Forecasting the Demand of Parts in an Assembly Plant Warehouse Using Time-Series Models

G. Rivera, R. Florencia-Juárez, J. P. Sánchez-Solís, V. García, and C. D. Luna

***Abstract*—Knowing the demand for products in advance would be ideal for companies that strategically relocate products in their warehouses to facilitate the picking process, which is the most expensive activity in a warehouse. In this sense, an assembly plant from Ciudad Juárez, Chihuahua, Mexico, that handles approximately 10,477 parts in its inventory, periodically relocates them in warehouse zones to facilitate the picking process. However, relocation is done empirically based on the total number of outbound inventory movements of each of the parts made in a given time. This paper describes the implementation of time-series models to forecast the demand for parts that could improve the relocation process. For this purpose, different Holt-Winters Seasonal and SARIMA models were implemented. For the implementation of the SARIMA models, the Box-Jenkins methodology was followed. The AIC and BIC metrics were used to identify the best Holt-Winters Seasonal model and the best SARIMA model. Tests were performed on the residual series to check that model is fit to the data. The RMSE and MAPE metrics were used to evaluate the performance of Holt-Winters Seasonal and SARIMA models. The results of the evaluation carried out indicate that the SARIMA model outperforms to Holt-Winters Seasonal model.**

***Index Terms*—Forecast demand, Holt-Winters Seasonal model, Seasonal Auto-Regressive Integrated Moving Average model, Time Series, Box-Jenkins methodology.**

1. Introduction

CIUDAD Juárez is a city whose economy is strongly based on the assembly plant industry. It has the largest number of assembly plants operating under the IMMEX (*Industria Manufacturera, Maquiladora y de Servicios de Exportación*) scheme in the entire state of Chihuahua with 330 assembly plants (65%), followed by the city of Chihuahua with 109 (22%) and the other municipalities with 66 (13%). At the national level, it is the second city with 330 assembly plants under this scheme. Among the different sectors, the Automotive, the Electronics, the Medical, the Plastics Metals, the Call Center, and the Packaging predominate [1].

In an assembly plant of this city dedicated to the manufacture of smoke detectors, fire alarms, among other products, the raw material used is stored in a warehouse within the same plant.

Currently, the warehouse is divided into four zones: A, B, C, and D. These zones are strategically organized to minimize the time for picking parts. The parts with the highest demand are stored in bins located in zone A, close to the production area. On the contrary, the parts with the lowest demand, either because they are obsolete or because they were never used, are located in zone D. It is important to note that, based on the parts demand, these are periodically relocated to the warehouse zones.

The criteria used by the assembly plant to locate each of the parts in the warehouse zones is based on the cumulative distribution of the demand they present. The demand is defined according to the outbound inventory movement of each of the parts. The parts that represent a cumulative distribution of 70% of the demand are located in zone A. The parts that are located in zone B represent 25% of the demand. In zone C are located the parts that represent 5% of the demand. Finally, in zone D, the parts that did not have movement are located.

Although the parts with the highest demand are identified in this empirical way, the trend of movements is not considered when assigning their location in the warehouse. For example, following this criterion, the part with the highest number of movements should be located in zone A. However, if the trend of its movements shows a considerable decrease in recent months, perhaps that part should not be located in zone A but zone B.

Failure to properly locate the parts in warehouse zones that minimize the picking time could cause the delivery time of these to the production area to be long; even the production lines could be stopped. For this reason, knowing the future demand of the parts could help improve the relocation process.

In this sense, time series models have been successfully applied in various sectors of the supply chain to forecast demand. In [2], different ARIMA models were implemented following the Box-Jenkins methodology to model and forecast the demand in a food company. Using the Akaike, Schwarz Bayesian, maximum likelihood, and standard error criteria, they identified the best model was an ARIMA (1,0,1). The authors note that forecasting future demand affects the supply chain and provides reliable guidelines for decision making.

In [3], different time series models applied to forecasting

Manuscript received on October 10, 2020, accepted for publication on November 22, 2020, published on December 30, 2020.

G. Rivera, R. Florencia-Juárez, J. P. Sánchez-Solís, and V. García are with the Department of Electrical and Computer Engineering, Autonomous University of Cd. Juárez, Cd. Juárez, 32310, Mexico (e-mail: {gilberto.rivera, rogelio.florencia, julia.sanchez, vicente.jimenez}@uacj.mx).

C. D. Luna is with the Applied Computer Science Graduate Program at Autonomous University of Cd. Juárez, Cd. Juárez, 32310, Mexico (e-mail: al199185@alumnos.uacj.mx).

The records in this database are made up of the attributes: *material name*, *description*, *type of movement*, *plant*, *location*, *quantity*, and *PostingDate*. Of these attributes, PostingDate was chosen as a period for the time-series model because it represents the date on which the movements were made. Additionally, outbound inventory movements of each part were counted to represent the value of the independent variable.

Due to the way the warehouse operates, the count of outbound inventory movements was grouped monthly to form the periods for each of the parts. The monthly movements of each of the parts were stored in different .csv files. It is worth mentioning that so far, the time-series models described in this paper have only been implemented in the part with the highest demand in the historical data. Figure 1 shows the histogram of movements of the part with the highest demand in the historical data.

