ARTICLE IN PRESS

Industrial Marketing Management xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Industrial Marketing Management

journal homepage: www.elsevier.com/locate/indmarman

A big data driven framework for demand-driven forecasting with effects of marketing-mix variables

Ajay Kumar[a,*], Ravi Shankar[b], Naif Radi Aljohani[c]

[a] Harvard Business School, Harvard University, Cambridge, MA 02138, United States
[b] Department of Management Studies, Indian Institute of Technology Delhi, New Delhi 110016, India
[c] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Keywords:
Big data analytics
Demand shaping and sensing
Fuzzy neural network
Market-mix modelling

ABSTRACT

This study aims to investigate the contributions of promotional marketing activities, historical demand and other factors to predict, and develop a big data-driven fuzzy classifier-based framework, also called "demand-driven forecasting," that can shape, sense and respond to real customer demands. The availability of timely information about future customer needs is a key success factor for any business. For profit maximization, manufacturers want to sense demand signals and shape future demands using price, sales, promotion and others economic factors so that they can fulfil customer's orders immediately. However, most demand forecasting systems offer limited insight to manufacturers as they fail to capture contemporary market trends, product seasonality and the impact of forecasting on the magnitude of the bullwhip effect. This paper aims to improve the accuracy of demand forecasts. In order to achieve this, a back-propagation neural network-based model is trained by fuzzy inputs and compared with benchmark forecasting methods on a time series data, by using historical demand and sales data in combination with advertising effectiveness, expenditure, promotions, and marketing events data. A statistical analysis is conducted, and the experiments show that the method used in the proposed framework outperforms in optimality, efficiency and other statistical metrics. Finally, some invaluable insights for managers are presented to improve the forecast accuracy of fuzzy neural networks, develop marketing plans for products and discuss their implications in several fields.

1. Introduction

It is well known that the importance of accurate forecasting of product sales has not gained enough attention from the research community. Past research shows that to achieve efficiency in e-commerce, several key factors come into play: (i) process efficiencies, (ii) demand aggregation, (iii) information sharing, (iv) web efficiencies, and (v) risk management. When forecasting managers plan the operation and marketing strategies for product sales, they consider the following four factors in the marketing model: (i) product, (ii) price, (iii) place, and (iv) promotion. If buyers can use exchangeability to work with both the long-term capacity suppliers and the spot market suppliers, they will be able to estimate the supply and demand more accurately through effective demand and supply chain risk pooling.

In recent years, several studies have been published regarding the supply chain risk management (Colicchia & Strozzi, 2012; Ho, Zheng, Yildiz, & Talluri, 2015; Rao & Goldsby, 2009; Sodhi, Son, & Tang, 2012; Tang, 2006; Tang & Nurmaya Musa, 2011). Juttner (2005) distinguished between the two types of risks commonly associated with supply chains: (i) supply and demand risks, that are internal to a supply chain; and (ii) environmental risks that are external to the supply chain, such as political, natural or societal uncertainties. Supply-side risks mainly reside in purchasing, supplier activities, and supplier relationships (Zsidisin, Panelli, & Upton, 2000), whereas demand-side risks may be due to disruptions to physical distribution or from mismatches between forecast and actual demand (Wagner & Bode, 2006).

Folinas and Rabi (2012) explained how poor forecasting can reduce the availability of products, change customer choice and have an impact on the working capital. In addition to that, distortions in demand forecasting also cause the bullwhip effect which can lead to inefficiencies, excessive inventory, stock-outs and backorders (Bhattacharya & Bandyopadhyay, 2011; Coppini, Rossignoli, Rossi, & Strozzi, 2010; Davino, De Simone, & Schiraldi, 2014; Delhoum & Scholz-Reiter, 2009; Geary, Disney, & Towill, 2006; Lee, Padmanabhan, & Whang, 1997a; Lee, Padmanabhan, & Whang, 1997b; Petrovic, 2001; Sodhi & Tang, 2011). Research by Lee et al. (1997b) identifies five

* Corresponding author.
E-mail addresses: ajaytomar.dce@gmail.com (A. Kumar), ravi1@dms.iitd.ac.in (R. Shankar), nraljohani@kau.edu.sa (N.R. Aljohani).

https://doi.org/10.1016/j.indmarman.2019.05.003
Received 2 November 2018; Received in revised form 24 March 2019; Accepted 7 May 2019
0019-8501/

major causes of the bullwhip effect: (i) demand forecasting, (ii) order batching, (iii) supply shortages, (iv) non-zero lead-time, and (v) price fluctuations. This is further complicated by the effects of seasonality, promotions, and product proliferation (Lu & Wang, 2010).

In the current complex global marketplace, traditional demand forecasting models based on simple analytical techniques are only capable to sense the demand signals with linear and exponential trends, cyclic behaviour, and seasonality. A historical demand data based method in which managers use past responses to predict future customer demands is no more an accurate and effective approach since some critical factors can get ignored in the process. In order to reduce the impact of environmental uncertainty many companies have developed closer relationships with their suppliers and customers and reduced their supply base to a smaller number of suppliers who are treated as partners (Hadaya & Cassivi, 2007). But organisations still need a method that can predict and paint an accurate picture of market demands to efficiently manage all the important aspects of the supply chain, namely production, inventory, distribution, and orders.

Forecasts of demand are the foundation of all planning activities (Haberleitner, Meyr, & Taudes, 2010; Sterman, 1989). The aim of demand planning is to achieve a synchronised flow of goods and services throughout the supply chain. In the hands of a forecasting manager, correct demand forecasts will work as a powerful tool for generating the whole demand picture. The two subfields of predictive analytics, demand and sales forecasting, influence all the companies and are used to remove the inefficiencies in the supply chain. Folinas and Rabi (2012) identified three stages in the evolution of the demand planning approach: (i) Collaborative Planning Forecasting and Replenishment (CPRF); (ii) demand driven value chains; and (iii) demand sensing. CPFR is “a cohesive bundle of business processes whereby supply chain trading partners share information, synchronised forecasts, risks, costs and benefits with the intent of improving overall SC performance through joint planning and decision making. Accordingly, CPFR enhances customer demand visibility and matches supply and demand with a synchronised flow of goods from the production and delivery of raw materials to the production and delivery of the final product to the end consumer” (Hollmann, Scavarda, & Thomé, 2015, p. 977). Demand sensing is also an equally important aspect for organisations as it helps to build a more responsive and agile supply chain. Chong, Ch'ng, Liu, and Li (2015) considered demand sensing by modelling the demand-forecasting problem using ‘big data’ analytics. They used a support vector machine (SVM) model to investigate the contribution of online promotional activities as predictors of product demand, but they did not consider the effect of demand shaping and advertising on consumer demands.

An accurate and efficient *demand, supply and price forecasting* model has a direct impact on customer satisfaction and inventory stock-out. To ensure proper functioning of a supply chain management system, companies must improve their demand forecasting models, so that they are in a better position to handle customer needs. We are living in an era of big data and companies are collecting data from multiple dimensions. They are moving from traditional forecasting techniques towards advanced data science methods since they know that historical sales data, intra-category, and inter-category marketing channels have a huge impact on forecasting accuracy. Companies are trying to map the patterns in customer behaviour so that they can optimize their marketing expenses, thereby improving their overall financial performance. By measuring the quantitative effect of their marketing campaigns through different channels, these companies are trying to calculate the return on investment (ROI) impact. This has ultimately led to the emergence of the new buzz phrase - demand-driven forecasting - which is the combination of demand shaping, demand sensing and responding to real consumer demands (Chase Jr, 2013, p. 13).

Demand-driven forecasting aims to sense demand signals and shape the future customer demands by utilizing advanced data mining techniques as illustrated in Fig. 1 (Chase Jr, 2013). These techniques use ‘big data’ analytics to measure the success of marketing strategies by identifying patterns in consumer behaviour. Demand-driven supply chains focus upon providing superior value to end users (Zokaei & Hines, 2007). This approach utilises closely connected supply networks to ensure that the production is always linked to the demand (Hadaya & Cassivi, 2007).

The first step of the demand-driven forecasting method is demand translation. Demand translation consists of translating actual and forecasted demand into information that can be used in supply planning. To accomplish this, sales information including product family identities, revenues, volume, and mix quantities are passed to the supply side production line and distribution planning databases (Ross, 2015, p. 299). Demand sensing, the second step in the process, “refers to sensing customer purchase behaviour or, more generally, customers' choice behaviour. The scope of demand sensing can range from estimation of the price a potential customer would be willing to pay for an existing or new product and identification of his economic segment, to understanding his latent consideration set, the set of new products or the set of new features in products that the customer will be interested in” (Ravikumar, Saroop, Narahari, & Dayama, 2005, p. 311). It utilises upstream data within the value chain to generate a more accurate demand forecast for the organization, which includes product seasonality and takes into consideration historical trends in buying patterns by supporting the identification of new business drivers that may affect demand patterns as well as manufacturing supply capabilities (Folinas & Rabi, 2012).

Demand shaping is the third step. It is an extension to demand sensing that links replenishment data and shipment history with the current point-of-sale information. In other words, demand shaping is a process in which all the available and accessible information (encompassing demand sensing data) is used to develop an optimized, well organised and steady plan of demand and supply so that the sales and profitability targets are met and customer satisfaction metrics are achieved (Chase Jr, 2013). The penultimate step, demand shifting “refers to the ability to promote another product as a substitute if the product originally demanded was not available. It is especially useful if demand patterns or supply capacity changes to steer customers from product A to product B" (Chase Jr, 2013, p. 46). There are two types of demand shifting: (i) at the point of sale, where customers are incentivised to purchase an alternative; and (ii) at the point of supply, where manufacturer negotiates with the sales team to shift demands in the future due to capacity constraints (Chase Jr, 2013, p. 46). Demand orchestration, which is the final step, is the balancing act between sensing and shaping the end-user demand. It “focuses on the development of demand plans that ensure that the expected trade-offs between demand opportunity and demand risk are optimized. Success is revealed in the level of performance of actual to expected demand and actual to expected operations costs” (Ross, 2015, p. 299).

In this paper, we have studied a back-propagation neural network-based demand-driven forecasting model. It is trained by fuzzy inputs and compared with benchmark forecasting methods on a time series data, by combining demand, supply, promotional campaigns, and practical sales data. This is followed by conducting a statistical analysis and experiment, the results of which show that the method used in the

Demand Translation → Demand Sensing → Demand Shaping → Demand Shifting → Demand Orchestration

Fig. 1. Demand-driven forecasting (Chase Jr, 2013).

proposed framework outperforms in optimality, efficiency and other statistical metrics. A further contribution is made by the adstock model of advertising research, which provides a mechanism for optimally allocating budgets, and also predicts the impact of different marketing channels (Beltran-Royo, Escudero, & Zhang, 2016; Broadbent, 1979; Chase Jr, 2013). Hence, by considering the role of promotional marketing activities as predictors combined with historical sales and demand, the present research modifies and tests the developed demand-driven forecasting and investigates the performance or efficiency of fuzzy neural network framework in evaluating efficiency, demand shaping and sensing effect, net-stock amplification and the bullwhip effect.

The current study offers interesting insights in three ways:

1. First, to the best of our knowledge, this study is the first one, which investigates the contributions of promotional marketing activities, historical demand and other factors, and develops a big data-driven framework for demand shaping, sensing and enhancing the forecasting accuracy of the time series model using fuzzy ANN based classifier in big data environment.
2. Second, our experimental outcomes indicate that the proposed model performs better than the traditional forecasting methods, when we simulate the impact of different marketing effectiveness plans and identify the most feasible option while maximizing revenue and ROI from each of the following sources: (i) marketing channel, (ii) the decay effect of adstock, which is denoted in terms of 'half-life' of advertising, and (iii) the impact of advertising on sales.
3. Third, our advanced analysis further indicates that the proposed big data-driven framework gives a better forecasting result after linking the demand forecasting with supply data. It also calculates the diminishing returns, regularly monitors the effectiveness of advertising and suggests optimized allocation of total marketing spends while maximizing the effectiveness.

