
ỨNG DỤNG MÔ HÌNH DỮ LIỆU TẦN SUẤT HỖN HỢP DỰ BÁO TĂNG TRƯỞNG KINH TẾ VIỆT NAM

Lê Mai Trang

Trường Đại học Thương mại
Email: lmtrang2000@tmu.edu.vn

Hoàng Anh Tuấn

Trường Đại học Thương mại
Email: hoanganhtuan@tmu.edu.vn

Đinh Thị Hà

Trường Đại học Thương mại
Email: hadinh@tmu.edu.vn

Nguyễn Thị Hiền

Trường Đại học Thương mại
Email: hiennguyen@tmu.edu.vn

Trần Kim Anh

Trường Đại học Thương mại
Email: trankimanh@tmu.edu.vn

Mã bài: JED-578

Ngày nhận: 13/03/2022

Ngày nhận bản sửa: 06/06/2022

Ngày duyệt đăng: 04/07/2022

Tóm tắt:

Dự báo tăng trưởng kinh tế luôn là mối quan tâm không chỉ của các nhà nghiên cứu mà còn của các nhà hoạch định chính sách của mỗi quốc gia trên thế giới. Đã có nhiều công trình nghiên cứu đưa ra các phương pháp khác nhau để dự báo tăng trưởng GDP, các phương pháp dự báo trước đây đều phân tích dựa trên bộ dữ liệu mà trong đó các biến quan sát phải đưa về cùng một tần suất, điều này có thể làm tăng sai số của ước lượng và bỏ sót những yếu tố quan trọng có tác động đến tăng trưởng kinh tế. Để sử dụng đầy đủ và hiệu quả thông tin kinh tế vĩ mô và tài chính, bài báo này ứng dụng mô hình phân tích dữ liệu với tần suất hỗn hợp MIDAS và mô hình MF-VAR để dự báo tăng trưởng GDP của Việt Nam dựa trên bộ số liệu thu thập trong giai đoạn 2006 – 2020. Kết quả thực nghiệm cho thấy rằng mô hình MIDAS cho kết quả dự báo tốt so với mô hình MF-VAR.

Từ khóa: Dự báo GDP, mô hình MIDAS, mô hình MF-VAR, tăng trưởng kinh tế.

Mã JEL: C53, E01, E17

Forecast Vietnam's economic growth by using a mixed-frequency data analysis model

Abstract:

Forecasting economic growth is always a concern not only of researchers but also of policymakers of each country in the world. There have been many studies providing different methods to forecast GDP growth, the previous forecasting methods were all based on data sets in which the observed variables return to the same frequency. This can increase the error of the estimate and miss important factors affecting economic growth. In order to fully and effectively use macroeconomic and financial information, this research employs a data analysis model with mixed frequency MF-VAR and MIDAS model to forecast the GDP growth of Vietnam based on the dataset collected in the period 2006 - 2020. The results show that forecasting using the MIDAS model has better results than the MF-VAR model.

Keywords: GDP forecast, economic growth, MIDAS model, MF-VAR model

JEL Codes: C53, E01, E17

1. Đặt vấn đề

Dự báo kinh tế vĩ mô, đặc biệt là dự báo tăng trưởng GDP, luôn là một nhiệm vụ quan trọng nhưng khó khăn đối với chính phủ của một quốc gia. Do tình hình kinh tế trong nước và quốc tế ngày càng trở nên phức tạp, đặc biệt là những bất ổn xảy ra gần đây bởi những cú sốc bên ngoài như dịch bệnh COVID-19, những biến động của thị trường tài chính toàn cầu và những điều chỉnh bên trong dẫn đến việc dự báo tăng trưởng GDP của Việt Nam ngày càng khó khăn hơn. Để đưa ra dự báo tăng trưởng GDP chính xác, các mô hình dự báo cần giải quyết hai vấn đề: một là lựa chọn hợp lý các yếu tố đưa vào mô hình dự báo; hai là sử dụng hiệu quả các dữ liệu về các chỉ số kinh tế đã được các cơ quan công bố. Ở Việt Nam, đã có rất nhiều các dự báo kinh tế được đưa ra bởi những tổ chức uy tín như Ngân hàng Thế giới (WB), Quỹ tiền tệ Quốc tế (IMF), Ngân hàng Phát triển Châu Á (ADB),... Những dự báo này thường không thống nhất với nhau do mỗi tổ chức sử dụng một mô hình dự báo riêng. Các mô hình này thường được áp dụng cho tất cả các quốc gia, nó có thể bỏ qua các đặc trưng kinh tế của từng quốc gia. Hơn nữa, các mô hình dự báo tăng trưởng GDP trước đây phân tích với các biến có cùng tần suất trong cùng một giai đoạn nghiên cứu. Vì vậy chưa khai thác tốt được bộ dữ liệu đã công bố ở các tần suất khác nhau, với các độ trễ công bố cũng khác nhau. Việc ứng dụng mô hình phân tích dữ liệu có tần suất hỗn hợp để dự báo tăng trưởng GDP sẽ khắc phục được những hạn chế đó.

Hiện nay các mô hình phân tích dữ liệu có tần suất hỗn hợp đang được ứng dụng rộng rãi do tính ưu việt, toàn diện và hiệu quả trong việc cải thiện độ chính xác của dự báo. Tuy nhiên, tính đến nay các nghiên cứu về tăng trưởng GDP của Việt Nam còn rất ít, đặc biệt là nghiên cứu liên quan đến dự báo tăng trưởng GDP bằng mô hình hồi quy dữ liệu tần suất hỗn hợp. Chính vì vậy, nghiên cứu này sẽ kế thừa những lý thuyết liên quan và phương pháp hồi quy dữ liệu có tần suất hỗn hợp để xây dựng mô hình dự báo tăng trưởng GDP của Việt Nam.

