ISSN 2616-5775 Vol. 6, Issue 8: 131-137, DOI: 10.25236/AJCIS.2023.060816

Research on material demand analysis of manufacturing industry based on time series model—ARIMA

Changsheng Li, Sujun Liu, Guoying Yang

College of Information Engineering, Lanzhou Petrochemical University of Vocational Technology, Lanzhou, 730060, China

Abstract: *In order to solve the problem of mismatch between material production plan and actual demand, this paper analyzes and forecasts the material demand of manufacturing industry. Firstly, the material demand frequency at different time points in the historical production data of a manufacturing industry is calculated by statistical method. Secondly, it quantitatively analyzes the change trend of the total sales volume of each material at different time points to the unit price of the material. Thirdly, with the quantity, frequency, total sales and unit price of materials as characteristic factors, through relevant statistical analysis, six kinds of materials are synthesized. Finally, according to the demand data of these six materials, the time characteristics are transformed into weekly characteristics, and the time series model ARIMA is applied to construct the weekly prediction model of material demand, and the performance of the prediction model is evaluated. The results show that the comparison between the forecast results and the actual values of the weekly forecast model passes the test and has a good application prospect.*

Keywords: *Time series model, Prediction Model, Feature selection, Correlation*

1. Introduction

With the deepening reform of manufacturing enterprises, enterprises are facing great opportunities and challenges in manufacturing management and cost control[1] .The current situation is At present, many manufacturing enterprises often face the problem of mismatch between the production plan of materials and their actual demand in the production operation; material demand planning refers to the detailed plan of production in the manufacturing industry based on the subordinate and quantitative relationship of items at each level of the product structure, according to the supply of materials and production demand, with the product completion period as the time base[2]. Therefore, in the process of manufacturing, it is crucial for the operation management of manufacturing enterprises to arrange material production scientifically and rationally, and to develop efficient and feasible material requirement plans. In this paper, we briefly analyze the advantages and shortcomings of material demand planning in manufacturing production, and introduce the time series model in the field of machine learning to build a weekly prediction model of material demand to assist the development of manufacturing production planning, hoping to provide some information reference for the relevant enterprise managers, so as to improve the production efficiency of enterprises.

2. Materials and Methods

2.1 Data sources and analysis

In the production operation of electronic product manufacturing enterprises often face the following problems: in the production of multi-species and small-lot materials, the actual demand for materials cannot be known in advance, resulting in production arrangements that may generate a large inventory, or more out-of-stock phenomenon, bringing economic and reputation losses to the enterprise. Therefore, how to reasonably plan and arrange material production is a hot topic of frequent concern for manufacturing enterprises.

In this paper, the production data of an electronic product manufacturing company provided in Question E of the 2022 Gaoxiong Cup National University Mathematical Modeling Competition are

ISSN 2616-5775 Vol. 6, Issue 8: 131-137, DOI: 10.25236/AJCIS.2023.060816

used for analysis and research. Its sample data table is shown in Table 1.

Table 1: Production data of electronic products

Date	Material Code	Demand	Unit selling price(¥)
2019-1-2 15:13:16	6004020080	1	3720.54
2019-1-2 15:13:16	6004020141	2	227.21
2019-1-2 15:13:16	6004020888	1	2354.07
2019-1-3 09:46:18	6004010255	1	1299.47
2019-1-3 14:55:18	6004010174	7	1387.69
2019-1-3 14:55:18	6004010174	14	1387.69
......	......	......	......

2.2 Data pre-processing

2.2.1 Focus on the selection of materials

Because of the wide variety of product materials in the historical data, this paper selects several key materials that are closely related to enterprise cost control for analysis. First, based on the data in the material code column of the data, the frequency of the demand for a total of 284 materials at each point in time in the historical data was statistically calculated, as shown in Table 2.

Table 2: Material frequency statistics

	Material Code	Frequency
0	6004020503	1224
1	6004010256	955
2	6004020375	794
3	6004020918	620
4	6004020374	612
...	...	...
282	6004020883	1
283	6004021000	1

At the same time, through the statistical analysis of the overall data, it is found that among various materials, the same code of materials at different points in time, their demand quantity and sales unit price are different, and the sales unit price will change, so to consider the trend of material demand, it is necessary to consider the trend of demand quantity and sales unit price change, this paper calculates the total sales of different materials at each point in time to the change of sales unit price of each material. The quantitative analysis is performed by calculating the total sales of different materials at each point in time. The details are shown in Table 3.

