Logistics Research (2022) 15:6
DOI_10.23773/2022_6

Forecasting Irregular Demand Using Single Hidden Layer Neural Networks

M. A. Hoffmann, R. Lasch, J. Meinig

Received: 31 May 2021 / Accepted: 15 June 2022 / Published online: 15 July 2022

ABSTRACT

The inventory management and production planning of parts with irregular demand patterns are challenging for manufacturing companies. These patterns often occur in the strategically critical spare parts sector, where the inventory and capital commitment costs are high. For this reason, an accurate forecast can improve service levels and ensure efficient stock keeping.

For this problem, time-series-based forecasting methods are often used to predict future demands. Furthermore, the research of recent years in terms of stochastic forecasting also focused on Artificial Intelligence (AI) methods, mainly Artificial Neural Networks (ANN).

In contrast to previous studies, this paper compares the prediction results of various ANN configurations and classical forecasting methods for all of the different demand categories according to Syntetos et al. [1], which means that erratic, lumpy, smooth, and intermittent demands are regarded separately. This study compares eleven statistical forecasting configurations with eight single hidden layer neural network configurations.

Furthermore, the influence of the number of hidden neurons on the prediction performance is investigated with the learning algorithms Backpropagation (BP) and Levenberg-Marquardt (LM) by evaluating them separately, which has not been covered in the context of all irregular demand categories yet. The study is based on actual demand data from 29 spare parts of a mechanical engineering company.

KEYWORDS: Irregular demand forecasting · Artificial Neural Network · Spare parts management · Artificial Intelligence

‘‘‘This article is part of a focus collection on ‘‘Logistics Management 2021: The German Perspective”.

Marcel André Hoffmann (corresponding)
e-mail: marcel.hoffmann@tu-dresden.de

Rainer Lasch
e-mail: rainer.lasch@tu-dresden.de

Johannes Meinig
e-mail: johannes.meinig@tu-dresden.de

Technische Universität Dresden, Chair of Business Management, esp. Logistics, 01062 Dresden, Germany;

1. INTRODUCTION

A demand forecast with minimum possible deviations is required to efficiently design production, sales planning, and warehousing. If no deterministic or causal determination of demand for the future is possible, demand data from the past often represent the best data basis for part-specific forecasts. When facing sudden spare part needs due to spontaneous failures, this stochastic determination of requirements is usually the best available option to balance high service and inventory levels [2]. This is even more crucial since spare parts are often urgent goods and need to be delivered as soon as possible to eliminate malfunctions and shutdowns in operations and production processes.

A wide range of statistical forecasting methods is considered suitable for different application areas. A moving average or first-order exponential smoothing is often used for more constant demand patterns, which are often found in primary products [2]. The forecast deviations are generally very low in such constant patterns, and good results can usually be achieved with these methods, while the optimization potential is low. As shown in the extensive review of Pinçe et al. [3], it becomes more problematic when demand patterns are irregular, as is often the case in the spare parts sector. There, strongly fluctuating demand levels and zero demand periods occur frequently, so that the forecast deviations in general and the requirements for an accurate forecast are considerably higher [4]. For this task, the method of Croston [5], the approximation of Syntetos and Boylan [6], or bootstrapping methods [7] are often used, which can usually achieve better results with irregular demand characteristics. Petropoulos et al. [8] investigated the effect of different input parameters and evaluated the accuracy of other forecasting methods for different demand types.

Apart from classical statistical forecasting methods, Artificial Intelligence (AI) approaches, especially Artificial Neural Networks (ANN), have shown great potential in optimizing particularly lumpy demand predictions in some studies [9]–[12]. The work of Pinçe et al. [3] suggests that ANN are the most relevant AI approaches for demand forecasting problems. Chen et al. [13] compared different neural network types to forecast critical spare part demands and found a moving fuzzy-neuron network to be superior. Lolli et al. [14] showed the excellent performance of the BP algorithm compared with Extreme Learning Machines to predict intermittent demand. Kourentzes [15] evaluated the application of ANN to forecasting intermittent demands regarding prediction accuracy and service level improvements. Carmo and Rodrigues [16] noted accuracy improvements with radial basis function networks compared to Croston's method. In other studies, these superior results of ANN could not be fully confirmed, as classical forecasting methods were partially or generally better [17], [18]. Furthermore, ANN calibration has been highlighted as a relevant task for prediction accuracy [19].

This paper aims to provide a quantitative comparison of classical demand prediction results and ANN approaches applied to time-series data of the different demand categories according to Syntetos et al. [1], which has only been partially considered in the study of Sahin et al. [17] at this point. In contrast to this paper, our study regards all four spare part demand patterns, while more classical forecasting methods such as second-order exponential smoothing or moving average are considered. In total, 19 configurations are compared, composed of 11 classical forecasting configurations and 8 ANN variants. To the best of the author's knowledge, the effect of different numbers of applied neurons in the hidden layer of the ANN has not yet been tested considering the forecasting performance of the different irregular demand categories. This approach is also motivated by the claim of hidden node optimization, which is stated by Lolli et al. [14].

The evaluation is based on real datasets of 29 different spare part demands of a medium-sized mechanical engineering company. The following research questions are considered:

RQ1: Which of the forecasting methods considered shows the best forecasting performance according to the demand categories of Syntetos et al. [1]?

RQ2: Can ANN improve the forecasting performance when applied to all demand categories?

RQ3: What implications for theory and practice can be drawn from the quantitative study and which further research directions occur in irregular demand forecasting with ANN?

The remainder of this paper is structured as follows: Section 2 describes the research design and introduces the field of spare parts management and the classical demand forecasting methods. In Section 3, the configuration of the applied ANN is explained. Section 4 continues with the quantitative analysis and compares the prediction results with different forecasting methods. In section 5, implications for theory, practice, and further research directions are discussed. Section 6 concludes with a summary and limitations of the study.

2. METHODOLOGY AND THEORETICAL FOUNDATIONS

2.1. Research design

The paper is grounded on the methodological approach of Chen et al. [13], which consists of the following three main steps: 1) data collection and analysis, 2) forecasting method evaluation, and 3) derivation and comparison of prediction results.

The dataset contains monthly demand values of 29 strategically relevant parts of the cooperating company covering 42 periods. Afterwards, these parts are assigned to the different demand categories according to Syntetos et al. [1] by calculating the squared coefficient of variation (CV^2) and the average inter-demand interval (ADI).

