

Ứng dụng các phương pháp phân tích chuỗi thời gian trong dự báo nhu cầu và lập kế hoạch sản xuất bê tông thương phẩm: Nghiên cứu điển hình tại Công ty CP Bê tông Vina 21

Nguyễn Anh Đức^{1*}, Lê Thanh Bình²

¹ Trường Đại học Xây dựng Hà Nội

² Công ty Cổ phần Bê tông Vina 21

TỪ KHOÁ

Bê tông thương phẩm
Dự báo chuỗi thời gian
Kế hoạch sản xuất
Doanh nghiệp vừa và nhỏ
Việt Nam

TÓM TẮT

Cùng với quá trình đô thị hóa và phát triển hạ tầng ngày càng mạnh mẽ tại Việt Nam, nhu cầu bê tông thương phẩm gia tăng đáng kể. Tuy nhiên, bài toán dự báo sản lượng bê tông vẫn là thách thức, đặc biệt đối với các doanh nghiệp vừa và nhỏ do tính biến động theo mùa vụ và thị trường. Nghiên cứu này so sánh bốn mô hình chuỗi thời gian—Seasonal Naive, SARIMA, Holt-Winters, và Prophet—để dự báo nhu cầu bê tông hàng ngày, áp dụng trên hai kịch bản huấn luyện: (1) sử dụng dữ liệu năm 2021–2023, trong đó năm 2021 có biến động do Covid-19 và (2) chỉ lấy năm 2022–2023. Kết quả cho thấy Prophet và Holt-Winters thường cải thiện đáng kể chỉ số RMSE khi khai thác toàn bộ dữ liệu, trong khi SARIMA thường không hội tụ hoặc dự báo kém khi thêm năm 2021. Ngược lại, nếu loại bỏ 2021, SARIMA có thể hoạt động nhưng sai số tuyệt đối vẫn cao hơn so với Prophet. Nhóm tác giả cũng thảo luận ứng dụng của các mô hình trong việc lập kế hoạch sản xuất, quản lý nguyên vật liệu, và đánh giá kinh tế tại trường hợp nghiên cứu điển hình. Qua phỏng vấn lãnh đạo và cán bộ kỹ thuật, nghiên cứu khẳng định tiềm năng áp dụng mô hình dự báo dữ liệu lớn và cho thấy tầm quan trọng của việc cân nhắc những năm bất thường trong bối cảnh ngành bê tông biến động mạnh. Kết quả nghiên cứu có thể được mở rộng và áp dụng cho các doanh nghiệp sản xuất bê tông thương phẩm hoặc sản xuất vật liệu xây dựng khác.

KEYWORDS

Ready-mixed concrete
Time series forecasting
Operation planning
SMEs
Vietnam

ABSTRACT

This paper presented a comparative study of four time series forecasting methods—Seasonal Naive, SARIMA, Holt-Winters, and Prophet—applied to real-world data from Vina 21, an SME producing ready-mixed concrete in Vietnam. By testing two different approaches to historical data (including or excluding the Covid-affected year 2021), the study provides insights into how atypical periods may influence model performance. Prophet and Holt-Winters show clear advantages when the full dataset (2021–2023) is used, displaying reduced RMSE and more robust handling of extended historical variability. In contrast, SARIMA performed poorly or did not converge under the heavier data load, likely due to the highly irregular demand patterns introduced by pandemic-related disruptions. However, SARIMA was more feasible when the 2021 data was omitted, albeit with a higher absolute error and only modestly improved MAPE relative to other approaches. From the perspective of Vina 21’s management team, adopting a forecast model that can handle significant swings in demand and maintain reasonable accuracy across diverse conditions is critical. Prophet emerged as a consistently strong candidate, especially for minimizing absolute forecast errors (RMSE), which directly supports production planning and raw material procurement decisions. Moreover, the findings indicate that integrating all available data generally benefits methods like Prophet and Holt-Winters, enabling them to capture broader patterns in seasonal and trend components. Future work should delve deeper into integrating external factors such as weather data, government construction policies, and project-specific timelines to further refine forecasting accuracy. Additionally, exploring ensemble approaches and advanced machine learning methods could yield further improvements. For SMEs like Vina 21, these results underscore the value of data-informed decision-making, validating that the choice of forecasting model should be aligned with the enterprise’s goals—whether it prioritizes minimizing absolute risk of under/overproduction or focuses on cost control and error ratios. The research results can be expanded to benefit the construction material manufacturer sector in Vietnam.

*Liên hệ tác giả: ducna@huce.edu.vn

Nhận ngày 06/01/2025, sửa xong ngày 15/02/2025, chấp nhận đăng ngày 17/02/2025

Link DOI: <https://doi.org/10.54772/jomc.01.2025.860>

1. Mở đầu

Nhu cầu bê tông thương phẩm tại Việt Nam đang có xu hướng gia tăng mạnh mẽ trong bối cảnh phát triển kinh tế và đô thị hóa nhanh chóng. Các công trình xây dựng quy mô lớn, dự án hạ tầng giao thông, khu đô thị hiện đại, cũng như hoạt động xây dựng dân dụng ngày càng trở nên phổ biến, tạo ra thị trường tiêu thụ bê tông thương phẩm rất sôi động [1]. Đặc biệt, các doanh nghiệp vừa và nhỏ (small and medium enterprises – SMEs) trong lĩnh vực sản xuất bê tông thương phẩm thường gặp thách thức cạnh tranh trực tiếp với các công ty lớn hơn về giá thành, chất lượng sản phẩm, cũng như mức độ chủ động trong chuỗi cung ứng nguyên vật liệu. Năng lực dự báo nhu cầu bê tông chính xác sẽ giúp các doanh nghiệp này tối ưu hóa kế hoạch sản xuất, quản trị nguồn lực, giảm thiểu lãng phí và nâng cao năng lực cạnh tranh [2].