In general, there is a lack of studies that have attempted to develop demand-driven forecasting model and showed how advertising and other promotional strategies alter the demand curve. Our primary goal is to develop a big data-driven demand-driven forecasting model and improve the forecast accuracy and interpretability of this model using an efficient and error-free fuzzy classifier. This will not only accelerate the process of learning but will also examine the role of promotional marketing activities as predictors combining historical demand and other factors. Further, in the conclusion section, we elaborate on some significant managerial insights, which would prove to be very effective for the supply chain-forecasting managers.

2. Literature review

Time series forecasting is an important problem and widely used to predict as accurate as possible future values in economic, finance, weather, energy, stock price etc. domains (Chang, 2017; Chang & Ramachandran, 2017; Jeon, Hong, & Chang, 2018; Khan and Shahidehpour, 2009; Verdouw, Beulens, Trienekens, & Van Der Vorst, 2011; Williams, Waller, Ahire, & Ferrier, 2014). Researchers have already developed several methods and categorized them into three categories: traditional time series methods, machine learning methods and fuzzy machine learning methods. The first forecasting category is traditional forecasting and it is largely relied on historical demand trends and patterns and assume that demand follows a time series pattern (Chen & Blue, 2010). Researchers have aimed to reduce the bullwhip effect by controlling the parameters of the models, which have included Moving Average (MA) (Chen, Drezner, Ryan, & Simchi-Levi, 2000), Exponentially Weighted Moving Average (EWMA) (Kone & Karwan, 2011), Exponential Moving Average (EMA) (Chen, Ryan, & Simchi-Levi, 2000) Autoregressive (AR) (Duc, Luong, & Kim, 2010; Li, Li, Li, & Shirodkar, 2012), Autoregressive Moving Average (ARMA) (Duc, Luong, & Kim, 2008; Moosmayer, Chong, Liu, & Schuppar, 2013), double exponential smoothing (Baharaeen & Masud, 1986; Chen, Ryan, & Simchi-Levi, 2000; Hansun & Subanar, 2016; Taylor, 2003; Taylor, De Menezes, & Mcsharry, 2006; Tsaur, 2003) and Autoregressive Integrated Moving Average (ARIMA) (Tangsucheeva & Prabhu, 2013). These widely used time series models (especially the ARIMA model) are generally applicable to linear modelling, but they do not capture the non-linearity inherent in time series data (Jaipuria & Mahapatra, 2014). These traditional forecasting methods can fail to forecast big data sets if they do not follow these assumptions like sample size, linearity, stationarity and distribution should be normal.

The second forecasting category is based on machine learning techniques such as artificial neural networks (ANN), support vector machine (SVM) etc. ANN incorporate non-linear models, and can provide superior forecasting model because they efficiently map non-linear relationships between input and output data (Aizenberg, Sheremetov, Villa-Vargas, & Martinez-Muñoz, 2016; Carbonneau, Laframboise, & Vahidov, 2008; Gashler & Ashmore, 2016). Jaipuria and Mahapatra (2014) developed an integrated approach for demand forecasting that used discrete wavelet transforms (DWT) analysis and an artificial neural network (ANN), which generally reduced the bullwhip effect. De Gooijer and Hyndman (2006) provided a comprehensive review of time series forecasting.

In the last few years, several machine-learning methods for accurate demand forecasting in various sectors have been introduced. Generally, demand is assumed to follow a time series pattern which is viewed as a sequence of observed value in a specific pattern. Hence, different types of time series machine learning models such as support vector machine (SVM) (Lu & Wang, 2010), mixed integer programming (Jula & Leachman, 2011), linear and logistic regression (Kone & Karwan, 2011; Moon, Simpson, & Hicks, 2013), artificial neural network (ANN) (Aizenberg et al., 2016; Chong et al., 2015; Doganis, Aggelogiannaki, & Sarimveis, 2008; Gashler & Ashmore, 2016; Jaipuria & Mahapatra, 2014; Moosmayer et al., 2013; Pang, Zhou, Wang, Lin, & Chang, 2018), temporal aggregation (Rostami-Tabar, Babai, Syntetos, & Ducq, 2013) and structural equation modelling (SEM) (Chae, Olson, & Sheu, 2014) have been suggested or proposed for enhancing the accuracy of demand forecasting and eventually remove bullwhip effect by controlling the parameters of the models. Therefore, it is important to enhance the demand forecasting accuracy in such a manner that bullwhip effect which is the key measure of supply chain performance, must be reduced.

Past research has shown that there is only a limited literature available in the market on the use of machine learning for demand-driven forecasting using big data analytics. Demand forecasting is concerned with estimating demand for future time periods, whereas demand sensing assesses the current state of various factors related to customers' choice behaviour by capturing different demand signals using real-time and historic data (Ravikumar et al., 2005). The process of demand sensing makes use of data analytics for supporting the selection, training as well as transformation of unstructured and structured type of data so that the demand and supply are synchronised. This includes diagnosing inconsistencies in demand to reduce demand volatility (Chase Jr, 2013; Folinas & Rabi, 2012; Larson & Chang, 2016). Doganis, Alexandridis, Patrinos, and Sarimveis (2006) developed a model for time series sales forecasting for short shelf-life food products based that used artificial neural networks and genetic algorithms. Aburto and Weber (2007) developed a hybrid approach that included ANN and ARIMA for making forecasting process in supply chain more accurate. Aggarwal, Saini, and Kumar (2009) combined the wavelet transform (WT) and multiple linear regression (MLR) techniques to forecast electricity price profiles using historical price and load data (Ali, Yohanna, Puwu, & Garkida, 2016; Efendigil, Önüt, & Kahraman, 2009; Sivaneasan, Yu, & Goh, 2017) developed fuzzy neural network-based forecasting models and added the demand shaping concept in their models but they did not consider the effect of demand sensing in big data environment. ANN has been already investigated in recent years for demand and sales forecasting problems (Adebiyi, Adewumi, & Ayo, 2014; Aizenberg et al., 2016; Mukhopadhyay, Solis, & Gutierrez, 2012) because of their modelling capability and parallel computing abilities (Wang, Zeng, & Chen, 2015). All these neural network-based forecasting models belongs to MLP, which is a class of feedforward ANN and suggests that demand should extend to time-delay if we use the MLP for demand forecasting case. Despite the several published articles that consider ANN for time series, only a few articles that uses the fuzzy ANN/SVM.

The third time series category is fuzzy set theory based fuzzy methods and they do not require those strict assumptions that are valid for traditional forecasting techniques but provide the remarkable performance in terms of forecasting accuracy. If we compare the ANN with fuzzy ANN, we found a few drawbacks, which are (i) problem in learning rate parameters (ii) problems in parameter tuning in hidden parameters and (iii) problem in setting of overfitting stopping criteria etc. Due to these problems, a faster and better algorithm is proposed by (Yu & Huarng, 2010), fuzzy neural network, and it has given the excellent results in several domains (Lolli et al., 2017). Over the recent years, several articles have been published on feedforward ANN with excellent results on time series dataset and the ability of excellent self-learning without distribution assumptions, but the problem comes after when we work on big dataset because lots of computation power required to optimize the networks, pre-processing and parameters tuning. An extreme learning algorithm has been developed to solve these issues after adding fuzzy inputs, but it has never been tested on demand forecasting problem. Therefore, it is worth to test the extreme learning fuzzy classifier in the demand-driven forecasting context due to their high computational power and excellent predictive power. In time series forecasting case, FNN gives the best performance because FNN combines the ability of the adjusting parameters of ANN with fuzzy logic adaptively and increase the inference's effectiveness. However, ANN and other traditional M/L algorithms have several applicability drawbacks that limit their settings in real environment including: time consuming and slow convergence process for tuning the input/output weights and parameters, errors in setting the learning rate parameters, problems in managing the stopping criteria by number of training/testing epochs and problems for trapping in local minima.

There is a lack of studies that have attempted to develop demand-driven forecasting model and showed how advertising and other promotional marketing strategies alter the demand curve. The main novelty of this article is the adoption of an efficient and error-free fuzzy classifier in this context for improving the forecast accuracy and examine the role of promotional marketing activities as predictors combining historical demand and other factors. Secondly, this proposed work represents an extension of previous demand-driven forecasting work (Chase Jr, 2013) in terms of modified architecture, with the aim of improving our understanding of the linear behaviour of fuzzy neural networks as a historical sales and demand predictor, and encouraging researchers and managers to implement them accurately in real world environment.

3. Proposed framework and model development

3.1. Framework of the research

Fig. 2 presents the proposed big data-driven forecasting framework and it uses a four-step process for predicting the future demand and sales using fuzzy artificial neural network (ANN) machine learning classifier. In the first step, demand signals are captured from historical downstream demand data and synchronize it with other promotional marketing activity data. The data is partitioned into training and validation sets and used to shape the future demand using fuzzy ANN and other traditional forecasting methods to calculate the unconstrained forecast error. In the next step, marketing effectiveness of each advertising medium is calculated using market-mix modelling (MMM) to shift the demand based on multiple capacity constraints collaboration with other important variables related to sales, marketing, operation and historical demand. The marketing-mix modelling is “a statistical approach where quantified marketing activities over time are mathematically linked to a dependent variable, such as sales or revenues”

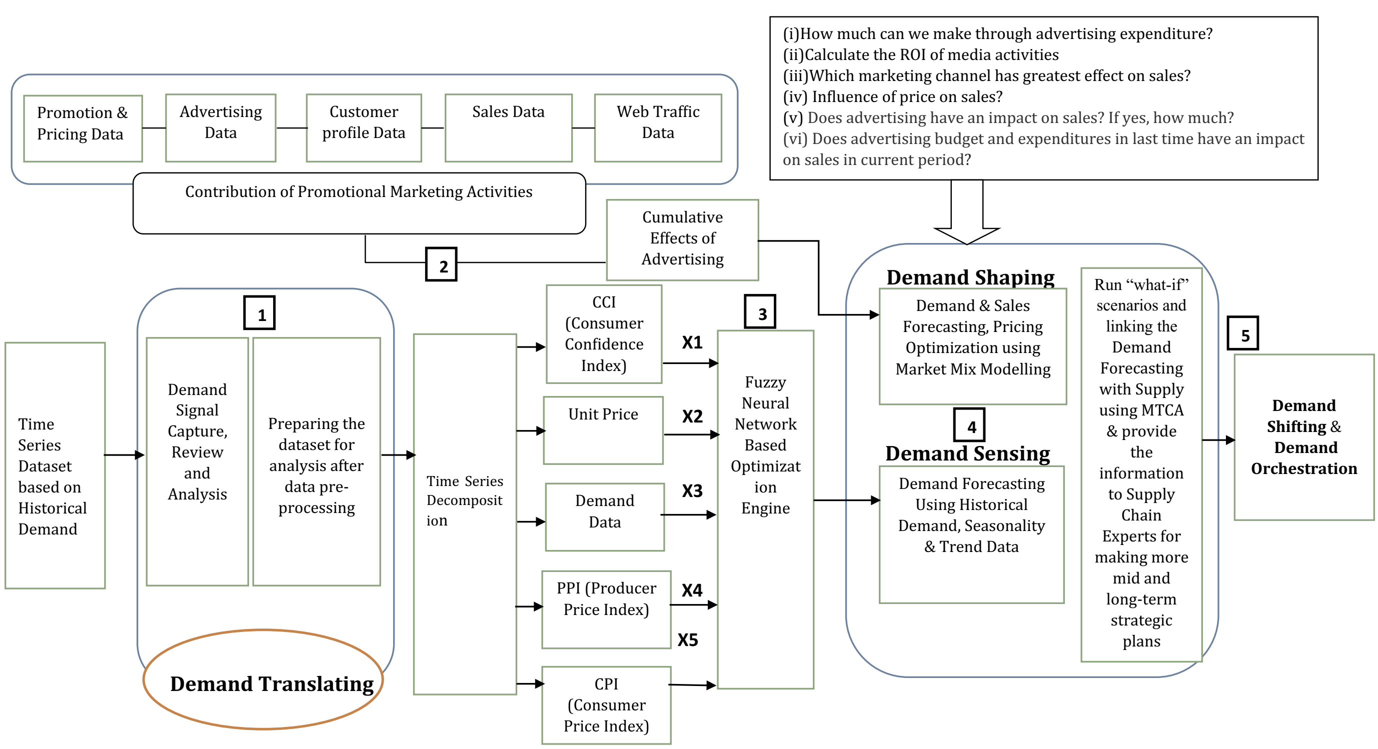

Fig. 2. Proposed big data-driven framework for demand-driven forecasting.