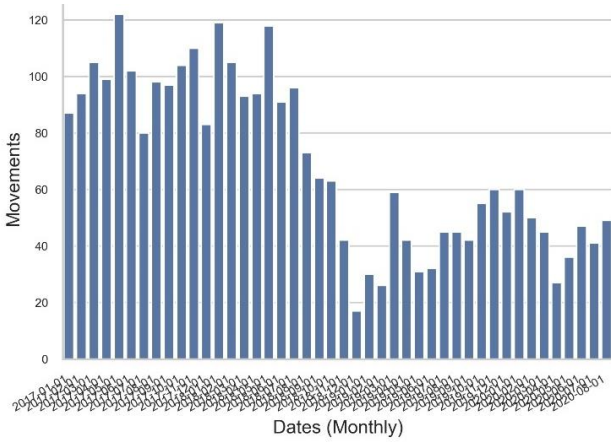


Fig. 1. Histogram of the part with the highest demand in the historical data.

4. DATA ANALYSIS

Before selecting the time-series model, it was determined whether the time series data exhibits trend and seasonal effects. To determine if the data exhibits trend and seasonal effects, the following analyzes were performed: (a) empirical visual analysis and (b) statistical tests.

4.1 Empirical visual analysis

Figure 2 shows the time series corresponding to the part with the highest number of movements in the warehouse inventory. At first glance, it is difficult to determine whether the data is seasonal.

In order to visually perform an exploratory analysis of the time series, it was decomposed into the trend (T_t), seasonality (S_t), and residual (E_t) components, shown in Figures 3 to 5, respectively. This decomposition allows better identification of the underlying patterns of the time series. A classical additive decomposition was used because the seasonal variation has the same magnitude over time where each value (Y_t) of the time series is the sum of these three components, as shown in Equation (1):

$$Y_t = T_t + S_t + E_t \quad (1)$$

In Figure 3 is observed that the trend increases or decreases based on time; that is, the mean changes over time. In Figure 4, a seasonal effect is perceived; that is, a repeated pattern is shown. Based on the above, it is suspected that the series could be non-stationary. Also, the residuals displayed in Figure 5 do not show a stable pattern and are not close to 0.

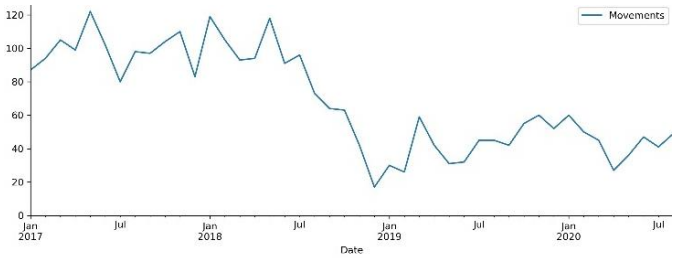


Fig. 2. Time series of the part with the highest number of movements.

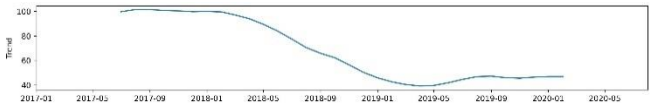


Fig. 3. Trend component.

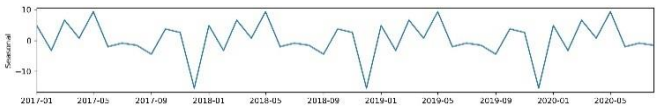


Fig. 4. Seasonal component.

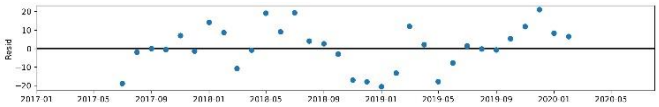


Fig. 5. Residual component.

4.2 Statistical tests to determine seasonality

A non-stationary time series shows seasonal effects and trends. If a time series is non-stationary, it means that the mean and the variance are not constant; that is, they change over time. Some forecasting models are susceptible to non-stationary time series. For this reason, before applying a time-series model to a series, it is necessary to know if the time series is non-stationary.

To reinforce the empirical analysis presented in Section 4.1, the Augmented Dickey-Fuller (ADF) test [9] was used to determine if a series is non-stationary. Additionally, the Kwai-towski-Phillips-Schmidt-Shin (KPSS) test [10] was also applied to determine if the series is stationary.

ADF is a statistical test known as the unit root test. The purpose of the test is to determine the strength with which a time series is defined by a trend. The null hypothesis of the

test is that there is a unit root in the time series, that is, that the series is not stationary. The alternative hypothesis of the test is that there is no unit root in the time series, that is, that the series is stationary.

If the *p-value* ≥ 0.05, the test fails to reject the null hypothesis (the series has a unit root, that is, it is non-stationary). If the *p-value* < 0.05, the test rejects the null hypothesis (the series does not have a unit root, that is, it is stationary).

After applying the ADF test, the computed *p-value* was 0.650, so fail to reject the null hypothesis, and it is confirmed that the time series is non-stationary.

On the other hand, the KPSS test is a statistical test used to analyze the stationarity of a series based on a deterministic trend. The null hypothesis of the test is that the time series is stationary. The alternative hypothesis of the test is that the time series is not stationary.