Tuy vậy, hoạt động dự báo trong ngành xây dựng ở Việt Nam phức tạp do tính mùa vụ đặc thù (mùa mưa, mùa khô, thời điểm Tết Nguyên đán, tháng 7 Âm lịch được xem là tháng xấu khiến nhiều chủ đầu tư trì hoãn khởi công...). Những yếu tố này tạo sự biến động lớn theo thời gian, không tuân tự trong năm, khiến dự báo tuyến tính dễ gặp sai số cao [3]. Bên cạnh đó, hành vi khách hàng, tiến độ công trình, và tính chu kỳ kinh tế cũng có thể tác động đến nhu cầu bê tông, đặc biệt giai đoạn sau dịch bệnh Covid-19 với những đứt gãy chuỗi cung ứng [4]. Để đối phó với sự bất ổn này, nhiều phương pháp dự báo dựa trên các mô hình chuỗi thời gian (time series) đã được đề xuất, có thể kể đến ARIMA, SARIMA, Holt-Winters (Exponential Smoothing), Prophet (trước đây là Facebook Prophet), ...[5].

Nghiên cứu này hướng đến hai mục tiêu chính. Thứ nhất, đánh giá hiệu quả của bốn mô hình dự báo phổ biến trong bối cảnh dữ liệu sản xuất bê tông thương phẩm ở Việt Nam, nơi có tính mùa vụ và biến động lớn. Thứ hai, lựa chọn mô hình phù hợp, giúp doanh nghiệp (điển hình là Công ty Vina 21) lên kế hoạch nguyên vật liệu, bố trí tài chính và nhân lực hiệu quả. Dữ liệu được thu thập từ bảng kê sản lượng bán ra hằng ngày, bao gồm nhiều mức bê tông, mức tiêu thụ trong năm 2021–2024, đồng thời có sự tham vấn chuyên gia về thông tin mùa vụ. Phần tiếp theo của bài báo được cấu trúc như sau: Phần 2 trình bày tổng quan nghiên cứu, bao gồm các lý thuyết và công trình liên quan; Phần 3 là Phương pháp nghiên cứu, giới thiệu chi tiết các mô hình thời gian dự báo, bao gồm cơ sở toán học và ưu nhược điểm; Phần 4 nêu và phân tích kết quả của bốn mô hình dự báo; Phần 5 thảo luận về khả năng áp dụng kết quả nghiên cứu trong dự báo và lập kế hoạch sản xuất; cuối cùng là phần kết luận của bài báo này.

2. Tổng quan nghiên cứu

2.1. Vai trò của dự báo trong ngành xây dựng

Trong bối cảnh phát triển kinh tế – xã hội, ngành xây dựng ở các quốc gia đang phát triển, bao gồm Việt Nam, chiếm tỉ lệ đóng góp đáng kể vào tổng sản phẩm nội địa [6]. Việc dự báo chính xác nhu cầu vật liệu xây dựng giúp doanh nghiệp giảm thiểu chi phí lưu kho, tránh tình trạng thiếu hụt nguyên vật liệu cục bộ, đảm bảo tiến độ thi công. Các

phương pháp dự báo nhu cầu bê tông, xi măng, sắt thép thường dựa vào mô hình kinh tế lượng, hồi quy đa biến, hoặc mô hình chuỗi thời gian [7]. Tuy nhiên, các biến động theo mùa (mùa mưa kéo dài hoặc tháng 7 Âm lịch) cũng như hành vi tâm lý thị trường (thí dụ nhu cầu tăng nhanh trước Tết) làm cho việc dự báo bằng mô hình tuyến tính đơn giản kém chính xác.

2.2. Nghiên cứu về dự báo chuỗi thời gian

Dự báo bằng chuỗi thời gian (time series forecasting) ra đời từ khá sớm, nhưng ngày nay được áp dụng rộng rãi nhờ sự phát triển của thống kê và máy tính [8]. Về mặt lý thuyết, có thể phân các mô hình chuỗi thời gian thành ba nhóm chính:

- Mô hình dự báo cổ điển: ARIMA (AutoRegressive Integrated Moving Average – Tự hồi quy trung bình trượt), SARIMA (Seasonal ARIMA – ARIMA có tính đến mùa vụ), Holt-Winters (Exponential Smoothing – san mũ) [9].
- Mô hình học máy: Random Forest, Gradient Boosting, Neural Network, RNN/LSTM [10].
- Mô hình lai (hybrid) hoặc mô hình Bayesian hiện đại (như Prophet) [11].

Trong phạm vi nghiên cứu này, nhóm tác giả lựa chọn bốn mô hình cổ điển và bán hiện đại (Prophet) để so sánh. Vài công trình tiêu biểu đã ứng dụng các mô hình này trong lĩnh vực xây dựng:

- ARIMA và SARIMA: Thường được sử dụng cho dữ liệu có tính dừng (stationary) hoặc dừng theo mùa, ví dụ lượng khách du lịch, nhu cầu xi măng, doanh số bán hàng tuần. ARIMA tỏ ra hiệu quả khi chuỗi không có đột biến lớn, còn SARIMA bổ sung thành phần mùa, phù hợp chu kỳ tuần/tháng/năm [12].
- Holt-Winters: Dựa trên nguyên lý san mũ (exponential smoothing) với ba thành phần: level, trend và seasonality. Mô hình này tỏ ra linh hoạt khi có mùa vụ ổn định, thường được các doanh nghiệp tổng thầu triển khai do đơn giản và có tính giải thích tốt [13].
- Prophet: Được Facebook (nay là Meta) đề xuất, dựa trên hàm Fourier để mô phỏng seasonality, kết hợp hồi quy Bayesian. Prophet thích hợp với dữ liệu có xu hướng (trend) dài hạn và nhiều thành phần mùa (tuần, tháng, năm), đồng thời xử lý tốt các ngoại lệ (ngày nghỉ, sự kiện, rủi ro cục bộ) [11].

2.3. Bối cảnh nghiên cứu tại Việt Nam

Các nghiên cứu dự báo nhu cầu vật liệu xây dựng tại Việt Nam còn khá rời rạc, thường tập trung vào xi măng hoặc sắt thép. Đối với bê tông thương phẩm, quy trình phối trộn và phân phối chịu ảnh hưởng lớn từ địa phương cung cấp với mạng lưới cung ứng phức tạp. Một số đơn vị lớn đã triển khai hệ thống quản lý hiện đại, nhưng các doanh nghiệp vừa và nhỏ còn hạn chế nguồn lực, dễ bị động về nguyên vật liệu, cạnh tranh giá.