(Fukawa & Zhang, 2015; Wolfe Sr & Crotts, 2011, p. 4). MMM is used to calculate the exposure to an advertisement and response curve relating advertising expenditures to sales data. Exposure is measured in terms of Gross Rating Points (GRP) (a purchase of 100 GRPs could mean that 100% of the target audience is exposed to an advertisement once, or that 50% of the target audience is exposed twice). Marketing Managers evaluate their campaigns and make buying decisions in terms of GRP rather than expenditure because it is not clear how much exposure can be purchased for a given budget (Beltran-Royo et al., 2016, p.28). The effects of advertising extend over several periods after the original exposure. This is referred to as advertising carry-over or 'adstock' (Broadbent, 1984). The effect of an advertisement decays over time and is usually expressed in terms of a half-life (Joseph, 2006). The adstock, model combines past and current advertising effects (Beltran-Royo et al., 2016). This paper studies how advertising and other promotional strategies altered the demand curve for a particular TV brand. A statistical model is developed in the third step adding with sales as a dependent variable and marketing channels, advertising expenditure through marketing channels, seasonality and macro factors as independent variables. In the last step, the big data-driven forecasting model tries to create the final demand response model using what-if analysis and multi-tiered causal analysis (MTCA), which is used to shape and sense the unconstrained demand. In a simple way we can say that if forecasting managers use this proposed big data-driven forecasting model, they can optimize the revenue, handle big datasets, improve forecasting accuracy, calculate the advertising impact on sales, linking the demand and supply data and minimize the marketing investment based on promotional activities and other marketing factors. The developed big data-driven forecasting framework could serve as a guide to give the directions to the forecasting managers for improving the contingency and strategic business planning using accurate and robust prediction of future demand and meaningful and an important initiative in big data supply chain management domain.

3.2. Dataset description and model validation

We tested our models on a time series data, similar to the one used by James Rawlins for a TV manufacturing company. The data set information covering sales, POS and demand data regarding daily TV sales/demand in the categories of demand, sales, supply, POS, advertising expenditure and marketing effectiveness GRP, was monitored over a period of 372 weeks from 2010 to 2017.

In this dataset, James utilized Gross Rating Points (GRP) to determine weekly advertising intensities of different channels and calculated the quantitative impact of marketing campaigns using the impact of adstock. We downloaded the data from James's Github repository (https://github.com/jamesrawlins1000/DDF-MMM–Data/blob/master/MMM_data.xlsx). Apart from information on demand and sales, the TV demand/sales forecasting data also contained different impressions recorded across multiple channels for calculating the advertising campaign's effectiveness such as Mobile SMS, Newspaper Ads, Radio, TV, Poster, Internet, etc. in the form of GRP in a city. This time series dataset had 2614 observations and 17 variables, and a description of each variable was also given in the dataset.

Gross Rating Points (GRPs) can be described as a weighted sum or grand total of the number of ads aired for a TV company or brand in a specific week. For instance, the weights are the rating points or RPs of a specific rating agency for the radio or television shows on or during which the ads were aired or shown. For a television ad, there is a telemeter that calculates the total time duration for which a television set remains on a particular TV channel during the airing of a specific advertisement or commercial on that particular channel in a house. The accuracy of this model is evaluated by checking or examining residuals for diagnosing if it has any repeated pattern that could be eliminated to enhance the chosen fuzzy ANN model.

Although this may appear to be the best model among other models, it is still important to perform a diagnosis to make sure that the big data driven proposed model is adequate and effective. It can be diagnosed by checking and verifying the PACF and ACF of residuals. The online activity metrics undertaken as a part of the process, can either be paid or earned or owned. The data is derived from various resources and the number of ways in which online behaviour can be identified is the clicks that are obtained on paid online ads, the search in the branded form, website visits that can be numbered and the conversations, both positive and negative, that are observed on social media.

While testing our model, we faced three major issues. (i) The first issue was the timing the measurements were to be taken. Online activities and their metrics can only be observed for a limited time-period as they are available for comparatively smaller time frames when compared to the sales or the surveys that are based on attitude metrics. To obtain the best results, we found a weekly level fit and through that, the best signal-to-noise ratio could be achieved. Also by using the integration of weekly intervals, a better control could be obtained. (ii) The second issue was the dynamic relations with respect to the marketing, metrics related to attitude and online activities along with sale. It has been argued by certain scholars that customer attitudes are reflected through their activities. This, with reference to Granger causality, suggests that brand preference and online activity are proportional in nature. Additionally, the increase in online activity is suggestive of the fact that there would be more online activity in the future, which will eventually lead to a rise in the activities associated with marketing and sales. (iii) Lastly, the timing of the effects was identified as the third issue that could not be controlled through any theory unlike marketing, economic, psychological and sociological. To resolve this, the lags were used to bring a balance between model complexities in addition to forecast accuracy under the time series analysis. Also, in order to capture the related results, impulse response function was integrated, and adjustments were done to take into consideration immediate and permanent effects when required. The GRP data collected via different media agencies may vary in levels and thus, it needs to be sliced and aggregated. This can create discrepancies in the accuracy of the data. The data can have certain veracity as the execution might vary from planned aspects like the differences in the planned advertising budget and GRPs.

4. Approaches used in demand-driven forecasting model

This section discusses different approaches adopted in the research to deal with demand forecasting. In addition, this section also provides some contextual information regarding big data, market-mix modelling and advertising adstock. A time series data set contain a set of observations which is generated sequentially in time. All types of organisations utilize the time series data set for forecasting of predicting next year's raw material demand and sales figures. Firstly, the time series model is used for understanding of underlying structure which produce the data, and secondly, to predict the future behaviour after model fitting. During data pre-processing and forecasting process, multiple machine learning forecasting methods are used and tested to develop each forecasting model and then best forecasting process is used in main model to predict the future value and uncover the seasonal patterns or trends. ARIMA (Autoregressive Integrated Moving Average Model) is a regression model which is used for calculating the ARIMA coefficient in stationary data that includes autocorrelation without trend or seasonality. There are four smoothing techniques available for time series forecasting: Exponential, Double Exponential, Moving Average and Holt-Winters. When seasonality is present we should not use the exponential, double exponential and moving average methods. In moving average method extra observations are forecasted by using the average of the previous observations. Double exponential method should be used when a trend is present but there is no seasonality available in data. The Holt-Winters is the most appropriate smoothing method for datasets featuring both trends and seasonality (Walley,

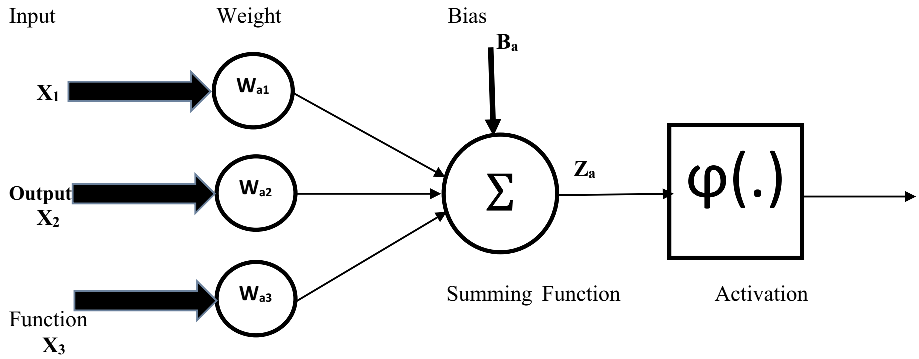

Fig. 3. Artificial Neural Network (ANN) model.

2013; Wang et al., 2015; Wang, Rivera, & Kempf, 2007; Zhang, 2003). MMM is focused on explaining and predicting the sales of the brand that are not just limited to the direct effects of the marketing but also the indirect effects. These have been captured using attitude metrics including awareness, consideration and preference based on surveys. The online activity metrics have also been undertaken as a part of the process and they can either be paid or earned or owned media. The data is derived from various resources and temperature is assumed as the key environmental control variable.

4.1. Fuzzy neural network

Artificial Neural Network (ANN) has several input nodes, hidden nodes and output nodes. Every node takes a weighted average of the output of the previous layer. ANN model consists of a set of synapses each of which is characterized by a weight, an activation function, an adder and a bias. If we choose the Backpropagation neural network then it can be divided into feedforward and backpropagation steps because backpropagation network is a multi-layered and feedforward neural network and considered one of the most popular methods used for multi-layered neural networks in supervised learning. The actual output value of the network is compared to the expected output and an error signal is computed for each of the output. Generally backpropagation network works by approximating the non-linear relationship between the output and input and also adjust the weight values internally according to the input (Garetti & Taisch, 1999; Yu & Huarng, 2010).

Mathematically a neuron can be described by:

$$\mu_a = \sum_{j=1}^{n} W_{aj} X_j \tag{1}$$

and

$$Ya = \phi(\mu_a + B_a) \tag{2}$$

Where μ_a = Linear combiner output due to input and $Za = \mu_a + B_a$.

There are three types of activation function in ANN: Threshold function, Piecewise-Linear function and Sigmoid function.

Threshold Function:

$$\phi(z) = \begin{Bmatrix} 1 & if\, z \geq 0 \\ 0 & if\, z < 0 \end{Bmatrix} \tag{3}$$

Piecewise Function:

$$\phi(z) = \begin{Bmatrix} 1, & z \geq \frac{1}{2} \\ z, & +\frac{1}{2} > z > -\frac{1}{2} \\ 0, & z \leq -\frac{1}{2} \end{Bmatrix} \tag{4}$$

Sigmoid Function:

$$\phi(z) = \frac{1}{1 + \exp(-az)} \tag{5}$$

Where a = slope parameter.

The fuzzy neural network model consists of three different layers namely L1, L2 and L3. The first L1 layer is input layer and the second layer L2 is used to implement the different fuzzy functions. We used a member function with five crisp value variables {very-low, low, medium, high, very-high} for layer 2 and layer 3 which are used jointly to represent the two-layers neural network with feed forward. The net-input and Gaussian activation function is given by

$$f(z; \sigma, a) = \exp\left\{\frac{(z-a)^2}{2\sigma^2}\right\} \tag{6}$$

Where a = mean value and σ = standard deviation for the fuzzy function. The output for L3 is given by

$$Output_j = \frac{1}{1 + \exp\{-(net_i + \varepsilon)\}} \tag{7}$$

These are the steps in fuzzy neural network (Stoeva & Nikov, 2000).

Steps in Fuzzy neural network

Step 1: Generate the weight w for the input and output hidden layer where each wji is a fuzzy number.

Step 2: Let assign value for α and η for training the fuzzy back propagation.

Step 3: Get next pattern set and compute hidden to artificial input and output neurons.

