 (Hình 4) ta nhận được bậc trung bình trượt bậc thường $q = 1$, bậc trung bình trượt theo mùa vụ bậc $Q = 2$;

Dependent Variable: D(D(LOG(GDP),0,4))
 Method: Least Squares
 Date: 01/16/07 Time: 08:20
 Sample (adjusted): 1992Q3 2006Q3
 Included observations: 57 after adjustments
 Convergence achieved after 128 iterations
 Backcast: 1990Q2 1992Q2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.836243	0.079977	10.45603	0.0000
SAR(4)	-0.881139	0.060500	-14.56420	0.0000
MA(1)	-0.961446	0.047411	-20.27880	0.0000
SMA(4)	1.669116	0.060279	27.68983	0.0000
SMA(8)	0.717489	0.051040	14.05738	0.0000

R-squared	0.491193	Mean dependent var	-2.47E-05
Adjusted R-squared	0.452054	S.D. dependent var	0.008597
S.E. of regression	0.006364	Akaike info criterion	-7.192756
Sum squared resid	0.002106	Schwarz criterion	-7.013541
Log likelihood	209.9936	Durbin-Watson stat	2.064099

Inverted AR Roots	.84	.69-.69i	.69+.69i	-.69+.69i
	-.69-.69i			
Inverted MA Roots	.96	.71+.65i	.71-.65i	.65+.71i
	-.65-.71i	-.65+.71i	-.65-.71i	-.71+.65i
	-.71-.65i			

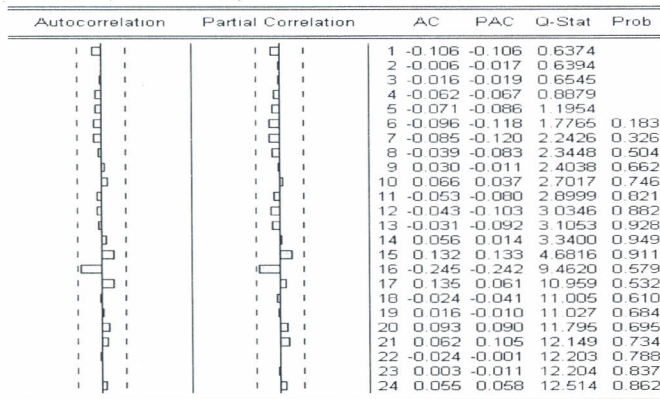
Phân tích giá trị PACF trong Lược đồ tương quan của GDP3 (Hình 4) ta cũng có: bậc tự hồi quy thường $p = 1$, bậc tự hồi quy theo mùa vụ $P = 1$;

Ước lượng mô hình $SARIMA(1, 1, 1) \times (1, 1, 2)_4$:

Với mô hình $SARIMA(1, 1, 1) \times (1, 1, 2)_4$ mặc dù rằng chỉ có 49,5% sự biến động của biến số GDP giải thích được từ mô hình, nhưng kiểm định Durbin-Watson cho kết quả là 2,06 nằm giữa (1,8 và 2,2) là khá tốt và chỉ ra rằng không tồn tại tự tương quan trong chuỗi GDP3; mặt khác các giá trị Inverted AR roots và Inverted MA roots đều nhỏ hơn 1, chứng tỏ quá trình AR và MA thoả mãn điều kiện khả nghịch nên cũng có thể tiếp tục tiến hành kiểm định phần dư để chấp nhận mô hình này.

Bước 3. Kiểm định phần dư

Kiểm định chuỗi phần dư của mô hình $SARIMA(1,1,1) \times (1,1,2)_4$ qua lược đồ tương quan chuỗi phần dư của mô hình (Hình 4), ta thấy phần dư là nhiễu trắng vì các giá trị ACF của chuỗi phần dư đều nằm giữa 2 đường kẻ vạch. Vì vậy có thể sử dụng mô hình $SARIMA(1,1,1) \times (1,1,2)_4$ để dự báo.



Hình 4. Lược đồ ACF-PACF của chuỗi phần dư

Bước 4. Dự báo, kiểm định so với thực tế

Dự báo giá trị GDP cho quý 4/2006 theo mô hình $SARIMA(1, 1, 1) \times (1, 1, 2)_4$ và sử dụng dữ liệu GDP thực hiện của quý 4/2006 để kiểm thử, ta nhận được kết quả sau:

Giá trị thực hiện: 125,327 nghìn tỷ; Giá trị dự báo: 125,044 nghìn tỷ. % Sai số tuyệt đối: $100 \times |125,327 - 125,044| / 125,327 = 0,23\%$.