2. Tổng quan nghiên cứu

2.1. Các nghiên cứu dự báo các chỉ tiêu kinh tế vĩ mô sử dụng dữ liệu cùng tần suất

Các chỉ tiêu kinh tế vĩ mô luôn được các nhà nghiên cứu quan tâm và tìm tòi những phương pháp sao cho kết quả dự báo các chỉ tiêu này có độ chính xác cao. Một số nhà nghiên cứu đã áp dụng các phương pháp khác nhau để dự báo tăng trưởng GDP và lạm phát của Việt Nam như Nguyễn Thị Thu Hằng & Vũ Phạm Hải Đăng (2012), Nguyễn Kim Anh (2015). Nếu như Nguyễn Thị Thu Hằng & Vũ Phạm Hải Đăng (2012) kết hợp giữa mô hình VAR và mô hình Bayesian VAR (BVAR) để dự báo thì Nguyễn Kim Anh (2015) sử dụng mô hình VAR và VECM. Nghiên cứu đã chứng minh được rằng mô hình BVAR có độ chính xác dự báo cao hơn mô hình VAR, cả khi dự báo lạm phát lẫn dự báo tăng trưởng GDP. Các sai số dự báo của mô hình BVAR nhỏ hơn nhiều so với mô hình VAR trong các dự báo từ một đến bốn kỳ về tương lai. Theo Nguyễn Kim Anh (2015), công tác dự báo lạm phát và tăng trưởng đã đóng góp quan trọng trong việc đạt được hiệu quả điều hành kinh tế vĩ mô của Chính phủ và các bộ, ngành. Việc sử dụng mô hình VAR và VECM để đưa ra dự báo tăng trưởng kinh tế và lạm phát với bộ số liệu tần suất quý từ I/2001 đến hết quý IV/2014 đã cho thấy tỷ lệ tăng trưởng GDP ở mức 6 - 6,2%; lạm phát dao động ở mức 4,4 - 5,5%. Mục tiêu tăng trưởng kinh tế của Chính phủ thuộc cận trên của miền dự báo. Để cung cấp các thông tin tiên nghiệm cho mô hình ước lượng BVAR, Nguyễn Đức Trung & Nguyễn Hoàng Chung (2017), sử dụng mô hình cân bằng động ngẫu nhiên tổng quát (DSGE). Mô hình được xây dựng và hiệu chỉnh với mục tiêu dự báo đối với các biến số vĩ mô của nền kinh tế như: Tăng trưởng sản lượng, lạm phát, lãi suất, biến động trong tỷ giá hối đoái và điều kiện thương mại. Kết quả nghiên cứu cho thấy sự phù hợp về phương pháp tiếp cận trong việc cân bằng và tương thích giữa lý thuyết và dữ liệu thực tế nhằm xây dựng một mô hình dự báo có ý nghĩa cho Việt Nam.

Những nghiên cứu đã nêu đều mang những ý nghĩa nhất định trong công tác dự báo kinh tế ở Việt Nam. Tuy nhiên, các phương pháp được sử dụng đều dựa trên tập dữ liệu có cùng tần suất (cùng tháng, cùng quý,...), trong khi đặc trưng của dữ liệu kinh tế, các chỉ số thường được công bố với các tần suất khác nhau, có thể theo ngày, tháng, quý hoặc năm và độ trễ khi công bố cũng không giống nhau. Vì thế, để có thể áp dụng được các mô hình truyền thống trong công tác dự báo kinh tế, các dữ liệu cần đưa về cùng một tần suất, với cách xử lý như vậy có thể làm giảm độ chính xác của kết quả dự báo. Nhằm khắc phục vấn đề trên, mô hình dữ liệu tần suất hỗn hợp được cho là phù hợp.

2.2. Các nghiên cứu dự báo các chỉ tiêu kinh tế vĩ mô sử dụng dữ liệu với tần suất hỗn hợp

Xu hướng tăng trưởng GDP được cho là bị ảnh hưởng bởi nhiều các biến số kinh tế, chẳng hạn như các biến số kinh tế vĩ mô (CPI, thất nghiệp, cán cân thương mại, chỉ số sản xuất công nghiệp...) và các biến số tài chính (cung tiền, dự trữ ngoại hối, chỉ số chứng khoán, tỷ giá hối đoái...). Việc xác định các yếu tố đầu vào để dự báo tăng trưởng GDP từ một lượng lớn các biến số kinh tế là không dễ dàng và đơn giản. Stock & Watson (1989, 2002) chỉ ra rằng các biến động kinh tế vĩ mô được chuyển giao và khuếch tán thông qua một loạt các hoạt động kinh tế thay vì một biến số kinh tế duy nhất, quá trình này diễn ra rất phức tạp. Do đó, họ cho rằng dự báo kinh tế vĩ mô nói chung và dự báo tăng trưởng GDP nói riêng cần tận dụng triệt để các biến số kinh tế có chứa nhiều thông tin, bao gồm các biến số như sản lượng công nghiệp, thu nhập bình quân đầu người, khối lượng thương mại, việc làm, v.v. và các biến số hàng đầu như lãi suất, tỷ giá hối đoái, lợi tức cổ phiếu, lợi tức trái phiếu dài hạn, v.v.

Mặc dù một số lượng nhỏ các yếu tố có thể được trích xuất từ một số lượng lớn các biến số kinh tế làm yếu tố dự báo, dự báo tăng trưởng GDP vẫn cần giải quyết vấn đề về tần suất dữ liệu khác nhau. Nhìn chung, dữ liệu GDP được công bố hàng quý và hàng năm; trong khi đó dữ liệu về chỉ số giá, tiêu dùng, đầu tư... được công bố hàng tháng; còn dữ liệu của thị trường chứng khoán, thị trường trái phiếu v.v. có thể được thu thập hàng ngày. Các mô hình dự báo truyền thống yêu cầu cùng tần suất dữ liệu và các phương pháp luận nhất định được áp dụng để chuyển dữ liệu hàng ngày và hàng tháng có tần suất cao hơn thành dữ liệu hàng quý có tần suất thấp hơn, chẳng hạn như phương pháp trung bình, phương pháp bắc cầu (Diron, 2008) và tổng hợp theo thời gian (Silvestrini & Veredas, 2008). Tuy nhiên, các phương pháp chuyển đổi tần suất này có thể làm mất một phần đáng kể thông tin có trong dữ liệu tần suất cao và do đó làm giảm hiệu quả sử dụng thông tin mẫu ở một mức độ nhất định.

Để tránh mất thông tin trong quá trình chuyển đổi tần suất, Ghysel & cộng sự (2002) và nhóm nghiên cứu của ông đã đề xuất mô hình phân tích trực tiếp trên bộ dữ liệu mẫu có tần suất hỗn hợp (MIDAS), mà không cần phải chuyển đổi chúng về cùng một tần suất. Sau đó nhóm nghiên cứu tiếp tục phát triển và xây dựng các mô hình MIDAS mở rộng. Hay như nghiên cứu của Rufino (2019) đã chứng minh tính khả thi của mô hình hồi quy MIDAS để giải quyết vấn đề tần suất hỗn hợp trong việc thực hiện “dự báo” tăng trưởng kinh tế của Philippin. Các biến thể khác nhau của mô hình MIDAS được ước tính bằng cách sử dụng dữ liệu GDP thực hàng quý và dữ liệu hàng tháng của lạm phát, chỉ số sản xuất công nghiệp và chỉ số giao dịch chứng khoán Philippines. Có thể thấy, phương pháp MIDAS rất phù hợp để dự báo tăng trưởng GDP dựa trên dữ liệu tài chính hàng ngày hoặc kinh tế vĩ mô hàng tháng (Ghysel & Wright, 2009; Andreou & cộng sự, 2011; Frale & Monteforte, 2011; Monteforte & Moretti, 2012).