Step 4: Compute the change of weights for input and output hidden layer.

Step 5: Update the weight for input-hidden layer.

Step 6: Update the weight for hidden-output layer.

Step 7: Calculate the output w' and w" for the final fuzzy membership weight sets.

Fuzzy Neural Network applies the fuzzy set theory and neural network together. Zadeh (1965) developed the fuzzy set theory concept and defined it by a membership function for capturing the vagueness in human thoughts. In fuzzy neural network, objective function allocates a membership grade between 0 and 1 to all objects and '∼' is used above a letter for the fuzzy values. In Fig. 3, we use TFN in fuzzy neural network as (*l, m, n*) to give the weights in fuzzy event while for a fuzzy event, parameters *l*, *m* and *n* denote smallest, most promising and the largest possible value respectively (Kumar, Shankar, Choudhary, & Thakur, 2016; Kumar, Shankar, & Debnath, 2015).

These are several definitions of fuzzy membership function discussed:

Definition 1. If $\tilde{A}_1 = (l_1, m_1, n_1)$and $\tilde{A}_2 = (l_2, m_2, n_2)$ then we can present the Boolean functions for two Triangular Fuzzy Numbers in the following way Zadeh (1965):

$$\tilde{A}_1 \ominus \tilde{A}_2 = (l_1, m_1, n_1) \ominus (l_2, m_2, n_2) = (l_1 - n_2, m_1 - m_2, n_1 - l_2) \tag{8}$$

$$\tilde{A}_1 \oplus \tilde{A}_2 = (l_1, m_1, n_1) \oplus (l_2, m_2, n_2) = (l_1 + l_2, m_1 + m_2, n_1 + n_2) \tag{9}$$

$$\tilde{A}_1 \otimes \tilde{A}_2 = (l_1, m_1, n_1) \otimes (l_2, m_2, n_2) = (l_1 l_2, m_1 m_2, n_1 n_2) \tag{10}$$

$$\tilde{A}_1 \oslash \tilde{A}_2 = (l_1, m_1, n_1) \oslash (l_2, m_2, n_2) = (l_1/n_2, m_1/m_2, n_1/l_2) \tag{11}$$

$$\lambda \otimes \tilde{A}_1 = (\lambda l_1, \lambda m_1, \lambda n_1) \ where \ \lambda > 0 \tag{12}$$

$$\tilde{A}_1^{-1} = (l_1, m_1, n_1)^{-1} = \left(\frac{1}{n_1}, \frac{1}{m_1}, \frac{1}{l_1}\right) \tag{13}$$

Definition 3. Membership function of a (TFN) $\tilde{A}$, represented by (l, m, n), is defined as

$$\mu_{\tilde{A}}(x) = \begin{cases} \dfrac{x-l}{m-l}, & l \le x \le m, \\ \dfrac{n-x}{n-m}, & m \le x \le n, \\ 0, & otherwise. \end{cases} \tag{14}$$

Degree of membership for a fuzzy number is:

$$\tilde{A} = (A^{L(y)}, A^{R(y)})$$

$$\tilde{A} = (l + (m-l)y, n + (n-m)y), \ y \in [0, 1] \tag{15}$$

4.2. Multi-tiered casual analysis (MTCA)

Multi-tiered casual analysis (MTCA) is a decision support system (DSS), which is used to model the push-pull effect and it connects the several regression models using a common variable "sales" to analyse the business prospective in entire supply chain. MTCA captures all phases of a supply chain process and integrates supply and POS data within a single framework with advertising strategies. MTCA uses two casual models for linking the demand data with supply: first casual model is used for forecasting the sales by focusing the advertising expenditure and promotion activities data and second model is used for shipment forecasting by taking time lag between point of sales data and other promotion data (Chase Jr, 2013). In the second phase, MTCA links the demand with shipment or supply data and develops another demand forecasting model by adding all the advertising campaigns, marketing channels and seasonal data in shipment forecasting process. In MTCA, if demand is:

$$\begin{aligned} D &= \alpha_0(\text{Const.}\ ) + \alpha_1(\text{Adverts Exp.}) + \alpha_2(\text{Promotional Activities}) + \alpha_3 \\ &\quad (\text{Sales}) \\ &+ \alpha_4(\text{Price}) + \alpha_5(\text{Seasonality}) + \ldots\ldots\alpha_n(\text{Other Variables}) \end{aligned} \tag{16}$$

Then shipment could be:

$$\begin{aligned} S &= \alpha_0(\text{Const.}\ ) + \alpha_1.\ \text{Demand (Lag 1 Period)} + \alpha_2 \\ &\quad (\text{Promotional Sales}) + \alpha_3(\text{Dealer Price}) \\ &+ \alpha_4(\text{Other Discounts}) + \alpha_5(\text{Seasonality}) \end{aligned} \tag{17}$$

4.3. Big data-driven market mix modelling, bullwhip effect and advertising adstock

Promotional activities and sales are the important driving factors of all successful businesses. Marketing managers use the promotional activities to increase the demand of a specific product. These marketing activities require additional costs in the form of advertising expenditures and discounts coupons. Effective marketing activities lead to increased sales because consumers buy more items as they are getting more discounts. Trapero, Pedregal, Fildes, and Kourentzes (2013), Ma, Fildes, & Huang, 2016 and Trapero, Kourentzes, and Fildes (2015) discussed about the performance of some forecasting promotional sales techniques and investigated how marketing channel discounts affect the buying behaviour of consumers when pricing conditions suddenly change in the market. Many research studies have been carried out for quantifying the effect of the promotional activities on sales. Also, various forecasting models have been proposed to moderate the quantitative effect of marketing channels (Kourentzes & Petropoulos, 2016; Ma et al., 2016; Sagaert, Aghezzaf, Kourentzes, & Desmet, 2018; Yan & Wang, 2012).

Market mix modelling method is playing an important role for calculating the quantitative effect of digital marketing channels because of the direct relationship between sales and cost dataset. Regression models generally fail to capture the quantitative effect of these digital media channels as they are not capable to measure the long-term effects on future demands. Advertising effectiveness and Return on Investment (ROI) is usually measured or evaluated using econometric models that compute the influence or impact of varying advertising GRPs on sales.

Lagged or Decay Effect and Diminishing Returns or Saturation Effect are the dimensions that are used to measure or compute advertising adstock. In our model, we compute the advertising adstock at different levels of alpha and include all of these values into dataset, but finally we select only one level of advertising adstock value for our model's development. In this advertising adstock calculation method, alpha is used to represent the campaign stickiness or effect of the campaign. Additionally, advertising dataset impact variable is used for showing the effect of sales and brand value score in quantitative domain. In the model, we see that the quantitative effectiveness of an advertisement decays with respect to time because decay rate depends on the advertising length, advertising impression and advertising frequency. So, if frequency of an advertisement is lesser, then the decay rate will be higher. Generally advertising effect of a particular advertisement becomes 0 after 3–4 weeks so we chose a window of 4 weeks because frequency of an advertisement may be high if customers are not watching at that time in a particular case.

Big Data Analytics can be defined as the process of assessing the massive, substantial datasets comprising of many different types of data. This process is used for uncovering as well as analysing hidden patterns, unknown correlations, customer preferences, market trends and other productive information. The main aim of this process is to help companies make more informed and strategic decisions related to their business. The process involves data researchers and scientists, who analyse massive data that may still be unexplored or unused, for developing better marketing strategies, increasing competitive edge, and enhancing customer service. The characteristics like variety, variability, volume, veracity and velocity can be used to describe big data.

Big data can be structured, semi-structured or un-structured. Structured data can be represented in the form of a table with rows and columns, for example, relational data and financial data. On the other hand audio, video, email, Facebook and Twitter data are the examples of un-structured dataset in the big data domain. Hadoop is used to handle this big dataset. It is an open source software framework used for parallel computing. Hadoop generally distributes both the dataset and processing across hundreds of servers. Hadoop Distributed File System (HDFS), which is a part of Hadoop, works as a storage system. There are multiple data nodes in HDFS. Map reduce, another Hadoop component, refers to two different tasks that Hadoop performs - mapper and reducer. Mapper takes data, breaks it into key/value pairs, and then transfers these key/value pairs to the reducer function (Chen, Chiang, & Storey, 2012; Kumar et al., 2016; Rehman, Chang, Batool, & Wah, 2016).

MMM or Marketing Mix Modelling involves three main steps: (i) assessing the past influence or impact, (ii) forecasting future influence that different marketing mix can have on sales, and (iii) predicting the impacts of future strategies. Marketing mix modelling is usually used for optimising the promotional strategies and advertising mix concerning sales and profits. In MMM, a historical model is estimated based on historical data, with sales as a dependant variable and other marketing activities, seasonality, price and other factors as independent variables. After applying the MMM we can give the answers to some of the following questions:

- How much have we made through advertising?
- What is the ROI of marketing effectiveness activities?
- What is the advertising impact on sales?
- Which marketing drivers have had the greatest effect on sales and demand?
- What are the diminishing returns of advertising on advertising effects?
- Are we optimally allocating our budget across products?
- Are we able to understand the economic, competitive, seasonal and operational factors that have an impact on sales?

To understand MMM, let us consider a regression model in which we take the sales value as the dependent variable and advertising through different channels as an independent variable. This can be represented as

$$X_t = \alpha + \beta Y_t + \zeta_t \tag{18}$$

where X = Dependent variableY = Advertisingα, β (Parameters)

So, the log-log model is one adaptation of this regression model that is necessary when we apply it to the market for measuring the quantitative effect of marketing campaigns through different channels.

In marketing mix modelling, we decompose sales into incremental components and base components. If we want more theoretical explanation of the market, then we use the log-linear regression model, otherwise we use the multiplicative model. Suppose the total value of volume is 16, a = 4 and b = 3 where base = 1. Now we can calculate the effect of a and b. If we delete the 'a' effect, we get the 16/4 = 4 volume. So, the effect of 'a' is 16–4 = 12 units. Similarly, we can calculate the effect of 'b' with 16/3 = 5 volume. So, the effect of 'b' is 16–5 = 11 units. Suppose the actual value of volume in a particular week is X_a and contribution of regression is X_p. Now we must distribute the difference (X_a - X_p) among the multiple factors. If we take three contributors (V_1, V_2 and V_3) with (C_1, C_2 and C_3) as individual contribution (Tellis, 2006), we can calculate the individual contribution as follows:

$$Y_1\!: \; C_1 + (C_1 * (X_a - X_p))/(C_1 + C_2 + C_3) \tag{19}$$

$$Y_2\!: C_2 + (C_2 * (X_a - X_p))/(C_1 + C_2 + C_3) \tag{20}$$

$$Y_3\!: C_3 + (C_3 * (X_a - X_p))/(C_1 + C_2 + C_3) \tag{21}$$

And the multiplicative model

$$X_t = \mathrm{Exp}(\alpha) * A_t^{\beta_1} * Q_t^{\beta_2} * \zeta_t \tag{22}$$

$$\mathrm{Log}\, X_t = \alpha + \beta_1 \log(A_t) + \beta_2 \log(Q_t) + \zeta_t \tag{23}$$

The exponential attraction models

$$M_i = \mathrm{Exp}(V_i)/\Sigma_j \mathrm{Exp}\, V_j \tag{24}$$

where M_i = Market share.

V_j = Marketing efforts of jth brand, thus

$$V_i = \alpha + \beta_1 A_i + \beta_2 P_i + \beta_3 R_i + \beta_4 Q_i + e_i \tag{25}$$

By substituting the value of Eq. (17) in Eq. (16)

$$M_i = \mathrm{Exp}(V_i)/\Sigma_j \exp\ V_j \tag{26}$$

$$M_i = \mathrm{Exp}(\Sigma_k \beta_k X_{ik} + e_i)/\Sigma_j \exp\,(\Sigma_k \beta_k X_{ik} + e_i) \tag{27}$$

where X_k = (0 - n) independent variables of market mix modelling and$\alpha = \beta_0$, $X_{i0} = 1$.