Như vậy, việc ứng dụng mô hình chuỗi thời gian để dự báo bê tông, nắm bắt mùa vụ năm, tìm mô hình phù hợp với đặc thù biến động

manh là nhu cầu cấp thiết. Nghiên cứu này được kỳ vọng đóng góp vào khoảng trống này, đồng thời cung cấp bài học thực tế cho các doanh nghiệp vừa và nhỏ ngành sản xuất vật liệu xây dựng.

3. Phương pháp nghiên cứu

Nghiên cứu được tiến hành theo các bước: (1) Thu thập và làm sạch dữ liệu; (2) Áp dụng bốn mô hình dự báo chuỗi thời gian: Seasonal Naive, SARIMA, Holt-Winters, Prophet (các mô hình này được phát triển từ cơ sở toán học thành các chương trình trên máy tính, sử dụng ngôn ngữ lập trình Python); (3) Đánh giá kết quả bằng các chỉ số sai số; (4) Thử nghiệm lựa chọn mô hình khả thi nhất phục vụ lập kế hoạch sản xuất và thu thập ý kiến tại trường hợp nghiên cứu điển hình.

3.1. Cơ sở toán học của bốn mô hình dự báo

3.1.1. Seasonal Naive (phương pháp tham chiếu)

Phương pháp Seasonal Naive không phải mô hình phức tạp, nhưng thường dùng làm cơ sở để so sánh (baseline). Ý tưởng: dự báo giá trị của ngày t bằng giá trị ngày $t - m$, trong đó m là độ dài mùa vụ. Ví dụ, nếu chu kỳ tuần $m = 7$, ta có:

$$\hat{y}_t = y_{t-m} \quad (1)$$

Phương pháp này nhanh, đơn giản, song hiệu quả rất hạn chế khi dữ liệu có nhiều biến động bất thường [14]. Tuy nhiên, Seasonal Naive hữu ích để chứng minh một mô hình phức tạp hơn phải vượt qua cơ sở đơn giản này.

3.1.2. SARIMA (Seasonal ARIMA)

ARIMA (AutoRegressive Integrated Moving Average – Tự hồi quy trung bình trượt) mô hình hóa chuỗi thời gian bằng cách kết hợp thành phần tự hồi quy (AR), sai phân (I) và trung bình trượt (MA) [11]. Công thức tổng quát:

$$ARIMA(p, d, q): \phi_p(L)\nabla^d y_t = \theta_q(L)\varepsilon_t \quad (2)$$

trong đó:

- ∇^d là toán tử sai phân bậc d .
- $\phi_p(L)$ là đa thức hồi quy bậc p .
- $\theta_q(L)$ là đa thức trung bình trượt bậc q .
- ε_t là nhiễu trắng (white noise).

SARIMA (Seasonal ARIMA) bổ sung thêm các bậc mùa vụ (P, D, Q, m) cho chu kỳ m , tạo thành:

$$SARIMA(p, d, q) \times (P, D, Q)_m: \phi_p(L)\Phi_P(L^m)\nabla^d y_t = \theta_q(L)\Theta_Q(L^m)\varepsilon_t \quad (3)$$

trong đó ∇_m^D là toán tử sai phân mùa vụ bậc D , Φ_P và Θ_Q là các đa thức AR/MA mùa vụ bậc P, Q . SARIMA đặc biệt phù hợp với chuỗi có tính mùa, ví dụ $m = 7$ (mùa vụ tuần) hay $m = 12$ (mùa vụ tháng) hoặc thậm chí $m = 365$ (mùa vụ năm).

3.1.3. Holt-Winters (Exponential Smoothing)

Phương pháp Holt-Winters dựa trên nguyên lý san mũ (exponential smoothing) cho ba thành phần: mức nền (level), xu hướng (trend) và mùa vụ (seasonality) [15]. Phiên bản Holt-Winters cộng (additive) có dạng:

$$\begin{aligned} l_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}), \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \\ s_t &= \gamma(y_t - l_t) + (1 - \gamma)s_{t-m}, \\ \hat{y}_t + h &= l_t + hb_t + st + h - m \cdot [(h - 1)/m] \end{aligned} \quad (4)$$

Trong đó:

- l_t là thành phần mức (level).
- b_t là xu hướng (trend).
- s_t là thành phần mùa vụ, độ dài chu kỳ là m .
- α, β, γ là các tham số san mũ ($0 < \alpha, \beta, \gamma < 1$).

Nếu dữ liệu có xu hướng nhân (multiplicative) thì ta dùng phiên bản multiplicative. Holt-Winters được ưa chuộng vì tính dễ diễn giải, thích hợp cho bài toán dự báo có tính mùa vụ tương đối ổn định [15].

3.1.4. Prophet

Prophet (Facebook Prophet) tiếp cận dự báo theo khuôn khổ hồi quy Bayesian [11], diễn tả chuỗi thời gian y_t như tổng của:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (5)$$

Trong đó:

- $g(t)$ là thành phần xu hướng (trend), thường mô phỏng dạng logistic (nhị phân) hoặc tuyến tính.
- $s(t)$ là các thành phần mùa vụ tuần, năm... được xấp xỉ bằng chuỗi Fourier.
- $h(t)$ là hiệu ứng ngày lễ, sự kiện đặc biệt, rủi ro cục bộ.
- ε_t là nhiễu.

Prophet ước lượng các tham số thông qua phương pháp Hamiltonian Monte Carlo (hoặc tối ưu tương đương), cho phép linh hoạt thêm bậc Fourier, thêm ràng buộc “changepoint” về xu hướng [11]. Lợi thế của Prophet là dễ tinh chỉnh các ngày lễ, thời gian cấm xây dựng, v.v. (ngày lễ, sự kiện bất thường), nên rất phù hợp với ngành xây dựng chịu ảnh hưởng đặc thù mùa vụ và tâm lý thị trường.