If the *p-value* ≥ 0.05, fail to reject the null hypothesis (the series is stationary). If the *p-value* < 0.05, reject the null hypothesis (the series is non-stationary).

After applying the KPSS test, the computed *p-value* was 0.069, so it fails to reject the null hypothesis, and it is confirmed that the time series is stationary.

Because the ADF test confirms that the series is non-stationary and the KPSS test indicates that it is stationary, the series had to be transformed as described in Section 5.2.

5. Time series model selection

Because the time series exhibits a pattern in the seasonal component, shown in Figure 4, forecasting models that consider seasonality were considered. The models considered were *Holt-Winters Seasonal* and *Seasonal ARIMA* (SARIMA). These models were implemented through the *Statsmodels* library in the Python language. Statsmodels is a suite of high-level statistical models that allows for statistical tests and data exploration [11].

To determine the best forecasting model, the metrics commonly used in classical time series, *Akaike Information Criteria* (AIC) [12] and *Bayesian Information Criteria* (BIC) [13], were considered.

5.1 Holt-Winters Seasonal model

The Holt-Winters model [14] that extends the Holt model is used to predict time series data that contain trends and seasonal components. It contains three smoothing equations for the: *level* of the series, *trend*, and *seasonal* components.

There are two versions of this model described in [7], [15], [16], whose difference lies in how the seasonality is modeled, *additive* and *multiplicative* versions. The additive version is suitable when variations are independent of the level and the multiplicative version when seasonal variations are changing proportionally to the level of the series.

Analyzing the seasonal fluctuations shown in Figure 2, it was not observed that these increase in magnitude with the level of the series. For this reason, the additive version was used. Equation (2) is applied to smooth the level of the series. Equation (3) to smooth the trend. Equation (4) to smooth the seasonal components. Equation (5) to forecast the next h values of the time series.

$$L_t = \alpha(Y_t - S_{t-M}) + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (2)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (3)$$

$$S_t = \gamma(Y_t - L_t) + (1-\gamma)S_{t-M} \quad (4)$$

$$Y_{t+h} = L_t + (h)T_t + S_{t-M-h} \quad (5)$$

where:

Y_{t+1} is the forecast value at time $t+1$.
L_t is the level at time t.
T_t is the trend value at time t.
S_t is the seasonal value at time t.
α, β, γ are smoothing parameters.
h is the forecast horizon.
M is the number of seasonal periods;

$0 \leq \alpha \leq 1; 0 \leq \beta \leq 1; 0 \leq \gamma \leq 1.$

Different combinations of values in the *smoothing_level* (SL), *smoothing_slope* (SSL), *smoothing_seasonal* (SSE), and *seasonal_period* (SP) parameters were tested using brute force to identify the best model. The values used in each parameter were 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1. The time used in this experimentation was 343.89 seconds using a computer with an Intel Corte i7-8650U @ 1.90GHz 2.11GHz processor, 16 GB RAM, and Windows 10 pro 64-bits operating system. Table 1 shows the three best models that obtained the lowest values in the AIC and BIC metrics. Of these, the best model was Holt-Winters Seasonal $(0.8, 0.1, 0.0)_{12}$ with AIC=212.880 and BIC=238.217 whose parameters were SL=0.8, SSL=0.1, SSE=0.0 and SP=12.

TABLE 1
Best Holt-Winters Seasonal Additive models

Model $(SL,SSL,SSE)_{SP}$	AIC	BIC
$(0.8, 0.1, 0.0)_{12}$	**212.880**	**238.217**
$(0.9, 0.1, 0.0)_{12}$	212.899	238.235
$(0.7, 0.2, 0.0)_{12}$	212.968	238.304

5.2 ARIMA models

In addition to smoothing techniques, ARIMA models were also used.

ARIMA (Auto-Regressive Integrated Moving Average) models are formed by three components: *AR* for the auto-regressive component, *I* for the integrated component, and *MA* for the moving average component. It is usually expressed with the notation: ARIMA (p, d, q) where p represents the number of correlated lags to be included in the AR component, d the number of times that the raw observations were differenced, and q the size of the moving average window [8]. The parameters p, d, and q are non-negative integers that

indicate the order of the different components of the model. The AR (p) and MA (q) components are predictors that explain the autocorrelation, and the I component indicates the order of differentiation that has been applied to the time series to leave the series stationary, since before including AR or MA terms, the series must be stationary.

SARIMA models are a modification of ARIMA to support a seasonal time series. They are usually expressed as ARIMA (p, d, q)×(P, D, Q)s where p, d, and q are the non-seasonal parameters, P, D, Q the seasonal component of the time series, and s is the seasonal periodicity. SARIMA models are defined from Equations (6) to (10).

$$\varphi_p(B)\phi_P(B^s)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (6)$$

$$\varphi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (7)$$

$$\varphi_p(B^s) = 1 - \phi_s B^s - \phi_{2s} B^{2s} - \cdots - \phi_{Ps} B^{Ps} \quad (8)$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (9)$$

$$\Theta_Q(B^s) = 1 - \Theta_s B^s - \Theta_{2s} B^{2s} - \cdots - \Theta_{Qs} B^{Qs} \quad (10)$$

where:

$\varphi_p(B)$: Polynomial of order p that represents the non-seasonal autoregressive component.