The effectiveness of advertising and ROI achieved from advertising is evaluated using econometric models that determine the impact of varying GRP levels on sales. As mentioned earlier, Lagged or Decay Effect and Diminishing Returns or Saturation Effect are the two dimensions that are used to measure or compute advertising adstock. Evaluation of advertising half-life allows managers to schedule advertising efficiently for maximizing the influence of every advertising campaign. Advertising saturation evaluation indicates current levels of advertising and helps managers identify the investment amount required for making advertising more productive (Beltran-Royo et al., 2016).

Bullwhip effect is caused by fluctuations in data or information which is supplied to companies that are further up the supply chain. The fluctuations in information or distorted data leads to inaccurate demand forecasting by companies.

$$\textit{Bullwhip Effect} = \frac{\textit{Variance of order}}{\textit{Variance of demand}} \tag{28}$$

$$\textit{Net stock Amplification} = \frac{\textit{Variance of net stock}}{\textit{Variance of demand}} \tag{29}$$

As we know if bullwhip effect = 1, it means we don't have variance amplification, so our order variance is equal to demand variance. If bullwhip effect < 1, it signifies that the orders are less than the demand and if bullwhip effect > 1, it is said that the bullwhip effect exists.

The bullwhip effect is either greater than less than one when we predict the demand by using double exponential and ARIMA methods. It is approximately equal to one when we predict the demand using our proposed fuzzy neural network-based model (Chen, Drezner, et al., 2000). Our proposed approach is very fast and calculates the accurate advertising adstock ratio with minimum iteration up to significant decimal value. We have focused on optimum adstock ratio which can be a reason of negative correlation with advertising sales data over saturated market and can be seen in model development with the following steps: (1) First, we take adstock rate and one coefficient of error, and then try to model the sales variable. (2) Second, we have five advertising channels and we use the nonlinear least squares method to calculate the optimal rates for all the advertising channels. (3) Then, we calculate the best fitting adstock rate of each advertising medium and identify the best advertising medium in our dataset, which is TV advertising.

In adstock function, we add some variables (advertising outcome), Advertisement (five advertising mediums in our case) and calculate the adstock rate of all advertising channels, base value and mean square error (MSE) of outcome. All companies are willing to measure the effectiveness of their marketing channels, but it is difficult due to the lag in a big dataset between customer response and advertisement exposure. In this case, we use the regression analysis taking TV, radio, online channels, print media etc. GRP points as the independent

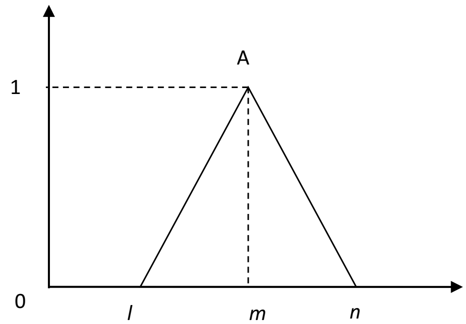

Fig. 4. Triangular fuzzy number.

variables and sales as the dependent variable to identify the best advertising channel, diminishing returns and impact of last week or month's advertising expenditure on current week or month.

5. Result and discussion

This big-data driven contextual investigation exercises is an application of demand shaping, demand sensing and accurate demand forecasting on a TV manufacturing supply chain data set. To begin with, we explored sales and advertising data and attempted to calculate the advertising impact on deals and best advertising channel to improve the model efficiency (demand sensing) and utilized this information for further tasks (demand shaping) to optimize the ROI and profitability. Six imperative questions have been discussed;

1. What techniques are utilized to compute the impact of advertisements in market mix modelling (demand sensing)?
2. How can we be able to test diminishing returns or adverting impact on a big data set (demand sensing)?
3. How can we determine the most important advertising channel that has the most noteworthy effect on sales (demand sensing)?
4. How can we be able to utilize this data for improving the demand forecasting process and overall profitability of our business (demand shaping)?
5. How can we enhance the accuracy of the model and machine learning classifier performance?
6. What sorts of strategies or scenarios can be utilized to link the demand forecasting with supply and profitability maximization?

Now we begin to build up the predictive models and reveal the hidden patterns on this data set. We examine the data set in Table 2 and set it up for predictive modelling after removing the outliers and missing values. We used the DataRobot for all data pre-processing tasks. DataRobot is an automated machine-learning tool, which automates many time series forecasting tasks, including: feature engineering, target variable transformation, detecting the stationarity and seasonality and implementing the back testing. We take a logarithm to distinguish the patterns and break down this data set in Figs. 4, 5 and 6 for recognizing the patterns and regularity. Then, we plot the ACF and PACF values after examining the weekly correlation. We get a direct relationship between supply data and demand data with the value of r coefficient (0.6010). The validation of this result is carried out by making use of a linear regression model with demand data and supply data as independent and outcome variables respectively and we found strong connection between demand and supply (p- value = 0.012 and R square value = 0.038). Now, ARIMA, ACF and PACF plots for residuals are used to inspect the model fitting and examine if ACF and PACF plots lie in the band of LCL and UCL or not. Provided this is true, the residuals are random, and the model is adequate. On the contrary, if the plots do not lie in the band, then the model needs to be improved. An explicit pattern of PACF depicts a trend in the dataset. In any case, since the examples don't rehash, we can conclude that the dataset does not have any regularity since autocorrelation diminishes as the number of lags increment.

Then, we continued on to fitting the model and got the data that we don't have any regularity in the data. This suggests that the autocorrelation is nearly static because it is decreasing as the number of lags are increasing and the PACF plot presented a large value, but successive lags had minimal plot values. It is usually hard to ascertain whether a time series is established from a nonlinear or linear underlying process or a single specific technique ends up being more successful than alternate strategies in the forecasting process. It is troublesome for the forecasting managers to select the right technique for the big data set, so they iterate with different methods and the best one with the most exact outcome is selected.

The paramount problem in an inquisitive project is to verify the model performance on validation data so all our forecasting models are built on 80% training data and the predicting power of the model is tested on the remaining 20% validation data. Table 3 shows the actual and forecasted values after applying ARIMA and other forecasting methods on validation demand data. In general, the fuzzy neural network-based forecasting model is sufficiently accurate with a MAPE of 4.49. The Lower and Upper values refer to the lower and upper limits of the confidence interval with the probability of the forecasted value in this range being 95%. The time plots in Figs. 7–11 depict how the model, which has been fitted using the training set, performed on the validation set.

The aptness of our model for prediction is validated by the fact that the forecasted and actual values lie in close proximity. Table 3 presents the results of predicted value after applying ARIMA, random forest, SVM, ANN, regression and our proposed fuzzy neural network approach. In this work, we used the fuzzy neural network specification suggested by (Jang, 1993). Fresno developed a package in R programming with ANFIS documentation based on the instructions suggested by (Jang, 1993).

Following Jang (1993), Stoeva and Nikov (2000) and Fresno et al. (2015), we divided the whole dataset into training and test datasets.

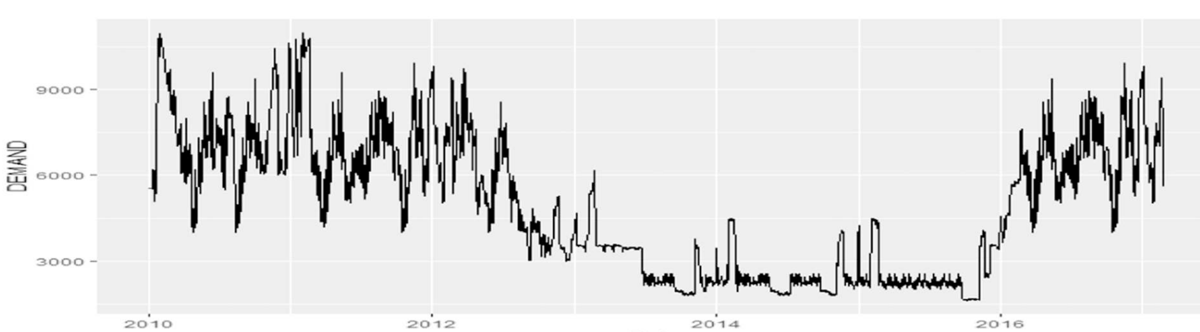

Fig. 5. Time series plot on demand data.

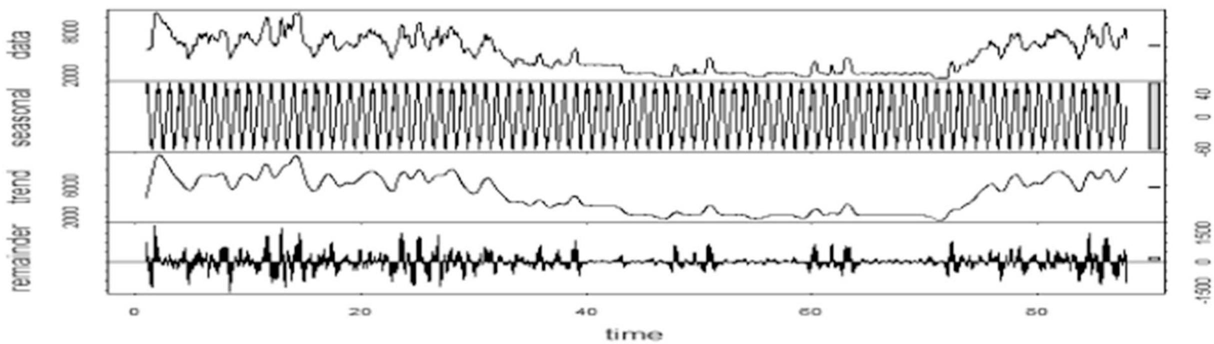

Fig. 6. Time series decomposition.

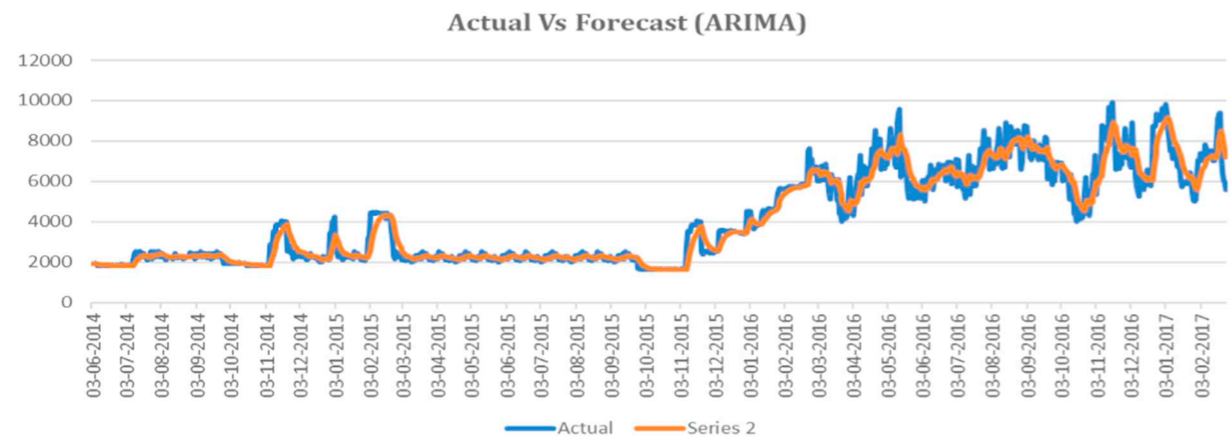

Fig. 7. ARIMA Model output.

