Machine Learning with Applications 8 (2022) 100302

Contents lists available at ScienceDirect

Machine Learning with Applications

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

Forecasting performance of wavelet neural networks and other neural network topologies: A comparative study based on financial market data sets

Markus Vogl [a,*], Peter Gordon Rötzel (LL.M) [b,a], Stefan Homes [a]

[a] University of Applied Sciences Aschaffenburg, Würzburger Stra"ze 45, 63743 Aschaffenburg (DE), Germany
[b] University of Stuttgart, Keplerstra"ze 17, 70174 Stuttgart (DE), Germany

ARTICLE INFO

JEL classification:
C22
C32
C52
C45
C53
G17

Keywords:
Wavelet neural networks
Wavelet decomposition
Financial forecasting
Forecast evaluation
Neural network topology
Financial markets

ABSTRACT

In this study, we analyse the advantageous effects of neural networks in combination with wavelet functions on the performance of financial market predictions. We implement different approaches in multiple experiments and test their predictive abilities with different financial time series. We demonstrate experimentally that both wavelet neural networks and neural networks with data pre-processed by wavelets outperform classical network topologies. However, the precision of conducted forecasts implementing neural network algorithms still propose potential for further refinement and enhancement. Hence, we discuss our findings, comparisons with "buy-and-hold" strategies and ethical considerations critically and elaborate on future prospects.

1. Introduction

Forecasting models tailored for financial time series are discussed frequently in business and science (Sezer et al., 2020). The development of computer-based methods has witnessed significant progress, which is well illustrated by, for example, Renaissance Technologies, led by James Simons. Renaissance Technologies has systematically outperformed market growth over many years through the execution and advanced analysis of algorithms and signals (Burton, 2016). However, the majority of actively managed funds[1] based on statistical analysis display severe underperformance owing to lower yields earned compared to the respective benchmark (market) indices (Otuteye & Siddiquee, 2019). In particular, said underperformance renders itself visible once trading and management fees are considered, which are compared with passive investments, such as buy-and-hold strategies (Otuteye & Siddiquee, 2019).

The rapid growth in the research field of artificial intelligence (AI) since 2010 highlights neural networks (NNs), particularly in terms of computer-aided methods (Sezer et al., 2020). Based on machine learning and pattern recognition algorithms in large amounts of data (such as time series), NNs exhibit a higher potential for producing more accurate predictions than conventional statistical methods (e.g. exponential smoothing, as shown by Hill et al., 1996) (Paliwal & Kumar, 2009). Applications to the field of financial market and risk management predictions are given in Petneházi (2021), among others, stating convolutional neural networks (CNNs) for value-at-risk predictions. However, according to Makridakis et al. (2018), NNs have not revealed their full potential, which can be seen in Peng et al. (2021), analysing several deep neural networks to assess empirical performance in terms of technical indicators. Furthermore, Peng et al. (2021) state that no strategy under analysis was able to outperform simple buy-and-hold strategies. Contrastingly, following, for example. Chalvatzis and Hristu-Varsakelis (2020) or Nobre and Neves (2019) state respective outperformance of buy-and-hold strategies already. The advantages of NNs as predictive models are owing to the significantly increased amount of data availability, as well as higher computational capacity (Jordan & Mitchell, 2015). Therefore, the methods and experiments discussed in this study examine, for example, the conjecture that larger and better data sets lead to a more accurate prediction of stock and index prices. The latter conjecture follows the well-known remark of

* Corresponding author.
E-mail addresses: markus.vogl@vogl-datascience.de (M. Vogl), peter.roetzel@th-ab.de (P.G. Rötzel), s150554@th-ab.de (S. Homes).

[1] Personal decisions are involved in the active process of investing, whereas passive management of funds is based solely on index allocation, for example, on market capitalisation (?).

[2] Non-periodic, localised wave function, which integral yields exactly zero value (?).

https://doi.org/10.1016/j.mlwa.2022.100302
Received 4 December 2021; Received in revised form 10 March 2022; Accepted 11 April 2022
Available online 18 April 2022

Peter Norvig, namely, that "more data beats clever algorithms but better data beats more data". Furthermore, the concept of wavelets,[2] taken from the field of signal processing, provides some interesting application possibilities in individual tests regarding the analysis of financial time series (Crowley, 2007). In particular, for the processing of time series with different periodicity, as well as for short- and long-term cycles, wavelet transforms offer advantages (e.g. for business cycle analysis) (Crowley, 2007). Although Sezer et al. (2020) state that novel methodologies in NN designs are well researched, it remains unresolved, whether a combination of NNs with signal processing techniques, such as previously mentioned wavelet functions, indicates a (positive) effect on the predictive performance in financial forecasting (Alexandridis & Zapranis, 2014).

