Hybrid ARIMAX-LSTM Model for Prediction of Thickness in Plate Mill Rolling:
A Time Series Analysis Approach

Tae Young Jung*
Shinsung University, Republic of Korea
Jae Geun Kim
Shinsung University, Republic of Korea
In Cheol Baek
Shinsung University and Inha University, Republic of Korea
Peter Yoo
State University of New York, Korea, Republic of Korea

Abstract

Data collection, analysis, and forecasting are increasingly vital in industrial settings, and various methods are being studied and utilized for effective data analysis and forecasting. Time series exhibits dynamic characteristics corresponding to manufacturing production processes and can provide valuable insights for process analysis and forecasting. This study proposes an approach to analyzing time series data on the thickness of steel plates in rolling processes using the ARIMAX (Autoregressive Integrated Moving Average with Exogenous variables) model, followed by further analysis and prediction of the residuals using the LSTM (Long Short-Term Memory) model. As the ARIMAX model alone may not fully capture the complex nonlinear patterns and long-term dependencies inherent in time series data, its performance in prediction can be limited. In this paper, we demonstrate that combining the ARIMAX model with LSTM improves the prediction of these intricate nonlinear characteristics observed in the time series data of the steel plate thickness during the rolling process, providing useful insights for various field applications.

Keywords: ARIMAX, LSTM, steel plate thickness, residuals
JEL: Q02

*Contact author: Email tyjung@shinsung.ac.kr, Phone: +82.41.350.1534

1. Introduction

Time series analysis and forecasting are widely used in various industries to analyze data patterns over time and predict future data changes. To analyze and forecast time series data, diverse probabilistic models are used, among which the ARIMA model is one of the most well-known and widely used. The ARIMA model, combining autoregressive (AR) and moving average (MA) models with differencing, is effective for analyzing the linear properties of basic time series data. The application of the ARIMA model for data forecasting has been widely studied, and the research by D. Aborass and A. H. Hassan demonstrates a predictive analysis using ARIMA. However, it has limitations when analyzing data with nonlinear characteristics. This limitation has been addressed in researches including those by P. G. Zhang (2003), P. T. Yamak (2019), P. Liu (2022) and A. A. Alsuwaylimi (2023). Therefore, another model capable of analyzing nonlinear characteristics is required for more accurate forecasting.

Artificial neural networks (ANNs) can be used to analyze and forecast such nonlinear characteristics. However, typical ANNs may experience issues like long-term dependency problems or vanishing gradients, which can reduce prediction accuracy. To address these limitations, deep learning-based methods like the Long Short-Term Memory (LSTM) model, designed to overcome issues with RNNs, are successfully applied to time series analysis and forecasting. However, because time series data often exhibit both linear and nonlinear characteristics, a hybrid ARIMA-LSTM model that analyzes the linear characteristics with an ARIMA model and applies LSTM to the residuals, which contain nonlinear characteristics, is expected to be more efficient and reliable by reducing slow learning times and overfitting risks associated with LSTM-only analysis. To analyze and predicted time series data with both linear and nonlinear characteristics, the integrated approach of ARIMA and LSTM has been researched by S. Siami-Namini (2018), D. Xu (2021), and P. Liu (2022), demonstrating improved results.

In this study, we aim to predict the thickness of steel plates in rolling processes by analyzing data related to rolling conditions and comparing the prediction accuracy to actual measured thicknesses. Plate steel, manufactured through hot rolling, is classified into categories like shipbuilding, structural, pressure vessel, and plate steel, each with critical dimensional quality requirements. The hot rolling process is a time-continuous process. In the hot rolling process, the rolling conditions are continuously changed within the same line's rolling operation to produce plates of varying thickness according to order specification. At the moment the rolling conditions changed, thickness deviations occur in the plates, which can lead to product quality defects. If it is possible to predict the thickness deviations that may occur during the process of changing rolling conditions, appropriate control measures (e.g., roll gap adjustments) considering the thickness deviations can be implemented to reduce the deviations.

This study applies the hybrid ARIMA-LSTM model to the time series data of rolling conditions to analyze and predict both linear and nonlinear characteristics.

2. Methodology

2.1. Description of Data

Steel plates are thick sheets (6mm or thicker) produced through hot rolling processes. Plate manufacturing can be classified into as-rolled (AR) and controlled rolling (CR), depending on the control of finishing rolling temperature (FRT). Controlled rolling allows for finer microstructure and improved strength and toughness by controlling the reduction rate in the non-recrystallized region. Differences in FRT and target tensile strength (TS) between these two methods affect plate thickness deviation. Additionally, changes in target thickness and rolling load can cause thickness deviation. Generally, in the hot rolling process, the thickness is controlled by adjusting the roll gap. If the thickness deviations caused by changed in rolling conditions can be predicted, the roll gap can be adjusted accordingly to account for the predicted deviation.

