

Improving intermittent demand forecasting based on data structure

S. Fatemeh Faghidian*, Mehdi Khashei** and Mohammad Khalilzadeh*

*Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

**Department of Industrial and systems Engineering, Isfahan University of Technology (IUT), Isfahan, Iran

**Corresponding Author: Khashei@cc.iut.ac.ir

Submitted: 21/09/2019

Revised: 07/02/2020

Accepted: 14/02/2020

ABSTRACT

Forecasting spare parts requirements is a challenging problem, because the normally intermittent demand has a complex nature in patterns and associated uncertainties, and classical forecasting approaches are incapable of modeling these complexities. The present study introduces a hybrid model that can impressively overcome the limitations of classical models while simultaneously using their unique advantages in dealing with the complexities in intermittent demand. The strategy of the proposed hybrid model is to use the three individual autoregressive moving average (ARMA), single exponential smoothing (SES), and multilayer perceptron (MLP) models simultaneously. Each of them has the potential of modeling a different structure and patterns of behavior among the data. The accuracy in forecasting ability is also increased by the suitable examination of these in the intermittent data. Croston's method is the backbone of the suggested model. The proposed hybrid model is based on CV^2 and ADI criteria, which improve its efficacy in examining inappropriate structures by reducing the cost of inappropriate modeling while increasing the prediction model accuracy. Using these results prevents the hybrid model from being confused or weakened in the modeling of all groups and reduces the risk of choosing the disproportionate model. The accuracy of prediction models was evaluated and compared using mean absolute percentage error (MAPE) by implementing an example, and promising results were achieved.

Keywords: autoregressive moving average model (ARMA); intermittent demand; multilayer perceptron (MLPs); parallel hybrid forecasting models; single exponential smoothing (SES).

INTRODUCTION

Inventory control of commodities with intermittent demand, such as spare parts, depends on accurate forecasts of demand. The high purchase cost and shortage costs increase risks in managing inventory. Receiving demand with high dispersion along with periods of no demand can be one of the characteristics identifying time series with intermittent demand (Croston, 1972). Items with intermittent demand need their own inventory control policies (Balugani et al., 2019). Croston's method is a popular forecasting method for intermittent demand (Rosienkiewicz et al., 2017).

The literature indicates that artificial neural networks (ANN) have also been widely utilized for demand forecasting; however, only a few studies have employed it for forecasting intermittent demand (Lolli et al., 2017; Rosienkiewicz et al., 2017; Kourentzes, 2013; Mukhopadhyay et al., 2012; Gutierrez et al., 2008; Pour et al., 2008). The evidence from these studies shows that ANN produces better results than other classical methods; the reason for this is that it has recently been favored by researchers in this field. ANN-based intermittent demand forecasting can be classified into three approaches. 1. The raw input pattern is utilized to forecast intermittent demand. For instance, Lolli et al.

(2017) used three input neurons to predict demand at different times (t). The first neuron represents demand at time $t-1$, the second neuron displays the number of periods separating the last two nonzero demand transactions at the end of the immediately preceding period, and the third neuron shows the cumulative number of successful periods with zero demand. The second and third neurons were introduced by Gutierrez et al. (2008) and Mukhopadhyay et al. (2012), respectively. Simultaneous use of these three neurons in an input set was adopted by Lolli et al. (2017). Different neural networks have been used to implement these models. 2. Intermittent demand can be predicted using Croston's method. The input set contains the values of two random variables defined using Croston's method. The difference between the models of this group and Croston's method is in applying ANN instead of SES (Kourentzes, 2013). 3. Forecast intermittent demand is divided into different stages, and ANN is used in one of these stages (Pour et al., 2008; Rosinkiewicz et al., 2017).

The difference between the present study and other recent studies using the ANN method is summarized as follows: For model inputs in the first group, a set of pattern structures is considered. By modifying them, results with varying accuracy are obtained. The inputs of the present research model and the second group are similar; the difference is in the type of forecasting model employed. What exists in the literature on this subject is the use of individual models to forecast intermittent demand. In the present study, however, the advantage of the combination theory is used to forecast and model irregular time series structures. Croston's variables are assumed to be independent and considered to be a specific type of probability density function (PDF). In this study, however, the independence of the variables is not accepted from the beginning; modeling is done in two ways, confirming and relaxing this hypothesis, and the results are compared. PDF restriction is not considered, and modeling is done as a free distribution.

The present study aimed to identify and model the different behavioral structures in the nature of intermittent demand, because the use of individual models may not achieve this aim. A hybrid rather than individual model is used to reduce the risks associated with selecting individual models, which is expected to lead to more accurate forecast results. This study is an extension of Croston's method. In this research, each time series was broken into its constituent elements, two separate time series and two random variables, including non-zero demand size Z_t and inter-demand intervals X_t . It is noteworthy that the actual demand time series had periods of zero values.

The main contribution of this paper is the modeling of the existing linear and nonlinear structures simultaneously by combining the results of the individual forecast models into a proposed hybrid prediction model to improve prediction results. The challenge was to select individual forecast models as components of the proposed hybrid model. There exist two significant points regarding the output forecast model. First, the output is one step ahead. Second, it forecasts the demand rate, not demand point, consistent with the nature of the data (Kourentzes, 2013). Non-zero values of demand are modeled in Croston's method. The output of the prediction models based on this method forecasts the demand rate.