The authors applied different ANN configurations within the second step and established statistical forecasting methods for demand prediction for the initial dataset. This is followed by the evaluation stage of the forecasting methods, which consists of the selection and adequate parameterization of the various forecasting methods to be considered. The influence of different ANN parameters was tested experimentally. The varied parameters were applied in all combinations, i.e., in the sense of a full factorial experiment.

The third step aims to obtain and compare the forecasting results achieved across all methods. Therefore, the prediction accuracy was evaluated for the most common statistical benchmark forecasting methods and the ANN configurations by applying a relative error measurement method to conclude the prediction performance and answer RQ1 and RQ2.

2.2. Characteristics and Classification of Spare Part Demands

Many parts, especially in the spare parts sector, show intermittent demand patterns [4]. Consequently, there are frequent periods of zero demand, and positive demands occur irregularly. Furthermore, in contrast to the principally constant demand for primary products, spare part demand per period is often characterized by

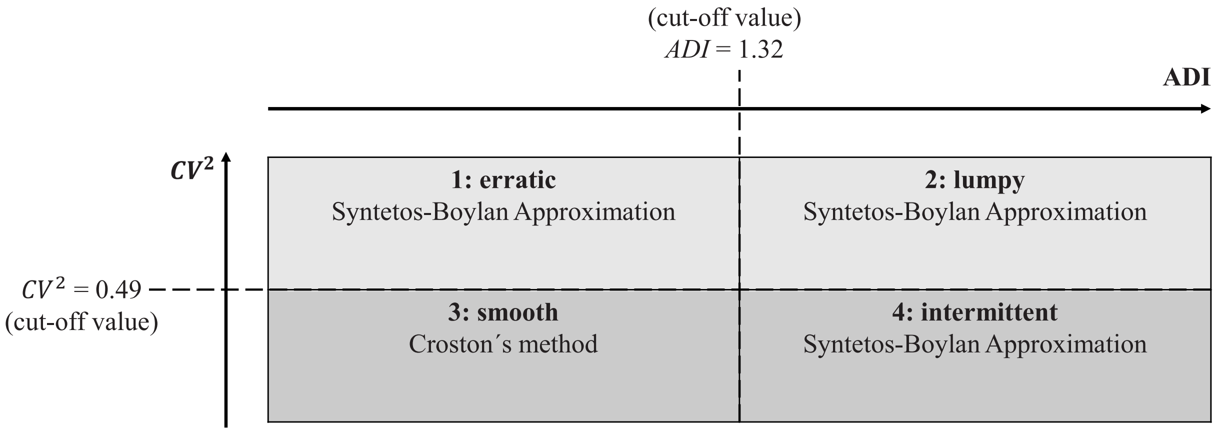

Fig. 1: Demand categories for spare parts according to Syntetos et al. [1]

considerable fluctuations. Syntetos et al. [1] proposed the categorization scheme shown in Figure 1 to quantify these characteristics and make them generally comparable. It includes the squared coefficient of variation (CV^2) to classify demand variation and the average inter-demand interval (ADI), also referred to as *p*, to quantify sporadicity. The more periods in which zero demand occurs, the higher the average inter-demand interval, and the more the demand level fluctuates, the larger the coefficient of variation becomes. Usually, one month is chosen as the unit for a period [1].

2.3. Classical Statistical Forecasting Methods

Several statistical methods are used in practice for time-series forecasting as they are easy to handle and automate. Prevalent methods are linear regression and moving average, whereas in the later the arithmetic averaging of the demand values from the past is conducted, whereby the period interval considered slides one period into the future with the expiry of a period [20]. Furthermore, 1st and 2nd order exponential smoothing is used for constant and trend shaped demand and apply, in contrast to the moving average, an exponentially descending weighting of the data points over time [21]–[23].

Apart from the more straightforward procedures mentioned above, some methods are applied for irregular demand forecasting. One of the most frequently used calculation methods for this was introduced by Croston in 1972 [5] and is based on first-order exponential smoothing. As zero-demand periods can often occur for spare parts, this approach follows the basic idea of separately forecasting the demand and the time of the demand. Syntetos and Boylan [6] modified the method of Croston to increase the forecasting performance as they stated a possible value distortion when the smoothing parameter is greater than $\alpha>0.5$. Therefore, they modified the formula for calculating the predicted demand. Another Croston modification was proposed by Levén and Segerstedt, which aims for a relatively simple approach and forecasts the demand and its occurrence simultaneously [24].

Bootstrapping methods are also recommended for forecasting irregular demand patterns [7]. Within this method, which was first introduced in 1979 [25], statistical features such as variance and mean of the demand distribution are predicted by repeated random sampling from available demand values. This method calculates the expected demand for a period, sets this as the mean value, and estimates the variance based on previous forecast errors [26]. A comprehensive overview and evaluation of bootstrapping methods can be found in Hasni et al. [27], [28].

Table 1 presents the applied statistical forecasting methods in this study.

3. CONFIGURATION OF THE APPLIED ANN MODELS

ANNs consist of a linked set of neurons representing mathematical processing units within which the input values are often processed in a non-linear activation function. The general architecture of an ANN consists of an input and output layer and at least one hidden layer with a varying number of neurons.

The ANN can derive a non-linear input-output relationship, for which it only needs demand data from the past for its training and testing process. Various structures are possible, but the already successfully implemented Multilayer-Perceptron (MLP) with feed-forward character is used in this paper, which is an ANN structure capable of solving non-linear regression and classification problems [29]. As shown in Figure 2, the feed-forward architecture implies that information only moves in one direction, from the input to the output layer. According to the examined literature, the ANN architecture should be kept simple to avoid overfitting. The use of around three hidden neurons has often been successfully applied to similar forecasting problems [9], [11], [14]. Based on these results, tests conducted in the

Table 1: Applied configurations of benchmark methods

Method	Specification	Smoothing parameter α	Abbreviation
Moving average	Latest 36 monthly data points	-	MA 3y
Moving average	Latest 12 monthly data points	-	MA 1y
Linear Regression	Method of least squares	-	Lin Reg
Exponential Smoothing	First-order	0.05	ES1 0.05
Exponential Smoothing	First-order	0.2	ES1 0.2
Exponential Smoothing	Second-order	0.05	ES2 0.05
Exponential Smoothing	Second-order	0.2	ES2 0.2
Croston's Method	24 months for initialization	0.05	CR 0.05
Croston's Method	24 months for initialization	0.2	CR 0.2
Syntetos-Boylan Approximation	24 months for initialization	0.05	SBA 0.05
Syntetos-Boylan Approximation	24 months for initialization	0.2	SBA 0.2

software MATLAB showed overfitting effects at five hidden neurons or more and the minimum validation error was achieved at four hidden neurons or less. Consequently, the authors decided to apply one to four hidden neurons in this study.