Gần đây, có khá nhiều các nghiên cứu được thực hiện để dự báo GDP bằng phương pháp MIDAS như các nghiên cứu Albert K. Tsui & cộng sự (2018), Habimana & cộng sự (2020), Gunay M. (2020), Mikosch, H., & Solanko, L. (2019), Kuck, K., & Schweikert, K. (2021), đã thực hiện dự báo tốc độ tăng trưởng tổng sản phẩm quốc nội (GDP) của Singapore, Thổ Nhĩ Kỳ và Rwanda. Theo Albert K. Tsui & cộng sự (2018), kết quả cho thấy hồi quy MIDAS sử dụng dữ liệu tài chính tần suất cao cho kết quả dự báo tốc độ tăng trưởng GDP tốt hơn so với các mô hình khác và hiệu suất dự báo tốt nhất đạt được bằng cách sử dụng lợi nhuận cổ phiếu hàng tuần. Habimana & cộng sự (2020) đã áp dụng các kỹ thuật dự báo, cụ thể là các phương trình cầu và một loạt các mô hình hồi quy lấy mẫu dữ liệu tần suất hỗn hợp (MIDAS) và mô hình vector tự hồi quy tần suất hỗn hợp (MF-VAR). Sau đó, so sánh khả năng dự báo của các phương pháp này từ mẫu. Chìa khóa cho những phát hiện của họ là sự kết hợp của phương trình cầu và mô hình U-MIDAS mang lại độ chính xác tốt nhất cho dự báo GDP quý hiện tại. Trong nghiên cứu của Kuck, K., & Schweikert, K. (2021), các tác giả đã thực hiện dự báo GDP với dữ liệu tần suất hỗn hợp của Đức bằng mô hình nhân tố động và mô hình MIDAS, và kết quả thực nghiệm cũng cho thấy mô hình MIDAS đơn giản nhưng cho kết quả dự báo tốt hơn mô hình nhân tố động. Hiệu quả của mô hình MIDAS cũng được chứng minh trong nghiên cứu của Mikosch, H., & Solanko, L. (2019) khi các tác giả đã so sánh hiệu quả của các mô hình MIDAS và U-MIDAS với mô hình cầu trong thực hiện dự báo tốc độ tăng trưởng GDP của Nga.

Những nghiên cứu trên cho thấy, dự báo chỉ số kinh tế vĩ mô nói chung và tăng trưởng GDP nói riêng đã thu hút được sự quan tâm lớn của nhiều nhà nghiên cứu trong và ngoài nước. Tuy nhiên, tính đến nay các nghiên cứu về tăng trưởng GDP của Việt Nam còn rất ít, đặc biệt là nghiên cứu liên quan đến dự báo tăng trưởng GDP sử dụng dữ liệu tần suất hỗn hợp. Trong khi đó, các công trình nghiên cứu ngoài nước đã sử dụng các phương pháp dự báo với dữ liệu có tần suất hỗn hợp để dự báo tăng trưởng GDP. Các phương pháp

này đều có những ưu điểm chung: (i) có thể sử dụng đầy đủ dữ liệu ở tần suất cao, tránh mất mát đáng kể thông tin mẫu, do đó nâng cao độ chính xác của dự báo; (ii) có thể thực hiện dự báo thời gian tức thời cho tăng trưởng GDP, sử dụng dữ liệu kinh tế và tài chính tần suất cao được công bố mới nhất. Do đó, nghiên cứu này sẽ kế thừa những lý thuyết liên quan và phương pháp hồi quy dữ liệu có tần suất hỗn hợp để xây dựng mô hình dự báo tăng trưởng GDP của Việt Nam.

3. Mô hình và Dữ liệu nghiên cứu

3.1. Mô hình nghiên cứu

Các mô hình dự báo trước đây chỉ áp dụng với các tập dữ liệu chuỗi thời gian có cùng tần suất. Để khắc phục nhược điểm của các mô hình phân tích có cùng tần suất và khai thác tốt nhất những dữ liệu công bố có được cho công tác dự báo, các mô hình xử lý và phân tích trên tập dữ liệu có tần suất khác nhau đã được đề xuất và ứng dụng. Trong đó, một số mô hình phổ biến như: mô hình VAR tần suất hỗn hợp MF-VAR (mixed frequency VAR), các mô hình bắc cầu (bridge equations), và các mô hình lấy mẫu tần suất hỗn hợp MIDAS (mixed data sampling). Trong nghiên cứu này, mô hình MIDAS và mô hình MF-VAR được ứng dụng để phân tích và dự báo chỉ số GDP của Việt Nam.

3.1.1. Mô hình MIDAS

Mô hình MIDAS là mô hình phân tích dữ liệu với tần suất hỗn hợp, được đề xuất bởi nhóm tác giả Eric Ghysels, Arthur Sinko & Rossen Valkanov năm 2002. Về cơ bản các mô hình MIDAS là các hồi quy dạng rút gọn được tham số hóa, liên quan đến các quá trình lấy mẫu ở các tần suất khác nhau. Trong đó, các biến giải thích có tần suất khác nhau và bằng hoặc cao hơn tần suất của biến phụ thuộc, và đối với các biến giải thích có tần suất cao hơn, các đa thức phân phối trễ được sử dụng để ngăn chặn sự gia tăng về số lượng tham số cũng như các vấn đề liên quan đến lựa chọn thứ tự trễ.

Mô hình dự báo MIDAS cơ bản cho biến phụ thuộc Tăng trưởng GDP tần suất quý tại h_q quý tiếp theo (h_q -step-ahead) với $h_q = h_m/m$ (h_m tính theo tháng) được xác định như sau:

$$y_{t_q+h_q} = y_{t_m+h_m} = \beta_0 + \beta_1 b(L_m; \theta) x_{t_m+w}^{(m)} + \varepsilon_{t_m+h_m}$$

Trong đó:

y_{t_q} là biến phụ thuộc Tăng trưởng GDP tần suất quý còn $x_{t_m}^{(m)}$ là biến giải thích có tần suất cao (tháng, tuần, ngày), $t_q = 1, 2, 3, \dots, T_q^y; T_q^y$ là quý cuối cùng mà dữ liệu GDP có sẵn; m là chỉ số xác định mức độ cao hơn về tần suất của biến độc lập so với biến phụ thuộc. Ví dụ nếu y có tần suất quý và x có tần suất tháng thì $m = 3$, còn nếu y có tần suất quý còn x có tần suất tuần thì $m = 12$.

- Tăng trưởng GDP biểu diễn theo tháng là y_{t_m} , $y_{t_m} = y_{t_q}$ với chỉ số t_m của y là $t_m = 3t_q$; $t_m = 3, 6, 9, \dots, T_m^y$ với $T_m^y = 3T_q^y$.

■ $b(L_m; \theta) = \sum_{k=0}^K c(k; \theta) L_m^k$ là đa thức trễ với L_m là toán tử trễ được xác định bởi:

$$L_m^k x_{t_m}^{(m)} = x_{t_m-k}^{(m)}$$

- Chỉ số t_m cho x_{t_m} là $t_m = 1, 2, 3, \dots, T_m^x$; là thời điểm cuối mà biến x có dữ liệu, nó lớn hơn T_m^y ; $w = T_m^x - T_m^y$;

- $c(k; \theta)$ là các tham số của các hệ số độ trễ của mô hình cần được ước lượng.

- β_0, β_1 là các hệ số của mô hình hồi quy.