This type of forecasting is suited to the nature of intermittent demand. For example, if the output of the prediction model is 0.25, it must be interpreted as the demand rate for each period. In other words, this value represents a unit of demand over the four periods (Kourentzes, 2013). The two criteria of mean absolute percentage error (MAPE) and mean error (ME) were used to measure the accuracy and bias of the model, respectively.

This paper is organized as follows. Section 2 includes a brief review of literature related to hybrid models and intermittent demand modeling. Section 3 presents the proposed hybrid model and its components. Section 4 explains the model's evaluation criteria. In Section 5, the characteristics of the data are clarified. Section 6 explains the parameters of the implementation of the model. The results of the implementation of models are then explained in Section 7. Finally, in Section 8, the results of modeling are discussed and conclusions are drawn.

HYBRID MODELS

Conventional statistical models do not have the potential to adequately model nonlinear trends in data. Evidence suggests that these nonlinear trends are better modeled by artificial intelligence (AI) models. However, AI models also demonstrate some disadvantages, including local optimum, low convergence rate, and overfitting (Versen et al., 2015).

Using linear or nonlinear models separately, inherent in intermittent demand, may lead to desirable results in modeling structures. The proposed model is introduced as a linear combination of individual models. These models are also known in the literature as parallel hybrid models.

The main motivation of presenting the hybrid model here is to identify and simulate the linear and nonlinear structures existing in the main data of intermittent demand. Because this feature is found in parallel hybrid structures, it has been employed in the present approach. The superior individual components of the proposed model were selected based on their strength and ability to model linear and nonlinear structures and the accuracy of the forecasted results.

Based on the results found in the existing literature, the ANN model for modeling nonlinear structures in intermittent time series was selected. This model can be regarded as data-based and nonparametric. The literature demonstrated that the integral part of hybrid prediction models for the linear modeling of linear and nonlinear hybrids is autoregressive moving average (ARMA) group models (Safari and Davallou, 2018; Han et al., 2017). For intermittent data with a trend or seasonality feature, (S)ARIMA models are employed (Lolli et al., 2014). Croston’s method is one of the oldest and most well-known models of intermittent demand forecasting. The performance of this model is based on the exponential smoothing method. Accordingly, this model is chosen for modeling linear structural in time series data.

Today, hybrid models are approaches for forecasting time series that reduce the risk of choosing an inappropriate model for data forecast structures, and they provide more accurate results (Hajirahimi and Khashei, 2018). Bates and Granger (1969) are pioneers in research in this field. Many researchers combined classical forecasting models with soft computing models to introduce a more efficient, better performing model (Safari and Davallou, 2018; Wang and Jiang, 2019; Han et al., 2017).

ARTIFICIAL NEURAL NETWORK

Artificial intelligence models can effectively extract nonlinear structures from time series data without limiting hypotheses. They provide the grounds for obtaining hidden structures in the data. Among artificial neural models, Multilayer perceptrons (MLPs) are widely used in feedforward network models. Gutierrez et al. (2008) utilized such a model with two input neurons and a hidden layer with two neurons to model lumpy demand forecasting. Mukhopadhye et al. (2012) used the same model with one input layer neuron. These models have found wide application in forecasting intermittent demand. Lolli et al. (2017) recently employed this network structure to provide an intermittent demand forecast model with the simultaneous consideration of the three input neurons introduced in the previous two studies. In the present study, one-step ahead forecasting is discussed, and feedforward is chosen. Model inputs include the past of lagged time series data.

ARMA MODEL

ARMA models are generally considered as linear models, and the linear relationship between the outputs and the inputs of the model can be explained using ARMA. The model ARMA (p,q) is formulated as follows:

$$\begin{aligned}
 &y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \varepsilon_t + B_1 \varepsilon_{t-1} + \dots + B_q \varepsilon_{t-q}, \quad \varepsilon_t \sim \text{i.i.d}(0, \delta_\varepsilon^2) \\
 &(1 - \sum_{j=1}^p \alpha_j L^j) y_t = \alpha_0 + (1 + \sum_{k=1}^q B_k L^k) \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d}(0, \delta_\varepsilon^2)
 \end{aligned}
 \tag{1}$$

where y_t is the actual value of the data, ε_t is the random error value, α_0 is an intercept, L^j is the j^{th} lag operator, and α_i ($i=1, 2, \dots, p$) are a finite set of parameters determined by linear regression. The steps followed in the ARMA (p, q) model were introduced and discussed by Box and Jenkins (1976). Before estimating the model parameters (p, q), the assumption of the stationary time series data should be questioned, for which the augmented Dickey–Fuller test (ADF) is usually used. If this assumption is not confirmed, the ARMA (p, q) will be changed to ARIMA (p, d, q), and the parameter I (d) will be added to the model as an integrated series. The parameter I (d) contains information on the number of differences converting the original series of data into stationary ones.

