

# Research on Sales Forecasting Method for Transmission Parts of Customized Production

Jun Cheng<sup>1</sup>, Tao Hong<sup>1</sup>, Xiaohong Li<sup>2</sup> and Honglin Wang<sup>1,\*</sup>

<sup>1</sup> University of Electronic Science and Technology of China, Chengdu 611731, China;

<sup>2</sup> Sichuan Deen Precision Technology Co., Ltd. 2, Meishan, Sichuan 620460, China

\*Corresponding author

**Abstract**—With the advancement of manufacturing technology and the diversification of customer needs, product sales forecasting is critical to the company in a small-volume, multi-variety production model. Based on the traditional ARIMA model and the deep neural network LSTM model, this paper constructs the ARIMA-LSTM hybrid model to predict the sales of transmission parts manufacturing enterprises. The model uses the LSTM neural network to describe the linear and nonlinear components of the time series rather than the simple additive relationship in the traditional hybrid model. Therefore, the model overcomes the problems of low prediction accuracy that may exist in the traditional hybrid model. This paper selects the historical sales data of a certain manufacturer of a large transmission parts manufacturer in Sichuan, and carries out experimental verification on the constructed ARIMA-LSTM hybrid model. The results show that the model has high prediction accuracy and can provide reference for enterprise decision-making. The research work of this paper is supported by Sichuan Science & Technology Program under Grant No.2017GZ0144 & 2017GZ0047.

**Keywords**—product sales forecasting; ARIMA model; LSTM model; ARIMA-LSTM hybrid model

## I. INTRODUCTION

Under the influence of the development trend of contemporary e-commerce, the traditional manufacturing sales model has undergone tremendous changes. The traditional offline sales model has gradually been replaced by the network personalized production model. E-mail is increasingly used in the production and sale of products. Online orders and other forms, order products are characterized by small batch sizes, multiple varieties, and high customization requirements. At the same time, with the advancement of manufacturing technology and the diversification of customer needs, contemporary manufacturing companies are under pressure to respond quickly to customer needs and reduce product delivery time. Under the new market situation and competitive pressure, it is crucial to accurately predict market demand. By predicting the company's ability to accurately grasp the future needs of the market and products, and on this basis, more scientific and effective development of various products. Production planning, optimizing product line layout, timely and rational adjustment of production plans, optimizing inventory strategies, and reducing sales costs to win customers and markets.

Mathematical models commonly used in the field of prediction include linear models and nonlinear models. The most widely used linear model is the ARIMA model, because its theory is mature and the model is simple and convenient to use.

The most widely used nonlinear model is the artificial neural network [1-2] because of its good nonlinear fitting ability. However, there are some shortcomings in the ANN [3]. For example, the training process will fall into the local minimum and the over-fitting phenomenon will affect the prediction accuracy of the model. With the development of deep learning, traditional neural networks are gradually replaced by deep neural networks. In the field of time series prediction, LSTM is a commonly used deep learning model, because it can utilize historical information, so it is very suitable for processing time series.

Although linear models and nonlinear models have good applications in their respective fields, in the face of complex time series problems, a single model does not achieve good prediction results. Therefore, the researchers combined the linear model with the nonlinear model, so that they can give full play to the respective advantages of the two models and improve the prediction accuracy.

The traditional model mixing method assumes that the linear component and the nonlinear component of the time series are simple additive relationships. This simplification has certain risks and may lead to a decrease in prediction accuracy. In this paper, the traditional model method is improved without first assuming the relationship between the two, but using the self-learning ability of the LSTM neural network to describe the complex relationship between the two. Finally, the hybrid model improved by this paper is used to predict the sales volume of a certain product in the mechanical transmission parts manufacturing enterprise. Finally, the experimental comparison is carried out to verify its feasibility.