3.2. Chỉ số đánh giá (Evaluation Metrics)

Nghiên cứu sử dụng hai chỉ số: RMSE (Sai số căn trung bình bình phương) và MAPE (Sai số bình phương tuyệt đối dạng phần trăm) [16]. RMSE phản ánh sai số tuyệt đối bình phương trung bình, nhạy với ngoại lệ lớn; MAPE đo sai số phần trăm trung bình, giúp đánh giá tính tương đối khi quy mô dữ liệu dao động.

Công thức:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2} \quad (6)$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \right| \quad (7)$$

4. Phân tích các phương pháp trong trường hợp nghiên cứu Công ty Cổ phần Bê tông Vina 21

Công ty Cổ phần Bê tông Vina 21 (dưới đây gọi tắt là Vina 21) là một doanh nghiệp vừa và nhỏ có trụ sở tại khu vực phía Tây Hà Nội. Vina 21 sở hữu năng lực trộn khoảng 200 tấn/giờ và đội xe bồn vận chuyển bê tông hơn 30 chiếc, cung cấp bê tông thương phẩm cho nhiều dự án xây dựng dân dụng và hạ tầng giao thông. Tuy có lợi thế địa lý và uy tín trong khu vực, công ty phải đối mặt với khó khăn trong việc dự báo nhu cầu bê tông – một yếu tố thiết yếu để lập kế hoạch sản xuất và quản trị tài chính.

Hoạt động sản xuất bê tông thương phẩm thường bị tác động bởi tính mùa vụ (mùa mưa, thời gian nghỉ Tết), cũng như các yếu tố kinh tế, chính sách đầu tư xây dựng. Đặc biệt, năm 2021 chịu ảnh hưởng của dịch COVID-19 khiến một số tháng trầm lắng, nhiều dự án tạm dừng, dẫn tới dữ liệu nhu cầu bê tông không còn điển hình như giai đoạn trước và sau. Điều này làm tăng thách thức cho bất kỳ mô hình dự báo chuỗi thời gian nào.

4.1. Miêu tả dữ liệu

Để tiến hành nghiên cứu, nhóm tác giả đã thu thập dữ liệu hằng ngày của Vina 21 cho bốn năm 2021, 2022, 2023, và 2024. Mỗi năm

được lưu trong một tệp CSV (2021.csv, 2022.csv, ...), với ví dụ được đưa ra trong Bảng 1.

Sau khi cộng dồn tất cả các ngày, thu được một chuỗi thời gian (time series) theo ngày, đi từ ngày 01/01/2021 đến 31/12/2024 trong phạm vi nghiên cứu. Năm 2021 đặc biệt có nhiều ngày khối lượng rất thấp hoặc không có giao dịch do giãn cách xã hội do dịch bệnh Covid-19.

4.2. Kết quả dự đoán bốn phương pháp: hai kịch bản so sánh

Nghiên cứu áp dụng bốn phương pháp thời gian: Seasonal Naive, SARIMA, Holt-Winters, và Prophet; với chu kỳ năm ($m = 365$). Mỗi mô hình có cách thiết lập tham số hoặc đặc trưng tính mùa vụ tương ứng. Điểm khác biệt chính nằm ở dữ liệu huấn luyện (training):

1. Kịch bản (a): Dùng 2021, 2022, 2023 để huấn luyện mô hình, kiểm tra trên 2024.

2. Kịch bản (b): Chỉ dùng 2022, 2023 (bỏ qua 2021) để huấn luyện, kiểm tra trên 2024.

Mục tiêu là so sánh hiệu quả dự báo khi có hoặc không có dữ liệu của năm 2021 – vốn biến động vì Covid-19 – và đánh giá mức độ thích hợp của từng mô hình.

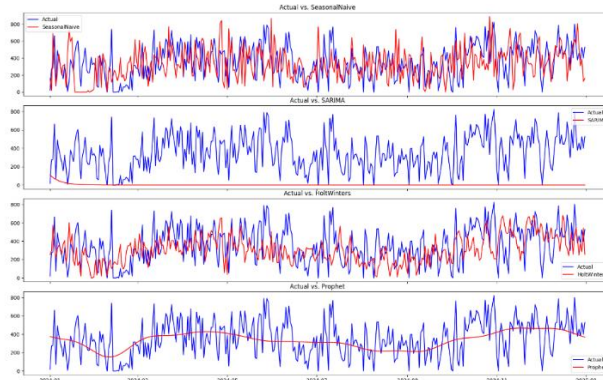
Bảng 1. Trích dữ liệu sản xuất theo các mức bê tông.

Ngày	Loại bê tông	ĐVT	Khối lượng	Đơn giá	Tổng thanh toán
...
01/01/2023	Bê tông thương phẩm mác 150	m ³	24.00	810,000	19,440,000
01/01/2023	Bê tông thương phẩm mác 200	m ³	8.00	860,000	6,880,000
01/01/2023	Bê tông thương phẩm mác 200	m ³	4.00	920,000	3,680,000
01/01/2023	Bê tông thương phẩm mác 350R7	m ³	68.50	1,110,000	76,035,000
...

- **Kịch bản (a):** Huấn luyện với 2021–2023, kiểm tra 2024

Nhóm tác giả tiến hành chạy bốn mô hình. Kết quả dự báo được thể hiện trên

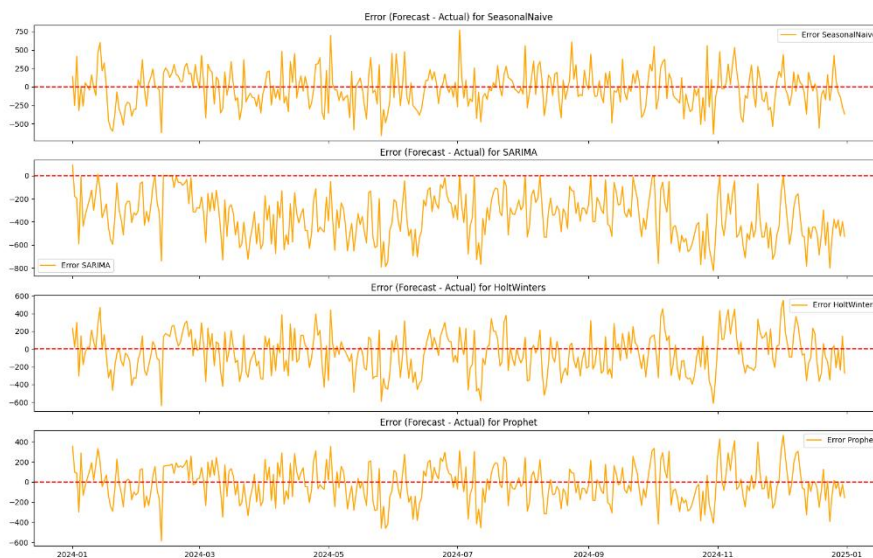
Hình 1.