Một trong các vấn đề chính của phương pháp MIDAS là tìm tham số phù hợp cho các hệ số trễ $c(k; \theta)$. Vì có tần suất cao hơn, việc mô hình hóa đầy đủ thường yêu cầu nhiều độ trễ trong phương trình hồi quy, điều này có thể dẫn đến tình trạng mô hình quá nhiều tham số khó để ước lượng. Một số lược đồ trọng số phổ biến để khắc phục hạn chế này là “Trễ Almon mũ” tương ứng với hàm trễ Almon. Cụ thể lược đồ Almon được biểu diễn như sau:

$$c(k; \theta) = \frac{\exp(\theta_1 k + \dots + \theta_Q k^Q)}{\sum_{k=1}^K \exp(\theta_1 k + \dots + \theta_Q k^Q)}$$

Với Q là số lượng tham số của θ , hay $\theta = (\theta_1, \theta_2, \dots, \theta_Q)$ là các tham số cần được ước lượng. Hàm này khá linh hoạt và có thể có nhiều hình dạng khác nhau chỉ với vài tham số. Chúng có thể là mô hình tăng dần, giảm dần hoặc lồi lõm. Ghysel & cộng sự (2007) đã sử dụng dạng hàm này với hai tham số, cho phép tính linh hoạt cao và xác định có bao nhiêu độ trễ được đưa vào hồi quy. Vì lược đồ trễ Almon được sử dụng phổ biến nhất và có tính linh hoạt cao nên trong nghiên cứu này nhóm tác giả sử dụng lược đồ trễ Almon để xác định các tham số phù hợp cho các hệ số trễ của mô hình.

3.1.2. Mô hình MF-Var

Mô hình MF-VAR là mô hình VAR phân tích dữ liệu chuỗi thời gian với tần suất hỗn hợp. Mô hình này được xây dựng dựa trên cách tiếp cận không gian trạng thái và được Mariani và Murasawa đề xuất năm 2010. Mô hình MF-VAR thường có nhiều tham số và các tham số này được ước lượng bằng một trong hai phương pháp sau. Phương pháp thứ nhất tiếp cận theo cổ điển là ước lượng bằng cách xây dựng hàm hợp lý cực đại. Phương pháp thứ hai là sử dụng suy diễn Bayes. Hai nghiên cứu tiêu biểu cho hai phương pháp này là: Mariano & Murasawa (2010) với cách tiếp cận cổ điển; Schorfheide & Song (2011) tiếp cận theo suy diễn Bayes.

Theo Mariano & Murasawa (2010), khi xem xét mô hình VAR trong không gian trạng thái, phân rã biến tần suất quý quan sát tại các thời điểm $t_m = 3, 6, 9, \dots, T^m$ thành biến với tần suất tháng không quan sát được được mô tả bởi phương trình tổng hợp sau:

$$y_{t_m} = \frac{1}{3}y_{t_m}^* + \frac{2}{3}y_{t_m-1}^* + y_{t_m-2}^* + \frac{2}{3}y_{t_m-3}^* + \frac{1}{3}y_{t_m-4}^*$$

Phương trình tổng hợp này xuất phát từ giả định rằng chuỗi y_{t_m} là trung bình của chuỗi và các biến trễ của nó. Với cách tiếp cận cổ điển thì quá trình VAR(p) được biểu diễn bởi phương trình sau:

$$\phi(L_m) \begin{pmatrix} y_{t_m}^* & \mu_y^* \\ x_{t_m} & \mu_x \end{pmatrix} = u_{t_m}$$

trong đó:

x_{t_m} và $y_{t_m}^*$ tương ứng là các biến tần suất cao.

u_{t_m} là sai số ngẫu nhiên có kỳ vọng bằng 0 và phương sai không đổi.

Các tham số trong mô hình được diễn giải và ước lượng dựa trên mô hình không gian trạng thái với phương pháp ước lượng hợp lý cực đại. Theo nghiên cứu của Vladimir Kuzin & cộng sự (2011) thì MF-VAR hiệu quả khi đưa ra các dự báo dài hạn hơn là dự báo tức thời.

3.2. Dữ liệu nghiên cứu

Nghiên cứu được thực hiện trên bộ dữ liệu gồm các chỉ số kinh tế vĩ mô, được thu thập với các tần suất khác nhau (theo quý, tháng, tuần) từ trang web của Tổng cục thống kê, IMF, ADB, Bloomberg, cơ sở dữ liệu CEIC... trong giai đoạn 2006 đến 2020. Lý do để nhóm tác giả lựa chọn giai đoạn 2006 – 2020 vì tính từ năm 2006 trở đi, các chỉ tiêu kinh tế vĩ mô của Việt Nam được thu thập và thống kê một cách đầy đủ nhất.

Trong nghiên cứu này, nhóm tác giả dự báo tốc độ tăng trưởng kinh tế của Việt Nam hàng quý bằng mô hình MIDAS cơ bản với toán tử trễ Almon với 3 mô hình trong đó các biến độc lập được lựa chọn theo tần suất khác nhau. Các phân tích được thực hiện trên phần mềm Eviews 10. Bộ dữ liệu được tách thành hai giai đoạn:

- Dữ liệu giai đoạn 2006 đến 2018 dùng để ước lượng các tham số trong các mô hình hồi quy.
- Dữ liệu giai đoạn 2019 đến 2020 được dùng để so sánh và đưa ra các dự báo.

Mục đích chính của việc chia tách dữ liệu thành hai giai đoạn là để so sánh và đối chiếu kết quả dự báo từ mô hình MIDAS (dựa trên kết quả hồi quy bộ dữ liệu giai đoạn 2006 - 2018 để dự báo cho giai đoạn 2019 - 2020; sau đó so sánh kết quả dự báo với kết quả thực tế để đánh giá độ chính xác của mức độ dự báo cũng như tính hiệu quả của mô hình).

Các biến sử dụng trong mô hình

Để dự báo tốc độ tăng trưởng kinh tế theo quý, nhóm nghiên cứu dựa trên bộ số liệu gồm 22 biến ứng

với 19 chỉ số kinh tế trong đó: 5 biến tần suất quý, 14 biến tần suất tháng và 3 biến tần suất tuần được mô tả trong Bảng 1.