Therefore, this gap warrants further investigation, leading to our first research question, namely, the elucidation of potential relations between the amount of input data and predictive time horizons. Furthermore, regarding prediction accuracy, it appears that various network topologies perform differently at variable time horizons (Tsantekidis et al., 2017). Our second research question elucidates the procedure of proportionally replacing the neurons of established network topologies with wavelons (i.e. a wavelet function as replacement of a sigmoidal activation function, refer to Alexandridis & Zapranis, 2014), thus, generating wavelet neural networks (WNNs) (Zhang & Benveniste, 1992). Alexandridis and Zapranis (2014), as well as other analysed publications (e.g. Billings & Wei, 2005; Zhang, 1997), reveal that certain wavelet functions are more suitable than other respective wavelet equations. A wavelet activation function (e.g. a Morlet-wavelet) can outperform a logistic function; however, additional analysis of further wavelet functions is required to achieve a respective generalisation of the latter superiority presupposition (Anjoy & Paul, 2017). The fundamental demand for combined models (i.e. of hybrid models), such as WNNs (see also Yang & Wang, 2021), is derived from the fact that (financial) time series contain various information components, such as time and frequency information (Chakrabarty et al., 2015). Frequency components can be extracted by applying transformations (e.g. Fourier transformation [FT]) to the data sets (Chakrabarty et al., 2015). Unfortunately, once the FT is applied, the required time information components are lost, whereas, the wavelet transformation preserves said time information (Chakrabarty et al., 2015). Additionally, NNs are advantageous over statistical models (e.g. exponential smoothing) regarding the processing capabilities of nonlinear functions (Alexandridis & Zapranis, 2013; Hill et al., 1996).

Therefore, within the academic literature, corresponding approaches are elaborated on from two different research streams, namely, intermediate and generalised. Moreover, a combination of NNs and wavelets is possible, applicable and reasonable for various scientific disciplines (Alexandridis & Zapranis, 2014). The intermediate research stream deals with the pre-processing of time series by applying wavelet decompositions (Alexandridis & Zapranis, 2014). Therefore, we discuss the effect of the additional input of time series processed with wavelet decomposition on a NN instead of applying only unprocessed time series. The experiment is focused on comparing the predictive success to well established NNs, such as multilayer perceptron (MLP) and long short-term memory (LSTM) networks without pre-processing the data. In addition, we aim to determine, whether performance deviations between different referring prediction periods exist. In response to the previously mentioned implication by Norvig, several publications elaborate on the "more data" hypothesis critically. For example, Walczak (2001) concludes that considering more than two years of data has no significant effect on the accuracy of forecasting models. We explore a WNN approach in which the activation functions of an NN (e.g. in the hidden layers) are proportionally replaced by wavelet functions (Zhang & Benveniste, 1992). Therefore, we investigate the change in the forecasting performance of an NN, in which a wavelet is applied as an activation function. More specifically, we intend to determine the most appropriate wavelet function in combination with NN topologies in terms of financial market prediction. Further, to substantiate our findings regarding both the research questions, we perform a respective back-test.

Moreover, we state other studies taken out of the academic literature (e.g. Kumbure et al., 2022 or Yang & Wang, 2021), proposing comparisons of WNNs or otherly NN comparisons to provide a holistic picture of stated groundwork determinations. Finally, we elaborate on the underlying topology that is best suited for this kind of methodology and discuss the results in terms of performance comparability with buy-and-hold strategies and ethical considerations critically.

2. Literature review

In the 1990s, many NN-projects[3] were carried out in the field of time-series analysis, and most of the latter projects present the basis for corresponding future research endeavours (Vellido et al., 1999). Nonetheless, according to Sezer et al. (2020), the area of NN research was neglected until 2011. During this period, many publications (e.g. Chen, 1994 or Vellido et al., 1999) reported problems with NNs, mainly owing to computers in the said period to provide very little computational power or respective memory (Lim et al., 2004). Furthermore, problems with successful implementation existed; hence, more simplistic models were considered sufficient at the time (Adya & Collopy, 1998). With the progression of the 2010s and driven by the extensive availability of data and inexpensive computational capacity, the field of NN-based AI experienced a rediscovery, visible in the vast dissemination of academic research (Jordan & Mitchell, 2015). The results show that NNs are suitable for producing acceptable predictions in the case of financial time series, even if only to a limited extent (Adya & Collopy, 1998). Nevertheless, the potential for improvement was recognised, including the application of hybrid models (e.g. Vui et al., 2013; Zhou et al., 2019). Hybrid models generally combine two or more well-established methods (e.g. classical NNs with signal processing techniques or stochastic elements) into a novel application (Dong et al., 2016). However, hybrid models should be tested carefully with different time series before deployment to obtain a reasonable understanding of their inherent functionality (Guresen et al., 2011; Yang & Wang, 2021).

First, following Alexandridis and Zapranis (2013, 2014), referring to the intermediate research stream mentioned earlier, we apply a hybrid model combined with a respective wavelet transform to transfer a time series into the frequency domain. Subsequently, based on Taspinar (2018), the signal is processed and subdued to a respective back-transformation to obtain data, serving as input for a classical NN. The second, more innovative strategy consists of partially replacing the neurons of the referred network with wavelons (Zhang & Benveniste, 1992). In addition, the resulting WNN is constructed from a three-layer MLP (Alexandridis & Zapranis, 2014). Henceforth, we refer to the previously mentioned basic structure applying wavelons as part of the generalised research stream, which differs from the intermediate research stream in terms of activation functions (Alexandridis & Zapranis, 2014). Wavelet activation functions are applied to the underlying network topology instead of the commonly implemented sigmoidal activation functions (Zhang et al., 1995).