Besides the ANN architecture, the applied learning algorithm strongly influences forecasting success [14]. Those algorithms automatically parameterize the connection weights (W) between the neurons to minimize an objective function with an overall error measure [29], [30].

The objective function within the training and testing process of an ANN in this study minimizes the Mean Squared Error (MSE), following [31]–[34]. It was decided to use the two most commonly applied learning algorithms. The Backpropagation learning algorithm was one of the first methods to be adopted for this purpose and is well established [29], [35], [36]. Furthermore, it has shown superior performance results for forecasting problems [11], [14]. However, as this algorithm is known for its higher calculation times, the more efficient Levenberg-Marquardt algorithm has been developed and also applied successfully to forecasting problems [37]–[39].

Consequently, eight ANN configurations with MLP architecture were applied to the forecasting task, using the training algorithms of LM and BP with a varying number of 1 to 4 neurons in the hidden layer.

In the literature, different split percentages between 64% and 82% training data and 18% to 36% validation data are applied [10], [13], [17], [41], [42]. The best results were found after testing different ratios based

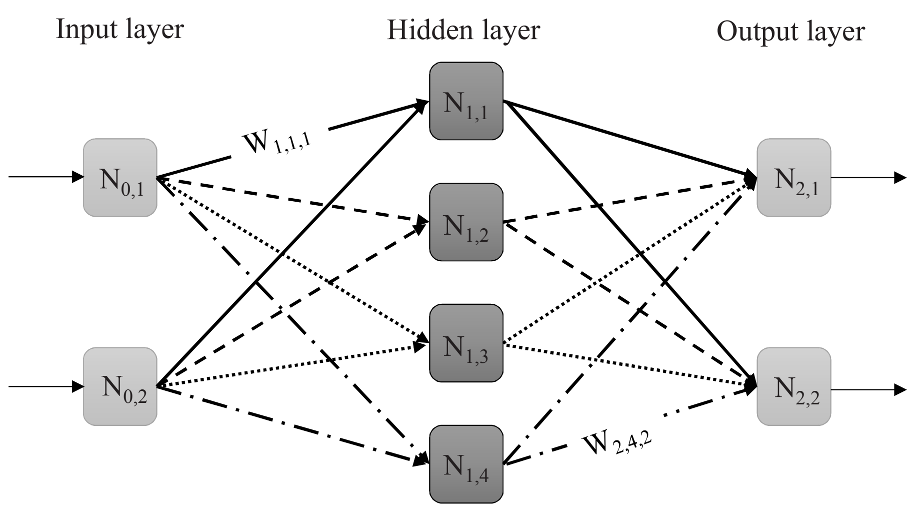

Fig. 2: MLP architecture with one hidden layer and four hidden neurons (N), adapted from [40]

Fig. 3: Applied split of monthly data for ANN training, testing, and performance evaluation

on the literature with 75% training data and 25% validation data. Another critical factor is splitting the data points into training and validation sets, which can be done serially or randomly. The character of the data division was chosen as serial to sustain possible serial correlations in the dataset [43], [44]. A sigmoid function was chosen as an activation function to ensure the network's non-linearity, realized through the function *fitnet*, which is a pre-implemented function in the MATLAB software [45].

4. DATA ANALYSIS

4.1. Characteristics of the quantitative study

All calculations are based on actual demand data covering 42 months. Therefore, the data were split into three serial blocks, as shown in Figure 3. The first 27 months were used for ANN training, whereas the following 9 months were applied for ANN testing [10], [13]. Consequently, the data points of the first 36 months are used to calculate the forecast values with all 19 applied methods, whereas negative predicted values were excluded. The demands of the last six months in the considered dataset were assumed to be future and unknown values upon which the forecast accuracies were subsequently measured over all applied forecasting methods. This assumption was necessary to consider a sufficiently long period to validate the forecast results and thus maximize the expressiveness of the study's results.

A relative error measure is needed to compare the forecasting quality across time series for every prediction method. The error measure should also be based on the absolute value of deviation to avoid balancing positive and negative deviations. The error measurement must also apply to intermittent demands with frequent zero demand periods, i.e., no division by zero may occur. For this reason, the frequently used Mean Absolute Percentage Error (MAPE) [46] does not apply to intermittent demands. Consequently, according to Gilliland [47], the modified MAPE was selected as the measure because it fulfills all criteria listed above. According to Gutierrez et al. [9], it is calculated as follows, where $\hat{y}_t$ represents the predicted demand, y_t the actual demand, e_t the forecast error of a period *t*, and *n* is the number of periods used for forecast performance evaluation:

$$\mathrm{MAPE}_{\mathrm{mod}} = \frac{\sum_{t=1}^{n} |e_t|}{\sum_{t=1}^{n} y_t} = \frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{\sum_{t=1}^{n} y_t} \quad (1)$$

This formula is only applied if at least one positive demand value in the period is under consideration. To allow a detailed comparison, it was necessary to select a sample from the industrial cooperation partner's product portfolio, consisting of >300 spare parts. Based on an ABC analysis, the 24 most strategically relevant objects were selected for forecasting. However, only one intermittent and four erratic parts according to Syntetos et al. [1] were included in this selection. This motivated the authors to consider the four parts with intermittent and one with an erratic pattern from category B for a more evenly distributed dataset and allow for more general conclusions of the results as the product portfolio predominantly consists of smooth demand patterns. The characteristics of all the 29 considered parts are listed in Table 2 below.