$\phi_P(B^s)^d$: Polynomial of order P that represents the seasonal autoregressive component.

$\theta_q(B)$: Polynomial of order q that represents the non-seasonal moving average component.

$\Theta_Q(B^s)$: Polynomial of order Q that represents the seasonal moving average component.

ε_t: An independently distributed random variable (white noise).

$(1-B^s)^D$: D[th] seasonal difference of season s.

To implement SARIMA models, the four steps of the Box-Jenkins methodology described in [17] were followed.

***Step 1*. Stationarity of the data**

Because the ADF test indicated that the series is not stationary and the KPSS test indicated the opposite, the Box-Cox transformation was applied to stabilize the variance of the time series. Subsequently, the series was differentiated to eliminate the trend and seasonality, resulting in a stationary time series, shown in Figure 6. Compared with Figure 2, the variance decreased, and the trend is constant; however, irregular fluctuations are observed between the months of November 2018 and March 2019.

When both tests were reapplied, the *p-value* of the ADF test was 0.000, so the null hypothesis was rejected, indicating that the series is stationary. The *p-value* of the KPSS test was 0.100, so the null hypothesis was not rejected, indicating that the series is stationary.

***Step 2*. Identification of the model**

Subsequently, the correlogram of the series was generated to obtain the autocorrelation function (ACF) and the partial autocorrelation function (PACF). ACF and PACF plots are used to determine the appropriate values of p and q and their seasonal equivalent P and Q of the possible candidate models since more than one model could be considered from these values. It should be mentioned that in this step, subjectivity is presented when selecting the best model since the selection is based on the interpretation of the ACF and PACF plots. Figures 7 and 8 show the ACF and PACF plots of the series, respectively.

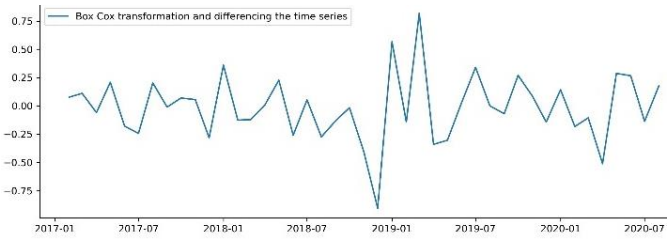


Fig. 6. Box-Cox transformation and differencing the time series.

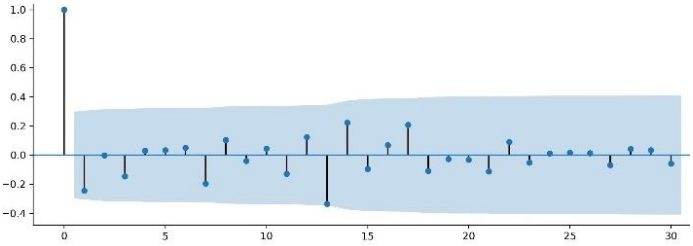


Fig. 7. Autocorrelation function plot.

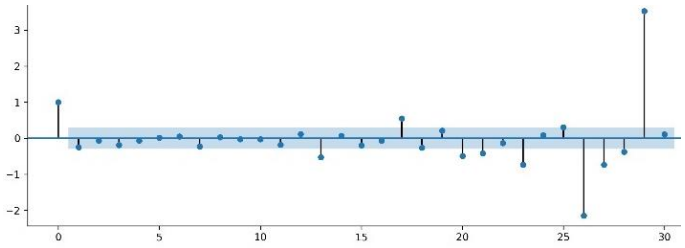
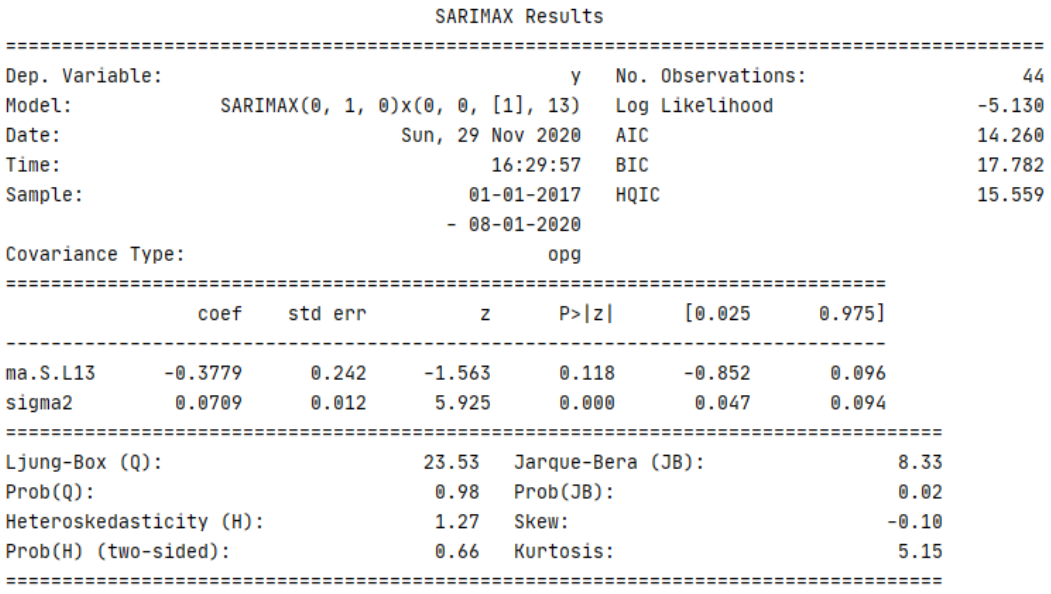


Fig. 8. Partial autocorrelation function plot.