SINGLE EXPONENTIAL SMOOTHING METHOD (SES)

The SES model is formulated as follows:

$$Y_t = \alpha d_t + (1 - \alpha)Y_{t-1} \quad 0 \leq \alpha \leq 1 \quad (2)$$

where α is the smoothing parameter, d_t is the actual value, and Y_t is forecasting.

EVALUATION CRITERIA

The present research is an interface between the forecast of intermittent demand and hybrid models, and accordingly, accuracy metrics will be used to evaluate the accuracy of the models. As the data introduced by Lolli et al. (2017) was the benchmark, the mean absolute percentage error (MAPE), mean error (ME), ME/A, and average demand (A) benchmarks were employed here. All of these criteria are of the nonscale type. The ME/A criterion structure represents the mean deviation of the forecasted values from the data.

DATASET

To implement intermittent forecast models, two indicators, CVs and ADIs (Willemuin et al., 1994), were investigated. CV is the measure of erratic behavior representing the coefficient of variations of non-zero demands. CV² was introduced and applied by Syntetos and Boylan (2001). ADI represents the average number of time periods between the two successive non-zero demands and is known as the intermittence levels in the data.

By means of these two criteria, demand can be divided into four groups: lumpy, intermittent, erratic, and smooth. Intermittent and lumpy demands are both common within the ADI benchmark in the interval (1.4-1.32), and their difference lies in CV² values. Values less than 0.49 are in the intermittent demand group, and those ranging from 0.49 to 0.8 are in the lumpy demand group. Similarly, smooth and erratic demands have an ADI less than 1.32 and correspond to intermittent and lumpy demands.

Data sets introduced by Lolli et al. (2017) were used to evaluate the proposed hybrid model. These data sets are twenty-four intermittent weekly time series of weekly demand for spare parts in the automotive industry. Specifications of those series include ADI, CV², CV, length, and the mean for each time series separately as shown in Table 1.

Table 1. Characteristics of the used time series data.

Series	1	2	3	4	5	6	7	8
Length	75	76	69	73	76	75	61	65
ADI	1.19	1.23	1.61	1.35	1.38	1.34	1.3	1.3
CV	0.71	0.57	0.44	0.57	0.62	0.96	0.63	0.65
CV2	0.5	0.32	0.19	0.32	0.35	0.92	0.4	0.42
Mean	22.23	3.07	0.87	3.89	2.61	12.01	13.34	4.23
Series	9	10	11	12	13	14	15	16
Length	84	73	84	74	72	74	83	82
ADI	1.91	1.92	1.87	1.85	1.5	2.85	1.38	1.41
CV	0.65	0.44	0.68	0.57	0.52	0.33	0.52	0.74
CV2	0.42	0.19	0.46	0.32	0.27	0.11	0.27	0.55
Mean	4.92	0.9	1.62	2.85	28.44	0.42	2.18	124.22
Series	17	18	19	20	21	22	23	24
Length	76	74	68	414	83	81	83	81
ADI	1.52	1.18	3.78	1.87	1.36	1.35	1.22	1.27
CV	0.70	0.83	0.42	0.77	0.87	0.57	0.60	0.57
CV2	0.49	0.69	0.18	0.6	0.76	0.32	0.36	0.32
Mean	6	3.74	1.24	14.02	12.55	128.68	104.25	30.98

PARAMETER SETTING

According to Lolli et al. (2017), T-th of the first data of each series is used to train the network, and its next S-th data is employed to test the data. Scholars in the field of intermittent demand forecasting have used different divisions for this purpose. To compare the results of the present study with those of Lolli et al. (2017), their classification was used. The partitioning ratio was assumed for a time series of 65-35%. The first 65% of each time series was used for the training, and the remaining 35% was used for testing. For the implementation of ANN and ARMA, SES models, MATLAB, and EVIEWS9 were used. To have one-step ahead forecasts with the MAPE measurement criterion, among the models employed in Lolli et al. (2017), the best performances are associated with the ANN model with three input neurons with back propagation (BP) learning and extreme learning with values of 0.79 and 0.76, respectively. The worst performances are associated with the improved Croston's model, SBA, equal to 0.81. The best and worst performances consider the MAPE criterion for the implemented models.

THE PROPOSED PARALLEL HYBRID MODEL

The proposed hybrid model was derived from the linear combination of individual models as follows:

$$\hat{f}_{\text{hybrid}} = \sum_{t=1}^m w_i \cdot \hat{f}_{i,t} = w_{\text{MLP}} \cdot \hat{f}_{\text{MLP},t} + w_{\text{ARMA}} \cdot \hat{f}_{\text{ARMA},t} + w_{\text{SES}} \cdot \hat{f}_{\text{SES},t} \quad t = 1, \dots, m \quad (3)$$

where \hat{f}_{hybrid} indicates the value of hybrid forecast, w_i is the corresponding weight of i-th component, and $\hat{f}_{i,t}$ designates the forecast value of the combined model at time t. For the hybrid models consisting of two ARMA and MLP components and a hybrid model consisting of two MLP and SES components, the above equation must be rewritten to fit its components. The hybrid parallel method is introduced as the linear hybrid of the values forecasted by different forecasting models.