Table 2: Spare parts sample characteristics

Object number	Mean demand [parts per month]	CV^2	ADI / p	Demand category
1	36.17	0.12	1.00	3: smooth
2	33.92	0.04	1.00	3: smooth
3	47.83	0.12	1.00	3: smooth
4	12.00	0.22	1.00	3: smooth
5	14.00	0.20	1.00	3: smooth
6	6.50	0.20	1.09	3: smooth

7	1.50	0.63	1.33	2: lumpy
8	26.33	0.18	1.00	3: smooth
9	10.67	0.10	1.00	3: smooth
10	1.25	2.04	2.00	2: lumpy
11	2.92	1.03	1.20	1: erratic
12	2.50	0.47	1.33	4: intermittent
13	30.08	0.15	1.00	3: smooth
14	1.33	1.91	2.00	2: lumpy
15	7.25	0.80	1.00	1: erratic
16	11.92	0.21	1.00	3: smooth
17	3.25	0.90	1.33	2: lumpy
18	8,00	0.75	1.33	2: lumpy
19	0.33	3.50	4.00	2: lumpy
20	3.83	0.85	1.20	1: erratic
21	2.00	1.29	1.50	2: lumpy
22	31.58	0.13	1.00	3: smooth
23	7.42	0.30	1.00	3: smooth
24	5.83	1.86	1.09	1: erratic
25	10.08	0.46	1.33	4: intermittent
26	5.89	0.84	1.06	1: erratic
27	5.06	0.47	1.33	4: intermittent
28	6.28	0.44	1.38	4: intermittent
29	8.08	0.39	1.44	4: intermittent

4.2. Comparison of the Applied Forecasting and ANN Methods

After determining the results of all described established statistical methods and ANN parameter combinations, the following overall results shown in Table 3 are calculated for a forecasting horizon of 6 months. To improve the overview of the average MAPE values, the values were ranked so that the method's configuration with the lowest average MAPE was assigned to the first rank, the one with the highest to the 19th rank, and so on. Identical MAPE values were given the same rank, with the values shown in Table 3 being rounded to three decimal places.

Regarding the arithmetic mean value of all modified MAPEs, the approximation of Syntetos and Boylan (with α=0.2) has the lowest average deviation, followed by Croston's method (α=0.2), which falls 1.9 percentage points behind. In third place is the moving average (with a past horizon of 1 year). The best neural network (LM1) only takes 12th place, with a difference of 10.1 percentage points compared to rank 1.

Since it was suspected that the high MAPE values for some individual parts strongly distort the arithmetic mean, the median was used as a further evaluation criterion. This leads to very different results. First place is now taken by the Croston method, followed by linear regression with a gap of 2.9 percentage points. Almost all of the results are very close when considering the mean MAPE values over the whole dataset. The ANN configurations perform comparatively poorly in almost all averages, with the best result achieved by applying BP1, which performs 6.5 percentage points worse than the best benchmark method. Therefore, a general application of these configurations to an entire parts portfolio is not advisable.

By isolating the results according to the demand categories of Syntetos et al. [1], the potential of ANN can be highlighted for specific demand patterns. A recommendation of the particular forecasting method is given in Figure 4. For erratic demands, the four best MAPE values considering both mean and median values are achieved by ANN configurations. The NN LM4 achieves the best results at the MAPE mean and the NN BP3 at the MAPE median. The second best in terms of the mean is BP2, and in terms of the median is LM4. This shows remarkable advantages in forecasting erratic demands with ANN. It is also worth noting that all configurations of Croston's method and the Syntetos-Boylan approximation perform relatively poorly when applied to erratic demands.

However, the Croston method and Syntetos-Boylan approximation can underpin their benchmark position for lumpy demands. Only BP3 slightly outperforms the benchmark methods in terms of the mean values by 2.4 percentage points. The remaining ANN configurations show comparatively high MAPE values.

The MAPE mean for smooth demand curves shows that the moving average with a 3-year past horizon performs best. This is followed by first-order exponential smoothing (0.05) and Croston's method (0.05). The best ANN configuration is the LM1, 2.3 percentage points behind the first place. Considering the median, the ES1 (0.05) performs best, followed by CR (0.05) and the SBA (0.05). The best neural network (BP1) shows a MAPE value only 0.6 percentage points higher than the best benchmark method. As expected for the smooth demand category, the MAPE values are considerably closer together. The NN BP4 achieved the best result within the intermittent demand category considering the mean MAPE values. This is followed by LM4, which is only 0.5 percentage points behind. The best statistical method is ES2 (0.2), being 14.9 percentage points worse than the benchmark. On the other hand, some ANN configurations also show very high deviations from the actual demand, and it has to be noted that the posterior ranks considering the intermittent demand are all located in the columns of ANN configurations, which applies to both mean and median values.

The number of applied neurons strongly influences the forecast performance in every demand category of this study. LM4 achieves the first rank within the erratic demands, whereas LM3 performs worst in the case of mean MAPE. Furthermore, some configurations perform very poorly across all demand categories, especially LM2 and LM3. Both ANN configurations