## II. INTRODUCTION TO ARIMA MODEL AND LSTM MODEL

### A. ARIMA Model

The basic theory of ARIMA is described in the literature[4-5][12]. The mathematical expression of the ARIMA model is as follows:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \delta + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q} \quad (1)$$

Where  $x_t$  represents the sequence value at the current time;  $\phi_1, \phi_2, \dots, \phi_p$  are the autoregressive coefficient;  $\delta$  is the constant term;  $u_t$  is the residual sequence;  $\theta_1, \theta_2, \dots, \theta_q$  are the moving average coefficient; the model has three parameters that

need to be artificially set, respectively, the autoregressive order Number ( $p$ ), number of differential smoothings ( $d$ ), moving average number ( $q$ ). The model is usually recorded as ARIMA( $p,d,q$ ).

The modeling steps mainly include data preprocessing, pattern recognition, parameter identification and model checking. For details, see the literature [6].

### B. LSTM Prediction Model

The basic theory of LSTM is described in [7]. Here, a simple network model consisting of two memory modules is used as an example to introduce the working principle of the LSTM neural network. The neural network model is expanded, and the internal detailed structure is as shown in FIG. 1. Like traditional neural networks, the LSTM neural network also includes an input layer, an implicit layer, and an output layer. The input layer consists of four nodes, the hidden layer consists of two memory modules, and the output layer consists of five nodes. The data flow of the neural network is: First, the data input to the  $x_t$  layer flows to the input gate, the forgetting gate, the output gate, and the cell input of the memory block in the hidden layer; the input data can be entered into the module after entering the hidden layer module A. The activation function and the computational component process; after that, the data has three flows: the flow of the first data is the input gate, the forgetting gate, the output gate, and the cell input for transmitting the data to the self module; the flow of the second information is Pass the data to module B's input gate, forgetting gate, output gate, and cell input; the last is to pass the data to output layer  $y_{t+1}$ . The above is the propagation process of one input data in the LSTM neural network.

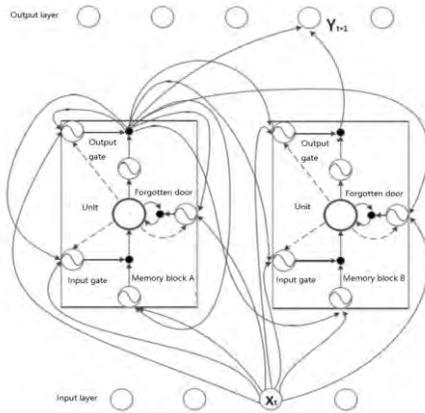


FIGURE I. LSTM NETWORK STRUCTURE OF THE MEMORY MODULE EXPANSION

## III. HYBRID MODELS AND THEIR IMPROVEMENTS

The mixed model [8-9] is a combination of different models in some way. It can synthesize the characteristics of each model, such as a hybrid model composed of linear model and nonlinear model. The precision of modeling and prediction of complex time series is higher than that of single model.

### A. Traditional Model Mixing Method

The construction of traditional hybrid models is mostly based on the idea that a complex time series consists of two parts, a linear component and a nonlinear component, with a simple linear summation between the two parts. Therefore, in the prediction, the time series of the research object can be decomposed first, and then different models obtained by the decomposition are predicted by different models. The final result is the superposition of the predicted values of different models.

The hypothetical time series  $y_t$  column based on the traditional hybrid approach can be expressed as follows:

$$y_t = L_t + N_t \quad (2)$$

$L_t$  represents a linear component in the time series;  $N_t$  represents a nonlinear component in the time series.

Many studies have used the following methods to construct a hybrid model.

1) *Modeling the time series using a linear model to obtain predicted values of linear components.* The residuals will contain nonlinear components of the time series. The mathematical expression is as follows:

$$e_t = y_t - \hat{L}_t \quad (3)$$

$\hat{L}_t$  represents the predicted value;  $e_t$  represents the residual, which contains the nonlinear components of the time series.

2) *Modeling the  $e_t$  sequence using a nonlinear model to obtain the predicted value of the nonlinear component.* The mathematical expression of the model established by the residual is as follows:

$$\hat{N}_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}) + \varepsilon_t \quad (4)$$

$f$  represents a relational function established by a nonlinear model, and  $\varepsilon_t$  represents a random error.