In this paper, two weighting methods were employed. One of the convenient weighing methods is to utilize the simple average (SA), wherein the weight of the components is assumed as $w_i = 1/n$, with n being the number of components of the parallel hybrid model. Alternatively, the weights of the single models are obtained by using the least squares (LS) method (Hajirahimi and Khashei, 2018). This method was used to determine the coefficients of the hybrid model based on the fundamental idea of linear regression model. The coefficients of the independent variables in the linear hybrid were determined based on the assumption that the model error value, defined as the square error in the sense of linear regression, should approach zero. Independent variables and values forecasted by different models were considered in order to search for the best linear fitting with the smallest distance from the real data values. For details of the method described above, see Hajirahimi and Khashei (2018).

The proposed hybrid seeks to model the linear structures among the data using both ARMA and SES models and nonlinear structures using the multilayer perceptron (MLP) or the single hidden layer feed forward network models.

The present study was based on the idea of Croston's method; thus, the main time series of the intermittent demand data were separated into its components Z_t and X_t , and therefore, the two time series comprise two variables, the non-zero demand size and the inter-demand intervals, respectively.

One assumption of Croston's model is the independence of the two variables Z_t and X_t from each other. This assumption was questioned in the current study. Therefore, the Hi-Duall and Hi-Rate models are presented as follows, and the results of the performance of the two models were compared.

Hi-Duall model: Similar to Croston's method, non-zero demand (Z_t) and inter-demand intervals (X_t) are individually forecasted using the hybrid forecast model of intermittent demand, and the obtained result was divided per the classical Croston's method.

Hi-Rate Model: The demand rate is forecasted directly, and there is no need to forecast the separation of the two time series Z_t, X_t . In this model, the variable $R_{i,t}$ is defined as $R_{i,t} = Z_{i,t}/X_{i,t}$, and it is used for modeling.

After the introduction of Croston's Method, researchers examined the bias model (Syntetos and Boylan, 2001). They showed that the results of Croston's method may have the deviation of the expected estimate of demand per time period. The literature contains modified Croston's methods. In most of these modified methods, the final prediction results are multiplied by a coefficient that is proportional to the assumptions of each model studied with the aim of solving the bias of the model. This bias issue is also discussed in Kourentzes (2013). The Hi-Duall model might suffer from inversion bias. To avoid the bias model, a data-driven debiasing coefficient C was introduced by Kourentzes (2013), which may be obtained by solving the following regression equation: $Z_t/X_t = C Y_t$.

The coefficient C is determined using the trained time series data, and its value is multiplied by the test time series data in order to eliminate the bias of the model. The aforementioned technique is used to solve the bias problem in the Hi-Duall model; however, this strategy is employed to avoid the results of forecasts in the method.

In the Hi-Rate model, to avoid biased forecasting results, rather than the separate forecasting of the Z_t, X_t time series, a time series R_t representing the rate of intermittent demand is used as the input of the hybrid forecasting models, which are not similar to Croston's method. This leads to the output of the proposed model as a forecast of the demand rate preventing the bias in the Hi-Rate model.

IMPLEMENTING THE PROPOSED HYBRID MODEL

To implement the models, the time series of the two variables $Z_{i,t}$ and $X_{i,t}$ should first be formed for each of the 24 sets of intermittent time series data. This information is utilized to run Hi-Dual models. However, to implement the Hi-Rate models, the opposite value $R_{i,t}$, defined as $R_{i,t} = Z_{i,t}/X_{i,t}$, is employed.

IMPLEMENTING THE MLP MODEL

In neural networks modeling (used in the present study), the number of neurons in the output layer and the type of activation function of that layer are assumed as a neuron with a linear activation function. The MLP model is employed for each variable. In each run, one of the following factors is changed. The number of neurons in the input is changed from 2 to 4, and afterwards, the number of intermediate neurons is changed from 1 to 3. The two activation functions tansing and logsig are also used for each of the listed cases. By changing each of the above items and keeping the other items constant, 18 models for each time series of the dataset should be implemented, and their performance should be compared. The MSE criterion is commonly considered and monitored to evaluate the performance of these models (Lolli et al., 2017); thus, the best performance is obtained.

As mentioned before, the data used in the present study was applied in Lolli et al. (2017), and in order to have comparable results, the same assumptions in the implementation of the neural network were used here, too. For instance, the values of the two constants, including learning rates and momentum, are assumed to be 0.1 and 0.9, respectively. In this paper, to implement MLP, the values of these two constants were taken into consideration, as in Lolli et al. (2017). These two values were also used to run all of its neural networks. Hence, benchmarking and, as the article reported, these numbers were also followed in this paper.

For the best results from the model performance, the following observations can be presented using the results obtained from the implementation of the proposed models for the time series belonging to the three random variables $X_{i,t}$, $Z_{i,t}$, and $R_{i,t}$. The number of input layer neurons was assumed to be 3 for both time series $R_{i,t}$ and $Z_{i,t}$, and 2 for the time series $X_{i,t}$. Given the $X_{i,t}$ time series structure for each of the four aforementioned groups, the logsig activation function along with the two neurons in the intermediate layer provided a better performance. This was similar for the two time series $Z_{i,t}$ and $R_{i,t}$. For smooth and erratic demands, the activation functions of logsig and 2 neurons in the intermediate layer are suggested. However, for the remaining lumpy and intermittent demands, due to the complexity of the data structure, the use of 3 neurons in the middle layer and the tansing activation function yielded more accurate results and a better performance.