The inputs of this dataset are normalized in arbitrary triangular fuzzy numbers for improving the performance of the simple neural network with new triangular fuzzy cases. We consider the five years data as the training data and 2090 observations (80%) are selected for this purpose and the next 523 observations (20%) are used for testing purpose to gain more accurate forecasting results. We divided the training data into five equal and smaller subsets and each set has 16% of the original training data set. Each subset contains 418 instances and these five folds participate in the cross-validation process during the time series-forecasting model building process. In the first cross-validation iteration, we find the subsets 1, 2, 3 and 4 as training data and subset 5 as the testing data. Then, we trained several forecasting models on subset 5 and calculate the performance indicators showing that how well our proposed model correctly identified outputs in the test set. We repeated this process five times, trained each model using five different combinations, and choose the model with best accuracy. These fuzzy models are developed using ANFIS package in R programming in neural network module. Arbitrary triangular fuzzy numbers and normalized gaussian function were used to calculate the weights in each iteration and to minimize the mean squared error respectively. Columns 8 and 10 in Table 3 represent the forecasted value and error after considering the demand shaping effect in this demand-driven forecasting model.

Table 5 presents the result when the optimal method is selected followed by it being trained by fuzzy neural network the results are shown in terms of MSE, MAPE and MAD. Results show that the proposed methodology gives slightly better conjectures for this specific big data-driven demand-driven forecasting model than contemporary models as ARIMA, neural network, random forest, MLP and other traditional forecasting methods and convincingly outperforms them when we link the demand with supply and consider the demand shaping

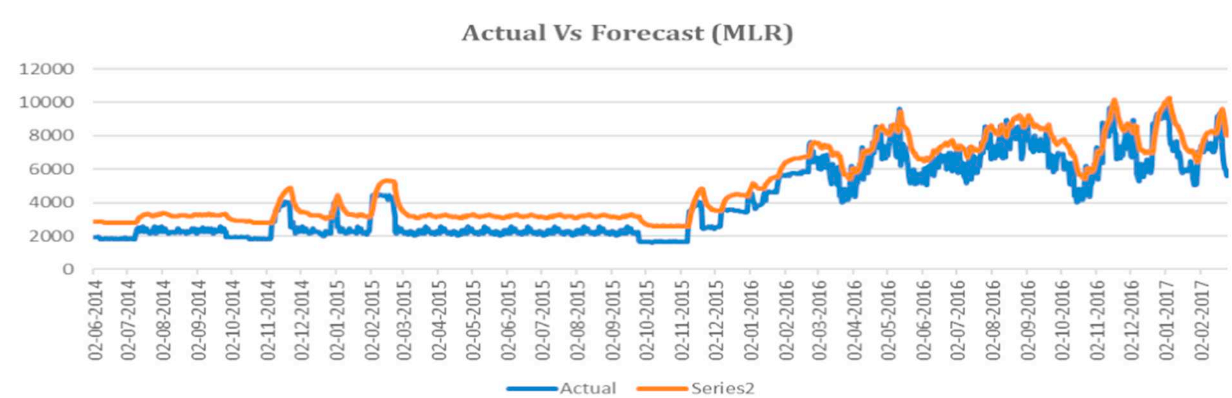

Fig. 8. Multiple linear regression (MLR) Model output.

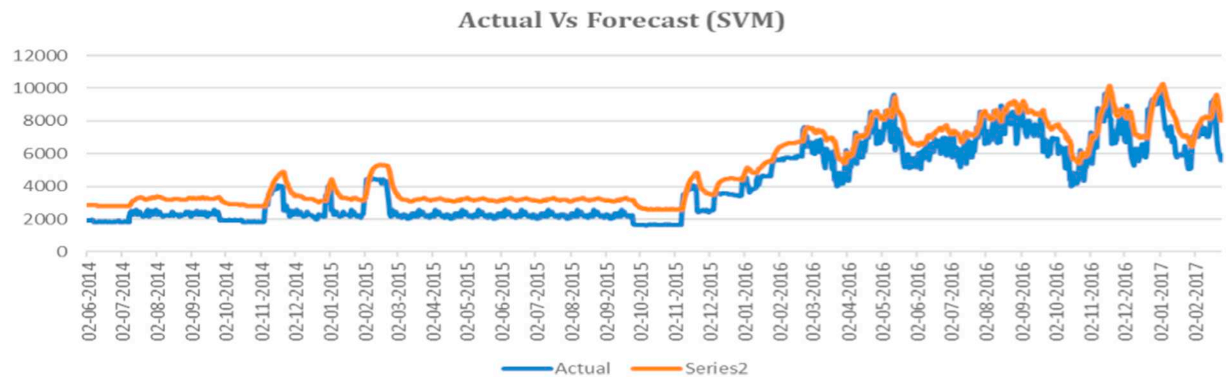

Fig. 9. SVM model output.

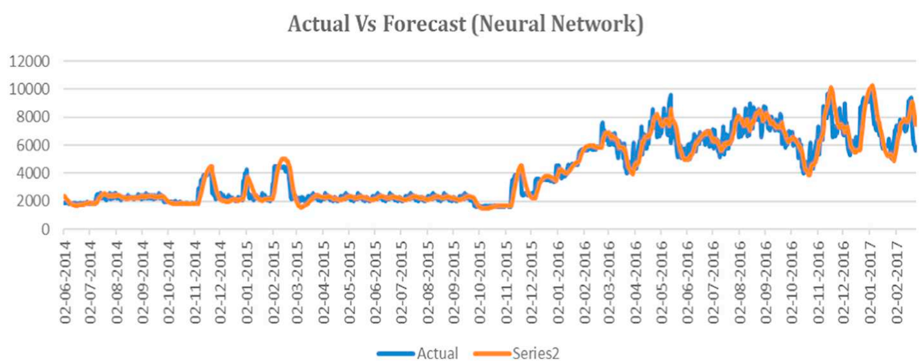

Fig. 10. Neural network model output.

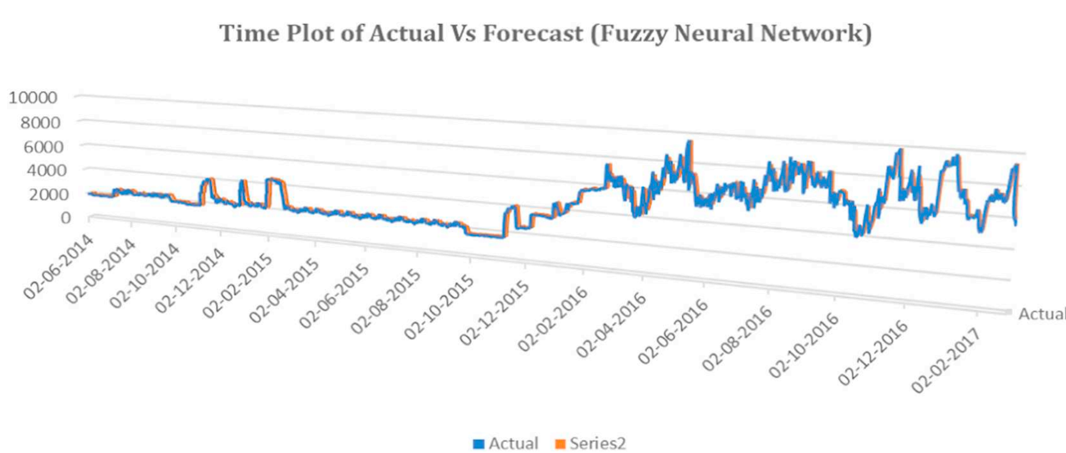

Fig. 11. Fuzzy neural network model output.

effects. The result of fuzzy neural network-based framework on the TV manufacturing, supply and demand dataset in the form of time plot has been shown in Fig. 11. Figs. 9 and 10 show the time plot between actual and forecasted values of neural network and random forest models on the training dataset. After gauging the predictive power of the developed model, we refit it on whole dataset. Then, we start to work on demand shaping and test multiple what-if scenarios taking one variable at a time and fixing others as constants. We develop and test six what-if conditions:

Table 1
What-if scenarios.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Predicted cases in year 2016	917,412	926,740	827,088	937,772	944,320	976,431
Demand cases in 2015	896,961	896,961	896,961	896,961	896,961	896,961
Change	20,451	29,779	−69,873	40,811	47,359	79,470
% Change	2.28%	3.32%	−7.79%	4.55%	5.28%	8.86%

Table 2
Dataset description and descriptive statistics.

	Demand	Consumer price index (CPI)	Consumer confidence index (CCI)	Producer price index (PPI)	Unit price ($)	POS/supply data	SALES ($)	Ad Exp-SMS($)
Count	2613.00	2613.00	2613.00	2613.00	2613	2613	2.613000e+03	2613.000000
Mean	5021.43	102.60	103.15	102.23	363.27	4522.97	1.641507e+06	60.388495
Std	2681.19	1.38	3.11	2.03	26.37	2603.99	9.416673e+05	13.557190
Min	1610.00	101.300	96.30	99.50	282.14	1510	4.627096e+05	37.916700
25%	2436.00	101.40	102.70	100.40	361.60	1776	6.717679e+05	47.554100
50%	4636.00	102.30	103.60	102.70	361.62	4412	1.605095e+06	61.267900
75%	6834.00	103.40	104.60	103.50	361.62	6266	2.267206e+06	71.012200
Max	18,565.00	106.50	107.90	107.20	400.10	16,482	5.960221e+06	89.728300

	Ad Exp-newspaper ads($)	Adv Exp-radio($)	Adv Exp-TV($)	Adv Exp-internet($)	GRP (newspaper ads)	GRP(SMS)	GRP(Radio	GRP(Internet)	GRP(TV)
Count	2613.000000	2613.000000	2613.000000	2613.000000	2613.000000	2613.000000	2613.000000	2613.000000	2613.000000
Mean	12.651179	88.074643	1324.501468	3079.184067	505.296876	30.618203	139.437260	286.228665	1146.114128
Std	1.117926	12.569956	123.677327	1520.891014	488.147782	31.570754	146.042432	138.406620	822.991163
Min	10.027128	62.968800	1067.155700	0.000000	5.659000	0.000000	66.863600	191.181800	697.636400
25%	11.885714	78.098100	1251.250000	2226.429000	114.957000	0.000000	95.136400	234.954500	849.545500
50%	13.186582	84.129200	1380.696200	3302.667000	221.528000	26.119600	109.636400	261.454500	928.545500
75%	13.437482	98.639000	1416.171400	4237.095000	854.310000	46.782000	126.045500	291.318200	1114.863600
Max	14.104193	118.467700	1479.456500	6354.571000	1791.183000	144.995100	1169.409100	1540.429400	7307.318200

1. First case considers no adjustments in advertising expenditures and other activities and assumes it on the same level as last year, 2015.
2. Second case considers 5% hike in the unit price and 10% hike in advertising expenditure in different marketing channels.
3. In Second case, 10% increase in the unit price and a 10% drop in advertising expenditures in different marketing channels is considered.
4. Third case reflects on a 10% surge in TV advertising expenditures and a 5% decrease in all other marketing channels with no increment in unit price.
5. Fourth case considers an increase of 10% in the unit price and TV advertising expenditures and a decrease of 10% in all other marketing channels' expenditures.
6. Fifth case considers an increase of 10% in the unit price, advertising expenditures of TV and print media both and a 10% drop in all other marketing channels' expenditures.

Scenario 6 is established as the optimal one in terms of profitability maximization when all what-if scenarios are examined on 2016 data as depicted in Table 1.

In the next step, we create the future demand-forecasting model based on scenario 6 demand data set. In this phase, we develop a model to link the supply data with demand forecast to complete the demand-driven forecasting model. We add the original demand and scenario 6 predicted demand in this model and develop the demand-shaping phase. Table 4 discusses the supply forecasting results on validation data set. After developing this forecasting model, we refit this model with the whole data set l as undertaken before on the demand data set. The future demand forecasts are based on scenario 6 in the demand-shaping phase. We consider all other variables of this data set at the level of a year ago and observe better supply forecasting results from 857,988 (2015) cases to 902,174 (2016) cases, an increase of 5.15%. This fuzzy neural network-based demand-driven forecasting framework is very effective in the case when we collect the data set from both demand, supply and marketing sides.