Như vậy mô hình $SARIMA(1, 1, 1) \times (1, 1, 2)_4$ cho kết quả dự báo là khá tốt và có thể chấp nhận mô hình đó để dự báo một số quý tiếp theo trong năm 2007.

c. Dự báo chính thức. Sử dụng mô hình $SARIMA(1, 1, 1) \times (1, 1, 2)_4$ để dự báo GDP quý 1 năm 2007. Kết quả dự báo với độ tin cậy trên 95%: GDP quý 1/2007 đạt mức thấp nhất là 86,942 nghìn tỷ đồng, cao nhất là 89,337 nghìn tỷ và mức trung bình là 88,14 nghìn tỷ.

Nếu so với GDP đạt được ở quý 1/2006 là 81,98 nghìn tỷ đồng thì tốc độ tăng trưởng GDP quý 1 năm 2007 so với cùng kỳ năm 2006 là 7,5%. So sánh với thực tế, theo Tổng cục Thống kê, tốc độ tăng trưởng GDP quý 1/2007 ở mức 7,69% so với quý 1/2006, nói cách khác GDP quý 1/2007 đạt khoảng 88,284 nghìn tỷ và như vậy % sai số tuyệt đối giữa GDP thực tế đạt được và GDP dự báo mức trung bình của quý 1/2007 là: 0,16%.

4.2. Dự báo chỉ số giá tiêu dùng theo tháng, quý

a. Đặc điểm dữ liệu

Tập số liệu về chỉ số giá tiêu dùng (CPI) Việt Nam theo tháng kể từ tháng 01/1995 đến tháng 12/2006 bao gồm 132 quan sát, CPI trung bình các tháng là 100,58 so với tháng trước là 100, độ lệch chuẩn là 0,96, CPI cao nhất là 103,8 và thường rơi vào tháng 2 dương lịch hàng năm, CPI thấp nhất khoảng 98,9 (thiếu phát), rơi vào tháng 3 dương lịch. Chỉ trong vòng một tháng, từ tháng 2 dương lịch cho tới tháng 3 dương lịch, CPI giảm đáng kể, thông thường dao động xung quanh giá trị trung bình khoảng +/-3

b. Xây dựng mô hình dự báo và kiểm nghiệm

Chia tập số liệu làm 2 tập: tập thứ nhất được sử dụng để xây dựng mô hình dự báo gồm số liệu CPI từ tháng 1/1995 đến tháng 10/2006; tập thứ 2 được sử dụng để kiểm thử chấp nhận mô hình gồm số liệu CPI tháng 11 và 12/2006.

Thực hiện một số vòng lặp 3 bước đầu trong quy trình 4 bước trong xây dựng mô hình dự báo trên tập dữ liệu thứ nhất và thực hiện các kiểm định thống kê cần thiết ta nhận được mô hình dự báo phù hợp của dữ liệu chuỗi thời gian CPI tháng là: $SARIMA(2, 1, 1) \times (1, 0, 1)_{12}$.

Sử dụng dữ liệu thực tế trong tập thứ hai để kiểm thử mô hình dự báo trên, kết quả như sau:

Thời gian	Tháng 11/2006	Tháng 12/2006
Giá trị thực tế	100,6	100,5
Giá trị dự báo với độ tin cậy trên 95%		
Ở mức cao nhất	101,50	100,99
Ở mức thấp nhất	99,55	100,00
Ở mức trung bình	100,53	100,77
% sai số tuyệt đối ở mức trung bình	0,069%	0,268%

c. Dự báo CPI quý 1 năm 2007

Ứng dụng mô hình $SARIMA(2, 1, 1) \times (1, 0, 1)_{12}$ cho chuỗi số liệu CPI từ 1/1995 đến 12/2006, để dự báo CPI cho cả 3 tháng thuộc quý 1/2007 theo phương pháp một bước ngoài mẫu, tức là lấy kết quả dự báo của tháng trước để xác định lại hệ số của mô hình và tính kết quả dự báo cho tháng sau, ta nhận được kết quả dự báo với độ tin cậy trên 95% như sau:

Thời gian	CPI ở mức thấp nhất	CPI ở mức trung bình	CPI ở mức cao nhất
1/2007	99,81	100,77	101,73
2/2007	100,99	101,95	102,91
3/2007	98,78	99,70	100,66

Điều này có nghĩa là với độ chính xác trên 95%, mức độ lạm phát CPI tháng 1/2007 ở các mức thấp nhất, trung bình và cao nhất tương ứng là -0,19%, 0,77% và 1,73%; tương tự đối với tháng 2/2007 là 0,99%, 1,95% và 2,91% và đối với tháng 3/2007 tương ứng sẽ là: -1,22%, -0,3% và 0,66%.