IMPLEMENTING THE ARMA MODEL

For the three time series $X_{i,t}$, $Z_{i,t}$, and $R_{i,t}$, the stationary assumption is first evaluated, and then, the best performance among the implemented models was selected. Considering the time series $R_{i,t}$, it can be concluded that the ARMA (1,1) model had a good performance for all time series except the 14th, 15th, and 21st ones. The appropriate model for the 14th and 15th series was the AR(2) model, and the AR(1) model was suitable for the 21st time series.

For the time series belonging to the random variable $X_{i,t}$, the ARMA (1,1) model was suitable, except for the 1st series where the appropriate model was MA(2), the 18th time series where the suitable model was MA(1), and the 3rd and 14th time series where the AR(2) model could be considered. The dispersion of the presented models for time series belonging to the random variable $Z_{i,t}$ was low. The general behavior of the ARMA (1,1) model continued with this time series too, except for the 21st and 18th time series where the AR(1) model was suitable, and the 14th time series where the MA(1) model led to a better performance. It should be noted that the value of $I(d)$ (which contains information on the number of differences converting the original series of data into stationary ones) for the time series R_{14} and Z_{14} is equal to 1, and for the rest of the time series, according to the calculations performed, it is equal to zero.

IMPLEMENTING THE SES MODEL

Similar to the previous two sections, the implementation of the SES model for the three time series $X_{i,t}$, $Z_{i,t}$, and $R_{i,t}$ is discussed in this section. The key parameter in the implementation of this model is to determine the optimal value of the parameter α , the coefficient of model smoothing. Instead of using a constant value for this parameter, each time series was estimated according to its data, and then the SES model was run for the time series using the estimated value. Babai et al. (2017) indicated that the parameter value α in the implementation of Croston's model is in the range of 0.05-0.2 with 0.05. Lolli et al. (2017) implemented Croston's model with different values of α in this range, compared the results, and chose the model that had the least accuracy criterion. α , which had the lowest MAPE, has not been stated. The present paper did not need to assess only the accuracy of the model. Rather, it required more predicted values to be used in the hybrid proposed model, and because the weights of the individual models were determined using the predicted values in the hybrid model, achieving the predicted values with higher accuracy could lead to better results in presenting the final hybrid model. As a result, in this paper, not only were the values used in Lolli et al. (2017), including the values of α (0.05, 0.1, 0.15, 0.2) implemented, but the α value was also estimated using Eviews. The estimated values for α for all-time series were in the range of 0.05–0.2.

IMPLEMENTING THE HYBRID PROPOSED MODEL

After implementing the above-discussed models, their weights were measured on each of the proposed hybrid models. The weights were calculated for the three-tailed and the two-dimensional models using the LS method. These models included a hybrid model with three components, MLP, ARMA, and SES (called MAS), and dual models including the hybrid MA model, which incorporates MLP, ARMA, and MS hybrid models with MLP and SES components, and the AS hybrid model with ARMA and SES components. The details of the evaluation of the models with the MAPE, ME/A criteria are divided into four groups of demand and overhaul in Tables 2–6.

Table 2. MAPE, ME/A performance of various hybrid models for intermittent demand.

Intermittent Model	Evaluation Criteria	Hi-Duall.LS	Hi-Bias Duall.LS	Hi-Duall.SA	Hi-Rate.SA	Hi-Rate.LS
(MAS)	MAPE	0.521471	0.613333	0.588236	0.575943	0.522124
(MA)		0.550011	0.620694	0.699135	0.636074	0.559377
(MS)		0.630784	0.624728	0.755034	0.6445	0.551161
(AS)		0.537605	0.609835	0.568301	0.553271	0.537001
(MAS)	ME/A	0.17671	-0.07263	0.102544	0.163067	0.122055
(MA)		0.132285	-0.0482	-0.0024	0.09204	0.138592
(MS)		0.203616	-0.09826	0.086229	0.091141	0.142369
(AS)		0.192626	-0.0529	0.177998	0.159554	0.157721

Table 3. MAPE, ME/A performance of various hybrid models for smooth demand.

Smooth Model	Evaluation Criteria	Hi-Duall.LS	Hi-Bias Duall.LS	Hi-Duall.SA	Hi-Rate.SA	Hi-Rate.LS
(MAS)	MAPE	0.564121	0.550369	0.526794	0.51771	0.506476
(MA)		0.565867	0.572055	0.534038	0.523372	0.515498
(MS)		0.568182	0.565998	0.532079	0.523051	0.505003
(AS)		0.546152	0.541091	0.552763	0.500078	0.486466
(MAS)	ME/A	0.117127	0.024958	0.208785	0.279896	0.238946
(MA)		0.103031	0.223553	0.210694	0.263086	0.235522
(MS)		0.101817	-0.0327	0.211367	0.259098	0.248502
(AS)		0.123732	0.036941	0.148995	0.267147	0.230216

Table 4. MAPE, ME/A performance of various hybrid models for lumpy demand.