6. Conclusion and future direction

Accurate forecasts and the right combination of forecasting activities are very important for an efficient inventory optimization model. This holds especially true for demand-driven forecasting case and how well we develop all statistical and machine learning models and pinpoint which model predicts the future demand with high accuracy, when backorder cost is high, holds paramount importance. We can break it down as the fact that the selection of a forecasting technique among various available methods should be based on following inquiries about the process (i) accuracy (ii) repeatability (iii) automated and capacity to handle big data set (iv) does it handle the question- why behind performance? By integrating machine learning and advanced analytics techniques into traditional demand forecasting, the managers can reap additional dividends with elevated accuracies, things that have eluded them in the past. With big data-driven framework, forecasting managers can optimize the inventory, improve profitability and launch specific region and customer-based data-driven sales plan which can also be used to predict new products demand.

This article improves the demand-driven forecasting model developed by (Chase Jr, 2013) to create a picture of accurate future demand prediction and calculates the diminishing returns of advertising effects. Demand-driven forecasting model uses the sales data along with demand and supply data and calculates the advertising impact on sales. High modelling capability, better computing abilities and flexibility make this proposed big data-driven forecasting model distinctive among contemporaries. In this research, after training a back-propagation neural network-based demand-driven forecasting model by fuzzy inputs, benchmarking is carried out on a time series data, by combining demand, supply, promotional campaigns, and practical sales data with

Table 3
Performance matrices result on demand data-set.

Date	Actual value	Predicted value by ARIMA	Predicted value by SVM	Predicted value by ANN	Predicted value by random forest	Predicted value by MLR	Predicted value by fuzzy neural network	Forecast error by neural network	Forecast error by fuzzy neural network
02-01-2017	9072	9087	9179	9098	9123	9229	9077	26	5
03-01-2017	9809	9824	9916	9833	9860	9966	9818	24	9
04-01-2017	9386	9401	9493	9416	9437	9543	9399	30	13
05-01-2017	8578	8593	8685	8631	8629	8735	8588	53	10
06-01-2017	8057	8072	8164	8112	8108	8214	8067	55	10
07-01-2017	7493	7508	7600	7555	7544	7650	7506	62	13
08-01-2017	7622	7637	7729	7699	7673	7779	7633	77	11
09-01-2017	7103	7118	7210	7207	7154	7260	7115	104	12
10-01-2017	7092	7107	7199	7156	7143	7249	7111	64	19
11-01-2017	7401	7416	7508	7499	7452	7558	7419	98	18
12-01-2017	7681	7701	7810	7789	7732	7870	7723	108	42
13-01-2017	6714	6734	6843	6791	6765	6903	6735	77	21
14-01-2017	7016	7036	7145	7143	7067	7205	7042	127	26
15-01-2017	6756	6776	6885	6889	6807	6945	6772	133	16
16-01-2017	6009	6029	6138	6133	6060	6198	6028	124	19
17-01-2017	5740	5760	5869	5812	5791	5929	5762	72	22
18-01-2017	5985	6005	6114	6098	6036	6174	6012	113	27
19-01-2017	5935	5955	6064	6056	5986	6124	5962	121	27
20-01-2017	5998	6018	6127	6097	6049	6187	6057	99	59

Table 4
Performance matrices result on supply data-set.

Date	Actual value	Predicted value By ARIMA	Predicted value by SVM	Predicted value by ANN	Predicted value by random forest	Predicted value by MLR	Predicted value by fuzzy neural network	Forecast error by neural network	Forecast error by fuzzy neural network
02-01-2017	9226	9241	9333	9252	9277	9383	9231	26	5
03-01-2017	9134	9149	9241	9833	9185	9291	9818	699	684
04-01-2017	9879	9894	9986	9416	9930	10,036	9399	−463	−480
05-01-2017	7212	7227	7319	8631	7263	7369	8588	1419	1376
06-01-2017	7245	7260	7352	8112	7296	7402	8067	867	822
07-01-2017	6280	6295	6387	7555	6331	6437	7506	1275	1226
08-01-2017	6240	6255	6347	7699	6291	6397	7633	1459	1393
09-01-2017	6482	6497	6589	7207	6533	6639	7115	725	633
10-01-2017	6542	6557	6649	7156	6593	6699	7111	614	569
11-01-2017	5759	5774	5866	7499	5810	5916	7419	1740	1660
12-01-2017	6841	6861	6970	7789	6892	7030	7723	948	882
13-01-2017	6456	6476	6585	6791	6507	6645	6735	335	279
14-01-2017	5453	5473	5582	7143	5504	5642	7042	1690	1589
15-01-2017	5884	5904	6013	6889	5935	6073	6772	1005	888
16-01-2017	5866	5886	5995	6133	5917	6055	6028	267	162
17-01-2017	5806	5826	5935	5892	5857	5995	5838	86	32
18-01-2017	5876	5896	6005	6098	5927	6065	6012	222	136
19-01-2017	5834	5854	5963	6056	5885	6023	5962	222	128
20-01-2017	5846	5866	5975	6097	5897	6035	6057	251	211

Table 5
Performance matrices result on demand data.

Model	MSE	MAPE	MAD	Bullwhip effect	Net stock amplification
ARIMA	743.66	40.46	202.21	1.57	1.78
Artificial neural network (ANN)	7.89	9.25	29.36	1.11	0.90
Support vector machine (SVM)	33.26	9.62	48.59	1.31	1.12
Random forest	18.29	11.57	52.37	1.06	0.45
Multiple linear regression (MLR)	117.34	27.36	189.55	1.14	0.63
Fuzzy neural network	6.71	4.49	21.95	0.99	0.37

respect to five standard models as; ARIMA, MLR, Random Forest, SVM and ANN. First, the study shows how the proposed model can be applied for efficient analysis of large data sets, and as a result helps predict products' future demand taking into account advertising expenditure, omnichannel, promotion and sales data for shaping the future demand. Second, our research also examines how much ground we gained through advertising and what is the ROI of our media activities and which marketing drivers have had the most significant impact on demand and sales. Our findings can assist researchers in understanding how advertising works for demand-driven forecasting model. It isn't sufficient to just perceive advertising as how it makes changes or modifications to a solitary parameter, for instance, price elasticity or flexibility of demand in equilibrium or parity. It is additionally critical for one to see and gauge how it changes the total appropriation of WTP i.e. willingness to pay in the population.

This ability can easily provide a marketing plan to marketing manager's planning process that is likely to exceed target revenue and will also assist additionally with the information of what marketing mix levels will increase or decrease present sales to meet or exceed target. Our analysis reveals that the proposed fuzzy neural network-based framework performs much better and provides better forecasting results because of having better adjustment and capturing linear behaviour of time series. This fuzzy ANN model neither requires stationary nature, nor any statistical information of data series and it is applicable to both nonlinear and linear data series. This integrated model's robustness is verified, and it can be easily adopted for enhancing supply chain performance. After considering the demand shaping effect and using our proposed approach, the performance of demand-driven forecasting model also improved exceptionally in the form of MSE and MAD in this study with MSE decreasing from 33.26 to 6.71 and MAD decreasing from 48.59 to 21.95. It is also found that the estimated bullwhip effect value and net stock value in our case is comparatively less than that of all other models. Lastly, this study shows that manufacturers can predict their product demands via online and offline market strategies. Online market is currently one of the fundamental hotspots for selling items and every single organization wishes to concentrate on improving product forecasts and diminishing bullwhip impacts by breaking down the information with advance data mining tools from their physical storefronts. Some possible future directions may include the investigation of the intermittent demand and optimization of the number of hidden layer and hidden nodes parameters in fuzzy neural network-based forecasting model. For future prescriptive, we can likewise include more dimensions in this dataset such as weather forecasts, shipping data, "buy-online and pick up in store" and customer profile data followed by scanning of the effectiveness of advertising on a regular basis and improving the demand-driven forecasting model accuracy by adding the DWT (discrete wavelet transforms) in fuzzy ANN and SVM predictive models.

References

Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing Journal, 7*(1), 136–144.

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics, 2*, 1), 1–7.

Aggarwal, S., Saini, L., & Kumar, A. (2009). Electricity price forecasting using wavelet domain and time domain features in a regression-based technique. *International Journal of Recent Trends in Engineering,* (2), 33–37.

Aizenberg, I., Sheremetov, L., Villa-Vargas, L., & Martinez-Muñoz, J. (2016). Multilayer neural network with multi-valued neurons in time series forecasting of oil production. *Neurocomputing, 175*, 980–989.

Ali, D., Yohanna, M., Puwu, M. I., & Garkida, B. M. (2016). Long-term load forecast modelling using a fuzzy logic approach. *Pacific Science Review A: Natural Science and Engineering, 18*(2), 123–127.

Baharaeen, S., & Masud, A. S. (1986). A computer program for time series forecasting using single and double exponential smoothing techniques. *Computers and Industrial Engineering, 11*(1–4), 151–155.

Beltran-Royo, C., Escudero, L. F., & Zhang, H. (2016). Multiperiod multiproduct advertising budgeting: Stochastic optimization modeling. *Omega, 59*, 26–39 (United Kingdom).

Bhattacharya, R., & Bandyopadhyay, S. (2011). A review of the causes of bullwhip effect in a supply chain. *International Journal of Advanced Manufacturing Technology, 54*(9–12), 1245–1261.

Broadbent, S. (1979). One way TV advertisements work. *Journal of the Market Research Society, 21*(3), 139–166.

Broadbent, S. (1984). Modeling with adstock. *Journal of the Market Research Society, 26*(4), 295–312.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184*(3), 1140–1154.

Chae, B. K., Olson, D., & Sheu, C. (2014). The impact of supply chain analytics on operational performance: A resource-based view. *International Journal of Production Research, 52*(16), 4695–4710.

Chang, V. (2017). Towards data analysis for weather cloud computing. *Knowledge-Based Systems, 127*, 29–45.

Chang, V., & Ramachandran, M. (2017). Financial modeling and prediction as a service. *Journal of Grid Computing, 15*(2), 177–195.

Chase, C. W., Jr. (2013). *Demand-driven forecasting: A structured approach to forecasting.* John Wiley & Sons.

Chen, A., & Blue, J. (2010). Performance analysis of demand planning approaches for aggregating, forecasting and disaggregating interrelated demands. *International Journal of Production Economics, 128*(2), 586–602.

Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. (2000). Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science, 46*(3), 436–443.

Chen, F., Ryan, J. K., & Simchi-Levi, D. (2000). The impact of exponential smoothing forecasts on the bullwhip effect. *Naval Research Logistics, 47*(4), 269–286.

Chen, H., Chiang, R., & Storey, V. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188.

Chong, A. Y. L., Ch'ng, E., Liu, M. J., & Li, B. (2015). Predicting consumer product demands via big data: The roles of online promotional marketing and online reviews. *International Journal of Production Research, 43*(2), 169–186.

Colicchia, C., & Strozzi, F. (2012). Supply chain risk management: A new methodology for a systematic literature review. *Supply Chain Management, 17*(4), 403–418.

Coppini, M., Rossignoli, C., Rossi, T., & Strozzi, F. (2010). Bullwhip effect and inventory oscillations analysis using the beer game model. *International Journal of Production Research, 48*(13), 3943–3956.

Davino, M., De Simone, V., & Schiraldi, M. M. (2014). Revised mrp for reducing inventory level and smoothing order releases: A case in manufacturing industry. *Production Planning and Control, 25*(10), 814–820.

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting, 22*(3), 443–473.