Bảng 1: Các biến số kinh tế

STT	Chỉ số kinh tế	Đơn vị	Ký hiệu	Tần suất	Nguồn dữ liệu
1	Tăng trưởng GDP	(%)	TTGDP	Quý	Tổng cục Thống kê
2	Tăng trưởng xuất khẩu	(%)	TTXK_Q	Quý	Tổng cục Thống kê
3	Cán cân thanh toán	(%)	CCTT	Quý	IMF
4	Đầu tư trực tiếp ròng	(%)	DTTR	Quý	IMF
5	Tài khoản vốn	(%)	TKV	Quý	IMF
6	Tăng trưởng doanh số bán lẻ	(%)	TTDSBL	Tháng	Tổng cục Thống kê
7	Tốc độ tăng trưởng SX công nghiệp	(%)	TTSXCN	Tháng	Tổng cục Thống kê
8	Tăng trưởng cung tiền	(%)	TTCT_M	Tháng	Cơ sở dữ liệu CEIC
9	Tỷ lệ dự trữ	(%)	TLDT_M	Tháng	Cơ sở dữ liệu CEIC
10	Chỉ số giá cổ phiếu tổng hợp	Index	CSGCP	Tháng	Bloomberg LP
11	Lãi suất cơ bản	(%)	LS	Tháng	ADB
12	Chỉ số tỷ giá hối đoái	Index	CSTGHD	Tháng	ADB
13	Tổng dự trữ quốc tế (Không tính vàng)	(%)	TDTQT	Tháng	IMF
14	Tăng trưởng nhập khẩu hàng hóa	(%)	TTNK	Tháng	Cơ sở dữ liệu CEIC
15	Cán cân thương mại	(%)	CCTM	Tháng	Tổng cục Thống kê
16	Tỷ lệ lạm phát	(%)	LP	Tháng	Cơ sở dữ liệu CEIC
17	Tỷ giá USD/VNĐ (Trung bình 1 tháng)	USD/VNĐ	USD_M	Tháng	vn.investing.com
18	Hợp đồng tương lai của Vàng	USD/OUNCE	V_M	Tháng	vn.investing.com
19	Hợp đồng tương lai của Dầu thô	USD/thùng	DT_M	Tháng	vn.investing.com
20	Tỷ giá USD/VNĐ (Trung bình 1 tuần)	USD/VNĐ	USD_W	Tuần	vn.investing.com
21	Hợp đồng tương lai của Vàng	USD/OUNCE	V_W	Tuần	vn.investing.com
22	Hợp đồng tương lai của Dầu thô	USD/thùng	DT_W	Tuần	vn.investing.com

Nguồn: Kết quả phân tích số liệu của nhóm tác giả.

4. Kết quả dự báo tăng trưởng kinh tế Việt Nam

4.1. Tình hình tăng trưởng kinh tế theo quý của Việt Nam trong giai đoạn 2006 – 2020

Sự phát triển của Việt Nam trong 15 năm qua là rất đáng ghi nhận thể hiện ở tốc độ tăng trưởng kinh tế rất ấn tượng. Nếu xét riêng trong giai đoạn này thì tăng trưởng GDP bình quân theo quý của Việt Nam là 6,4% (Bảng 2), thuộc nhóm các nước tăng trưởng cao trong khu vực và thế giới.

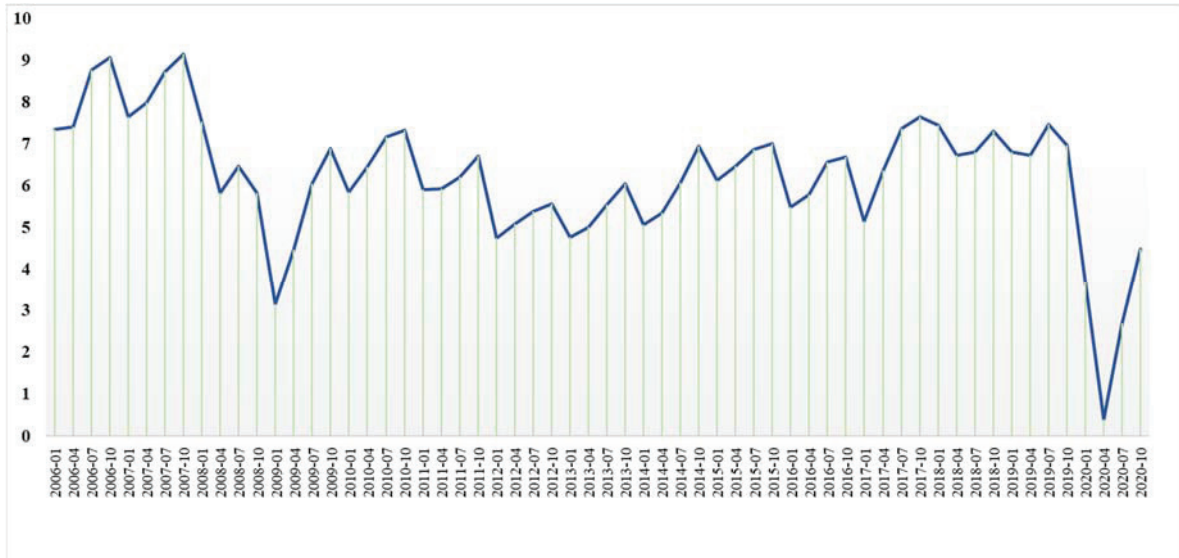
Biểu đồ Hình 1 cho thấy, mức tăng trưởng kinh tế hàng quý của nước ta liên tục thay đổi theo thời gian. Trong những năm đầu của giai đoạn nghiên cứu, mức tăng trưởng kinh tế của Việt Nam khá cao, đạt đỉnh

Bảng 2: Thống kê mô tả tăng trưởng kinh tế theo quý của Việt Nam giai đoạn 2006 – 2020

Mean	6,2408
Median	6,4550
Maximum	9,1600
Minimum	0,3900
Std. Dev.	1,5188

Nguồn: Kết quả phân tích số liệu của nhóm tác giả.

Hình 1: Tăng trưởng kinh tế theo quý của Việt Nam giai đoạn 2006 – 2020



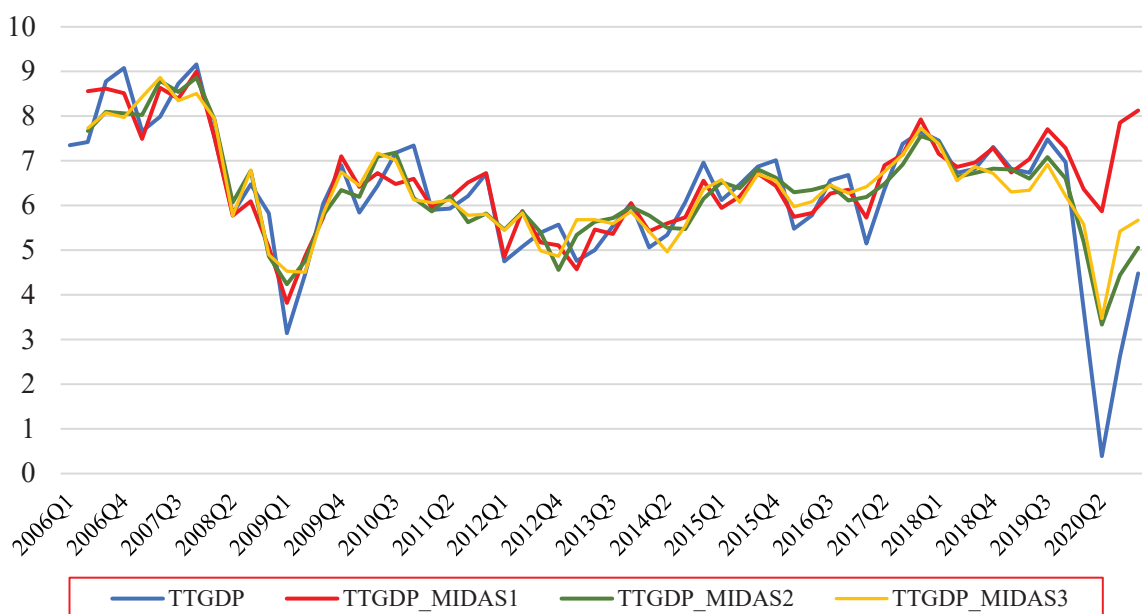
Nguồn: Kết quả phân tích số liệu của nhóm tác giả.