It can be seen in the ACF plot shown in Figure 7 that the first lag is significant, so MA (1) component is suggested, that is, set q = 1. In the PACF plot shown in Figure 8, it can be observed that the first lag is significant, so AR (1) component is suggested, that is, set p = 1. Since the series was differentiated only once, set d = 1.

To define the seasonal component, it can be observed in the ACF plot shown in Figure 7 that lags 0 and 13 are significant, so set Q = 1 is suggested. In the PACF plot shown in Figure 8, it can be observed that only lags 1, 13, and 26 are significant, so set P = 1 is suggested. Finally, set D = 0 and set the seasonal period s = 12 or 13. Therefore, SARIMA $(1,1,1)x(1,0,1)_{12}$ or SARIMA $(1,1,1)x(1,0,1)_{13}$ was identified as an initial model.



SARIMAX Results

Dep. Variable:	y	No. Observations:	44
Model:	SARIMAX(0, 1, 0)x(0, 0, [1], 13)	Log Likelihood	-5.130
Date:	Sun, 29 Nov 2020	AIC	14.260
Time:	16:29:57	BIC	17.782
Sample:	01-01-2017 - 08-01-2020	HQIC	15.559
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ma.S.L13	-0.3779	0.242	-1.563	0.118	-0.852	0.096
sigma2	0.0709	0.012	5.925	0.000	0.047	0.094

Ljung-Box (Q):	23.53	Jarque-Bera (JB):	8.33
Prob(Q):	0.98	Prob(JB):	0.02
Heteroskedasticity (H):	1.27	Skew:	-0.10
Prob(H) (two-sided):	0.66	Kurtosis:	5.15

Fig. 9. Summary of the SARIMA model $(0,1,0) \times (0,0,1)_{13}$.

Based on the values determined from the ACF and PACF plots, different SARIMA models were tested to find the best fitting model by varying the values of their parameters between 0 and 2 for seasonal periods s = 12 and 13. This was because calculating all the possible combinations of values of the parameters of a SARIMA model by brute force can be computationally costly.

Table 2 shows the three best SARIMA models that obtained the lowest values in the AIC and BIC metrics. Of these, the best model was SARIMA $(0,1,0) \times (0,0,1)_{13}$ with an AIC = 14.260 and BIC = 17.782. This model does not contain non-seasonal components, nor does it contain seasonal autoregressive components, so it only consists of one seasonal moving average component.

TABLE 2
BEST SARIMA MODELS

Model	AIC	BIC
$(0,1,0) \times (0,0,1)_{13}$	**14.260**	**17.782**
$(0,1,0) \times (1,0,0)_{13}$	14.487	18.010
$(0,1,1) \times (0,0,1)_{13}$	14.644	19.928

***Step* 3. Parameter estimation**

In this step, the coefficient of the seasonal moving average component of the SARIMA model $(0,1,0) \times (0,0,1)_{13}$ was estimated. The estimation was performed with the *Summary* function of the trained model, which is an object of the *SARIMAX* class from the *statsmodels.tsa.statespace.sarimax* library. The parameter was estimated using the *log-likelihood* for the *maximum likelihood* estimation. The calculated coefficient was $\theta_1 = -0.3779$. Figure 9 presents the summary of the model.

***Step* 4: Diagnostic Checking**

After adjusting the SARIMA model, tests were applied to identify if the residual series is white noise, that is, that the residual errors do not show autocorrelation with time and that they are normally distributed. This gives evidence that the fitted model is suitable for forecasting. If the tests gave evidence that the residual series was not white noise, it would indicate the need to improve the model further.

For this purpose, the *plot_diagnostics* function was used, which presents four graphs: a) standardized residual, b) histogram, c) normal Q-Q, and d) correlogram. Figures 10 to 13 show the plots generated by this function.

The standardized residual plot in Figure 10 indicates the absence of trend and constant variance with a mean approaching zero. The histogram plot in Figure 11 shows that the residuals follow the bell curve distribution (normal distribution). The normal Q-Q plot in Figure 12 indicates that most residuals fit a straight line which suggests that they are normally distributed. Finally, the ACF plot in Figure 13 indicates that the residuals are not autocorrelated with time.

Additionally, the Ljung-Box Q [18] and Heteroskedasticity test were applied. The null hypothesis of the Ljung-Box Q test is that the data is distributed independently. The alternative hypothesis is that the data is not distributed independently. After applying the Ljung-Box test, the computed *p-value* was 0.98, so fail to reject the null hypothesis, and it is confirmed that the data is distributed independently. The null hypothesis of the Heteroskedasticity test is that the variance of the residual series is constant, that is, not heteroskedasticity. The alternative hypothesis is that heteroskedasticity exists. After applying the Heteroskedasticity test, the computed *p-value* was 0.66, so fail to reject the null hypothesis, and it is confirmed no heteroskedasticity.