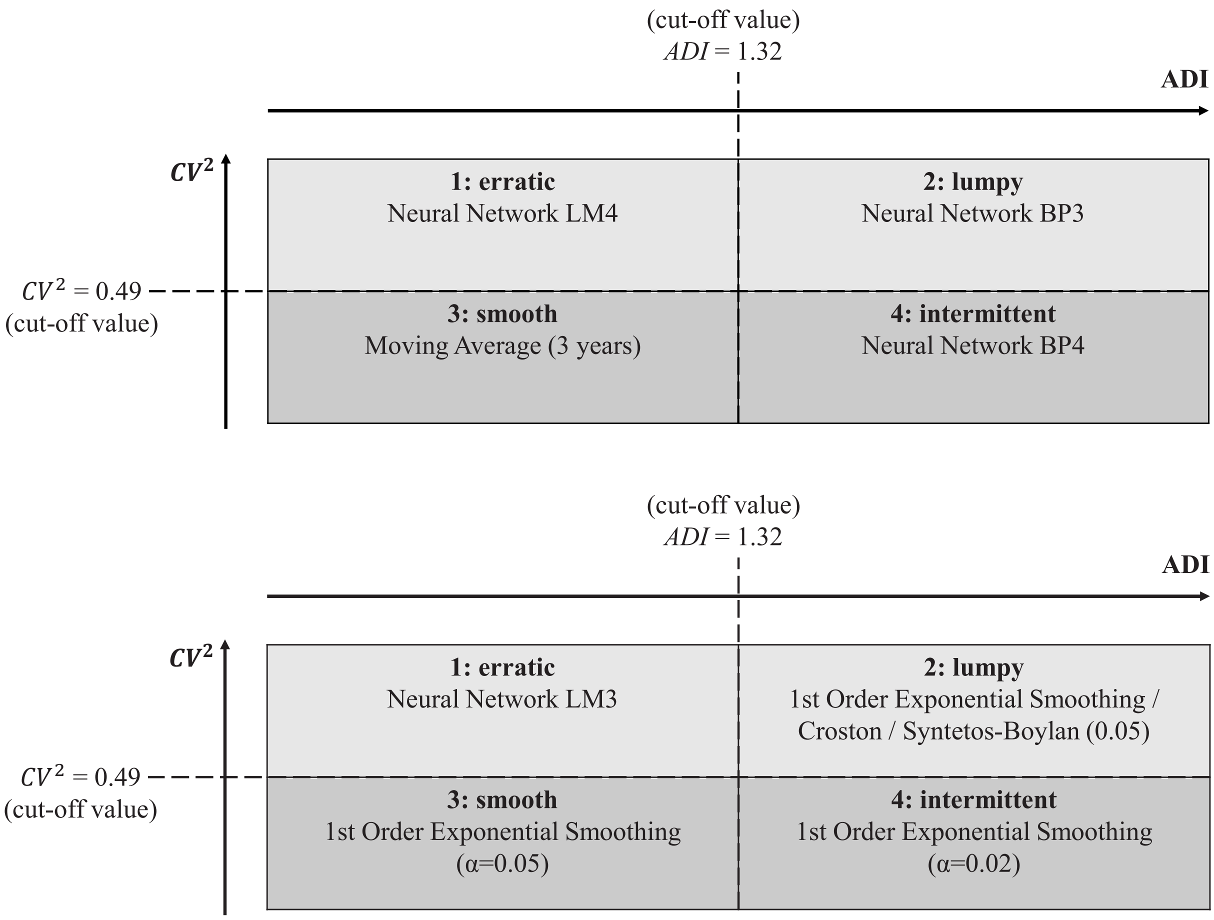

Fig. 4: Own categorization scheme for different demand patterns considering the arithmetic mean (top) and median (bottom), based on Syntetos et al. [1]

Table 3: MAPE Values (interm: intermittent, AM: Arithmetic Mean, M: Median)

Demand category	Criterion	MA 3y	MA 1y	Lin Reg	ES1 0.05	ES1 0.2	ES2 0.05	ES2 0.2	CR 0.05	CR 0.2	SBA 0.05	SBA 0.2	NN BP1	NN BP2	NN BP3	NN BP4	NN LM1	NN LM2	NN LM3	NN LM4
all	AM	0.847	0.796	0.809	0.842	0.805	0.810	0.841	0.842	0.795	0.835	0.776	1.126	1.289	1.116	1.057	0.877	1.415	1.464	0.987
all	Rank (AM)	11	3	5	9	4	6	8	10	2	7	1	16	17	15	14	12	18	19	13
all	M	0.648	0.582	0.579	0.629	0.606	0.602	0.649	0.640	0.550	0.649	0.584	0.615	0.816	0.760	0.893	0.655	0.868	0.879	0.855
all	Rank (M)	10	3	2	8	6	5	11	9	1	11	4	7	15	14	19	13	17	18	16
erratic	AM	1.116	1.050	1.032	1.144	1.111	1.030	1.209	1.150	1.148	1.139	1.098	1.102	0.838	0.859	1.302	0.962	1.171	1.639	0.649
erratic	Rank (AM)	11	7	6	13	10	5	17	15	14	12	8	9	2	3	18	4	16	19	1
erratic	M	0.900	0.726	0.681	0.865	0.754	0.750	0.876	0.960	0.903	0.953	0.910	0.620	0.959	0.473	0.674	0.801	1.039	1.081	0.489
erratic	Rank (M)	12	6	5	10	8	7	11	17	13	15	14	3	16	1	4	9	18	19	2
lumpy	AM	1,065	1.012	1.068	1.031	1.046	1.078	1.102	1.044	0.989	1.034	0.963	2.065	2.211	0.939	1.502	1.108	2.377	1.275	1.811
lumpy	Rank (AM)	9	4	10	5	8	11	12	7	3	6	2	17	18	1	15	13	19	14	16
lumpy	M	0.940	0.979	0.950	0.939	0.953	0.940	1.000	0.939	0.939	0.939	0.944	1.000	1.187	0.977	1.174	0.947	1.048	1.000	1.000
lumpy	Rank (M)	5	12	9	1	10	5	13	1	1	1	7	16	19	11	18	8	17	13	13
smooth	AM	0.428	0.445	0.462	0.433	0.445	0.443	0.489	0.434	0.443	0.436	0.456	0.474	0.582	0.740	0.800	0.451	0.525	0.524	0.722
smooth	Rank (AM)	1	7	11	2	7	5	13	3	5	4	10	12	16	18	19	9	15	14	17
smooth	M	0.409	0.427	0.440	0.399	0.436	0.436	0.491	0.404	0.432	0.404	0.447	0.405	0.582	0.676	0.710	0.411	0.444	0.468	0.758
smooth	Rank (M)	5	7	11	1	9	9	15	2	8	2	13	4	16	17	18	6	12	14	19
interm	AM	1.280	1.081	1.057	1.256	1.028	1.098	0.953	1.231	1.016	1.210	0.960	1.399	2.149	2.523	0.804	1.491	2.449	3.810	0.809
interm	Rank (AM)	13	8	7	12	6	9	3	11	5	10	4	14	16	18	1	15	17	19	2
interm	M	0.654	0.537	0.542	0.629	0.512	0.525	0.560	0.640	0.516	0.649	0.542	0.546	0.788	0.954	0.849	0.672	0.689	0.760	0.786
interm	Rank (M)	12	4	5	9	1	3	8	10	2	11	5	7	17	19	18	13	14	15	16

achieve comparatively bad results and high MAPE values. This leads to the fact that ANN configuration and parameterization are crucial tasks. Furthermore, the general application of a fixed number of neurons, such as in [9]–[11], seems less than ideal and requires testing and evaluation for forecasting optimization.

5. IMPLICATIONS FOR THEORY, PRACTICE, AND FURTHER RESEARCH

This quantitative study shows that the specific configurations of the ANN have a strong influence on the achieved prediction performance. For instance, the difference in prediction performance of LM3 and LM4 is 300 percentage points when considering the mean MAPE values of intermittent demands. The only difference in both ANN is the number of neurons in the hidden layer. While the general fluctuation of the ANN outputs is hard to substantiate, this behavior appears to be caused by a very different weighting of the latest data points. Figure 5 exemplarily shows the disparate behavior of BP2 and BP4 when applied to the data of part number 20. The BP2 algorithm forecasts negative values for periods 37-42, whereas the BP4 predicts strongly growing demands compared to the periods of training and testing (1-36). This pattern requires further investigation. Another important aspect is a possible under- and overfitting effect in the ANN configuration. While the number of applied neurons was based on previous studies and literature findings, the influence of more than four neurons on the prediction performance should also be examined in further studies.