To the best of our knowledge, some publications discuss the mentioned approaches; however, elaborate exclusively on their exploration in general or aim directly at specialised applications (e.g. see Yang & Wang, 2021 for a comparison in terms of energy series) (Alexandridis & Zapranis, 2014). This is due to increased design and testing requirements for WNNs (Oussar & Dreyfus, 2000). Nevertheless, according to Alexandridis and Zapranis (2013), a framework for building WNNs is available, which discusses construction and training algorithms in detail. However, most publications with an economically

[3] For example, see Hill et al. (1996), Kaastra and Boyd (1996), and Zhang et al. (1995) for further reference.

relevant alignment focus solely on the intermediate research stream. In particular, Bao et al. (2017), which combines accuracy and profitability measurements, can be considered in this regard. Further, Bao et al. (2017) present reasonable and reproducible values for three different network topologies (including LSTM). The referring data sets that Bao et al. (2017) investigate are various economic indices, including S&P500 and DJIA100, as well as four other Asian indices.

Even if the literature provides some successful predictions with NNs, as mentioned before, successful forecasts are not guaranteed because of the stochastic functionality of NNs, which is based on the learning algorithm implemented in each NN (Kaastra & Boyd, 1996). The latter learning algorithm performs a gradient descent operation as part of the training process and can compute an acceptable solution (i.e. local error minimum). Unfortunately, the results do not necessarily represent the optimal solution (i.e. the global error minimum) of the approximation of the respective training function (Kaastra & Boyd, 1996). Furthermore, some wavelet functions in the field of application with NNs are more commonly implemented than others (Alexandridis & Zapranis, 2014). Depending on the application, different wavelet functions are recommended, mainly the family of Gaussian wavelets (e.g. Billings & Wei, 2005 or Zhang, 1997) (Alexandridis & Zapranis, 2014). Moreover, wavelets can be custom-designed if required, i.e., specifically tailoring said functions to the respective application and its objectives (Misiti et al., 2007). This study focuses on two Gaussian wavelets, namely, the first and second derivatives of the Gaussian bell curve as depicted in Barlow (1983). The latter is also labelled as the Mexican Hat function, owing to its shape.

Different application frequencies are also observable in the choice of topology for NNs. Therefore, we present the three most applied topologies from our sampled literature, as shown in Table 1. This selection is also favoured to set up benchmark networks in the experiment with MLP and LSTM as bases. Since all other topologies are built on or derived from MLPs, the MLP is the most widely researched network topology in the field of financial market prediction and is repeatedly taken as a referring benchmark (Sezer et al., 2020). According to De Faria et al. (2009), this simple NN predicts the positive or negative sign for yields in the Brazilian stock market with 60% accuracy. The aforementioned second and ongoing wave of AI research has shown significant progress since the 2010s, especially, with newly developed NN topologies (Jordan & Mitchell, 2015). Furthermore, it would be of interest to elaborate on subsequently the aforementioned hybrid models (Sezer et al., 2020; Vui et al., 2013). However, such an experiment does not necessarily lead to predictive success, as demonstrated through the combination of generalised autoregressive conditional heteroscedasticity models and MLP, which are highly influenced by existing noise effects in the data under analysis by Guresen et al. (2011).

By contrast, Tsaih et al. (1998) find the hybrid extended futures forecast model to be promising in terms of trading S&P500 futures. In addition, a similar methodology can be found in Zhang et al. (2001), who also analyse futures trading and show that hybrid models consisting of wavelets and NNs lead to more profitable trading results than otherwise implemented MLPs, although they are dependent on outliers in different market situations. Therefore, we will also elaborate on the general performance of NN solutions in comparison with simplistic "buy-and-hold" strategies and elucidate for which market actors under which dynamical presupposition these methodologies can be favourable (see Section 6.2).

3. Research hypotheses

We further elaborate on hybrid models by focusing on the application of different NNs with wavelet components, since the correct specification of network topology as well as wavelet-function is a key problem according to Chen et al. (2006). Both of the aforementioned research questions aim to improve the performance of NNs in financial market predictions. The first research question claims that the quantity and quality of the input data are crucial for the precision of prediction capability and the quality of an NN. We tested this experimentally by developing several NNs according to the intermediate approach and comparing the respective results based on the quantity of input data as well as with two common NN topologies, namely, MLP and LSTM. The hypothesis of the second question is that the involvement of a WNN by a generalisation approach fundamentally improves the success of predictions. Furthermore, we demonstrate that not all wavelet functions serve equally well as activation functions. Therefore, further classical NNs are constructed with different wavelets as activation functions.

Table 1
The most commonly implemented network topologies as depicted in our sampled literature. The results are in line with the findings of Sezer et al. (2020) as a point of reference.

Network topology	Percentage of analysed publications
LSTM	31%
MLP	21%
Hybrid	8%
Other specifications	40%

Notes: LSTM: long short-term memory; MLP: multilayer perceptron.