Lumpy Model	Evaluation Criteria	Hi-Duall.LS	Hi-Bias Duall.LS	Hi-Duall.SA	Hi-Rate.SA	Hi-Rate.LS
(MAS)	MAPE	0.531572	0.550954	0.552782	0.53143	0.517833
(MA)		0.54375	0.54867	0.556157	0.535964	0.529202
(MS)		0.578437	0.598174	0.564875	0.533103	0.520864
(AS)		0.557056	0.561339	0.56119	0.540575	0.536808
(MAS)	ME/A	0.324427	0.013597	0.364966	0.371052	0.326002
(MA)		0.3205	0.112751	0.361737	0.336138	0.319642
(MS)		0.397736	0.075773	0.357906	0.336581	0.323614
(AS)		0.355581	0.107521	0.329382	0.369439	0.368473

Table 5. MAPE, ME/A performance of various hybrid models for erratic demand.

Erratic Model	Evaluation Criteria	Hi-Duall.LS	Hi-Bias Duall.LS	Hi-Duall.SA	Hi-Rate.SA	Hi-Rate.LS
(MAS)	MAPE	0.214524	0.205179	0.205662	0.21593	0.212296
(MA)		0.209033	0.218162	0.209123	0.212873	0.207235
(MS)		0.212006	0.208594	0.205846	0.218125	0.210926
(AS)		0.21632	0.206467	0.208789	0.223889	0.217294
(MAS)	ME/A	0.330687	0.127393	0.238282	0.327883	0.306015
(MA)		0.237179	-0.06096	0.154167	0.297702	0.278091
(MS)		0.308843	0.049956	0.205234	0.3576	0.325008
(AS)		0.34039	0.120926	0.277491	0.310612	0.312606

Table 6. MAPE, ME/A performance of various hybrid models for overhaul demand.

Overhaul	Evaluation Criteria	Hi-Duall.LS	Hi-Bias Duall.LS	Hi-Duall.SA	Hi-Rate.SA	Hi-Rate.LS
(MAS)	MAPE	0.457922207	0.479958721	0.468368528	0.4602533	0.439682176
(MA)		0.467165201	0.489895104	0.499613273	0.477070873	0.452828047
(MS)		0.497352153	0.499373602	0.514458368	0.479694622	0.446988689
(AS)		0.464283329	0.479683027	0.472760606	0.454453493	0.444392374
(MAS)	ME/A	0.237238	0.023329	0.228644	0.285474	0.248255
(MA)		0.198249	0.056785	0.181049	0.247241	0.242962
(MS)		0.253003	-0.00131	0.215184	0.261105	0.259873
(AS)		0.253083	0.053123	0.233466	0.276688	0.267254

CONCLUSION AND DISCUSSION

A set of hybrid models for forecasting intermittent demand were introduced in the present research to comprehensively identify all existing structures and patterns (linear and nonlinear) in data and to improve the forecasting performance of the model, and the following results were observed. The proposed hybrid model was divided into two categories, Dual and Rate. As the interactions between the two non-zero demand and inter-demand interval variables show, a more precise framework for forecasting can be proposed. The MAPE criterion was considered here to examine the proposed hybrid models. In the present study, acceptance of the hypothesis regarding variable independence is contingent on the findings obtained from the hybrid models. Evaluation of numerical results for overhaul models indicated higher accuracy of the Rate group models than their corresponding Dual group models. Therefore, considering the interrelationships between variables led to a comprehensive understanding of the behavioral patterns in the data. If the variables of the model are assumed to be independent, modeling using the LS weighting method will have the precise implementation. Combining the same components and using the MAPE criterion, the percentage of improvement of the Hi-Rate.LS group models compared to the Hi-Duall.LS group models for MAS, MA, MS, and AS models was 3.99%, 3.06%, 10.12%, and 4.29%, respectively (these values were obtained using Table 6).

The data set used in this study included all four categorized groups. The degree of complexity in the data was one of the factors most strongly affecting the choice of the beneficial hybrid forecasting model in accordance with the time

series. In the classification, that is, the two mentioned criteria, the smooth group was the first examined. One feature of this group is the lower level of the two dispersion criteria. It can be deduced that the degree of complexity is low in the structure of the time series. According to the numerical results obtained from the implementation of the proposed models in the smooth group, the precise performance in the Hi-Rate.LS group with AS components and a MAPE of 0.486466 is reported. The difference in the performance of the AS model between the Hi-Rate.LS and Hi-Rate.SA groups is explained by the application of different weighting methods in the hybrid model and the superiority of the LS method. This is also the case with the Duall model.

Examining the models indicated that applying the hybrid model of the MLP type cannot lead to appropriate documentation. It also suggested that models with two SES and ARMA components can be impressively utilized in the smooth grouping to reduce computational costs and time and increase the efficiency of the hybrid. For example, the percentage of improvement of the AS model compared with those of the MAS, MA, and MS models in the Hi-Rate.LS group was 3.95%, 5.63%, and 3.67%, respectively.