3) *Finally, add  $L_t$  and  $\hat{N}_t$  to get the final result. which is*

$$\hat{y} = \hat{L}_t + \hat{N}_t \quad (5)$$

### B. Units Improvements to the Traditional Model Hybrid Approach

If the linear and nonlinear components in the time series are not simply additive relationships but other more complex relationships such as product relationships, the use of traditional model blending will reduce the prediction accuracy by

simplifying this relationship. Even the predicted accuracy will be lower than a single model.

This paper assumes that an unknown function can be used to represent the relationship between linear and nonlinear components in a time series, mathematically expressed as follows:

$$y_t = f(L_t, N_t) \quad (6)$$

Determined by the nonlinear model, represented by the LSTM model in this paper.

In this paper, based on the formula (6), the method of establishing autoregressive moving average [10] and long-short and short-term neural network combined forecasting model is as follows:

Firstly, an autoregressive moving average model is established, in order to process the linear components contained in the sequence through the model. The residuals obtained from the original time series processed by the linear model are bound to contain nonlinear correlations that cannot be extracted from the sequence using the autoregressive moving average model.

$$L_t = \left[ \sum_{i=1}^p \phi_i z_{t-i} - \sum_{j=1}^q \theta_j \varepsilon_{t-j} \right] + e_t = \hat{L}_t + e_t \quad (7)$$

Where  $\hat{L}_t$  is the sequence value of the ARIMA model at time  $t$ ,  $z_t$  is the differential stationary sequence, and  $e_t$  is the residual of time  $t$ .

When there is a significant nonlinear correlation in the residual, the autoregressive moving average model cannot handle it. The residuals in the ARIMA model will be used in the second step of building a hybrid model.

The second step, the main goal is nonlinear modeling. Long- and short-term neural networks (LSTM models) are used to model the nonlinear relationships in the residuals.

$$N_t = f_1(e_{t-1}, \dots, e_{t-n}) \quad (8)$$

In the formula,  $f_1$  is a nonlinear function determined by long- and short-term neural networks.

Finally, an LSTM model is used to fit the relationship between linear component  $L_t$  and nonlinear component  $N_t$ . The mathematical expression of the final hybrid prediction model is as follows:

$$y_t = f(L_t, N_t) \quad (9)$$

We can find that the way of constructing hybrid model in this paper is quite different from that of traditional model in comparison (9) and (5). In this paper, we abandon the hypothesis

that a sequence is a linear sum and a nonlinear component is a linear sum in the latter theory. Instead, we use a nonlinear model with self-learning ability, (LSTM), to fit the relationship between them objectively. That is to say, (9) can describe the relationship between different components of time series more than formula (5). Therefore, it should be more accurate than traditional methods to construct hybrid model by applying this hybrid idea to forecast field theoretically.

#### IV. MODEL VERIFICATION AND COMPARISON

##### A. Experimental Data

This paper adopts a total of 1092 days of historical product sales data of a certain large-scale mechanical transmission parts manufacturing company from December 17, 2014 to December 12, 2017, as the data research object of this paper. At the same time, in order to avoid the impact of special data on the conclusion of the experiment, the 1092 historical product sales data is divided into two samples for empirical analysis. Sample 1 selects 546 days of product sales observations from December 17, 2014 to June 14, 2016; sample 2 selects 546 days of product sales observations from June 15, 2016 to December 12, 2017. In this paper, the forecast period of both samples is set to 10 days, so the product sales data of the last ten days is used as test data. In the experimental part, the test data will be used to calculate the prediction error and verify the performance of the model.

##### B. Experimental Plan

The experiment uses the method of comparative analysis to verify the correctness of the algorithm. A total of four experiments are carried out as follows:

###### Traditional ARIMA Model

ARIMA is one of the components of the hybrid model constructed in this paper, which is used to fit the linear properties of the sequence. The reason why it is used as a separate model for sales forecasting experiments [11] is to compare the advantages and disadvantages of single and mixed models. The experimental demonstration method in this paper is to use the single group that constitutes the hybrid model to conduct the modeling prediction experiment, and then compare it with the prediction effect of the hybrid model to demonstrate the advantages of the hybrid model. Finally, based on different ideas, the hybrid model is built to predict the experiment. In this way, the mixed model prediction effect composed of different mixing modes under the condition of the same constituent component is compared.