4. Methodology, experimental design, and data description

For the experiments, we create eight different NNs in accordance with the criteria defined in the research questions (see Section 3) and examine the latter for respective predictive capabilities. Therefore, we initially examine and implement two well-established network topologies, then, based on the results, we implement the other six experimental NNs. This procedure aims to create the same foundation for both research questions, thus, minimising the implementation effort, as well as the risk of potential errors, while concurrently improving comparability. MLP-based NNs can only handle one time series simultaneously and can only predict a single value. By contrast, the other LSTM topologies are capable of an n-step-ahead-forecast and accept multiple time series as input. Therefore, LSTMs allow predictions over longer periods, which exceed just one-day-ahead forecasts.

In addition, we intend to derive statements about whether more input data leads to better results, referring to the initially stated and academically critically discussed proposition of Walczak (2001) (see Section 1). To study the research stream of WNNs, we first implement four topologies followed by two more topologies to study subsequently the effects of wavelet decomposition.

4.1. Wavelet neural networks

Like classical NNs, a WNN generally consists of three different layers: input, hidden, and output layers (Alexandridis & Zapranis, 2014). Owing to theoretical construction possibility of feedforward NNs in terms wavelet decompositions (see Pati & Krishnaprasad, 1993 or Zhang & Benveniste, 1992 for early reference), WNNs represent an alternative to cope with NN weaknesses (such as randomised starting values in training algorithms) (Alexandridis & Zapranis, 2013). Furthermore, WNNs represent a generalisation of radial basis function networks (Alexandridis & Zapranis, 2013). In contrast to well-established networks, the hidden layer of a WNN does not contain neurons, but wavelons that fulfil the same task as neurons. The only difference lies in the respective activation function specifications, namely, incorporating a wavelet function instead of a sigmoidal function representation (Alexandridis & Zapranis, 2014). To be more detailed, the nodes (i.e. the wavelons) of a WNN are the wavelet coefficients of the function expansion, yielding a significant value (Alexandridis & Zapranis, 2013). Moreover, multidimensional wavelets preserve the "universal approximation" property that is characteristic for NNs (Alexandridis & Zapranis, 2014). Furthermore, reasons to conduct said alterations lay within the characteristics of wavelets (refer to Bernard et al., 1998),

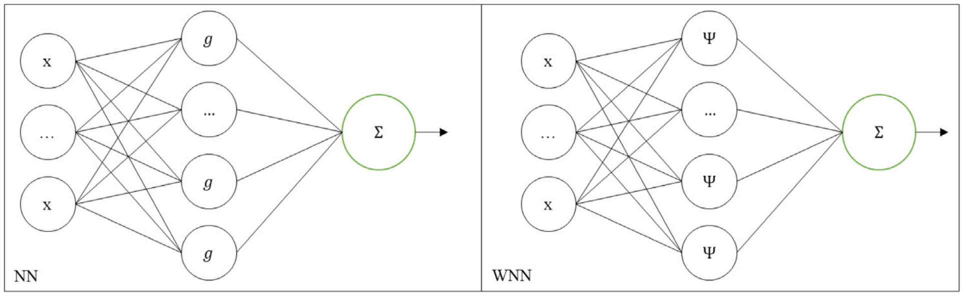

Fig. 1. Structure of a three-layer wavelet neural network (WNN), consisting of input, hidden, and output layers. Note that the activation function of a WNN (right) differs from those of standardised neural networks (left), namely, the WNN incorporates wavelons, which represent themselves via respective wavelet functions.

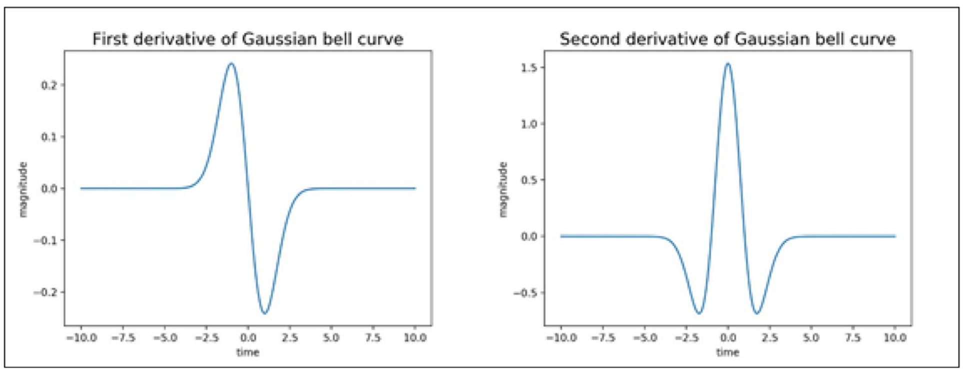

Fig. 2. Wavelet functions applied as wavelon activation functions, namely, the Gaussian wavelet activation functions: Gauss1 (left) and Gauss2 (right).