Due to the low degree of complexity in the data, the application of the MLP model may lead to confusion of the model; consequently, inefficient outcomes ensue. The percentage of improvement of the hybrid MAS model in the Hi-Rate.LS group, indicating the addition of individual MLP to the hybrid model with ARMA and SES components compared with the AS model in the same group, was -4.11%. This indicates the inappropriate performance of the MAS model compared to the AS model.

Contrary to the smooth demand, there is lumpy demand, wherein both parameters have the highest degree of dispersion; hence, the complexity and intricacy in the structure are high. The performance of the hybrid model with three MLP, ARMA, and SES components in the Rate and Duall groups by LS weighting method is more accurate using MAPE criteria, the values of which were 0.517833 and 0.531572, respectively. Due to the high complexity of demand, the application of a three-component hybrid model does not result in confusion of the model, and the existing structures in the data can be effectively modeled and identified.

Although these three individual models are appropriate for modeling structures in this group, applying the weighting method, and considering the hypothesis of the independence of variables, affects the performance of the hybrid model with these three components. The percentage of improvement of the MAS model in the Hi.Rate.LS group compared to the MAS model in the Hi.Rate.SA group was 2.55%. Moreover, the percentages of improvement of the MAS model in the Hi.Rate.LS group compared with the MAS model in the Hi. Duall.SA and Hi. Duall.LS groups were 6.32% and 2.58%, respectively.

As the numerical evaluations in the lumpy demand group revealed, the worst performer in the Hi-Duall.LS group was the MS hybrid model. The poor outcome of this model can be explained by the incorrect selection of components in the hybrid model. The complexity and ambiguity of the data in this group were such that the existence of a model with high potential in identifying and modeling nonlinear structures (MLP) with a less robust model (SES) in identifying linear structures is unable to estimate and recover patterns in it. Therefore, it can be stated that the existence of a robust model in the structure of the hybrid model does not guarantee proper performance (by applying precision criteria).

The degree of structural complexity in the intermittent group was of lower severity than that of the lumpy group, but it was still high in complexity compared with that of the other two groups. The performance of models in this demand group showed the superiority of the three-component hybrid model MAS in the Hi-Rate.LS and Hi-Duall.LS groups. The reported MAPE benchmark values were 0.522124 and 0.521471, respectively, indicating that this model has lower MAPE values compared with the other models.

As experimental records indicate, an efficient hybrid model still requires a three-component hybrid model to forecast the complexities of the data structure in this group. The performance of the three-component hybrid model, compared with the other hybrid models, confirmed the superiority of the three-component hybrid models over two-component hybrid models. Documents from the implementation of two-component hybrid models indicate their inefficiency in identifying and modeling data structure in this group. The percentage improvements of the MAS model in the Hi-Rate.LS group compared with the other hybrid models in this group with MA, MS, and AS components were

6.65%, 5.27%, and 2.78%, respectively. The percentage improvements of the MAS model in the Hi-Duall.LS group for the hybrid models with MA, MS, and AS components were 5.18%, 17.32%, and 3.00%, respectively.

The poor proof of the two-component hybrid models compared to the three-component hybrid models in the intermittent group is explained by the selection of inadequate individual models as components of the hybrid model. Consequently, these individual models cannot identify all the patterns. This inadequacy of the hybrid models is reflected in their MAPE benchmark.

The degree of complexity of data in erratic demand is not as high as that of the lumpy or the intermittent groups, nor the lowest like that of the smooth group. Experimental evidence obtained from the performance of the proposed hybrid models in this demand group showed that the lowest MAPE values for the MA model in the Hi-Rate.LS and Hi-Duall.LS groups were 0.207235 and 0.209033, respectively. The values of this criterion for the hybrid MS and MAS models in these two groups were also very close to the values mentioned for the MA model. The proposed hybrid model requires the use of the MLP model for modeling nonlinear structures alongside the ARMA or SES models to identify linear structures. As a result, the numerical values of MAPE were very close to those of both MA and MS models. In this demand group, the choice of the hybrid MAS model in both the Rate and Duall groups increased the cost and time of modeling, while not significantly improving the MAPE value compared with the performances of the MA and MS models.

This research introduces a useful model for each demand group by adopting different strategies in selecting the components of the hybrid model and comparing the accuracy rates of their implementation. Finally, by analyzing the results for each of the four demand groups, hybrid models with the indicated components are proposed in Table 7.

Table 7. Introduced efficient model, for each demand group.

Demand group	Model components	Hybrid type	MAPE
Smooth	AS	Rate.LS	0.486466
Lumpy	MAS	Rate.LS	0.517833
Intermittent	MAS	Rate.LS	0.522124
Erratic	MA	Rate.LS	0.207235

Ultimately, it can be stated that the structure and complexities existing in the data play an important role in determining which hybrid model to use for each of the four demand groups. These implications are affected by the two criteria of CV^2 and AID, the values of which are predetermined, and changes in them can be ignored from one dataset to another, such that they are not dominant in changing the totality of the conclusion. Using these results prevents the hybrid model from being confused or weakened in the modeling of all groups and reduces the risk of choosing the disproportionate model. Because combination modeling in forecasting intermittent demand is new, and also given that the present research model is presented in parallel, it is suggested that series modeling or parallel modeling with components that differ from those of the present research be used, and the results be compared with those of this research.