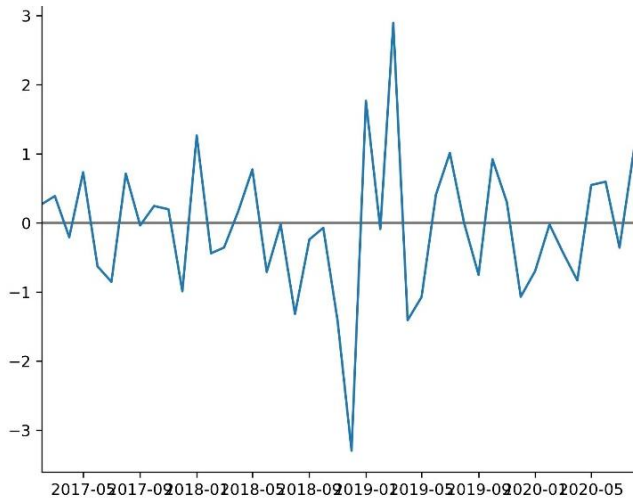


Fig. 10. Standarized residual plot of the SARIMA model $(0,1,0) \times (0,0,1)_{13}$.

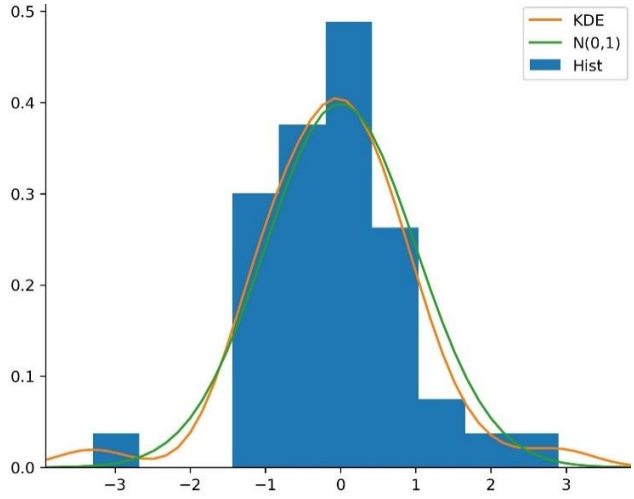


Fig. 11. Histogram plot of the SARIMA model $(0,1,0) \times (0,0,1)_{13}$.

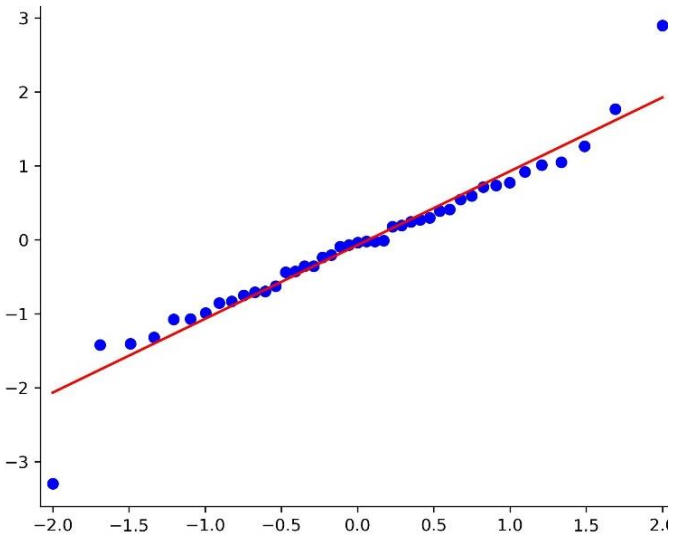


Fig. 12. Normal QQ plot of the SARIMA model $(0,1,0) \times (0,0,1)_{13}$.

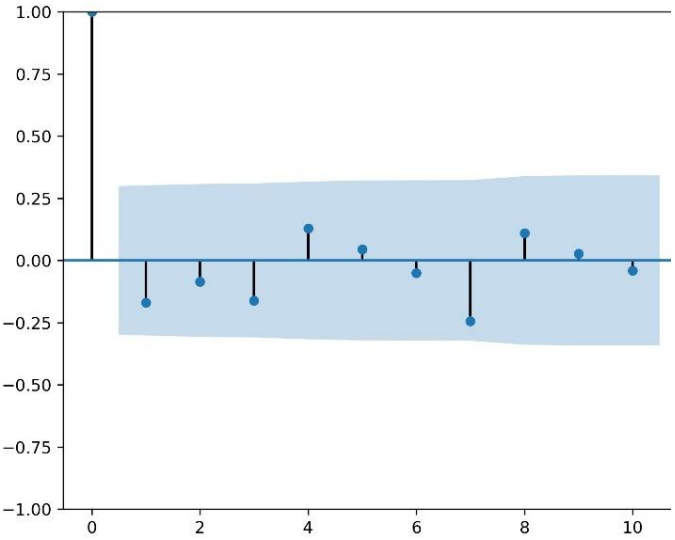


Fig. 13. Correlogram plot of the SARIMA model $(0,1,0) \times (0,0,1)_{13}$.