Như vậy với độ chính xác là trên 95%, dự báo mức độ lạm phát CPI quý 1/2007 theo các mức độ thấp nhất, trung bình và cao nhất tương ứng sẽ là: -0,42%, 2,42% và 5,4%. Nói chung mức độ lạm phát CPI quý 1/2007 xoay quanh mức 2,42%.

So sánh với thực tế: Mức độ lạm phát CPI do Tổng cục Thống kê công bố tương ứng tháng 1/2007 là 1,05%, tháng 2 là 2,17%.

Bây giờ nếu bổ sung CPI thực tế của tháng 1/2007 là 101,05 vào tập dữ liệu để xác định lại các hệ số trung bình trượt và tự hồi quy của mô hình $SARIMA(2,1,1) \times (1,0,1)_{12}$, thì với độ tin cậy trên 95%, kết quả dự báo CPI các tháng 2/2007 và 3/2007 như sau:

Thời gian	Thấp nhất	Cao nhất	Trung bình	Thực hiện	% Sai số tuyệt đối
2/2007	101,2	103,1	102,18	102,17	0,0097%
3/2007	98,91	100,81	99,86	99,8	0,0601%

Có nghĩa là dự báo mức độ lạm phát CPI tháng 2/2007 ở mức trung bình là 2,18% và thiếu phát vào tháng 3/2007 là - 0,14%, và dự báo mức độ lạm phát CPI cho quý 1/2007 ở mức trung bình là: 3,09%.

Thực tế theo Tổng cục Thống kê, mức độ lạm phát CPI trong tháng 2/2007 là 2,17% và tháng 3/2007 là: -0,2% và cho cả quý 1/2007 là 3,02%.

Kết quả phân tích trên cho thấy độ chính xác của dự báo bằng sử dụng mô hình *SARIMA* (hoặc *ARIMA*) sẽ càng cao nếu thời gian dự báo càng gần.

5. ĐỀ XUẤT ỨNG DỤNG MÔ HÌNH ARIMA TRONG DỰ BÁO NGẮN HẠN VỀ KINH TẾ VĨ MÔ

5.1. Ưu nhược điểm của mô hình ARIMA, SARIMA trong dự báo kinh tế ngắn hạn

Ưu điểm:

- Cơ sở toán học của mô hình ARIMA cũng như SARIMA là lý thuyết về quá trình ngẫu nhiên, việc tính toán được đảm bảo về mặt toán học thống kê. Các mô hình đó cung cấp một công cụ dự báo vững chắc về mặt lý thuyết, có sức thuyết phục và dễ dàng được chấp nhận bởi các nhà nghiên cứu và ứng dụng.

- Các mô hình *ARIMA*, *SARIMA* được xây dựng một cách tổng quát, do vậy nó có thể được sử dụng cho cả dữ liệu chuỗi thời gian dừng cũng như không dừng; với các dữ liệu chuỗi thời gian có hoặc không có thành phần mùa vụ.

- Quá trình thích nghi mô hình với dữ liệu chuỗi thời gian đã cho rất linh hoạt, thông qua nhiều cách kết hợp các giá trị tham số p, d, q cũng như P, D, Q nếu dữ liệu chuỗi thời gian có tính thành phần mùa vụ. Từ đó theo những tiêu chuẩn kiểm định thống kê, có thể rút ra được một mô hình dự báo thích hợp nhất, phản ánh khách quan nhất quy luật dữ liệu và cho kết quả dự báo có độ chính xác khá cao...

- Mô hình *SARIMA* đặc biệt thích hợp cho các dự báo ngắn hạn và cho dữ liệu có yếu tố mùa vụ. Về phương diện này *SARIMA* hiện được xem là mô hình "tốt nhất".

Nhược điểm:

Mô hình *ARIMA*, *SARIMA* đòi hỏi người sử dụng phải có một trình độ toán học nhất định, và nhiều khi rất phức tạp trong xử lý, ước lượng các tham số và kiểm định thống kê. Do đó, ứng dụng phương pháp này thường tốn nhiều thời gian và công sức, cũng như các chi phí khác.

Mô hình tuy rất khách quan song vẫn có phần mang tính chủ quan ở một số khâu trong việc nhận định mô hình thích hợp, bởi lẽ ở khâu này phụ thuộc chủ yếu vào khả năng, kinh nghiệm của người làm công tác dự báo.

Nếu khoảng dự báo lớn hơn một thời kỳ dự báo thì sẽ cho một kết quả dự báo kém hơn vì trong phương trình dự báo chưa thể sử dụng hết các giá trị cần có.