In this study, we predict plate thickness based on actual thickness data as the output and factors affecting thickness deviation, such as FRT, target thickness, tensile strength, and rolling load as inputs.

In addition, the hot rolling process is a time-continuous process, and therefore, the data mentioned above can be regarded as time series data.

2.2. ARIMA and ARIMAX Model

The ARIMA model, combining the AR (Autoregressive) and MA (Moving Average) models, is widely used for time series forecasting. The AR model constructs a time series model based on the autocorrelation of data, predicting future values through a linear combination of past and current values. Meanwhile, the MA model linearly derives values from past error terms that affect the current value. The combination of AR and MA models can be expressed as follows:

$$y_t = c + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad (1)$$

Here, p is the order used in the AR(p) representation, with $\emptyset_1$, $\emptyset_2$, and $\emptyset_p$ representing the AR model coefficients according to the order p. The terms y_{t-1}, y_{t-2}, and y_{t-p} are the past observations over periods. The parameter q is used to represent the MA(q) form, where θ_1, θ_2, and θ_q are the MA model coefficients, and ϵ_{t-1}, ϵ_{t-2} and ϵ_{t-q} represent past error terms over q periods. The ARIMA model combines the AR and MA models, adding differencing of order d. The ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) model extends the ARIMA model by including exogenous variables, which are external factors affecting the time series data.

2.3. LSTM Model

The Long Short-Term Memory (LSTM) model is a specialized form of the Recurrent Neural Network (RNN), designed to overcome certain limitations of RNNs. RNNs are artificial neural networks intended for deep learning with time series or sequential data. Their structure includes hidden states and loops, allowing previous computations to influence future outcomes. However, RNNs face challenges, particularly in retaining long-term dependencies, as information from earlier data often fades over time. To address this issue, the LSTM model was introduced.

LSTM enhances the basic RNN structure by incorporating a cell state, which enables the flow of information to be retained across time steps, thus addressing the vanishing gradient problem. The LSTM model consists of three main gates: Forget Gate, Input Gate, and Output Gate, which control the cell state's information flow:

- **Forget Gate** decides whether to retain or discard past information. Using the sigmoid activation function, the output approaches 0 to discard information or 1 to retain it.
- **Input Gate** determines how much current information to store in the cell state. This gate uses both sigmoid and tanh activation functions to calculate the gate's output.
- **Output Gate** calculates what value to output from the cell, adjusting it based on the current information and the cell state value, allowing relevant information to be

reflected in the output.

2.4. Hybrid ARIMA-LSTM Model

The ARIMAX model is effective in forecasting when data exhibits linear characteristics; however, real-world data often contains both linear and nonlinear properties. Using the ARIMAX model alone may therefore fall short in predicting the nonlinear aspects of the data. Specifically, the residuals, which represent the error between the predicted and observed values, may contain nonlinear characteristics that are difficult to analyze using a purely linear structure.

In the equation (2)

$$Re = D_{ob} - D_{L-est} \tag{2}$$

where D_{ob} denotes the actual observed value and D_{L-est} represents the predicted value generated by the ARIMAX model, Re is the residual, or the difference between the observed and predicted values. This residual, being what the ARIMAX model cannot account for, reflects the nonlinear component that remains after the linear elements have been modeled.

Since the ARIMAX model cannot model nonlinear characteristics effectively, it is advantageous to analyze and predict these residuals using an artificial neural network (ANN) model, particularly the LSTM. The hybrid ARIMAX-LSTM approach proposed in this study leverages both the linear and nonlinear attributes of the data. First, the ARIMAX model analyzes the data's linear components to form the initial predictions. Next, the residuals, or the difference between the observed and ARIMAX-predicted values, are calculated and further analyzed with the LSTM model to capture the nonlinear characteristics.

By applying the hybrid ARIMAX-LSTM model, this method combines predictions from both linear and nonlinear analyses, ultimately improving the overall forecasting performance.

3. Data Analysis and Prediction

3.1 Data Sets

The data used in the simulation consists of measurements from actual steel industry operations, specifically related to changes in plate thickness due to the rolling process. This dataset includes output data representing the thickness of plates and key input variables that are considered to influence thickness deviations in the rolling process, such as Finishing Rolling Temperature (FRT), tensile strength, target thickness, and rolling load. These thickness deviation factors are treated as exogenous variables in the ARIMAX model, with 50% of the data used for training and the remaining 50% reserved as test data for validating the prediction results.

For the LSTM model, the actual observed plate thickness serves as the output. Inputs include the rolling process parameters (FRT, tensile strength, target thickness, rolling load) and residuals calculated from the ARIMAX model’s predictions. The LSTM model uses these inputs to learn and enhance the accuracy of the overall forecasting.