Further research is required in ANN parameterization and minimization of implementation effort in practice. Also, combination possibilities with information types other than past demand, such as *installed base information,* should be explored with more intensity to fully exploit the optimization potentials of ANN in spare part demand forecasting. A comprehensive view of the entire warehousing system could produce even more meaningful results, especially concerning the actual advantage in operational practice. This could be achieved by evaluating service levels, inventory, and shortage costs.

Strong fluctuations characterize irregular demands, leading to very high and unexpected values in specific periods due to internal (e.g. promotions) or external (e.g. weather, disasters) aspects. The effect of outlier elimination in this context on improving forecasting performance and accuracy might be another aspect worth investigating.

As a managerial implication, general forecasting is recommended for companies of every size. Applying statistical methods for this purpose is a conventional approach and still shows a respectable performance. However, ANNs can improve the prediction performance, as shown in Section 4. Based on this study's results, a general application of an ANN to the whole product portfolio without testing and evaluation is not recommended. The presented ANN configurations did show comparatively good results for specific demand categories, especially for erratic and intermittent. The BP3 and LM4 configurations could improve the prediction performance substantially. Hence, the authors suggest an approach that tests the performance of statistical and ANN forecasting methods based on historical data and uses the best performing method for future demand prediction. This is accompanied by the application of key performance indicators to constantly monitor the accuracy of demand predictions. In this study, the $MAPE_{mod}$ was inserted to evaluate the prediction performance of the used forecasting methods. Several other error measurement methods, such as the Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE), are also frequently applied for this task.

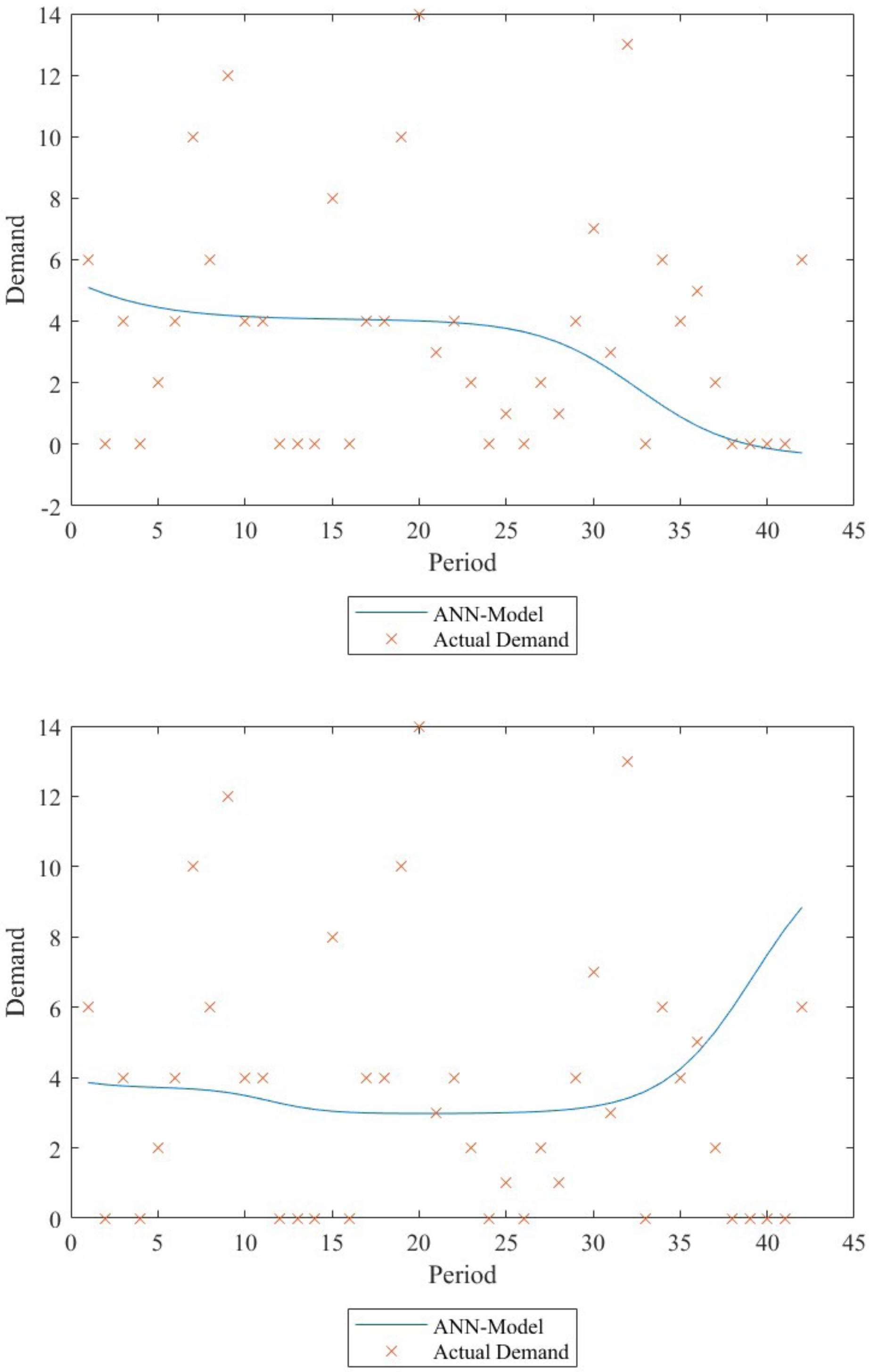

Fig. 5: BP2 (top) and BP4 (bottom) model of part 20

6. CONCLUSIONS AND LIMITATIONS

In this quantitative study, the application of neural networks shows a high potential for forecasting irregular demands, which applies especially to erratic and intermittent demands. The forecasting accuracy could be improved by up to 38.1 percentage points by using an ANN (LM4) compared to the best benchmarking method (ES2) when considering the mean MAPE values of erratic demands. However, the application of an ANN does not show a general improvement in prediction performance in this study.

According to research question 1, an ANN improves forecast precision for erratic, intermittent, and lumpy demands in terms of MAPE mean consideration compared to the applied benchmark methods. This also underpins the previous research results of good ANN forecasting for lumpy [9], [11], [17] and intermittent demands [17]. However, the moving average dominates for smooth demand patterns. Considering the MAPE median, the statistical methods of 1st order exponential smoothing and Croston showed very good results, except for erratic demands where ANN improved prediction accuracy compared to the benchmark methods.