Table 3: Total sales statistics of each material at different time points

Date	Material Code	Demand	Unit selling price(¥)	Total sales(¥)
2019-01-02 11:10:19	6004020664	5	3249.15	16245.73
2019-01-02 11:10:20	6004020849	1	2978.15	2978.16
2019-01-02 11:10:20	6004020859	3	2828.26	8484.78
2019-01-02 11:13:19	6004020709	1	16810.05	16810.05
...	...	...	...	...
2022-05-21 16:24:12	6004021112	2	5018.19	10036.38
2022-05-21 16:24:12	6004021141	1	4652.99	4652.99

Finally, using the quantity, frequency, total sales and unit sales price of material demand as the characteristic factors, after correlation statistical analysis, their respective correlations are shown in the Figure 1.

In summary, the four characteristics: quantity, frequency, total sales and sales unit price of material demand for a comprehensive statistical ranking, filtered out 6 key concern materials, and its material code is shown in Table 4.

ISSN 2616-5775 Vol. 6, Issue 8: 131-137, DOI: 10.25236/AJCIS.2023.060816

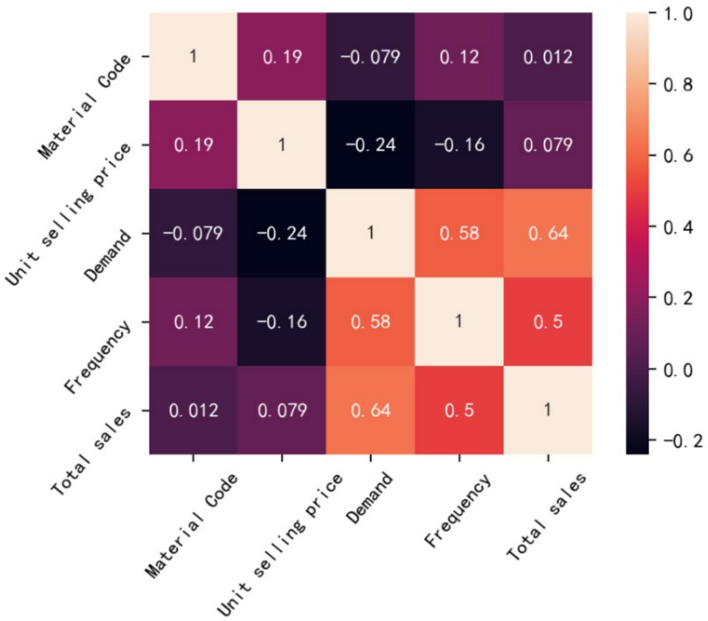

Figure 1: Correlation analysis of material characteristics

Table 4: 6 key materials of concern

Material Code
6004020503
6004020918
6004010372
6004010256
6004010174
6004021055

2.2.2 Constructing the weekly time of material requirements

In building the weekly forecast model of material demand, this paper first calculates the total number of weeks from January 2, 2019 to May 21, 2022 based on the sales date, and corresponds each date to the number of weeks one by one, and the results are shown in the Table 5.

Table 5: Correspondence table of dates and weeks

Date	Number of weeks	Material Code	Demand	Unit selling price(¥)
2019/1/3 14:55	1	6004010174	7	1387.68
2019/1/3 14:55	1	6004010174	14	1387.68
2019/1/3 14:55	1	6004010174	10	1387.68
2019/1/7 10:34	2	6004020918	1	2308.11
2019/1/8 09:22	2	6004020918	2	2325.43
...	...	...	...	...
2022/5/20 13:30	177	6004020503	2	224.33
2022/5/21 14:20	177	6004020503	1	224.33
2022/5/21 14:20	177	6004020503	2	224.33

Then the time series model ARIMA in machine learning algorithm is applied to build a weekly forecasting model of material demand with week as the basic time unit.

3. Model building and solving

3.1 Time series model - ARIMA

The full name of the ARIMA (Autoregressive Integrated Moving Average) model is Differential Integrated Moving Average Autoregressive Model[3-5]. It is also known as the integrated moving autoregressive model. ARIMA *(p,d,q) is* a common type of statistical model used for time series forecasting. The forecasts are expressed as a linear function consisting of the nearest true value and the nearest forecast error.

$$Y_t = C + Y_{t-1} + \cdots + a_p Y_{t-p} + e_t + \beta e_{t-1} + \cdots + \beta_q e_{t-q} \tag{1}$$

where: C is the constant term Y_t is the predicted value at time point t, $\boldsymbol{a}$ is the weighting factor. The three important parameters of the ARIMA model [6,7] : *p, d, q*.

p: Represents the lag of the time series data itself used in the prediction model, also called AR/Auto-Regressive term

d: Represents how many orders of differencing are needed for the timing data to be stable, also called the Integrated term.

q: Represents the number of lags of the prediction error used in the prediction model (lags), also called MA/Moving Average term

The ARIMA (*p,d,q*) model implies that the time series is differenced *d* times and that each observation in the series is represented by a linear combination of the past *p* observations and *q* residuals. The prediction is "error-free" or integrated to achieve the final prediction.