namely, high compression capabilities in computing the value at a single point or updating the functional estimate from a novel local measure, which involves only a small subset of all given coefficients, respectively (Alexandridis & Zapranis, 2013, 2014). Thus, WNNs allow for constructive procedures, which efficiently initialise network parameters, i.e. convergence to the global minimum of the referring cost function (Alexandridis & Zapranis, 2013). To be more detailed, the initial weight vector of a WNN[4] is present into close proximity of the global minimum and, therefore, drastically reduces training times, as also stated in Alexandridis and Zapranis (2013) and Yang and Wang (2021). Thus, the general idea of a WNN is given by the aim to adapt the corresponding wavelet basis to the respective training data (Alexandridis & Zapranis, 2014). In case of multi-wavelet NNs the according activation function is given by a linear combination of wavelet bases and combinable with discrete wavelet transforms (DWTs) or principal component analysis (PCA) algorithms (Alexandridis & Zapranis, 2013). A conceptual illustration of this is shown in Fig. 1. The only significant change in the implementation of a WNN compared with other NNs is the inclusion of the manually defined activation function, namely, the resulting and previously noted wavelons. Two wavelet functions are applied, namely, the first (Gauss1) and second (Gauss2) derivatives of the Gaussian bell curve, as illustrated in Fig. 2. The two standard network topologies, MLP and LSTM, incorporate the previously mentioned Gaussian wavelets, resulting in four test networks, which are depicted in Table 2.

Table 2
Comparison of well-established neural networks and experimental wavelet neural networks elaborated on during the empirical investigation of this study.

Established network topology	Experimental network
MLP	WNN (MLP & Gauss1) WNN (MLP & Gauss2)
LSTM	WNN (LSTM & Gauss1) WNN (LSTM & Gauss2)

Notes: MLP: multilayer perceptron; LSTM: long short-term memory.

4.2. Wavelet decomposition neural networks

In contrast to WNNs, the NN itself is not changed within the wavelet decomposition neural network (WDNN) research; however, more time series are added to the input amount (Alexandridis & Zapranis, 2014). First, regarding the research field of signal processing, it is common to conduct an initial decomposition of the signal (e.g. of a respective stock price series) by applying a wavelet transform (Mallat, 1989). Second, following Taspinar (2018), the results are processed in the respective frequency domain. Finally, the signal is transformed back to the time domain by executing a back-transformation (Mallat, 1989). One disadvantage is that signal decomposition by transformation leads to overlaps at the edges of the respective time–frequency series and is amplified by larger choices of the decomposition level (Williams & Amaratunga, 1997). Unfortunately, said edge effects cannot be completely avoided due to the finite nature of real-world time-series data (Torrence & Compo, 1998). Therefore, in accordance with Anjoy and Paul (2017), we select decomposition levels, which do not exceed the value of five. The additional data to be analysed within the WDNNs are pre-processed by applying the previously mentioned wavelet decomposition and being fed into the LSTM topology accordingly.

The wavelet transform displays the property of a low-pass filter,[5] which we intend to exploit (Taspinar, 2018). Furthermore, we focus on two NNs with different levels of decomposition. The first WDNN performs one level of pre-processing and the second performs five,

[4] For an in-depth mathematical treatise refer to Alexandridis and Zapranis (2013, 2014) and Zapranis and Alexandridis (2008, 2009).

[5] Denoising of the input data is conducted (Torrence & Compo, 1998).

Table 3
Comparison of well-established neural networks and experimental wavelet decomposition neural networks (WDNNs). The WDNNs decompose respective signals or series through wavelet transformation within the frequency domain before conducting a back-transformation into the time domain, respectively.

Established network topology	Experimental network
LSTM	WDNN (Level 1) WDNN (Level 5)

Notes: LSTM: long short-term memory; WDNN: wavelet decomposition neural network.

which results in an increase in the number of input time series by one and five, respectively. Both WDNNs are based on a basic LSTM topology. Hence, following the recommendation of Lahmiri (2014), Daubechies wavelet 4 (DB4) is the most feasible for decomposition purposes. Hereinafter, we generate the experimental NNs, as shown in Table 3.

4.3. Basic network structure, fitting procedure and accuracy intervals

The aforementioned network structure of the three layers is applied for all NNs built on an MLP basis. Further, regarding networks with an LSTM basis, additional layers are added to enable the function of n-step-ahead forecasts, which are not available for the MLPs under consideration. The output layer corresponds to the number of features (i.e. the number of time series) being fed into the NN. Nevertheless, our experiment intends only to predict the daily (-adjusted) closing prices. In terms of network structure, the hidden layer remains sufficiently variable such that the number of neurons can be adjusted, if necessary. Following the guidelines proposed by Guresen et al. (2011), the number of learning epochs (see Rashid, 2016) of an NN should be larger than the number of weights and, thus, of the neuron connections in the network (Guresen et al., 2011). However, respecting this guideline leads to overfitting in the present experiment (Guresen et al., 2011). The application of excessive amounts of learning epochs ensures that the memory capacity of the network approximates the training data perfectly, which negatively affects the prediction performance of unknown data (Lawrence et al., 1998). Therefore, for all topologies, it is important to determine the number of learning epochs that lead to the best forecasting results before overfitting sets in. To simplify this step, we implement an early stopping mechanism, as proposed in Caruana et al. (2000). This ensures that the training process is stopped, as soon as no further significant improvement in fitting can be achieved (Kaastra & Boyd, 1996). The mean squared error (MSE) can be calculated as a measure of goodness of fit (GoF) for a given (stock) time series, which is approximated by the referring network (Alexandridis & Zapranis, 2014). The MSE indicator describes the average of the squared deviations at each given value in the time series of the approximated curve to the real value of the training data curve (Alexandridis & Zapranis, 2014). The MSE is supposed to be minimised in the NN without causing overfitting (Alexandridis & Zapranis, 2014). The final value to be predicted must not be trained as an input data point in the NN (Adya & Collopy, 1998; Kaastra & Boyd, 1996). To test the repeated predictions of the NNs, we apply an 80–20 split of the data as proposed by Kaastra and Boyd (1996).