tại quý 4 năm 2007 với mức tăng trưởng ấn tượng khoảng 9,2%. Sang giai đoạn 2008 đến 2010, kinh tế của Việt Nam có nhiều thay đổi, tăng trưởng kinh tế lao dốc mạnh từ quý 1 năm 2008, chạm đáy quý 1 năm 2009 với mức tăng trưởng khoảng 3,1%, và bắt đầu hồi phục ngay sau đó, đây cũng là giai đoạn kinh tế Việt Nam chịu sự ảnh hưởng của cuộc khủng hoảng tài chính toàn cầu. Trong giai đoạn từ 2010 đến 2019, đây là giai đoạn đánh dấu sự tăng trưởng với tốc độ có phần ổn định hơn so với các giai đoạn còn lại, mức tăng trưởng thấp nhất là 4,75% (quý 1/2012). Do ảnh hưởng của đại dịch Covid-19, từ quý 4/2019 tăng trưởng kinh tế của Việt Nam bắt đầu có xu hướng giảm nhanh chóng, đặc biệt là quý 2 năm 2020 giảm sút nghiêm trọng với tốc độ tăng trưởng chỉ đạt mức 0,39%. Tuy nhiên, trung bình của toàn giai đoạn 2006 đến 2020 tăng trưởng kinh tế bình quân mỗi quý của nước ta tăng khoảng 6,2%.

4.2. Dự báo Tăng trưởng kinh tế Việt Nam bằng mô hình MIDAS

Mô hình MIDAS có rất nhiều biến thể khác nhau, trong nghiên cứu này trước hết chúng tôi ứng dụng mô hình MIDAS cơ bản với các tham số trong mô hình được chọn. Nhóm nghiên cứu cũng xem xét ba mô hình

Hình 2: Kết quả dự báo tăng trưởng GDP theo Quý của 3 mô hình MIDAS



Nguồn: Kết quả phân tích số liệu của nhóm tác giả.

sao cho mỗi mô hình gồm 16 biến trong đó có 15 biến độc lập.

- Mô hình 1: 15 biến độc lập gồm 4 biến tần suất quý (TTXK_Q, TKV, CCTT, DTTR) và 11 biến tần suất tháng.

- Mô hình 2: 15 biến độc lập gồm 1 biến tần suất quý (TTXK_Q) và 14 biến tần suất tháng (11 biến cũ + 3 biến mới là: V_M, USD_M, DT_M).

- Mô hình 3: 15 biến độc lập gồm 1 biến tần suất quý (TTXK_Q), 11 biến tần suất tháng và 3 biến tần suất tuần (V_W, USD_W, DT_W).

Do đặc thù của mô hình MIDAS là dự báo nhanh tăng trưởng kinh tế trong kỳ khi dữ liệu chưa được công bố dựa trên những thông tin có được đến thời điểm gần nhất. Do đó, nhóm nghiên cứu áp dụng cho dự báo GDP của Việt Nam trong các quý từ quý 1/2019 đến quý 4/2020 với dữ liệu tần suất cao được thu thập đến tháng 12/2020. Cách thức dự báo được tiến hành như sau: Nhóm tác giả sẽ tiến hành dự báo cho quý tiếp theo dựa trên số liệu của quý trước đó và so sánh với giá trị số liệu thực tế của quý đó khi công bố (Ví dụ: sử dụng kết quả của quý 4/2018 để dự báo cho kết quả của quý 1/2019 và so sánh kết quả với số liệu thực tế công bố của quý 1/2019; tiếp tục sử dụng kết quả công bố của quý 1/2019 để dự báo cho quý 2/2019 và so sánh kết quả với số liệu công bố thực tế của quý 2/2019, quy trình dự báo tiếp tục diễn ra tương tự cho các quý còn lại). Kết quả dự báo tăng trưởng GDP của Việt Nam theo quý được trình bày ở Hình 2 và Bảng 3.

Đánh giá sai số dự báo

Bảng 3: Đánh giá sai số dự báo của 3 mô hình MIDAS

Thời gian	TTGDP_Q (Thực tế)	MIDAS1 (Dự báo)	e	MIDAS2 (Dự báo)	e	MIDAS3 (Dự báo)	e
2019Q1	6,82000	6,73753	0,08247	6,80520	0,01480	6,30211	0,51789
2019Q2	6,73000	7,03864	0,30864	6,59847	0,13153	6,33969	0,39031
2019Q3	7,48000	7,70656	0,22656	7,08092	0,39908	6,91769	0,56231
2019Q4	6,97000	7,28560	0,31560	6,60436	0,36564	6,21178	0,75822
2020Q1	3,68000	6,36761	2,68761	5,18949	1,50949	5,58077	1,90077
2020Q2	0,39000	5,86972	5,47972	3,33584	2,94584	3,46911	3,07911
2020Q3	2,62000	7,85323	5,23323	4,43976	1,81976	5,42282	2,80282
2020Q4	4,48000	8,12709	3,64709	5,05209	0,57209	5,67228	1,19228
R2		0,88291		0,76375		0,75937	
RMSE		0,41053		0,58313		0,58851	
MAE		0,59289		0,52779		0,60778	

Nguồn: Kết quả phân tích số liệu của nhóm tác giả

Đồ thị của các đường dự báo thu được từ 3 mô hình MIDAS trong Hình 2 và các giá trị sai số dự báo trung bình và dự báo điểm ở Bảng 3 cho thấy, nhìn chung các mô hình MIDAS cơ bản với các trọng số Almon cho kết quả dự báo tăng trưởng GDP hàng quý của Việt Nam khá tốt với sai số nhỏ, cả 3 mô hình đều có sai số tuyệt đối trung bình dưới 1%. Trong đó mô hình 1 và 3 có tới 4 quý, mô hình 2 có tới 5 quý có sai số dự báo dưới 1%, tuy nhiên mô hình 1 có 2 dự báo quý 3, quý 4 năm 2020 lên tới hơn 5% (nguyên nhân là do trong giai đoạn này dịch Covid-19 diễn biến phức tạp, khó lường dẫn đến sai số lớn hơn trong các mô hình dự báo trên). Nếu xét theo tiêu chuẩn sai số MAE thì mô hình 2 có kết quả dự báo tốt hơn hai mô hình còn lại (với sai số dưới 3%). Kết quả này chứng tỏ việc phát triển các mô hình này là đúng hướng, các thông tin đầu vào được lựa chọn phù hợp.