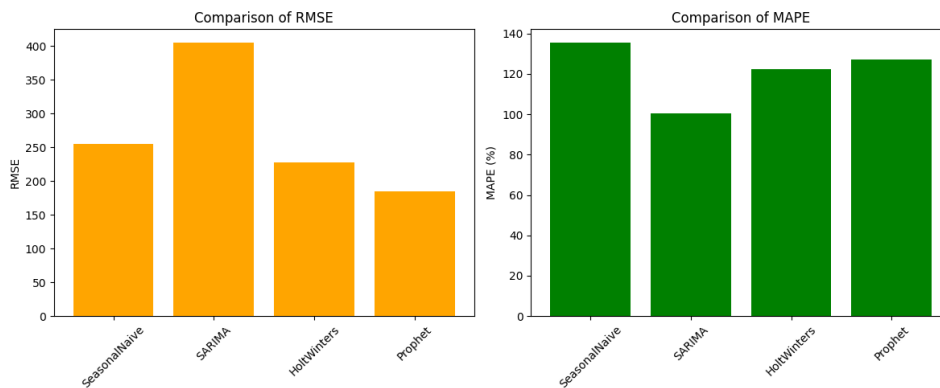
Hình 1. Kết quả dự đoán của các mô hình trường hợp a (dữ liệu huấn luyện bao gồm 2021-2023).

Sai số giữa dự đoán và thực tế được thể hiện trên

Hình 2.



Hình 2. Sai số giữa dự báo và thực tế trong trường hợp a.



Hình 3. Sai số RMSE và MAPE của bốn phương pháp trong trường hợp a.

Kết quả có thể tóm lược:

1. Prophet: Với chuỗi ba năm (2021–2023), Prophet thu nhận nhiều biến động, xu hướng. Kết quả mô hình “trung hòa” khá tốt, RMSE – thấp nhất trong các mô hình, nhưng MAPE vẫn tương đối cao.

2. Holt-Winters: Nhờ chuỗi dữ liệu dài, Holt-Winters ước lượng được biên độ biến thiên tốt hơn. RMSE, MAPE (~122 %).

3. SARIMA: Khi đưa thêm năm 2021, trong quá trình tính toán, mô hình SARIMA gặp nhiều vấn đề hội tụ và ngốn tài nguyên bộ nhớ. Ở một số cấu hình, SARIMA không thể hoàn tất vòng lặp hoặc cho ra kết quả một đường nằm ngang (tức toàn bộ tham số bị triệt tiêu hoặc mô hình ước lượng sai phân bằng 0). Kết quả cuối cùng RMSE lên đến ~404, MAPE ~100, nhưng cốt lõi là mô hình thất bại về mặt đường dự báo (nằm ngang, hoặc giá trị rất bé).

4. Seasonal Naive: Phương pháp cơ sở này dự đoán “hôm nay = hôm nay - 365 ngày”, RMSE ~255, MAPE ~135.

Hình 1 minh họa đường dự báo, thể hiện Prophet, Holt-Winters, Seasonal Naive vẫn dao động sát thực tế, còn SARIMA gần như phẳng hoặc sai số quá lớn do không hội tụ. Tương tự,

Hình 2 là biểu đồ sai số, trong đó SARIMA có sai số rất lớn ở nhiều ngày, còn Prophet và Holt-Winters phân bố quanh 0.

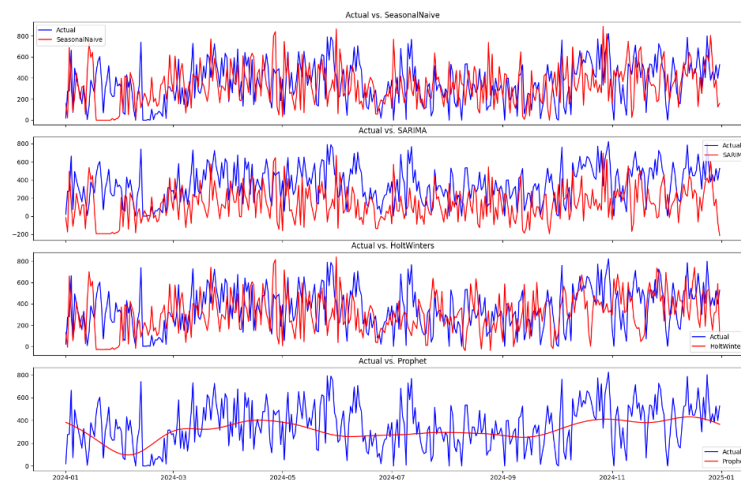
Hình 3 là biểu đồ so sánh RMSE, MAPE, xác nhận Prophet có RMSE thấp nhất, Holt-Winters thấp thứ hai, Seasonal Naive đứng thứ ba, SARIMA kém nhất về RMSE.

Nguyên nhân: Dữ liệu 2021 quá biến động (nhiều ngày 0 m³, giãn cách) khiến SARIMA, vốn nhạy cảm với tính chất “dừng theo mùa”, khó xử lý. Quá trình tối ưu mô hình (có tham số mùa 365) gây tốn tài nguyên, nhiều vòng lặp không hội tụ. Kết quả là SARIMA đưa ra đường nằm ngang hoặc số nhỏ dần về 0.