###### LSTM Neural Network Model

LSTM is another component of the hybrid model constructed in this paper, which is used to fit the nonlinear properties of the sequence. Similarly, building it into a single model for predictive experiments is also a demonstration of the need to build a hybrid model. Comparing the prediction results of a single LSTM model with a hybrid model composed of ARIMA and LSTM, the advantages of the hybrid model over a single model and the necessity of its construction are demonstrated.

###### Hybrid Model

The third group and the fourth group of experiments used two same components to form a mixed model for sales prediction. The difference is that the mixed model of the two groups of experiments is constructed in different ways. The purpose of setting up these two groups of experiments is twofold: first, comparing the prediction effect of the first two groups of single models to prove the necessity of building a mixed model; second, The prediction results of the two hybrid models are compared between the traditional construction method and the improved construction method proposed in this paper.

V. ANALYSIS OF EXPERIMENTAL RESULTS

The four curves in Figures 2 and 3 are the error curves of the four experimental models under the ten-day prediction period. According to the error graph, it can be seen that the error value of the ARIMA model changes the most with time, and the prediction performance is the most unstable. Compared with the ARIMA model, the LSTM model has a small early error and is relatively stable, but the late fluctuations also start to increase. The reference value is lost because the error is too large. The late model error of the hybrid model constructed based on the traditional method is more stable than the single model. The ARIMA-LSTM model based on this method is the most ideal, the trend of error curve is the most stable, and the relative error is the smallest. The hybrid model constructed in this paper has better prediction performance than the traditional hybrid model.

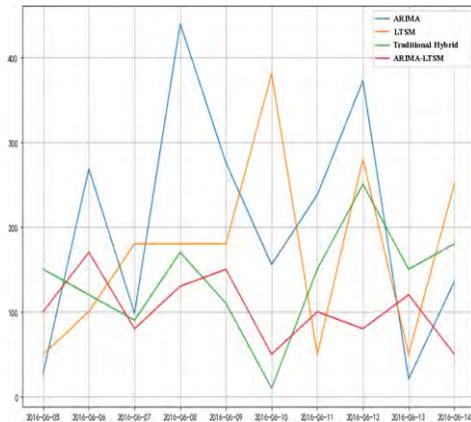


FIGURE II. COMPARISON OF FOUR MODEL ERRORS (SAMPLE 1)

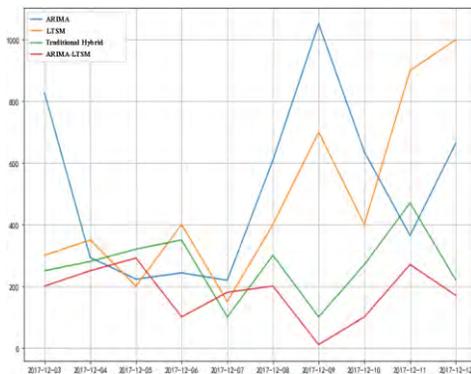


FIGURE III. COMPARISON OF FOUR MODEL ERRORS (SAMPLE 2)

The mean absolute error (MAE) and the mean absolute percentage error (MAPE) were used as evaluation indicators to make a qualitative evaluation of the prediction effects of different models. See Table 1 and Table 2 for the sales forecast performance indicators.

TABLE I. EVALUATION INDICATORS FOR PREDICTIVE EFFECTS (SAMPLE 1)

	MAE	MAPE
ARIMA-LSTM hybrid model	103.0	0.1746
Traditional hybrid model	138.0	0.2668
ARIMA	203.0	0.3354
LSTM	170.0	0.3107

TABLE II. FORECAST PERFORMANCE EVALUATION INDICATORS (SAMPLE 2)

	MAE	MAPE
ARIMA-LSTM hybrid model	177	0.4787
Traditional hybrid model	266	0.9311
ARIMA	512	3.3294
LSTM	480	2.8756

Based on the experimental results of the four groups of models, the hybrid model proposed in this paper is the best in predicting the performance of the prediction results.