Therefore, all network topologies are based on only 80% of the data and subsequently predict values from the isolated 20% test dataset (Kaastra & Boyd, 1996). Finally, the specification of the results and their precision measures requires a formal discussion. Even if the well-accepted accuracy is a measure representing the GoF in terms of time-series approximation, it does not necessarily imply anything about how well an NN can predict future value developments outside the training data set (Cristea et al., 2000). Assuming overfitting, for instance, produces very high accuracy based on the training data, yet, displays no predictive capability (Caruana et al., 2000). Especially in the case of hybrid models, such as WNNs, it is not certain that the accuracy represents a performance measure that evaluates the predictive power of (financial) time series in a meaningful and correct manner (Alexandridis & Zapranis, 2014; Kumbure et al., 2022). Furthermore, we determine a potential gap within the academic literature; namely, only a few studies address this difficulty properly. Therefore, we disregard accuracy as a measure and apply another measure that is more practicable and more descriptive in terms of evaluating forecasts, as described in the following (Alexandridis & Zapranis, 2014). In the final back-test described in Section 4.3, we performed several runs and statistically evaluate individual predictions. Subsequently, we present the results as a 95% confidence interval around the true value to be predicted and as percentage values, following Altman and Bland (2005). Therefore, the individual prediction results are averaged over a time horizon. Further, twice the standard deviation of the mean value is subtracted and then added to obtain the accuracy interval, which we apply within our empirical setting (Altman & Bland, 2005).

4.4. Data description

Following Halevy et al. (2009), one of the most important factors for solving both research questions is represented by the data under investigation. Therefore, we extract all relevant data sets from the renowned Refinitiv Eikon Datastream (formerly Thomson Reuters), which are publicly available as adjusted-values, that is, adjusted for corporate actions, such as stock splits (Refinitiv Limited, 2019). Another factor to be considered in the choice of data is the respective stock exchange from which the time series are obtained. At this point, we favour the New York Stock Exchange (NYSE) because of its high trading volume (Statista Research Department, 2019). If the NYSE is not capable of providing sufficient data quality, the trading venue with the largest trading volume in the home country of the respective corporation is selected as the relevant source. All time series are provided in the related national currency of the stock exchange.[6] The recorded datasets were available until the end of February 2020. We examine 20 datasets, of which 15 are public companies and five are indices, as displayed in Table 4. The stocks analysed are based on eight of the largest companies within the MSCI World index in terms of market capitalisation. As these positions in the MSCI World index represent American stocks almost exclusively, we supplement the analysis with six of the largest positions within the Euro STOXX50 index (Ticker SX5T), in a manner that European stocks are also included in our experiment (STOXX, 2021). These datasets are supplemented ultimately by the stock of the Asian Alibaba Group.

The indices analysed are the NASDAQ100, DJIA100, DAX30, EURO STOXX50 and CSI300. If required, we clean the datasets of recording errors, i.e., in some cases, missing or erroneous values are averaged from the previous and subsequent values (Kaastra & Boyd, 1996). Refinitiv datasets are available on a daily basis and are provided as ".csv"-files. The procedure for time-series data originating from indices is handled analogously. The first date recorded in each case is the day when a share has been tradable on the referring stock exchange and, therefore, can differ from the respective initial public offering date. For all financial time series under consideration, the opening and closing prices, as well as the highest and lowest values within the daily data frequency, are recorded. Each dataset is split into training and test sets (refer to Section 4.3). The common division of 80% for the training set and 20% for the test set is applied (Kaastra & Boyd, 1996).

In this experiment, we implement NNs as regressors to output the prices of stocks and, consequently, are, in principle, not obligated to scale the input data (Albon, 2018). However, as observed in the datasets under analysis, many stock prices rise sharply over the years considered. To consider the latter insight, we apply a scaling function

[6] US-Dollar for American companies and Alibaba Group, Euro for European companies, as well as index points for indices.

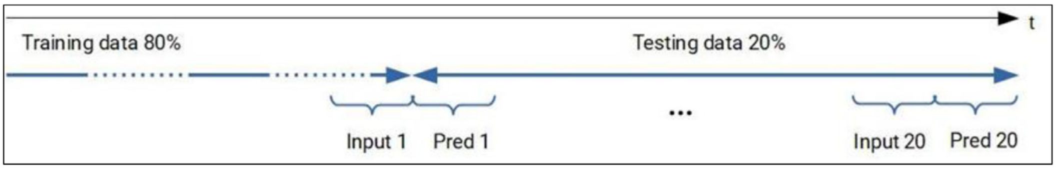

Fig. 3. Overview of the predictability logic of our empirical analysis, namely, iterative counts per prediction, per data set, as illustrated.