Delhoum, S., & Scholz-Reiter, B. (2009). The influence of decision patterns of inventory control on the bullwhip effect based on a simulation game of a production network. *Production Planning and Control, 20*(8), 666–677.

Doganis, P., Aggelogiannaki, E., & Sarimveis, H. (2008). A combined model predictive control and time series forecasting framework for production-inventory systems. *International Journal of Production Research, 46*(24), 6841–6853.

Doganis, P., Alexandridis, A., Patrinos, P., & Sarimveis, H. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. *Journal of Food Engineering, 75*(2), 196–204.

Duc, T. T. H., Luong, H. T., & Kim, Y. D. (2008). A measure of bullwhip effect in supply chains with a mixed autoregressive-moving average demand process. *European Journal of Operational Research, 187*(1), 243–256.

Duc, T. T. H., Luong, H. T., & Kim, Y. D. (2010). Effect of the third-party warehouse on bullwhip effect and inventory cost in supply chains. *International Journal of Production Economics, 124*(2), 395–407.

Efendigil, T., Önüt, S., & Kahraman, C. (2009). A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Systems with Applications, 36*(3), 6697–6707.

Folinas, D., & Rabi, S. (2012). Estimating benefits of demand sensing for consumer goods organisations. *Journal of Database Marketing & Customer Strategy Management, 19*(4), 245–261.

Fukawa, N., & Zhang, Y. (2015). Profit-sharing between an open-source firm and application developers – Maximizing profits from applications and in-application advertisements. *Industrial Marketing Management, 48*, 111–120.

Garetti, M., & Taisch, M. (1999). Neural networks in production planning and control. *Production Planning and Control, 10*(4), 324–339.

Gashler, M. S., & Ashmore, S. C. (2016). Modeling time series data with deep Fourier neural networks. *Neurocomputing, 188*, 3–11.

Geary, S., Disney, S. M., & Towill, D. R. (2006). On bullwhip in supply chains – Historical review, present practice and expected future impact. *International Journal of Production Economics, 101*(1), 2–18.

Haberleitner, H., Meyr, H., & Taudes, A. (2010). Implementation of a demand planning system using advance order information. *International Journal of Production Economics, 128*(2), 518–526.

Hadaya, P., & Cassivi, L. (2007). The role of joint collaboration planning actions in a demand-driven supply chain. *Industrial Management and Data Systems, 107*(7), 954–978.

Hansun, S., & Subanar, S. S. (2016). H-wema: A new approach of double exponential smoothing method. *Telkomnika (Telecommunication Computing Electronics and Control), 14*(2), 772–777.

Ho, W., Zheng, T., Yildiz, H., & Talluri, S. (2015). Supply chain risk management: A literature review. *International Journal of Production Research, 53*(16), 5031–5069.

Hollmann, R. L., Scavarda, L. F., & Thomé, A. M. T. (2015). Collaborative planning, forecasting and replenishment: A literature review. *International Journal of Productivity and Performance Management, 64*(7), 971–993.

Jaipuria, S., & Mahapatra, S. S. (2014). An improved demand forecasting method to reduce bullwhip effect in supply chains. *Expert Systems with Applications, 41*(5), 2395–2408.

Jang, J. S. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on, 23*(3), 665–685.

Jeon, S., Hong, B., & Chang, V. (2018). Pattern graph tracking-based stock price prediction using big data. *Future Generation Computer Systems, 80*, 171–187.

Joseph, J. V. (2006). Understanding advertising adstock transformations. Retrieved from http://mpra.ub.uni-muenchen.de/7683/.

Jula, P., & Leachman, R. C. (2011). Long- and short-run supply-chain optimization models for the allocation and congestion management of containerized imports from asia to the United States. *Transportation Research Part E: Logistics and Transportation Review, 47*(5), 593–608.

Juttner, U. (2005). Supply chain risk management: Understanding the business requirements from a practitioner perspective. *International Journal of Logistics Management, 16*(1), 120–141.

Khan, A. A., & Shahidehpour, M. (2009). One day ahead wind speed forecasting using-wavelets. *IEEE/PES power systems conference and exposition. 2009. IEEE/PES power systems conference and exposition* (pp. 1–5). Seattle, WA: PSCE '09.

Kone, E. R. S., & Karwan, M. H. (2011). Combining a new data classification technique and regression analysis to predict the cost-to-serve new customers. *Computers and*

Industrial Engineering, 61(1), 184–197.

Kourentzes, N., & Petropoulos, F. (2016). Forecasting with multivariate temporal aggregation: The case of promotional modelling. *International Journal of Production Economics, 181*, 145–153.

Kumar, A., Shankar, R., Choudhary, A., & Thakur, L. S. (2016). A big data mapreduce framework for fault diagnosis in cloud-based manufacturing. *International Journal of Production Research,* 1–14.

Kumar, A., Shankar, R., & Debnath, R. M. (2015). Analyzing customer preference and measuring relative efficiency in telecom sector: A hybrid fuzzy AHP/DEA study. *Telematics and Informatics, 32*(3), 447–462.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management, 36*(5), 700–710.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997a). The bullwhip effect in supply chains. *Sloan Management Review, 38*(3), 93–102.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997b). Information distortion in a supply chain: The bullwhip effect. *Management Science, 43*(4), 546–558.

Li, B., Li, J., Li, W., & Shirodkar, S. A. (2012). Demand forecasting for production planning decision-making based on the new optimised fuzzy short time-series clustering. *Production Planning and Control, 23*(9), 663–673.

Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T., & Gucci, S. (2017). Single-hidden layer neural networks for forecasting intermittent demand. *International Journal of Production Economics, 183*, 116–128.

Lu, C. J., & Wang, Y. W. (2010). Combining independent component analysis and growing hierarchical self-organizing maps with support vector regression in product demand forecasting. *International Journal of Production Economics, 128*(2), 603–613.

Ma, S., Fildes, R., & Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra-and inter-category promotional information. *European Journal of Operational Research, 249*(1), 245–257.

Moon, S., Simpson, A., & Hicks, C. (2013). The development of a classification model for predicting the performance of forecasting methods for naval spare parts demand. *International Journal of Production Economics, 143*(2), 449–454.

Moosmayer, D. C., Chong, A. Y. L., Liu, M. J., & Schuppar, B. (2013). A neural network approach to predicting price negotiation outcomes in business-to-business contexts. *Expert Systems with Applications, 40*(8), 3028–3035.

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*, 721–735.

Pang, X., Zhou, Y., Wang, P., Lin, W., & Chang, V. (2018). An innovative neural network approach for stock market prediction. *The Journal of Supercomputing,* 1–21.

Petrovic, D. (2001). Simulation of supply chain behaviour and performance in an uncertain environment. *International Journal of Production Economics, 71*(1–3).

Rao, S., & Goldsby, T. J. (2009). Supply chain risks: A review and typology. *The International Journal of Logistics Management, 20*(1), 97–123.

Ravikumar, K., Saroop, A., Narahari, H., & Dayama, P. (2005). Demand sensing in e-business. *Sadhana, 30*(2–3), 311–345.

Rehman, M. H.u., Chang, V., Batool, A., & Wah, T. Y. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management, 36*(6), 917–928.

Ross, D. F. (2015). *Demand management. Distribution planning and control.* Springer245–306.

Rostami-Tabar, B., Babai, M. Z., Syntetos, A., & Ducq, Y. (2013). Demand forecasting by temporal aggregation. *Naval Research Logistics, 60*(6), 479–498.

Sagaert, Y. R., Aghezzaf, E.-H., Kourentzes, N., & Desmet, B. (2018). Tactical sales forecasting using a very large set of macroeconomic indicators. *European Journal of Operational Research, 264*(2), 142–155.

Sivaneasan, B., Yu, C. Y., & Goh, K. P. (2017). Solar forecasting using ANN with fuzzy logic pre-processing. *Energy Procedia, 143*, 727–732.

Sodhi, M. S., Son, B. G., & Tang, C. S. (2012). Researchers' perspectives on supply chain risk management. *Production and Operations Management, 21*(1), 1–13.

Sodhi, M. S., & Tang, C. S. (2011). The incremental bullwhip effect of operational deviations in an arborescent supply chain with requirements planning. *European Journal of Operational Research, 215*(2), 374–382.

Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science, 35*(3), 321–339.

Stoeva, S., & Nikov, A. (2000). A fuzzy backpropagation algorithm. *Fuzzy Sets and Systems, 112*(1), 27–39.

Tang, C. S. (2006). Perspectives in supply chain risk management. *International Journal of Production Economics, 103*(2), 451–488.

Tang, O., & Nurmaya Musa, S. (2011). Identifying risk issues and research advancements in supply chain risk management. *International Journal of Production Economics, 133*(1), 25–34.

Tangsucheeva, R., & Prabhu, V. (2013). Modeling and analysis of cash-flow bullwhip in supply chain. *International Journal of Production Economics, 145*(1), 431–447.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society, 54*(8), 799–805.

Taylor, J. W., De Menezes, L. M., & Mcsharry, P. E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting, 22*(1), 1–16.

Tellis, G. J. (2006). Modeling Marketing Mix. *The Handbook of Marketing Research* (pp. 506–522). California: Sage Publication, Inc.

Trapero, J. R., Kourentzes, N., & Fildes, R. (2015). On the identification of sales forecasting models in the presence of promotions. *Journal of the Operational Research Society, 66*, 299–307.

Trapero, J. R., Pedregal, D. J., Fildes, R., & Kourentzes, N. (2013). Analysis of judgmental adjustments in the presence of promotions. *International Journal of Forecasting, 29*, 234–243.

Tsaur, R. C. (2003). Forecasting by fuzzy double exponential smoothing model. *International Journal of Computer Mathematics, 80*(11), 1351–1361.

Verdouw, C. N., Beulens, A. J. M., Trienekens, J. H., & Van Der Vorst, J. G.a. J. (2011). A framework for modelling business processes in demand-driven supply chains. *Production Planning and Control, 22*(4), 365–388.

Wagner, S. M., & Bode, C. (2006). An empirical investigation into supply chain vulnerability. *Journal of Purchasing and Supply Management, 12*(6 SPEC. ISS), 301–312.

Walley, P. (2013). Does the public sector need a more demand-driven approach to capacity management? *Production Planning and Control, 24*(10 − 11), 877–890.

Wang, L., Zeng, Y., & Chen, T. (2015). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems with Applications, 42*(2), 855–863.

Wang, W., Rivera, D. E., & Kempf, K. G. (2007). Model predictive control strategies for supply chain management in semiconductor manufacturing. *International Journal of Production Economics, 107*(1), 56–77.

Williams, B. D., Waller, M. A., Ahire, S., & Ferrier, G. D. (2014). Predicting retailer orders with pos and order data: The inventory balance effect. *European Journal of Operational Research, 232*(3), 593–600.

Wolfe, M. J., Sr., & Crotts, J. C. (2011). Marketing mix modeling for the tourism industry: A best practices approach. *International Journal of Tourism Sciences, 11*(1), 1–15.

Yan, R., & Wang, K.-Y. (2012). Franchisor–franchisee supply chain cooperation: Sharing of demand forecast information in high-tech industries. *Industrial Marketing Management, 41*(7), 1164–1173.

Yu, T. H. K., & Huarng, K. H. (2010). A neural network-based fuzzy time series model to improve forecasting. *Expert Systems with Applications, 37*(4), 3366–3372.

Zadeh, L. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.

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

Zokaei, K., & Hines, P. (2007). Achieving consumer focus in supply chains. *International Journal of Physical Distribution and Logistics Management, 37*(3), 223–247.

Zsidisin, G. A., Panelli, A., & Upton, R. (2000). Purchasing organization involvement in risk assessments, contingency plans, and risk management: An exploratory study. *Supply Chain Management, 5*(4), 187–197.