6. Results and Discussion

The forecast values of the Holt-Winters Seasonal $(0.8, 0.1, 0.0)_{12}$ and SARIMA $(0,1,0) \times (0,0,1)_{13}$ models were compared. To compare performance, the error metrics *Mean Absolute Percentage Error* (MAPE) and *Root Mean Square Error* (RMSE) were used, expressed in Equations (11) and (12), respectively.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \widehat{y_i}}{y_i}\right|, \tag{11}$$

$$RSME = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \widehat{y_i})^2}{N}}, \tag{12}$$

where:

y_i is the actual value.
$\widehat{y_i}$ is the forecast value.
N is the number of fitted points.

The dataset was divided into 80% for training and 20% for testing the models. The forecast values estimated by each model were compared against the real values of the test set, and the RMSE and MAPE metrics were used to measure the error. Table 3 shows the comparison of the results of the RMSE and MAPE metrics. Figures 14 and 15 show the forecast values generated by the Holt-Winters Seasonal and SARIMA models, respectively.

TABLE 3
Comparative results

Model	RMSE	MAPE	AIC	BIC
SARIMA $(0,1,0) \times (0,0,1)_{13}$	**7.55**	**15.09**	**14.260**	**17.782**
Holt-Winters $(0.8,0.1,0.0)_{12}$	40.03	95.88	212.880	238.217

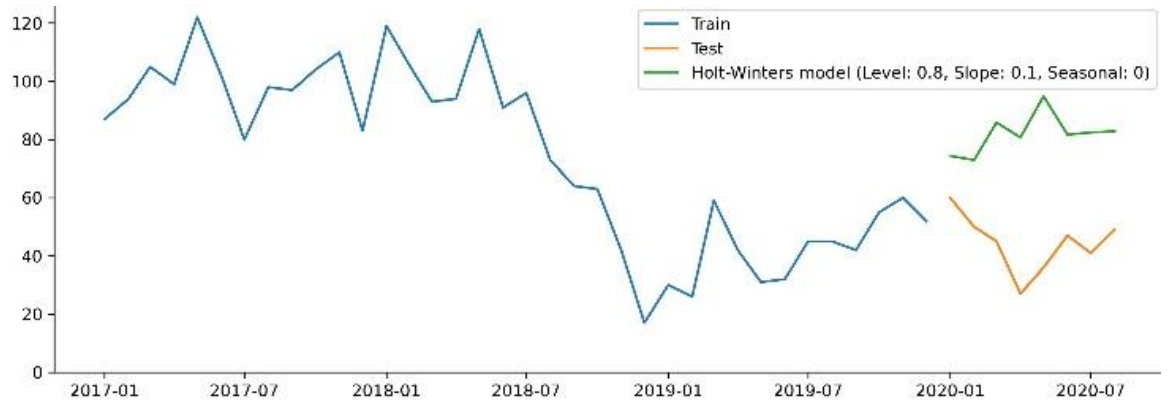


Fig. 14. Forecast values by Holt-Winters Seasonal $(0.8,0.1,0.0)_{12}$ model.

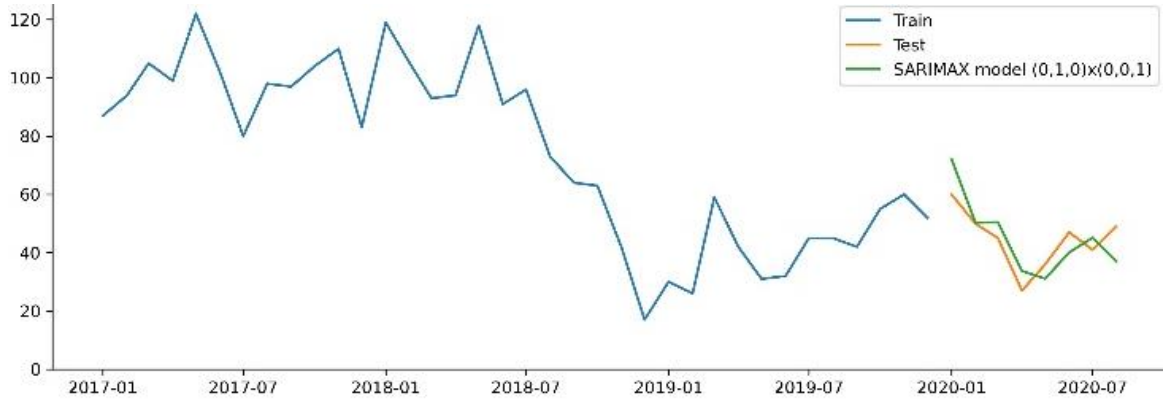


Fig. 15. Forecast values by SARIMAX $(0,1,0) \times (0,0,1)_{13}$ model.

As can be seen in Table 3, the SARIMA $(0,1,0) \times (0,0,1)_{13}$ model had the best performance since it obtained the lowest values in the RMSE and MAPE metrics. Additionally, it is worth mentioning that this model also obtained the lowest values in the AIC and BIC metrics with 14.260 and 17.782, respectively, against 212.880 and 238.217 obtained by the Holt-Winters Seasonal model.