3.2 ARIMAX Model Setting

The comparison of settings for the AR(p) and MA(q) orders for the ARIMA or ARIMAX model, including exogenous variables, is shown in Table 3.1:

Table 3.1

RMSE Comparison of AR(p), Differencing(d) and MA(q) orders

AR(p) order	Differencing(d) order	MA(q) order	RMSE
1	0	1	5.1771
1	1	1	335.8647
2	0	2	5.1745
3	0	2	5.1766
3	0	3	5.1725
4	0	4	5.1752

As indicated in Table 3.1, the optimal configuration for the ARIMAX model for forecasting the data is when the RMSE (Root Mean Squared Error) value is minimized. The best setting is **p=3, d=0, q=3**, with the lowest RMSE of **5.1725**.

3.3 LSTM Model Setting

The main settings that impact the performance of the LSTM model include the number of hidden units, the maximum number of epochs, and the mini-batch size. These settings were adjusted, and the RMSE (Root Mean Squared Error) values were compared to finding the configuration with the lowest RMSE. Additionally, an initial learning rate of 0.01 was set, and a dropout rate of 50% was applied to improve model generalization and prevent overfitting.

Table 3.2 below shows how changes in these key settings affect the RMSE values. The configuration that achieved the minimum RMSE of 4.3724 consisted of 25 hidden units, a maximum of 1400 epochs, and a mini-batch size of 25. This RMSE reflects the difference between the LSTM model's predictions and the actual residuals derived from the ARIMAX predictions versus observed values.

Table 3.2

RMSE Comparison of setting parameters of LSTM

No. Hidden Units	Max. Epochs	Mini-Batch Size	RMSE
25	1400	25	4.3718
25	1800	25	4.4061
25	1400	10	4.3791
50	1400	25	4.4071
50	1400	20	4.4029
100	1200	25	4.3804
100	1400	25	4.3868

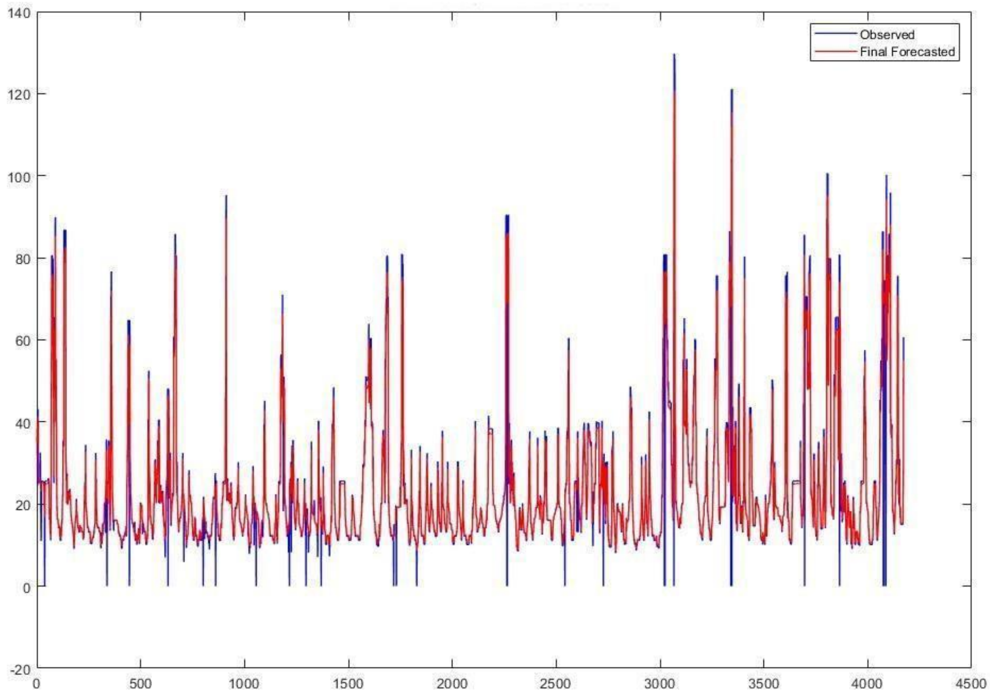
3.4 Hybrid ARIMAX-LSTM Model Forecasting

The ARIMAX model, which incorporates external variables influencing plate thickness deviation, is used to predict plate thickness in the rolling process. After obtaining predictions from the ARIMAX model, the residuals (errors between the predicted and actual values) are then predicted using the LSTM model. Finally, the ARIMAX and LSTM predictions are integrated, and the combined forecast is compared to the actual observed values.

To assess the effectiveness of this hybrid approach, the RMSE (Root Mean Squared Error) is calculated for both the standalone ARIMAX model and the integrated ARIMAX-LSTM model. This comparison demonstrates the improvement in prediction accuracy achieved by the hybrid model over the ARIMAX model alone.