- **Kịch bản (b):** Huấn luyện với 2022–2023, kiểm tra 2024

Để kiểm chứng vai trò của năm 2021, các tác giả lược bỏ toàn bộ dữ liệu 2021, chỉ huấn luyện mô hình trên hai năm 2022–2023 rồi kiểm thử trên năm 2024. Tương tự, các kết quả được thể hiện trên

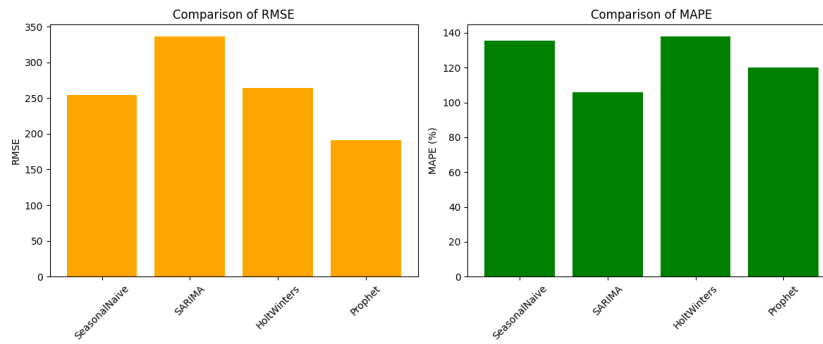
Hình 4,
Hình 5, và
Hình 6.



Hình 4. Kết quả dự đoán của các mô hình trường hợp b (dữ liệu huấn luyện bao gồm 2022-2023).



Hình 5. Sai số giữa dự báo và thực tế trong trường hợp b.



Hình 6. Sai số RMSE và MAPE của bốn phương pháp trong trường hợp b.

Nhận xét về kết quả:

1. Prophet: Khi giảm còn 2 năm huấn luyện, Prophet vẫn duy trì RMSE ~191, MAPE ~120. Vẫn là mô hình tốt nhất về sai số tuyệt đối, dù kém hơn so với kịch bản (a) (do thiếu bớt một năm dữ liệu).

2. Holt-Winters: RMSE ~264, MAPE ~138, kém hơn so với kịch bản (a) (227 & 122). Rõ ràng Holt-Winters mất lợi thế khi bị lược bớt một năm dữ liệu đầu vào.

3. SARIMA: Một quan sát thú vị là SARIMA có thể chạy xong mà không gặp lỗi “quá tải” hay “đường ngang”. Việc bớt dữ liệu cũ của năm 2021 khiến quá trình tính mùa vụ (365 ngày) ít nhiều. Mặc dù vậy, RMSE vẫn khá cao (336), MAPE ~106. SARIMA vẫn có xu hướng văng ở những ngày cao, nhưng không đổ vỡ như kịch bản (a).

4. Seasonal Naive: Giữ RMSE ~255, MAPE ~135, tương đối gần kịch bản (a) vì cách “hôm nay = hôm nay - 365” không thay đổi.

Hình 4 cho thấy SARIMA không bị phẳng, mà có đường dao động song vẫn sai khác lớn. Prophet và Holt-Winters tiếp tục dao động quanh thực tế.

Hình 5 minh họa sai số, trong đó SARIMA không nằm ngang nữa, nhưng vẫn có mức lỗi khá cao.

Hình 6 so sánh RMSE, MAPE, kết luận Prophet vẫn dẫn đầu RMSE, Holt-Winters đứng thứ hai, SARIMA xếp thứ ba (hoặc kém nhất về RMSE nếu so sánh chung) nhưng hoạt động được (không treo) vì ít dữ liệu nhiều từ 2021.

• Nhận xét về dữ liệu 2021

Có thể thấy việc dùng dữ liệu của năm 2021 có cả bất lợi và thuận lợi:

- Thuận lợi: Mô hình như Prophet, Holt-Winters (san mũ) ưu tiên nhiều dữ liệu quan sát, do đó có thể trích rút được khuynh hướng dài hạn tốt hơn, dẫn đến RMSE được cải thiện.

- Bất lợi: Đối với SARIMA, năm 2021 ghi nhận nhiều ngày 0 m³, giãn cách xã hội, rất chênh lệch với 2022–2023. Mô hình ARIMA (có tham số mùa 365) không hội tụ hoặc tiêu thụ nhiều bộ nhớ của máy tính, kết quả là một đường nằm ngang (các tham số bị triệt tiêu).

Từ đó, ta rút ra hai cách tiếp cận:

1. Nếu doanh nghiệp muốn mô hình đơn giản hơn, chấp nhận bỏ qua một năm bất thường, thì SARIMA vẫn chạy được, MAPE đôi lúc khả quan.

2. Nếu muốn tận dụng tối đa dữ liệu cho Prophet, Holt-Winters, thì giữ năm 2021; SARIMA có thể bị loại khỏi danh sách do quá tải trong quá trình chạy mô hình.

Như vậy, tùy mục tiêu và năng lực tính toán, người dùng có thể cân nhắc:

- Phương án 1: Dùng trọn bộ 2021–2023, bỏ qua SARIMA (hoặc buộc SARIMA bậc rất thấp), còn Prophet/Holt-Winters sẽ tối ưu.

- Phương án 2: Loại 2021, để SARIMA hoạt động, chấp nhận hy sinh một năm thông tin và Prophet, Holt-Winters vẫn có sai số tạm được (RMSE cao hơn một chút).

Trong trường hợp Công ty Vina 21, nếu ưu tiên sai số tuyệt đối (RMSE), Prophet vẫn là lựa chọn an toàn trong cả hai kịch bản. Holt-Winters cũng tương đối khả quan, miễn có đủ dữ liệu. Còn SARIMA, trừ khi công ty muốn đặc biệt quan tâm MAPE (kiểm soát tỉ lệ sai lệch), thì nên thận trọng với những năm đột biến như 2021.