3.2 Construction of 6 focused material demand trend models

In this paper, we designed a Python program to design an ARIMA model using the ARIMA algorithm in the machine learning library Sklearn: 70% of the six focused material demand datasets were then used as the training set of the time-series model, and the other 30% were used as the validation set of the model to construct an ARIMA model for material demand forecasting.

The steps to build an ARIMA model include:

(1) Ensure that the timing is smooth;

(2) Finding a reasonable model (or models) (and selecting possible p- and q-values);

(3) Fitting the model;

(4) Evaluating models from the perspectives of statistical assumptions and prediction accuracy;

(5) Forecast.

Then, the time series model was trained on the historical data of the six selected material requirements of focus, and the date time series and week time series graphs were plotted, as shown in Figure 2.

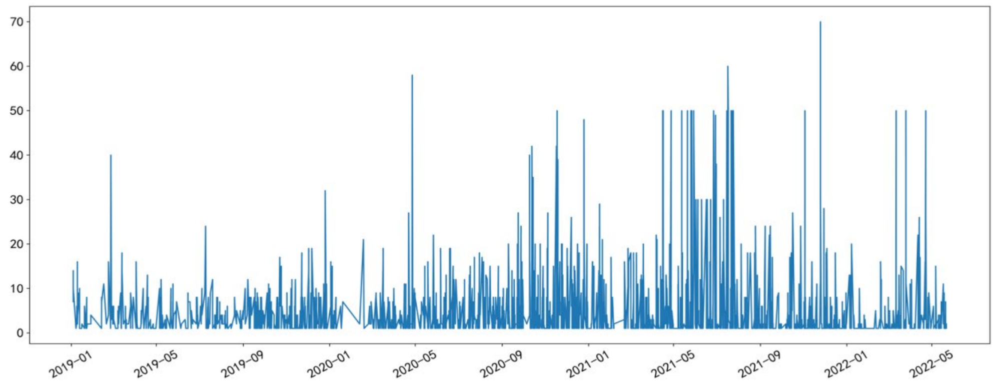

Figure 2: Time-series diagram of the date of the demand for the six key materials of interest

After the transformation of the interval January 2, 2019 to May 21, 2022, date and week by week, respectively, the total demand of each material per week is calculated as the demand of this material for this week, and the transformation of the date and week is carried out, and the weekly time series chart of the demand of the six key concern materials is shown in Figure 3.

ISSN 2616-5775 Vol. 6, Issue 8: 131-137, DOI: 10.25236/AJCIS.2023.060816

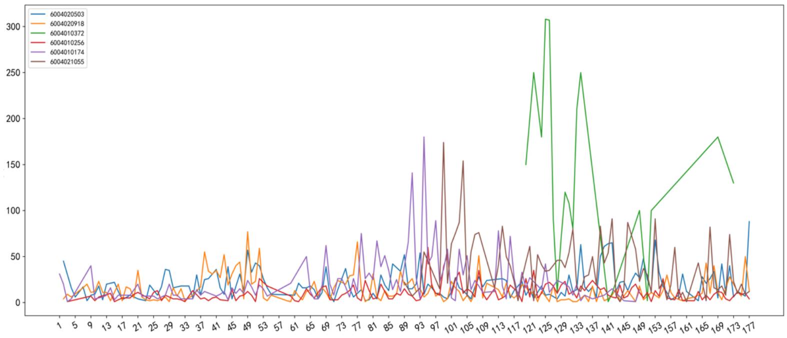

Figure 3: Weekly time series chart of 6 key material requirements

The analysis of the weekly time series chart of the demand of the six key materials[8] reveals that among the six materials, the demand of the material coded “6004010372” was only available from the 120th week; the demand of the material “6004021055” was only available from the 89th week, and there was no demand data of these two materials in the previous period; while the demand of the materials coded “6004020918”, “6004010256” showed a relatively stable trend throughout the observation time from week 1 to week 177. The demand for the material coded “6004010174” and “6004020503” was more volatile, especially around the 91st to 105th weeks showed a more obvious trend of fluctuation.