Mô hình *ARIMA*, *SARIMA* chỉ sử dụng cho một dữ liệu chuỗi thời gian đơn lẻ, do đó không thể phát hiện các nhân tố ảnh hưởng nhân quả đến các giá trị quan sát của dữ liệu chuỗi thời gian.

Để dự báo chính xác mô hình đòi hỏi nhiều quan sát (khoảng từ 50 quan sát trở lên) và nó không thích hợp đối với vấn đề phân tích tác động của cơ chế chính sách.

5.2. Đề xuất ứng dụng mô hình ARIMA, SARIMA trong dự báo ngắn hạn về kinh tế - xã hội Việt Nam

Dự báo đóng vai trò quan trọng trong việc điều hành nền kinh tế của Nhà nước, định hướng chiến lược kinh doanh của các doanh nghiệp. Nếu dự báo dài hạn hỗ trợ các quá trình xây dựng chiến lược, quy hoạch, kế hoạch 5 năm về phát triển kinh tế-xã hội thì dự báo ngắn hạn hỗ trợ quá trình ra quyết định, các hoạt động chỉ đạo điều hành nhằm thực hiện thắng lợi mục tiêu chiến lược đề ra.

Bài báo này tuy chỉ tập trung trình bày việc ứng dụng mô hình SARIMA để dự báo GDP Việt Nam theo quý, CPI theo tháng, nhưng với thực tiễn số liệu kinh tế-xã hội có được hiện nay ở nước ta, theo giải pháp được đề xuất hoàn toàn có thể dự báo được những chỉ tiêu kinh tế-xã hội quan trọng sau:

Dự báo một số chỉ tiêu kinh tế - xã hội chủ yếu theo quý: như GDP, GDP theo các ngành; Giá trị xuất nhập khẩu và chỉ số giá xuất nhập khẩu theo tháng, quý; đầu tư trực tiếp nước ngoài theo quý; tình hình giải ngân vốn ODA theo quý; lạm phát chỉ số giá tiêu dùng và tỷ giá hối đoái theo tháng, quý;...

Dự báo giá cả hoặc chỉ số giá của những mặt hàng xuất, nhập khẩu chủ yếu của nền kinh tế Việt Nam theo tháng; giá cả hoặc chỉ số giá những mặt hàng thiết yếu của đời sống dân sinh theo tháng;

Dự báo tăng trưởng vốn, lao động, việc làm,..., trong các loại hình doanh nghiệp thuộc các thành phần kinh tế và dự báo lực lượng lao động, lực lượng lao động được đào tạo nghề toàn quốc theo năm.

Dự báo nhiều hiện tượng biến động của thời tiết, môi trường khác,

6. MỘT SỐ KẾT LUẬN

Về mặt lý thuyết, ARIMA là mô hình thích hợp cho việc dự báo ngắn hạn một số hiện tượng kinh tế - xã hội mà số liệu đo các hiện tượng đó tạo thành dữ liệu chuỗi thời gian và trong trường hợp dữ liệu chuỗi thời gian có yếu tố mùa vụ thì mô hình SARIMA được phát triển tiếp từ mô hình này sẽ là mô hình thích hợp.

Bài báo này cho thấy rằng, với thực tiễn số liệu phản ánh nền kinh tế Việt Nam hiện có, có thể sử dụng mô hình ARIMA và SARIMA để dự báo ngắn hạn về tăng trưởng kinh tế. Những vấn đề kinh tế-xã hội nào nữa ở Việt Nam có thể được dự báo bằng mô hình này cũng được đề xuất trong bài báo.

Mô hình SARIMA xuất hiện khoảng 10 năm gần đây, ở Việt Nam còn khá mới và chưa được tổ chức nghiên cứu ứng dụng một cách có hệ thống. Bài báo này đã trình bày một cách tóm tắt quy trình dự báo những chỉ tiêu kinh tế - xã hội quan trọng dựa trên dữ liệu có được trong quá khứ và hiện tại bằng mô hình SARIMA.

Như đã biết mô hình ARIMA, SARIMA là mô hình đơn biến, tức là nó chỉ dựa trên bản thân một dữ liệu chuỗi thời gian. Trong trường hợp có nhiều dữ liệu chuỗi thời gian không có yếu tố mùa vụ và trễ của chúng có quan hệ với nhau thì làm cách nào có thể dự báo dữ liệu tương lai từ dữ liệu quá khứ và hiện tại của chúng. Mô hình tự hồi quy véc tơ VAR mà thực chất là mô hình ARMA nhiều biến có thể sẽ giúp trả lời câu hỏi này. Trường hợp các dữ liệu chuỗi thời gian đó có yếu tố mùa vụ và các trễ của chúng có quan hệ với nhau thì câu trả lời vẫn còn đang mở. Việc xây dựng, phát triển những mô hình phù hợp cho những tình huống như vậy đang được quan tâm nghiên cứu, phát triển.