Because the historical data provided by the warehouse was grouped monthly as mentioned above, the dataset corresponding to the time series used was generated with 45 observations. Despite being few observations, it can be seen from the results that the SARIMA models shown in Table 2 are better than those obtained by the Holt-Winter Seasonal models shown in Table 1. The possible cause of the Holt-Winter Seasonal models could not achieve a better performance could be that they could not adequately capture the seasonality of the time series due to the few observations in the dataset, which could indicate that the SSE value was set to 0.

7. Conclusions

In this work, the case of an assembly plant located in Cd. Juárez, Chihuahua, to forecast the future demand of one of the parts of the inventory was presented. Historical data between January 2017 and September 2020 was used.

Different Holt-Winters Seasonal and SARIMA models were implemented to forecast demand since the data exhibits trend and seasonal effects.

For the implementation of the SARIMA models, the Box-Jenkins methodology was used. As selection criteria for the best Holt-Winters Seasonal and SARIMA model, the AIC and BIC metrics were used. The selected models were Holt-Winters Seasonal $(0.8,0.1,0.0)_{12}$ and SARIMA $(0,1,0) \times (0,0,1)_{13}$.

RMSE and MAPE metrics were used to determine the model that makes the most accurate forecasts.

The metrics indicate that the SARIMA model has a better performance than the Holt-Winters Seasonal model since it manages to better fit the predicted data of the time-series of the used part.

The results provide evidence the SARIMA model fits the data well when forecasting the demand for inventory parts.

Based on the above, applying the SARIMA model to all the parts to know their future demand could lead to a better strategic relocation in the warehouse and, thus, facilitate the picking process.

References

[1] "Index Juárez Asociación de Maquiladoras, A.C.," *Información Estadística Mensual*. https://indexjuarez.com/ (accessed Dec. 05, 2020).

[2] J. Fattah, L. Ezzine, Z. Aman, H. El Moussami, and A. Lachhab, "Forecasting of demand using ARIMA model," *Int. J. Eng. Bus. Manag.*, vol. 10, pp. 1–9, 2018, doi: 10.1177/1847979018808673.

[3] T. van Gils, K. Ramaekers, A. Caris, and M. Cools, "The use of time series forecasting in zone order picking systems to predict order pickers' workload," *Int. J. Prod. Res.*, vol. 55, no. 21, pp. 6380–6393, Nov. 2017, doi: 10.1080/00207543.2016.1216659.

[4] A. T. Bon and T. K. Ng, "An Optimization of Inventory Demand Forecasting in University Healthcare Centre," in *IOP Conference Series: Materials Science and Engineering*, Feb. 2017, vol. 166, no. 1, p. 012035, doi: 10.1088/1757-899X/166/1/012035.

[5] T. Mendes Dantas, F. L. Cyrino Oliveira, and H. M. Varela Repolho, "Air transportation demand forecast through Bagging Holt Winters methods," *J. Air Transp. Manag.*, vol. 59, pp. 116–123, Mar. 2017, doi: 10.1016/j.jairtraman.2016.12.006.

[6] C. Chatfield, *Time-Series Forecasting*. CRC press, 2000.

[7] R. a Yaffee and M. McGee, "An introduction to time series analysis and forecasting: with applications of SAS® and SPSS®." Elsevier, 2000.

[8] A. T. Jebb, L. Tay, W. Wang, and Q. Huang, "Time series analysis for psychological research: examining and forecasting change," *Front. Psychol.*, vol. 6, no. JUN, p. 727, Jun. 2015, doi: 10.3389/fpsyg.2015.00727.

[9] D. A. Dickey and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *J. Am. Stat. Assoc.*, vol. 74, no. 366a, pp. 427–431, Jun. 1979, doi: 10.1080/01621459.1979.10482531.

[10] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root," *J. Econom.*, vol. 54, pp. 159–178, 1992.

[11] "StatsModels." https://www.statsmodels.org/stable/index.html (accessed Dec. 05, 2020).

[12] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, 1974, doi: 10.1109/TAC.1974.1100705.

[13] G. Schwarz, "Estimating the Dimension of a Model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, Mar. 1978, doi: 10.1214/aos/1176344136.

[14] P. R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages," *Manage. Sci.*, vol. 6, no. 3, pp. 324–342, Apr. 1960, doi: 10.1287/mnsc.6.3.324.

[15] H. A. Engelbrecht and M. Van Greunen, "Forecasting methods for cloud hosted resources, a comparison," in *Proceedings of the 11th International Conference on Network and Service Management, CNSM 2015*, Dec. 2015, pp. 29–35, doi: 10.1109/CNSM.2015.7367335.

[16] M. Markovska, A. Buchkovska, and D. Taskovski, "Comparative study of ARIMA and Holt-Winters statistical models for prediction of energy consumption," 2016.

[17] S. Makridakis and M. Hibon, "ARMA Models and the Box-Jenkins Methodology," *J. Forecast.*, vol. 16, no. 3, pp. 147–163, May 1997, doi: 10.1002/(SICI)1099-131X(199705)16:3<147::AID-FOR652>3.0.CO;2-X.

[18] G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series models," *Biometrika*, vol. 65, no. 2, pp. 297–303, Aug. 1978, doi: 10.1093/biomet/65.2.297.