4. Parameter optimization and model validation of ARIMA model

4.1 Parameter optimization of ARIMA model

To improve the forecasting performance of the model, the *p, q* and *d* parameters in the model are determined separately. For the parameter optimization of the ARIMA time series model, this paper finds that it is not a smooth time series through the weekly time series plot of demand. So it needs to be further processed. Using ARIMA model to transform the non-stationary series into a stationary series[9,10] , through the analysis of the autocorrelation coefficient ACF of the demand for this material, as follows in Figure 4.

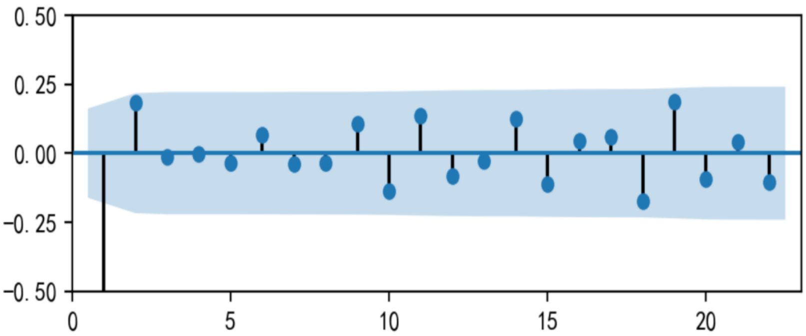

Figure 4: ACF analysis of autocorrelation coefficient

Table 6: ARIMA (0, 1, 2)

Dep. Variable.	D2.Demand		No. Observations.			97
Model.	ARIMA(0, 1, 2)		Log Likelihood			-386.118
Method.	css-mle		S.D. of innovations			12.427
Date.	Sun, 18 Sep 2022		AIC			780.237
Time.	00:23:42		BIC			790.536
Sample.	2		HQIC			784.401
	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0168	0.077	0.217	0.828	-0.135	0.169
ma.L1.D. Demand	-0.8501	0.095	-8.986	0	-1.035	-0.665
ma.L2.D. Demand	-0.0503	0.097	-0.521	0.603	-0.24	0.139

For a time series, the correlation between the present value and its past value is determined: if the correlation is positive, it indicates that the existing trend will continue, as well as the partial autocorrelation coefficient PACF analysis finds that it its data after the 1st order is already in the confidence interval within the confidence interval [11] and the value of d can be determined as 1.

Then, the lag p value of the time series data itself and the q value after the prediction error were optimized by the grid search algorithm to determine the optimal parameters of the time series model ARIMA(0, 1, 2) .

The details of the final model constructed are shown in Table 6.

4.2 Model Testing

In order to verify the feasibility and validity of the ARIMA model constructed in this paper for weekly demand forecasting of manufacturing materials, the data set of the 110th to 177th weeks in the original material demand data set is used as the test data for the ARIMA model for weekly demand forecasting of materials in this paper. The prediction is carried out by ARIMA (0, 1, 2) model and the actual values of model prediction and original data are compared and analyzed as follows Figure 5.

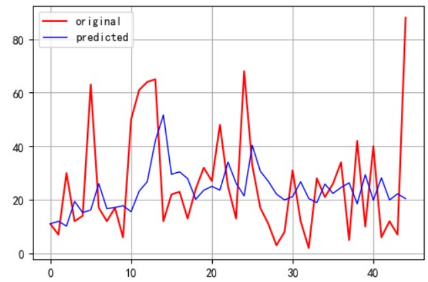

Figure 5: Comparison analysis between predicted and actual values

The comparison between the predicted and actual values in the prediction results revealed that there was a certain error between the predicted and the true results. In order to further analyze the performance of the ARIMA(0, 1, 2) model, after the ARIMA(0, 1, 2) model is used for prediction, the residual sequence is calculated by calculating the difference between the true value and the predicted value, and the residual plot and kernel density map of the model are drawn as Figure 6 and Figure 7.