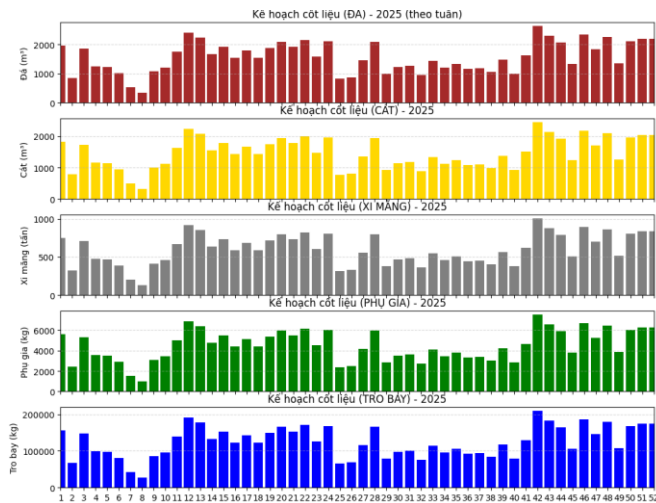
5. Thảo luận về tính ứng dụng trong dự báo và lập kế hoạch sản xuất bê tông

Trong giai đoạn cuối của nghiên cứu, nhóm tác giả đã trình bày kết quả trên cho lãnh đạo Công ty CP Bê tông Vina 21, gồm giám đốc điều hành, trưởng phòng kỹ thuật, trưởng bộ phận bán hàng, và các tổ trưởng ca sản xuất. Tất cả đều đánh giá cao khả năng tận dụng dữ liệu nội bộ để xác định kế hoạch nguyên vật liệu, nhân công, và dòng tiền. Đặc biệt, việc so sánh bốn mô hình (Seasonal Naive, SARIMA, Holt-Winters, Prophet) giúp công ty hiểu rõ tính chất “phù hợp” hay “thiếu ổn định” của từng phương pháp.

Khi đã chọn lọc được mô hình dự báo tối ưu (ví dụ như Prophet nếu doanh nghiệp coi trọng sai số tuyệt đối), Vina 21 có thể triển khai quy trình lập kế hoạch sản xuất theo tuần, theo tháng hoặc theo năm. Ví dụ,

Hình 7 minh họa cách sử dụng Prophet để dự báo nhu cầu nguyên vật liệu cho năm 2025. Để tính toán, các tác giả dự đoán tách riêng từng

mác bê tông (mác 200, 250, 300...) và áp dụng định mức cấp phối, từ đó tính được khối lượng xi măng, cát, đá, phụ gia... cho từng tuần.



Hình 7. Dự báo nguyên vật liệu cho 2025 (dự kiến 7 ngày dự trữ).

Các thông tin dự báo được ban lãnh đạo Vina 21 đánh giá là rất giá trị, và có thể được sử dụng với các mục đích:

1. Quản lý kho nguyên vật liệu: giảm tồn kho quá mức, tránh thiếu hụt bất ngờ.
2. Đặt hàng ổn định từ nhà cung cấp: đảm bảo giá vật tư tốt, không bị động khi thị trường biến động.
3. Bố trí ca sản xuất: phân chia ca trộn, ca vận chuyển phù hợp với cao điểm nhu cầu, tránh lãng phí năng lực máy móc.
4. Quản trị tài chính: dự toán chính xác hơn cho dòng tiền (mua vật liệu, duy trì vận hành, dòng thu từ khách hàng).

Nhờ nghiên cứu này, lãnh đạo Vina 21 nhận thấy rằng, dù dữ liệu năm 2021 có thể nhiễu do Covid-19, việc đưa đầy đủ vào mô hình Prophet hoặc Holt-Winters vẫn đem lại cái nhìn đầy đủ hơn về xu hướng. Những ngày cao đột biến cũng dần được dự báo chặt chẽ hơn so với phương pháp cũ hoàn toàn thủ công. Ngoài ra, trong tương lai, công ty dự kiến tích hợp thêm yếu tố thời tiết, thông tin dự án xây dựng lớn, và lịch cấm đường... để mô hình dự báo trở nên sát thực tế hơn.

Từ kinh nghiệm trên, các tác giả gợi ý rằng các doanh nghiệp vừa và nhỏ trong ngành bê tông thương phẩm có thể bắt đầu với mô hình Prophet hoặc Holt-Winters (để thực thi, có độ tin cậy cao khi có chuỗi thời gian đủ dài). Bên cạnh đó, nhà quản lý cần hiểu rõ chỉ số RMSE, MAPE, và xác định ưu tiên: nếu quan tâm lỗi tương đối (MAPE), SARIMA đôi lúc tỏ ra hấp dẫn, nhưng nếu quan tâm lỗi tuyệt đối (RMSE) để tránh thiếu/thừa sản xuất nghiêm trọng, Prophet hoặc Holt-Winters vượt trội về khả năng dự báo. Phân hồi tích cực từ giám đốc, trường phòng kỹ thuật, và bộ phận bán hàng tại Vina 21 khẳng định giá trị của việc ứng dụng mô hình dự báo hiện đại, đảm bảo tính minh bạch và chủ động trong kế hoạch kinh doanh cho năm tới.