VI. SUMMARY

Since ARIMA is a rigorous linear prediction method, most product sales time series are a combination of linear and nonlinear, so the application is often not accurate. In order to overcome this shortcoming, it must be combined with the nonlinear prediction method, which not only retains the linear prediction ability of the model itself but also compensates for its lack of nonlinear prediction. Therefore, hybrid models are already a mainstream approach to time series prediction.

The traditional model mixing method assumes that the linear component and the nonlinear component of the time series are simple additive relationships. This simplification has certain risks and may lead to a decrease in prediction accuracy. Therefore, this paper abandons the hypothesis of the traditional model hybrid approach, without presupposing the relationship between the two, but using the self-learning ability of the LSTM neural network to describe the complex relationship between the two. On this basis, this paper uses the historical data of large-scale mechanical transmission parts manufacturing enterprises to verify the model. In this paper, four sets of model experiments are set up, namely ARIMA model, LSTM model, ARIMA-LSTM model constructed by traditional methods, and ARIM-LSTM model after improved hybrid method. Experiments show that the hybrid model can compensate for the shortcomings of the single model to a certain extent to improve the prediction accuracy. In particular, the ARIM-LSTM after the hybrid method has certain advantages over the single and traditional hybrid ARIM-LSTM models.

## ACKNOWLEDGEMENTS

The research work of this paper was supported by Sichuan Science & Technology Program under Grant No.2017GZ0144 & 2017GZ0047.

## REFERENCES

- [1] Reyes J, Morales-Esteban A, Martínez-Álvarez F. Neural Networks to Predict Earthquakes in Chile[J]. *Applied Soft Computing*, 2013, 13(2): 1314-1328.
- [2] Singhal D, Swarup K S. Electricity Price Forecasting Using Artificial Neural Networks[J]. *International Journal of Electrical Power & Energy Systems*, 2011, 33(3): 550-555.
- [3] Dong Z, Yang D, Reindl T, et al. Satellite Image Analysis and a Hybrid ESSS/ANN Model to Forecast Solar Irradiance in the Tropics[J]. *Energy Conversion and Management*, 2014(79): 66-73.
- [4] Zhang G P. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model[J]. *Neurocomputing*, 2003(50): 159-175.
- [5] Faruk D Ö. A Hybrid Neural Network and ARIMA Model for Water Quality Time Series Prediction[J]. *Engineering Applications of Artificial Intelligence*, 2010, 23(4): 586-594.
- [6] Cheng Jun. Research and application of sales forecasting method for mechanical transmission parts manufacturing enterprises based on ARIMA-LSTM hybrid model [D]. University of Electronic Science and Technology of China, 2018.
- [7] Graves A. Long Short-Term Memory[J]. *Neural Computation*, 1997, 9(8):1735-80.
- [8] Yue Jiguang, Yang Yuming, Sun Qiang et al. Hybrid prediction model for interval time series[J]. *Control and Decision*, 2013, 28(12): 1915-1920.
- [9] Zhang Wei. Application of Mixed Model in Economic Time Series Prediction[J]. *Computer Simulation*, 2011, 28(6): 354-356.
- [10] Zhang Lijie, Qi Jizhen, Li Minqiang et al. China Cotton Price Forecast Based on Autoregressive Moving Average and Support Vector Machine[J]. *Statistics and Decision*, 2013(6): 30-33.
- [11] Xiang Yun, Hou Ting, Li Zhendong. Application of ARIMA Model in GDP Forecast of Yunnan Province[J]. *Time Financial*, 2014(3): 87-91.
- [12] Zhang Min. Research on combined model problem based on ARIMA [D]. Dalian Maritime University, 2015.