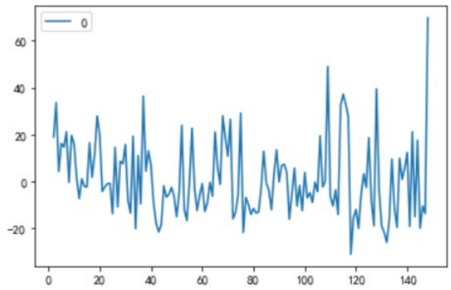

Figure 6: Residual analysis of model predicted values

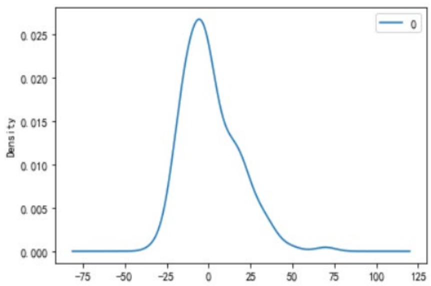

Figure 7: Kernel density analysis of model predicted values

The analysis of the residual plot and kernel density plot of the model shows that the residual error of the ARIMA(0, 1, 2) model fluctuates around the zero mean and has uniform variance; meanwhile, its kernel density plot also shows a standard normal distribution trend, which indicates that the residuals of the model are not correlated and the model prediction tends to the zero mean normal

ISSN 2616-5775 Vol. 6, Issue 8: 131-137, DOI: 10.25236/AJCIS.2023.060816

distribution trend, so the time-series model ARIMA(0, 1, 2) established in this paper is fully feasible.

5. Conclusion

In this paper, to address the problem of mismatch between production plan and its actual demand, a time series model ARIMA in the field of machine learning is introduced to construct a weekly forecasting model of manufacturing material demand to assist in the formulation of material demand plan in manufacturing production. In the model prediction, in order to improve the accuracy of the constructed weekly prediction model, this paper firstly analyzed and checked the parameters of ARIMA model by autocorrelation coefficient ACF and partial autocorrelation coefficient PACF, then optimized three important parameters by grid search algorithm, and determined the optimal parameter set of weekly prediction model ARIMA(0, 1, 2); then, by the residual plot of the model, the Then, the stability and robustness of the model were verified by further comparison and analysis on historical data and predicted data through the residual plots and kernel density plots of the model; finally, it was determined that the weekly material demand forecasting model constructed in this paper has stable performance, accurate prediction and strong generalization ability, and has good prospects for promotion and application in manufacturing material production.

Acknowledgements

Fundamental Project: Textbook Research Project of Lanzhou University of Petrochemical Technology (Project No. KJ2020-2), Gansu Provincial Education Science "Fourteenth Five-Year Plan" 2022 Special Project on School Safety, Stability and Emergency Work (Project No. GS[2022] GHBZX239).

References

[1] Hu Xinbing. An introduction to cost control in manufacturing enterprises [J]. China Township Enterprise Accounting, 2020(02):121-122.

[2] Chang Fang, Xinbao Liu, Jun Pei, et al. Optimal production planning in a hybrid manufacturing and recovering system based on the internet of things with closed loop supply chains[J].Operational Research, 2016.DOI:10.1007/s12351-015-0213-x.

[3] Li Changsheng, Yang Guoying, Liu Zongcheng. Research on course data analysis based on time series model and clustering algorithm [J]. Information and Computer (Theory Edition), 2021, 33(14): 79-82.

[4] Fan Yingbing, Lu Bin. Research on China's gold reserve forecast based on time series model [J]. Software Guide, 2021, 20(04):96-100.

[5] Shan R, Wang SHH, Gao DFL, Gao JH. Research on combined prediction model based on time series model and gray model [J]. Journal of Yanshan University, 2012, 36(01):79-83.

[6] Lin Z. Analysis Research on Joint Development of Regional Manufacturing and the Logistics Industry——Based on Coupling System of Supply and Demand[J]. Technoeconomics & Management Research, 2016.

[7] Zhang Wanlin. Forecasting analysis of Ningbo port logistics demand based on GM(1, 1)-ARIMA combined model[J]. Logistics Engineering and Management, 2023, 45(01):32-35.

[8] Guo Shunsheng, Wang Lei, Huang Kun. Research on predicting automobile sales based on time series model [J]. Mechanical Engineer, 2013(05):8-10.

[9] Lan H, Liao C-M, Zhao Y. Power output prediction of photovoltaic power plants based on ARMA model [J]. Electrical Measurement and Instrumentation, 2011, 48(02):31-35.

[10] Bolotova Y V, Novakovic A M. An Empirical Analysis of Pricing in the United States Cheese Industry: 2000-2010[J].Ssrn Electronic Journal, 2011.DOI:10.2139/ssrn.1865314.

[11] Yao JH, Zou JJ. Construction and numerical simulation of SVM-ARIMA model for CPI forecasting [J]. Statistics and Decision Making, 2022, 38(21):48-52.