Table 4
Selected publicly data sets for this study, consisting of the company or respective index name, the ticker symbol, and the country of origin. The data sets are extracted from Refinitiv Eikon Datastream (formerly Thomson Reuters) on a daily frequency until February 2020.

Name	Symbol	Country
Stocks		
Alibaba Group	BABA.K	China
L'Oréal S.A.	OREP.PA	France
LVMH SE	LVMH.PA	
Allianz SE	ALV	Germany
Linde plc	LIN	
SAP SE	SAP.N	
Siemens	SIEGn.DE	
Apple Inc.	AAPL.O	USA
Alphabet Inc. C	GOOGL.O	
Amazon Inc.	AMZN.O	
Facebook Inc.	FB.O	
Johnson & Johnson	JNJ	
Johnson & Johnson	JNJ	
JP Morgan Chase	JPM	
Microsoft Corp.	MSFT.O	
Visa Inc.	V	
Indices		
China Securities Index 300	CSI300	China
EURO STOXX 50	STOXX50	Europe
DAX 30	GDAXI	Germany
Dow Jones Industrial Average	DJI	USA
NASDAQ 100	NDX	

that transfers the input variables into a range between zero and one. Additionally, we note the benefits of the efficiency of the respective NN by employing a respective scaling operation (Kaastra & Boyd, 1996). Moreover, we perform back-tests over a period of two years, for example, a training and test set split of 80–20 results in a training data period of eight years for a test period of two years, assuming that data for exactly 10 years are available. For stocks that are listed for a shorter period, the test period is reduced, whereas the 80–20 ratio remains constant. The NNs predict 30 days of future stock prices for each test iteration, which repeatedly records predictions from 20 successive inputs, as illustrated in Fig. 3. In the case of shorter datasets, we retain the 30 day prediction periods, yet, with a shorter iteration count than 20, such as Facebook Inc. (13 samples) or Alibaba Group (9 samples).

5. Empirical results

Before elaborating on respective results, we note the stochastic nature of NNs to impose a great influence on the magnitude deviations between the predictions (Kaastra & Boyd, 1996). The before mentioned implication originates from the existence of many predictions, which enable the prediction of only small fractions of a percent from the intended value. In addition, however, a similar number of predictions deviate from the true value by several percent, according to non-optimal gradient descent algorithms, as stated in Kaastra and Boyd (1996). In total, we record up to 600 predictions[7] from each stock or index and LSTM topology, less for shorter data sets,[8] respectively. Furthermore, up to 20 predictions are added with NNs on MLP basis for each data set, thus, we obtain a total number of 58,446 predictions, which are documented and evaluated according to the back-test described previously. As suspected, the examination of the prediction results provides a differentiated picture of the quality of predictions regarding the different selected NNs. Therefore, it is favourable to address different perspectives within the respective evaluations. The first perspective envisages the results with a focus on the general performance of an NN over the time horizon from one to 30 days. Therefore, we consider all available results of each network, separated per time horizon but averaged across all data sets. The second perspective elucidates the performance within the predicted stocks and indices. For each stock or index, the best network is determined for each time horizon. Further, it can be deduced whether an NN performs better than others regarding certain data sets and time-horizons.

Hence, we present the results in more detail. The various result tables display the prediction accuracy around the true value by calculating the confidence interval described in Section 4.3, resulting in more narrow intervals for better predictions than wide intervals. The intervals are given at 95% confidence; hence, individual predictions of individual networks are much more accurate. Because two standard deviations are selected, the intervals can also yield negative values. Nevertheless, due to the stochastic nature of NNs, more accurate outcomes are not guaranteed in every prediction. The variation ranges in the results across all stocks and indices examined are, in the best case, approximately 9% around the true value and about 44% in the worst case. The best case is presented as a one-day-ahead forecast stated in Table 5, indicating the best NNs per prediction period. The worst-case interval is approximately 5% wider than the best network displayed for a 30 day-ahead prediction (approximately 39% interval width) and, therefore, is not shown in Table 5. Hence, there are large differences in the respective fluctuation ranges of the accuracy intervals, as stated in Table 5. A possible explanation is provided by the predicted time horizon.

Following Nguyen and Chan (2004) it is found that the longer the time horizon, the wider the interval around the true value and the worse the long-term forecast. The best case mentioned above is represented by the forecast for the next day, while the worst case is a 30 day forecast. Regarding the evaluation of individual stocks, the referring intervals are sometimes more precise, namely, up to 3.21% for the Johnson & Johnson stock, yet, mostly within the range of the values stated in Table 5. In individual cases, such as within the Apple Inc. dataset, however, these are significantly less accurate, resulting in extreme cases in terms of fluctuation ranges of up to 89.60%, rendering predictions futile. To provide a holistic representation of all results, the complete overview of our empirical findings are given in the supplementary material of this study.

In general, many cases of the researched networks (i.e. WNNs and WDNNs) indicate better performance than the basic topologies, despite a few limitations, as shown in Tables 6 and 7. These tables provide a ranking of the best NNs with respect to individual datasets. The first ranking (see Table 6) displays results for a period of one day and the second ranking (see Table 7) for longer-term predictions. For example,

[7] 20 times 30 days.