Hình 2 cho thấy các mô hình hầu như có thể dự đoán được các giai đoạn thay đổi lớn của tốc độ tăng trưởng GDP, cụ thể là giai đoạn 2008 – 2009 khi kinh tế bị suy thoái do ảnh hưởng của khủng hoảng tài chính toàn cầu, giai đoạn 2016 – 2017 khi kinh tế Việt Nam chuyển sang tăng trưởng nhanh hơn, hay như giai đoạn ba quý đầu năm 2020, khi đại dịch Covid-19 khiến GDP của Việt Nam giảm tốc đột ngột. Phương pháp này chứng tỏ sự ưu việt hơn so với các mô hình truyền thống (VAR, SVAR, BVAR, VECM) trong việc

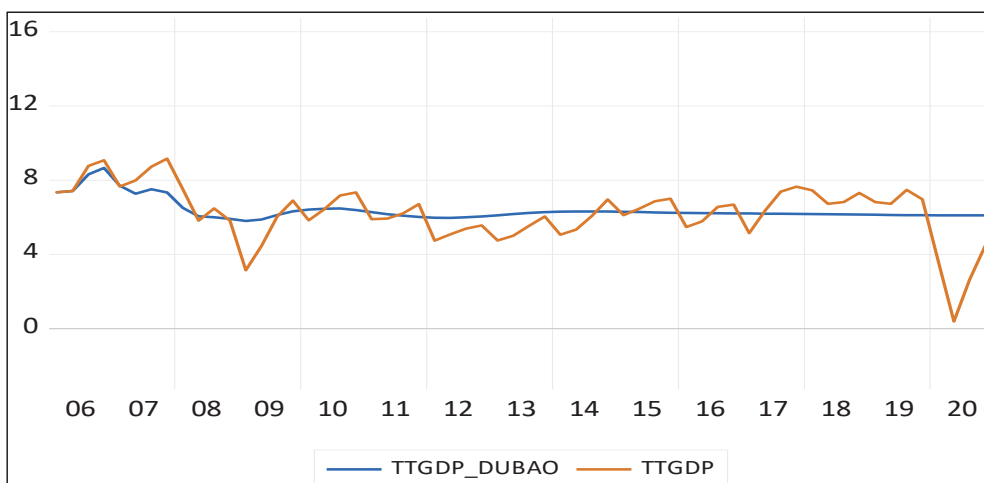
nắm bắt những thay đổi nhanh chóng và đột biến của nền kinh tế do lợi thế sử dụng dữ liệu tần suất cao trong mô hình. Các mô hình truyền thống thường bị giới hạn do sử dụng dữ liệu dạng cân bằng (ví dụ như mô hình tháng chỉ dùng dữ liệu tháng, mô hình quý chỉ dùng dữ liệu quý) nên tính linh hoạt sẽ không cao như mô hình MIDAS. Mặt khác, mô hình truyền thống bị giới hạn bởi một số trọng số (tham số) nhất định trong khi mô hình MIDAS có thể sử dụng nhiều phương pháp khác nhau để xác định các trọng số phù hợp với điều kiện dữ liệu sẵn có.

Trong số các mô hình MIDAS được xây dựng (kết quả Bảng 3), mô hình MIDAS2 có khả năng dự tính tốc độ tăng trưởng GDP trong kỳ của Việt Nam tốt nhất (với sai số dự báo thấp nhất). Phương pháp này xác định trọng số của MIDAS dựa trên trẻ phân phối đa thức Almon cũng đã được chứng minh là phát huy hiệu quả mạnh khi dự báo GDP tại các nghiên cứu quốc tế. Xếp sau MIDAS2 là mô hình MIDAS3, mô hình MIDAS1 dự báo tăng trưởng GDP với sai số lớn nhất trong số các mô hình MIDAS. Việc bổ sung các biến có tần suất cao hơn (tháng và tuần) trong hai mô hình MIDAS2 và MIDAS3 đã cho kết quả dự báo tốt hơn với sai số thấp hơn chứng tỏ các biến lựa chọn đưa vào mô hình có vai trò quan trọng quyết định đối với kết quả và độ chính xác của dự báo. Sử dụng các biến có tần suất càng cao (quý, tháng, tuần thậm chí biến ngày) thì kết quả dự báo của biến phụ thuộc (có tần suất thấp) càng tốt và chính xác.

4.3. Dự báo tăng trưởng kinh tế Việt Nam bằng mô hình MF-VAR

Để dự báo Tăng trưởng GDP bằng mô hình MF-VAR, chúng tôi chạy mô hình MF-VAR gồm 16 biến, với biến phụ thuộc là TTGDP theo quý và 15 biến độc lập gồm: 4 biến tần suất quý (TTXK_Q, TKV, CCTT,

Hình 3: Kết quả dự báo tăng trưởng GDP theo Quý của mô hình MF-VAR



Nguồn: Kết quả phân tích số liệu của nhóm tác giả.

DTTR) và 11 biến tần suất tháng. Kết quả thu được thể hiện trong Hình 3 và Bảng 4 sau đây:

Hình 3 thể hiện mô hình MF-VAR cho kết quả dự báo khá sát so với giá trị TTGDP thực tế, phù hợp đưa ra dự báo cho giá trị trung bình của TTGDP trong giai đoạn nghiên cứu. Tuy nhiên đường cong dự báo khá ổn định, ít biến thiên nên khó nắm bắt được xu thế của của TTGDP.

Đánh giá sai số dự báo

Bảng 4 cho thấy mô hình MF-VAR có kết quả dự báo khá tốt trong ngắn hạn, trong 4 quý đầu của năm 2019 các sai số dự báo đều rất nhỏ, với mức sai số trung bình dao động quanh 0,8%; trong khi 4 quý của năm 2020 có sự chênh lệch sai số dự báo khá lớn. Kết quả này cũng cho thấy mô hình MF-VAR có hiệu quả dự báo trong ngắn hạn giống như mô hình MIDAS, nhưng kết quả dự báo của mô hình MIDAS phản ánh xu thế và các giá trị biến thiên tốt hơn và bám sát kết quả thực tế hơn so với mô hình MF-VAR. Nhóm nghiên cứu cũng đã chạy thử mô hình MF-VAR 2, MF-VAR 3, (gồm các biến như mô hình MIDAS 2, MIDAS 3), tuy nhiên kết quả thu được từ các mô hình này không có ý nghĩa thống kê.

Bảng 4: Đánh giá sai số dự báo của mô hình MF-VAR

Năm	TTGDP (Thực tế)	TTGDP (Du báo)	Sai số (%)
2019-01	6.82	6.14	0.68
2019-04	6.73	6.13	0.6
2019-07	7.48	6.12	1.36
2019-10	6.97	6.12	0.85
2020-01	3.68	6.11	2.43
2020-04	0.39	6.11	5.72
2020-07	2.62	6.11	3.49
2020-10	4.48	6.11	1.63
R ²	0.83		
RMSE	1.252		
MAE	0.87		

Nguồn: Kết quả phân tích số liệu của nhóm tác giả.