According to research question 2, applying ANN to the whole data set with no consideration of the individual demand category does not show a benefit compared to benchmark forecasting methods. The best results are achieved with the method of Croston and the Syntetos-Boylan approximation, which underpins their benchmark position in forecasting irregular demands. Thus, no overall superiority of a classical statistical forecasting method or an ANN can be derived from the results. A differentiation by parts characteristics, especially their demand patterns, is required to obtain the best results.

According to research question 3, it could be shown that the application of ANN methods in demand forecasting offers high potential for sales and production planning in practice. When using ANNs to predict irregular demands, the network's architecture and the learning algorithm are fundamental. When implementing them in practice, one should focus on these parameters. Nevertheless, the implementation and calculation effort for forecasting with ANNs is usually considerably higher and carries the risk of overfitting. Furthermore, the separate consideration of hidden neurons reveals that this number strongly influences the prediction performance, which underpins the assertion of Lolli et al. [14] to focus on this parameter for ANN configuration. An overall superior performance of three neurons in the hidden layer, as stated in several previous studies, could not be confirmed.

The authors recommend introducing a constant evaluation of forecasting performance, such as the $MAPE_{mod}$, based on existing data. In this way, those methods that offer the highest precision for the individual component can be used for the forecast. This is particularly advisable because both demand pattern and category can change over time and, thus, a rigid and one-time choice of method can lead to a loss of forecasting precision.

The results presented should be validated in a more extensive study, as the sample size was, similarly to Lolli et al. [14] and Gutierrez et al. [9], limited to 29 strategically relevant spare parts of the cooperating company that showed an uneven distribution considering the demand categories. Furthermore, the number of applied ANN algorithms was limited as the focus was on the most frequently used architecture and learning algorithms. The obtained prediction results from the ANN models showed high fluctuations in some cases, limiting the generalizability of the study's findings.

Further research is required regarding the parametrization of ANN for irregular demand forecasting, which is especially true for the applied number of neurons in the hidden layer. The effect of partially high fluctuating prediction results of the ANN requires further investigation. The presumption is that the number of input data points is too small for the ANN to calculate consistent output values. In this study, every part was calculated separately without considering any further input data than previous demand data. A conceivable approach is to build a model that simultaneously takes the whole product portfolio as input data to exploit possible dependencies between specific demands.

This study reveals strong optimization potential for logistical processes regarding demand forecasting and planning, which is crucial for ensuring a high service level with an appropriate inventory level. Especially in the spare parts sector, where failures and breakdowns cause downtimes and urgent deliveries, an optimized stock is essential to ensure a response to the market and functioning logistical processes.

REFERENCES

[1] A. A. Syntetos, J. E. Boylan, and J. D. Croston, "On the categorization of demand patterns," *J. Oper. Res. Soc.*, vol. 56, no. 5, pp. 495–503, 2005.

[2] P. Schuh, D. Schneider, L. Funke, and K. Tracht, "Cost-optimal spare parts inventory planning for wind energy systems," *Logist. Res.*, vol. 8, no. 1, pp. 1–8, 2015.

[3] Ç. Pinçe, L. Turrini, and J. Meissner, "Intermittent demand forecasting for spare parts: A Critical review," *Omega (United Kingdom)*, vol. 105, 2021.

[4] J. E. Boylan and A. A. Syntetos, "Spare parts management: A review of forecasting research and extensions," *IMA J. Manag. Math.*, vol. 21, no. 3, pp. 227–237, 2010.

[5] J. D. Croston, "Forecasting and Stock Control for Intermittent Demands.," *Oper. Res. Q.*, vol. 23, no. 3, pp. 289–303, 1972.

[6] A. A. Syntetos and J. E. Boylan, "The accuracy of intermittent demand estimates," *Int. J. Forecast.*, vol. 21, no. 2, pp. 303–314, 2005.

[7] T. R. Willemain, C. N. Smart, and H. F. Schwarz, "A new approach to forecasting intermittent demand for service parts inventories," *Int. J. Forecast.*, vol. 20, no. 3, pp. 375–387, 2004.

[8] F. Petropoulos, S. Makridakis, V. Assimakopoulos, and K. Nikolopoulos, "'Horses for Courses' in demand forecasting," *Eur. J. Oper. Res.*, vol. 237, no. 1, pp. 152–163, 2014.

[9] R. S. Gutierrez, A. O. Solis, and S. Mukhopadhyay, "Lumpy demand forecasting using neural networks," *Int. J. Prod. Econ.*, vol. 111, no. 2, pp. 409–420, 2008.

[10] A. Nasiri Pour, B. Rostami Tabar, and A. Rahimzadeh, "A hybrid neural network and traditional approach for forecasting lumpy demand," *World Acad. Sci. Eng. Technol.*, vol. 40, no. 4, pp. 384–390, 2008.

[11] S. Mukhopadhyay, A. O. Solis, and R. S. Gutierrez, "The accuracy of non-traditional versus traditional methods of forecasting lumpy demand," *J. Forecast.*, vol. 31, no. 8, pp. 721–735, 2012.

[12] M. Rosienkiewicz, "Artificial intelligence methods in spare parts demand forecasting," *Logist. Transp.*, vol. 18, no. 2, pp. 41–50, 2013.

[13] F. L. Chen, Y. C. Chen, and J. Y. Kuo, "Applying Moving back-propagation neural network and Moving fuzzy-neuron network to predict the requirement of critical spare parts," *Expert Syst. Appl.*, vol. 37, no. 9, pp. 6695–6704, 2010.

[14] F. Lolli, R. Gamberini, A. Regattieri, E. Balugani, T. Gatos, and S. Gucci, "Single-hidden layer neural networks for forecasting intermittent demand," *Int. J. Prod. Econ.*, vol. 183, pp. 116–128, 2017.

[15] N. Kourentzes, "Intermittent demand forecasts with neural networks," *Int. J. Prod. Econ.*, vol. 143, no. 1, pp. 198–206, 2013.

[16] J. L. Carmo and A. J. Rodrigues, "Adaptive forecasting of irregular demand processes," *Eng. Appl. Artif. Intell.*, vol. 17, no. 2, pp. 137–143, 2004.