TÀI LIỆU THAM KHẢO

- [1] Boris Kovalerchuk, Evgenii Vityaev, *Data mining in finance, Advances in relational and hybrid method*, Kluwer Academic Publishers, 2001.
- [2] Đỗ Văn Thành, Cơ sở dữ liệu chủ đề về kinh tế vĩ mô, *Tạp chí Tin học và Điều khiển* **24** (1) (2008).
- [3] Đỗ Văn Thành, Dự báo KT-XH theo cách tiếp cận của Phát hiện tri thức từ Cơ sở dữ liệu và một số vấn đề cần triển khai, “Báo cáo Hội nghị Ứng dụng Toán trong Công nghiệp lần 2”, Hà Nội, 12-2005.
- [4] Đỗ Văn Thành, Phát hiện luật kết hợp, “Bài giảng Trường thu hệ mở và Ứng dụng”, Viện Toán học, Hà Nội 8-2002.
- [5] Đỗ Văn Thành, Cơ sở dữ liệu chủ đề phục vụ phân tích và dự báo, *Tạp chí Thông tin và Dự báo KT-XH*, Tháng 42-48, 2005; LAM CHO RO!
- [6] Đỗ Văn Thành, “Ứng dụng mô hình *ARIMA* trong dự báo ngắn hạn về tăng trưởng kinh tế”, Đề tài NCKH cấp Bộ, Bộ Kế hoạch và Đầu tư, 1/2007.
- [7] Greene, H. William, *Econometric Analysis*, Fourth edition, New York, Prentice Hall, 2000.
- [8] Kennedy, Peter, *A Guide to Econometrics*, Fourth edition, Cambridge, The MIT Press, 1998.
- [9] Hossein Arsham, Time-Critical Decision Making for Business Administration, in <http://www.mirror-service.org/sites/home.ubalt.edu/ntsbarsh/>
- [10] D.R. Osborn, A survey of seasonality in UK macroeconomic variables, *International Journal of Forecasting* **6** (1990) 327–336.
- [11] Liên minh châu Âu và Bộ Kế hoạch Đầu tư, Đo lường nền kinh tế, Tập 1+2+3, dự án: ”Hỗ trợ xây dựng Hệ thống thông tin Quốc gia phục vụ lập kế hoạch kinh tế -ALA/VIE/95/12”, 12/2000.
- [12] Liên minh châu Âu và Bộ Kế hoạch Đầu tư, *Cẩm nang về Báo cáo Tổng quan Kinh tế Quý*, dự án: “Hỗ trợ xây dựng Hệ thống thông tin Quốc gia phục vụ lập kế hoạch kinh tế -ALA/VIE/95/12”, 12-2001.
- [13] Nguyễn Khắc Minh, *Các phương pháp phân tích và dự báo trong kinh tế*, NXB Khoa học và Kỹ thuật Hà Nội, 2002.
- [14] Tạ Mạnh Cường, “Dự báo chuỗi dữ liệu phụ thuộc thời gian theo mùa vụ bằng mô hình Holt-Winter”, Luận văn thạc sỹ, Khoa CNTT, Đại học Công nghệ, Đại học Quốc gia Hà Nội, 11-2006.
- [15] Thomas Reinart, *Stages of the Discovery Process, in Handbook of Data Mining and Knowledge Discovery* (Ed. by Willi Klossgen and Jan M. Zytkow), Oxford 2002.
- [16] Trần Văn Thái, “Phát hiện tri thức theo mùa vụ từ cơ sở dữ liệu chuỗi thời gian”, Luận văn thạc sỹ, Khoa CNTT, Đại học Công nghệ, Đại học Quốc gia Hà Nội, 12-2005.
- [17] Philip Hans Franses, “Forecasting Seasonal Time Series”, Econometric Institute, Erasmus University Rotterdam (2004).
- [18] Ronald Bewley, “Time Series Forecasting, UNSW”, Science Department, LGRC University of California (2000).
- [19] Shashank Shekhar, “Recursive methods for forecasting short-term traffic flow using seasonal ARIMA time series model”, North Carolina State University, 2004.
- [20] Willi Klossgen and Jan M. Zytkow, *Handbook of Data Mining and Knowledge Discovery*, Oxford, 2002.

Nhận bài ngày 22 - 3 - 2007
 Nhận lại sau sửa ngày 27 - 11 - 2007