5. Kết luận

Nghiên cứu sử dụng mô hình phân tích dữ liệu với tần suất hỗn hợp để dự báo tăng trưởng GDP theo quý của Việt Nam, kết quả của dự báo cho thấy tính khả thi và độ tin cậy của mô hình. Các mô hình MIDAS cơ bản với các trọng số Almon cho kết quả dự báo tăng trưởng GDP hàng quý của Việt Nam khá chính xác với sai số tuyệt đối trung bình dưới 1% so với GDP công bố thực tế, đồng thời cho thấy các mô hình này hầu như có thể dự đoán được các giai đoạn thay đổi lớn của tốc độ tăng trưởng GDP. Trong các mô hình dự báo thì mô hình MIDAS2 và MIDAS3 cho ra kết quả dự báo chính xác nhất với sai số lớn nhất dưới 3%, điều này cho thấy sự quan trọng của các biến số dự báo đưa vào mô hình; các biến độc lập có tần suất càng cao (tuần, tháng, quý) thì khả năng dự báo và kết quả thu được càng tốt. Do vậy, mô hình có thể được dùng để dự báo tăng trưởng kinh tế của Việt Nam cho các kỳ tiếp theo, một khi dữ liệu được cập nhật. Trong khi đó kết quả phân tích cho thấy MF-VAR tuy có kết quả dự báo khá tốt nhưng về mức sai số dự báo đều lớn hơn các mô hình MIDAS đồng thời mức độ biến thiên ổn định không phản ánh tốt các xu thế tăng giảm so với giá trị thực tế. Nếu xét theo tiêu chuẩn sai số MAE thì mô hình MF-VAR có kết quả dự báo kém hơn tiêu chuẩn MAE sai số trung bình các mô hình MIDAS (với sai số 0,87% so với trung bình 0,6%).

Các mô hình phân tích dữ liệu tần suất hỗn hợp được quan tâm và ứng dụng nhiều trong lĩnh vực kinh tế và tài chính trong những năm gần đây, đặc biệt là trong công tác dự báo các chỉ số kinh tế vĩ mô của quốc gia, tuy nhiên ở Việt Nam hầu như chưa có nghiên cứu trong lĩnh vực này. Kết quả nghiên cứu mô hình MIDAS trên phần nào đã cung cấp thông tin thiết thực và là cơ sở tham khảo cho các nhà nghiên cứu, các nhà hoạch định chính sách, các đơn vị, vụ, viện nghiên cứu của Nhà nước liên quan đến công tác thống kê và dự báo, Ngân hàng nhà nước, Tổng cục thống kê... trong việc tìm ra một phương pháp thích hợp nhất để phục vụ công tác dự báo tăng trưởng GDP của Việt Nam. Ngoài ra, sử dụng mô hình phân tích dữ liệu với tần suất hỗn hợp không chỉ giúp dự báo GDP mà còn hữu ích trong việc ứng dụng vào dự báo các chỉ số kinh tế vĩ mô quan trọng khác của quốc gia như lạm phát, xuất khẩu...

Tài liệu tham khảo

- Albert K. Tsui; Cheng Yang Xu & Zhaoyong Zhang, (2018), 'Macroeconomic forecasting with mixed data sampling frequencies: Evidence from a small open economy', *Journal of Forecasting*, 37 (6), 666-675.
- Andreou, E., Ghysels, E., & Kourtellis, A. (2011), 'Forecasting with mixed-frequency data', In *The Oxford handbook of economic forecasting*.
- Diron M. (2008), 'Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data', *Journal of Forecasting*, 27 (5), 371-390.

-
- Frale, Cecilia & Monteforte, Libero (2011), 'FaMIDAS: A Mixed Frequency Factor Model with MIDAS Structure', Bank of Italy Temi di Discussione (Working Paper) No. 788.
- Ghysels, E., & Wright, J. H. (2009), 'Forecasting professional forecasters', *Journal of Business & Economic Statistics*, 27(4), 504-516.
- Ghysels, E., Santa-Clara, P., Valkanov, R. (2002), 'The MIDAS touch: mixed data sampling regression models', Working paper, UNC and UCLA.
- Ghysels, E., Sinko, A., & Valkanov, R. (2007), 'MIDAS regressions: Further results and new directions', *Econometric reviews*, 26 (1), 53-90.
- Gunay, M. (2020), *Nowcasting Turkish GDP with MIDAS: Role of Functional Form of the Lag Polynomial* (No. 2002).
- Habimana, Olivier and Tabaro, Didier and Kalisa, Thierry (2020), 'Nowcasting Rwanda's Quarterly GDP Using Mixed-Frequency Methods'. Available at SSRN: <https://ssrn.com/abstract=3653136> or <http://dx.doi.org/10.2139/ssrn.3653136>
- Kuck, K., & Schweikert, K. (2021), 'Forecasting Baden-Württemberg's GDP growth: MIDAS regressions versus dynamic mixed-frequency factor models', *Journal of Forecasting*, 40(5), 861-882.
- Kuzin, V., Marcellino, M., & Schumacher, C. (2011), 'MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area', *International Journal of Forecasting*, 27(2), 529-542.
- Libero Monteforte & Gianluca Moretti (2012), 'Real-Time Forecasts of Inflation: The Role of Financial Variables', *Journal of Forecasting*, 32 (1), 51-61.
- Mariano, R. S., & Murasawa, Y. (2010), 'A coincident index, common factors, and monthly real GDP', *Oxford Bulletin of economics and statistics*, 72(1), 27-46.
- Mikosch, H., & Solanko, L. (2019), 'Forecasting quarterly Russian GDP growth with mixed-frequency data', *Russian Journal of Money and Finance*, 78(1), 19-35.
- Nguyễn Đức Trung & Nguyễn Hoàng Chung (2017), 'Mô hình dự báo cho nền kinh tế nhỏ và mở của Việt Nam: Phương pháp tiếp cận BVAR-DSGE', *Tạp chí Phát triển Kinh tế*, 28 (10), 05-38.
- Nguyễn Kim Anh (2015), 'Dự báo tăng trưởng và lạm phát của Việt Nam năm 2015', *Tạp chí Kinh tế và Phát triển*, 213(3), 34-41.
- Nguyễn Thị Thu Hằng & Vũ Phạm Hải Đăng (2012), 'Dự báo tăng trưởng kinh tế và lạm phát ở Việt Nam bằng mô hình vector tự hồi quy Bayesian', working paper, trường đại học Kinh tế và Kinh doanh.
- Rufino, C. (2019), 'Nowcasting Philippine Economic Growth Using MIDAS Regression', *DLSU Business & Economics Review*, 29 (1), 14-23.
- Schorfheide, F., & Song, D. (2011), *Real-time forecasting with a mixed-frequency VAR. Federal Reserve Bank of Minneapolis* (No. 701), Working Paper.
- Silvestrini, A. & Veredas, D. (2008), 'Temporal aggregation of univariate and multivariate time series models: A survey', No 685, Temi di discussione (Economic working papers), Bank of Italy, Economic Research and International Relations Area.
- Stock, J. H., & Watson, M. W. (1989), 'New indexes of coincident and leading economic indicators', *NBER macroeconomics annual*, 4, 351-394.
- Stock, James H. & Mark W. Watson (2002), 'Forecasting Using Principal Components from a Large Number of Predictors', *Journal of the American Statistical Association*, 97 (460), 1167.