[17] M. Sahin, R. Kizilaslan, and Ö. F. Demirel, "Forecasting Aviation Spare Parts Demand Using Croston Based Methods and Artificial Neural Networks," *J. Econ. Soc. Res.*, vol. 15, no. 2, pp. 1–21, 2013.

[18] M. Z. Babai, A. Tsadiras, and C. Papadopoulos, "On the empirical performance of some new neural network methods for forecasting intermittent demand," *IMA J. Manag. Math.*, vol. 31, no. 3, pp. 281–305, 2020.

[19] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," *34th Int. Conf. Mach. Learn. ICML 2017*, vol. 3, pp. 2130–2143, 2017.

[20] C. Chatfield, "Time-series forecasting," *Significance*, no. 2, 2005.

[21] J. G. De Gooijer and R. J. Hyndman, "25 Years of Time Series Forecasting," *Int. J. Forecast.*, vol. 22, no. 3, pp. 443–473, 2006.

[22] T. Miller and M. Liberatore, "Seasonal exponential smoothing with damped trends," *Int. J. Forecast.*, vol. 9, no. 4, pp. 509–515, 1993.

[23] M. Goodman, "A New Look at Higher-Order Exponential Smoothing for Forecasting," *Oper. Res.*, vol. 22, no. 4, pp. 880–888, 1974.

[24] E. Levén and A. Segerstedt, "Inventory control with a modified Croston procedure and Erlang distribution," *Int. J. Prod. Econ.*, vol. 90, no. 3, pp. 361–367, 2004.

[25] B. Efron, "Bootstrap Methods: Another Look at the Jackknife," *Ann. Stat.*, vol. 7, no. 1, pp. 1–26, 1979.

[26] R. H. Teunter and L. Duncan, "Forecasting intermittent demand: A comparative study," *J. Oper. Res. Soc.*, vol. 60, no. 3, pp. 321–329, 2009.

[27] M. Hasni, M. S. Aguir, M. Z. Babai, and Z. Jemai, "On the performance of adjusted bootstrapping methods for intermittent demand forecasting," *Int. J. Prod. Econ.*, vol. 216, pp. 145–153, 2019.

[28] M. Hasni, M. S. Aguir, M. Z. Babai, and Z. Jemai, "Spare parts demand forecasting: a review on bootstrapping methods," *Int. J. Prod. Res.*, vol. 57, pp. 4791–4804, 2019.

[29] E. Alpaydin, *Introduction to Machine Learning*, 2nd ed. Cambridge: MIT Press, 2010.

[30] D. E. Rumelhart and G. E. Hintont, "Learning Representations by Back-Propagating Errors," *Cogn. Model.*, no. 2, pp. 3–6, 2019.

[31] H. Jin and G. Montúfar, "Implicit bias of gradient descent for mean squared error regression with wide neural networks," *arXiv:2006.07356*, 2020.

[32] J. Qi, J. Du, S. M. Siniscalchi, X. Ma, and C. H. Lee, "On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression," *IEEE Signal Process. Lett.*, vol. 27, pp. 1485–1489, 2020.

[33] L. Zhang, "Optimizing ANN training performance for chaotic time series prediction using small data size," *Int. J. Mach. Learn. Comput.*, vol. 8, no. 6, pp. 606–612, 2018.
[34] R. J. Frank, N. Davey, and S. P. Hunt, "Time series prediction and neural networks," *J. Intell. Robot. Syst. Theory Appl.*, vol. 31, no. 1–3, pp. 91–103, 2001.
[35] D. E. Rumelhart and G. E. Hintont, "Learning Representations by Back-Propagating Errors," *Cogn. Model.*, no. 2, pp. 3–6, 1986.
[36] MathWorks, "Resilientbackpropagation," 2006. [Online]. Available: https://de.mathworks.com/help/deeplearning/ref/trainrp.html. [Accessed: 25-Feb-2021].
[37] H. Yu and B. M. Wilamowski, "Levenberg-Marquardt Training," in *The Industrial Electronics Handbook – Intelligent Systems*, 2nd ed., B. M. Wilamowski and J. D. Irwin, Eds. Boca Raton: CRC Press, 2011.
[38] MathWorks, "Levenberg-Marquardt backpropagation," 2006. [Online]. Available: https://de.mathworks.com/help/deeplearning/ref/trainlm.html?searchHighlight=trainlm&s_tid=srchtitle. [Accessed: 25-Feb-2021].
[39] S. Moshiri and N. Cameron, "Neural network versus econometric models in forecasting inflation," *J. Forecast.*, vol. 19, no. 3, pp. 201–217, 2000.
[40] J. Gijsbrechts, R. N. Boute, J. A. Van Mieghem, and D. J. Zhang, "Can Deep Reinforcement Learning Improve Inventory Management? Performance on Lost Sales, Dual-Sourcing, and Multi-Echelon Problems," *Manuf. Serv. Oper. Manag.*, pp. 1–35, 2019.
[41] M. Ghiassi, D. K. Zimbra, and H. Saidane, "Urban Water Demand Forecasting with a Dynamic Artificial Neural Network Model," *J. Water Resour. Plan. Manag.*, vol. 134, no. 2, pp. 138–146, 2008.
[42] N. Khoa, V. Truong, S. Sangmun, V. T. Nha, and K. Ichon, "Intermittent demand forecasting by using Neural Network with simulated data," *Proc. 2011 Int. Conf. Ind. Eng. Oper. Manag.*, pp. 723–728, 2011.
[43] MathWorks, "Divide Data for Optimal Neural Network Training," 2020. [Online]. Available: https://de.mathworks.com/help/deeplearning/ug/divide-data-for-optimal-neural-network-training.html. [Accessed: 30-Sep-2021].
[44] R. S. Tsay, "Testing serial correlations in high-dimensional time series via extreme value theory," *J. Econom.*, vol. 216, no. 1, pp. 106–117, 2020.
[45] MathWorks, "Function fitnet neural network," 2010. [Online]. Available: https://de.mathworks.com/help/deeplearning/ref/fitnet.html?searchHighlight=fitnet&s_tid=srchtitle. [Accessed: 25-Feb-2021].
[46] S. Makridakis, M. Hibon, and C. Moser, "Accuracy of Forecasting: An Empirical Investigation," *J. R. Stat. Soc. Ser. A*, vol. 142, no. 2, p. 97, 1979.
[47] M. Gilliland, "Is forecasting a waste of time?," *Supply Chain Manag. Rev.*, vol. 6, no. 4, pp. 16–23, 2002.