ACKNOWLEDGMENT

The authors would like to acknowledge the valuable comments and constructive suggestions made by an anonymous referee, which were greatly beneficial in improving the manuscript.

REFERENCES

- Balugani, E., Lolli, F., Gamberini, R., Rimini, B. & Babai, M.Z. 2019.** A periodic inventory system of intermittent demand items with fixed lifetimes. *International Journal of Production Research*, (Jan).<https://doi.org/10.1080/00207543.2019.1572935>.
- Bates, J.M. & Granger, C.W.J.1969.** The hybrid of forecasts. *Operational Research Society (1970-1977)* **20**: 451-68. <https://doi.org/10.2307/3008764>.
- Box, P. & Jenkins, G.M. 1976.** *Time Series Analysis: Forecasting and Control*. Holden-day Inc, San Francisco, CA.
- Croston, J.D.1972.** Forecasting and stock control for intermittent demands. *Journal of the Operational Research Society (1970-1977)* **23**(3): 289-303. <https://doi.org/10.1057/jors.1972.50>.
- Gutierrez, R.S., Solis, A.O. & Mukhopadhyay, S.2008.** Lumpy demand forecasting using neural networks. *International Journal of Production Economics*, **111**: 409-20. <https://doi.org/10.1016/j.ijpe.2007.01.007>.
- HajiRahimi, Z. & Khashei, M. 2018.** A least squares-based parallel hybridization of statistical and intelligent models for time series forecasting. *Computers & Industrial Engineering*, **118**: 44-53. <https://doi.org/10.1016/j.cie.2018.02.023>.
- Han, Q. Meng, F. Hu, T. & Chu, F. 2017.** Non-parametric hybrid models for wind speed forecasting. *Energy Conversion and Management*, **148**: 554-68. <https://doi.org/10.1016/j.enconman.2017.06.021>.
- Kourentzes, N.2013.** Intermittent demand forecasts with neural networks. *International Journal of Production Economics*, **143**: 198-206. <https://doi.org/10.1016/j.ijpe.2013.01.009>.
- Lolli, F., Gamberini, R., Regattieri, A. & Rimini, B.2014.** Application of tramo-seats automatic procedure for forecasting intermittent demand patterns, in: *OPT-I 2014 – 1st International Conference on Engineering and Applied Sciences Optimization*. National Technical University of Athens, Kos Island, Greece:1435-45.
- Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatosb, T. & Gucci, S. 2017.** Single-hidden layer neural networks for forecasting intermittent demand. *International Journal of Production Economics*, **183**: 116-28. <https://doi.org/10.1016/j.ijpe.2016.10.021>.
- Mukhopadhyay, S., Solis, A.O. & Gutierrez, R.S. 2012.** The accuracy of non-traditional versus traditional methods of forecasting lumpy demand. *Journal Of Forecasting*, **31**: 721-35. <https://doi.org/10.1002/for.1242>.
- Pour, A.N., Tabar, B.R. & Rahimzadeh, A. 2008.** A Hybrid Neural Network and Traditional Approach for Forecasting Lumpy Demand, *Proceedings of World Academy of Science: Engineering & Technology*, **42**: 384-90.
- Rosienkiewicz, M., Chlebus, E. & Detyna, J. 2017.** A Hybrid Spares Demand Forecasting Method Dedicated to Mining Industry. *Applied Mathematical Modeling*, **49**:87-107. <https://doi.org/10.1016/j.apm.2017.04.027>.
- Safari, A. & Davallou, M. 2018.** Oil price forecasting using a hybrid model. *Energy*: **148**: 49-58. <https://doi.org/10.1016/j.energy.2018.01.007>
- Syntetos, A.A. & Boylan, J.E. 2001.** On the bias of intermittent demand estimates. *International Journal of Production Economics*, **71**(1-3): 457-66. [https://doi.org/10.1016/S0925-5273\(00\)00143-2](https://doi.org/10.1016/S0925-5273(00)00143-2).
- Syntetos, A.A., Boylan, J.E. & Croston, J.D. 2005.** On the categorization of demand patterns. *Journal of the Operational Research Society*, **56**(5): 495-503. <https://doi.org/10.1057/palgrave.jors.2601841>.
- Versen, E.B., Morales, J.M., Moller, J.K. & Madsen H.2015.** Short-term probabilistic forecasting of wind speed using stochastic differential equations. *International Journal Forecast*, **32**: 981-90. <https://doi.org/10.1016/j.ijforecast.2015.03.001>.
- Wang, Q. & Jiang, F. 2019.** Integrating linear and nonlinear forecasting techniques based on grey theory and artificial intelligence to forecast shale gas monthly production in Pennsylvania and Texas of the United States. *Energy*, **178**: 781-803. <https://doi.org/10.1016/j.energy.2019.04.115>
- Willemain, T.R., Smart, C.N., Shockor, J.H. & DeSautels, P.A. 1994.** Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal Forecast*, **10**(4): 529-38. [http://dx.doi.org/10.1016/0169-2070\(94\)90021-3](http://dx.doi.org/10.1016/0169-2070(94)90021-3).