Figure 3.1

ARIMAX-LSTM prediction of plate thickness



The graph displays the predicted plate thickness in the rolling process using the ARIMAX-LSTM hybrid model, configured with AR(3) and MA(3) for the ARIMAX component, alongside the actual observed values. The standalone ARIMAX model yields a prediction RMSE of 5.1766 when compared to the observed values, while the integrated ARIMAX-LSTM model achieves an improved RMSE of 4.3724. This indicates that the hybrid model provides a more accurate prediction, significantly reducing the error compared to the ARIMAX model alone.

Table 3.3

RMSE Comparison of ARIMAX and ARIMAX-LSTM

	RMSE
ARIMAX (p, d, q = 3, 0, 3)	**5.1766**
ARIMAX(3,0,3)+LSTM	**4.3718**
	15.5471(%) error reduction

As shown in Table 3.3 above, comparing the RMSE values indicates that the ARIMAX-LSTM model achieves approximately a 15.5% reduction in error compared to the ARIMAX model alone. This demonstrates that the ARIMAX-LSTM hybrid model performs better in predicting plate thickness, providing improved accuracy.

4. Conclusion

In the time series data analysis and forecasting, the ARIMA model has shown strong predictive capabilities. However, due to its nature of only analyzing and predicting the linear portion of data, it faces limitations when dealing with errors and nonlinear characteristics. To effectively analyze and predict both the linear and nonlinear properties of such data, the ARIMAX-LSTM hybrid model is applied. As the prediction results indicate, the RMSE of the standalone ARIMAX model (5.1766) decreases to 4.3718 when using the ARIMAX-LSTM hybrid model, demonstrating an error reduction of approximately 15.5471%.

By leveraging the strengths of the ARIMAX model for linear features and the LSTM model for nonlinear features, the integrated ARIMAX-LSTM model proves to be an effective method for improving prediction performance.

If applied to an online model in an actual production site in the future, the predicted thickness deviations can be reflected in roll gap adjustments, which is expected to reduce quality costs caused by dimensional defects in steel plate when rolling conditions change rapidly.

Reference

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*(3), 159–175.

Alsuwaylimi, A. A. (2023). Comparison of ARIMA, ANN, and Hybrid ARIMA-ANN Models for Time Series Forecasting. *Information Science Letters*, *12*(2), 1003-1016.

Rahayu, W., et al. (2023). Hybrid ARIMA-ANN Model for Solving Nonlinearity in Time Series Data. *Advances in Computer Science Research*, *109*.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. *IEEE International Conference on Machine Learning and Applications*.

Ma, C., et al. (2024). Predicting Stock Prices Using Hybrid LSTM and ARIMA Model. *IAENG International Journal of Applied Mathematics*, *54*(3), 424-432.

Liu, P. (2022). Time Series Forecasting Based on ARIMA and LSTM. *Proceedings of ICEMED*.

Baek, I.-C., & Jung, T.-Y. (2024). Prediction of thickness accuracy in Plate Mill rolling via artificial neural network. *Korean Society of Mechanical Technology*, *26*, 338–345.

Wang, L., Zou, H., Su, J., Li, L., & Chaudhry, S. (2013). An ARIMA-ANN Hybrid Model for Time Series Forecasting, *Systems Research and Behavioral Science*, 30(3), 244 -259

Xu, D., Zhang, Q., Ding, Y., & Zhang, D. (2021). Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting, *Environmental Science and Pollution Research*, 59, 1239-1259

Ciaburro, G. (2024). MATLAB for Machine Learning. 2nd Edition, Packt Publishing

Hochreiter, S., Schmidhuber, J., (1997). Long Short-Term Memory. *Neural Computation 9 (8), 1735-1780*

Nielsen, A. (2019). Practical Time series Analysis. O'Reilly

Khashei, M., Bijari, M. (2010). An Artificial neural network (p, d, q) model for time series Forecasting. *Expert System with Applications, 37, 479-489*

Yamak, P. T., Yujian, L., Gadosey, P. K. (2019). A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting. *2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*

Aborass, D., Hassan, A. H., Sahalash, I., Al-Rimmawi, H. (2022). Application of ARIMA Models in Forecasting Average Monthly Rainfall in Birzeit, Palestine. *Computers and Industrial Engineering, 35(1-2), 213-216*

Shi, S., Xu, L., Liu, B. (1996). Applications of Artificial Neural Networks to the Nonlinear Combination of Forecasts. *Expert System 13, 195-201*

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61, 85-117*

Schuster, M., Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing, 45(11), 2673-2681*

Shao, G. (2023). Prediction of Stock Prices Based on the LSTM Model. *Proceedings of the 8th International Conference on Financial Innovation and Economic Development (ICFIED 2023), 377-387*

Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., Zhang, H. (2018). Deep Learning with Long Short-Term Memory for Time Series Prediction, *IEEE Communications Magazine, 57(6), 114-119*