6. Kết luận

Bài báo đã trình bày một nghiên cứu so sánh bốn mô hình dự báo chuỗi thời gian—Seasonal Naive, SARIMA, Holt-Winters, và Prophet—trên dữ liệu thực tế từ Công ty Cổ phần Bê tông Vina 21. Nhóm tác giả triển khai hai kịch bản huấn luyện: (a) dùng trọn bộ dữ liệu 2021–2023 (trong đó 2021 là năm biến động do Covid-19) và (b) chỉ dùng hai năm 2022–2023. Kết quả thực nghiệm cho thấy: Prophet và Holt-Winters đạt sai số tuyệt đối (RMSE) thấp nhất khi dữ liệu lịch sử được giữ đầy đủ (2021–2023), nhờ khả năng nhận diện xu hướng và mùa vụ trên tập dữ liệu dài. SARIMA thường không hội tụ hoặc cho kết quả “đường nằm ngang” nếu đưa thêm năm 2021 nhiều nhiễu, tiêu tốn nhiều tài nguyên bộ nhớ. Trong trường hợp loại bỏ năm 2021, SARIMA có hoạt động nhưng vẫn mang sai số RMSE cao hơn Prophet. Seasonal Naive đóng vai trò mô hình tham chiếu, dự báo đơn giản, có hiệu quả trung bình và ít thay đổi khi có hay không có năm 2021. Về mặt ứng dụng thực tiễn tại Vina 21, kết quả phỏng vấn Ban Giám đốc và các phòng ban kỹ thuật cho thấy việc lựa chọn mô hình phù hợp (đặc biệt Prophet nếu ưu tiên giảm thiểu sai số tuyệt đối) mang lại giá trị rõ rệt trong việc lập kế hoạch nguyên vật liệu, phân bổ ca sản xuất và quản trị tài chính. Khai thác dữ liệu năm 2021 “bất thường” có thể giúp Holt-Winters, Prophet nâng cao năng lực dự báo, nhưng lại gây “quá tải” cho SARIMA khi chu kỳ mùa vụ thiết lập dài (365 ngày). Doanh nghiệp cần cân nhắc mục tiêu và khả năng tính toán của mình, cũng như thảo luận với chuyên gia, để quyết định bỏ qua hay giữ lại những năm chứa nhiễu điểm ngoại lệ. Trong tương lai, nghiên cứu có thể mở rộng với nhiều yếu tố bổ sung (dữ liệu thời tiết, lịch thi công dự án trọng điểm, chính sách nhà nước...), hoặc thử nghiệm các mô hình học sâu. Từ góc độ Công ty Vina 21 nói riêng và các doanh nghiệp vừa và nhỏ ngành sản xuất vật liệu nói chung, kết quả này thể hiện tầm quan trọng của dữ liệu chuỗi thời gian trong việc chủ động quyết định sản xuất và tối ưu hóa nguồn lực trong bối cảnh nhu cầu xây dựng tại Việt Nam sôi động nhưng vẫn còn nhiều biến động.

Lời cảm ơn

Nhóm tác giả chân thành cảm ơn sự cộng tác của ban lãnh đạo Công ty Cổ phần Bê tông Vina 21 đã cung cấp dữ liệu sản xuất kinh doanh của công ty và cùng tham gia vào quá trình nghiên cứu.

Tài liệu tham khảo

- [1]. N. T. Lâm, N. N. Linh, T. V. Nam, V. D. Kiên, T. V. Khải, and P. Đ. Hiếu, “Ảnh hưởng của tro bay thay thế một phần xi măng đến tính chất của bê tông thương phẩm,” *Tap chí Khoa học Công nghệ Xây dựng (KHCCND) - ĐHXD*, vol. 14, no. 4V, 2020, doi: 10.31814/stce.nuce2020-14(4v)-09.
- [2]. Z. Liu, Y. Zhang, and M. Li, “Integrated scheduling of ready-mixed concrete production and delivery,” *Autom Constr*, vol. 48, 2014, doi: 10.1016/j.autcon.2014.08.004.
- [3]. R. Adamek, S. Smeeke, and I. Wilms, “Lasso inference for high-dimensional time series,” *J Econom*, vol. 235, no. 2, 2023, doi: 10.1016/j.jeconom.2022.08.008.

- [4]. N. V. Tâm, N. B. Ngọc, N. Q. Toàn, and L. V. Quý, “Đánh giá ảnh hưởng của đại dịch COVID-19 đến các doanh nghiệp xây dựng quy mô vừa và nhỏ tại Việt Nam,” *Tạp chí Khoa học Công nghệ Xây dựng (KHCVND) - ĐHXD*, vol. 15, no. 2V, 2021, doi: 10.31814/stce.nuce2021-15(2v)-13.
- [5]. K. S. Sahoo, S. Krishana, and M. Bhuyan, “Combining Block Bootstrap with Exponential Smoothing for Reinforcing Non-Emergency Urban Service Prediction,” in *Proceedings - 2023 IEEE International Conference on Big Data, BigData 2023*, 2023. doi: 10.1109/BigData59044.2023.10386377.
- [6]. Tổng cục Thống kê, *Niên giám thống kê*. Nhà xuất bản Thống kê, 2020.
- [7]. A. K. Abinesh, “Forecast of Ready Mix Concrete Operation By Neuralnetwork,” *SSRN Electronic Journal*, 2019, doi: 10.2139/ssrn.3509682.
- [8]. D. L. Danilov, “Principal components in time series forecast,” *Journal of Computational and Graphical Statistics*, vol. 6, no. 1, 1997, doi: 10.1080/10618600.1997.10474730.
- [9]. A. Gautam and V. Singh, “Parametric versus non-parametric time series forecasting methods: A review,” 2020. doi: 10.25103/JESTR.133.18.
- [10]. Y. Elfahham, “Estimation and prediction of construction cost index using neural networks, time series, and regression,” *Alexandria Engineering Journal*, vol. 58, no. 2, 2019, doi: 10.1016/j.aej.2019.05.002.
- [11]. E. F. Agyemang, J. A. Mensah, E. Ocran, E. Opoku, and E. N. N. Nortey, “Time series based road traffic accidents forecasting via SARIMA and Facebook Prophet model with potential changepoints,” *Heliyon*, vol. 9, no. 12, 2023, doi: 10.1016/j.heliyon.2023.e22544.
- [12]. B. Hoła, M. Topolski, I. Szer, J. Szer, and E. Blazik-Borowa, “Prediction model of seasonality in the construction industry based on the accidentality phenomenon,” *Archives of Civil and Mechanical Engineering*, vol. 22, no. 1, 2022, doi: 10.1007/s43452-021-00348-7.
- [13]. T. Moon and D. H. Shin, “Forecasting Construction Cost Index Using Interrupted Time-Series,” *KSCE Journal of Civil Engineering*, vol. 22, no. 5, 2018, doi: 10.1007/s12205-017-0452-x.
- [14]. J. Waller and E. I. Svensson, “The measurement of selection when detection is imperfect: How good are naïve methods?,” *Methods Ecol Evol*, vol. 7, no. 5, 2016, doi: 10.1111/2041-210X.12498.
- [15]. O. Trull, J. C. García-Díaz, and A. Troncoso, “Initialization methods for multiple seasonal holt-winters forecasting models,” *Mathematics*, vol. 8, no. 2, 2020, doi: 10.3390/math8020268.
- [16]. E. Ostertagová and O. Ostertag, “Forecasting using simple exponential smoothing method,” *Acta Electrotechnica et Informatica*, vol. 12, no. 3, 2013, doi: 10.2478/v10198-012-0034-2.