[8] For example, Facebook Inc. (13 times 30 days) or Alibaba Group (9 times 30 days).

the MLP continues to be the best NN for one-day-ahead forecasts for 40% of the datasets examined. The WNN with the first Gaussian derivative as the activation function ranks as the second best in terms of predictability performance and is the best NN on more than one-third of the datasets. Thus, the basic MLP and WNN (MLP & Gauss1) are almost equivalent. Even though the performance is slightly lower than that of the MLP, the performance is better than the LSTM topology. This over-performance can still be seen in long-term forecasts (see Table 7). The WNN (LSTM & Gauss1) is the best predictor for 8 out of 20 datasets and, thus, is superior to LSTM. By contrast, the WNN (LSTM & Gauss2) does not represent the best performer for any dataset, but displays a similar performance to the established LSTM. Even if the WDNNs provide the best results with 15% and 10% for a quarter of all data sets, only a closer examination reveals further important indications.

Moreover, adding time series to the input, which is pre-processed with wavelet decomposition, does not always produce favourable results and does not lead to an improvement in the performance of the predictions.

For example, regarding the topology WDNN (Level 5), some predictions perform worse than all other NNs. However, few cases exist in which the WDNN (Level 5) reveals over-performance that is, actually representing the best results within a dataset. Compared to the other topologies, the WDNN (Level 1) provides the best performance over all data sets for a period of three days. Furthermore, the accuracy interval of [−7.35%; 6.65%] displays a width of 14% around the true stock price. Regardless of the comparisons, the best prediction of the WDNN (Level 1) is found with a two-day prediction period and a fluctuation interval of 13.4%. For single datasets, the WDNN (Level 1) is best suited for short-term prediction, namely, for Microsoft Corp., Apple Inc., JP Morgan Chase, Allianz SE, and the CSI300 Index. An example is shown in Fig. 4. Please note that due to the high volume of forecasts, only selective graphical displays are possible since a holistic display would render itself unsuitable. The aforementioned WDNN (Level 5) causes said topology to be at a disadvantage over the total set of predictions in most cases. In addition, the WDNN also provides a systematic under-estimation of the stock price at the 95% significance level of about 20% on average, which is reflected in an accuracy interval width of 40%.

However, an exceptional perspective on the WDNN (Level 5) emerges while regarding individual stock datasets separately. The WDNN (Level 5) is the best topology for long-term predictions of the stocks of Apple Inc., Facebook Inc., and Alibaba Group, as well as the STOXX50 Index. The WNNs implemented in this experiment provide the most accurate predictions for long-term time horizons out of all topologies and datasets under consideration. With a prediction period ranging from four to nine days, the WNN with the first Gaussian derivative as an activation function provides the best predictions. Further, from the 10th to the 30th prediction, the WNN with the Mexican Hat activation function is more advantageous. This answers the question of which wavelet of the two is better suited from this perspective. The Mexican Hat wavelet displays better results, but with little difference compared to the first Gaussian derivative. The accuracy intervals for both WNNs range from approximately 15% (four days) to 29% (30 days) around the true value. Following Fig. 5, we present an example of a 30 day forecast employing the WNN (LSTM & Gauss2). The basic topology, which is more adequate for WNNs, is represented by the MLP for one-day-ahead price forecasts. The first Gaussian derivative as a wavelet is slightly better suited (interval [−6.12%; 3.71%]) than the Mexican Hat function (interval [−4.89%; 5.21%]). For longer periods, only LSTM is favourable because of the possibility of n-step-ahead forecasts. Contrary to the one-day-ahead forecast, the Mexican Hat wavelet reveals a better performance than the first Gaussian derivative, with respect to time horizons of at least 10 days.

Finally, the third perspective elucidates the number of almost flawless predictions. Now we examine the amount of predictions of each network topology, which lie within a specified interval. Table 8, shows that the MLP as an established topology also presents the best precision with one-day-ahead predictions. Of all predictions, 5% lie within a tenth of a percent frame around the value to be achieved, and two-thirds of all predictions are between −2% and +2%. Further, the three experimental NNs achieve a more accurate "hit rate"[9] than the basic LSTM model.

Furthermore, we note that the studied MLP-based NNs offer a more accurate performance than those based on LSTM across all networks with WDNN (Level 1) as the only exception. A combination of the results regarding the accuracy interval and precision suggests the application of wavelets as an activation function to reduce the scatter of the predictions, but it does not necessarily improve the accuracy itself. However, if the wavelet functions are operationalised for input processing, a larger variation is observed, with an equally inconsistent increase in flawlessness. Furthermore, we present a complete overview of the one-day versus multiple-day forecasting topologies for each dataset in Table 9.

Table 5
Evaluation table referring to all data sets, stating the best network topology per period as well as respective accuracy intervals.

Prediction period in days	Accuracy interval	Best topology per period
1	[−4.94%; 4.05%]	MLP
2	[−6.28%; 6.02%]	LSTM
3	[−7.35%; 6.65%]	WDNN (Level1)
4	[−10.11%; 5.46%]	
5	[−10.59%; 5.85%]	
6	[−11.27%; 6.34%]	WNN (LSTM & Gauss1)
7	[−11.56%; 6.69%]	
8	[−11.80%; 7.01%]	
9	[−12.29%; 7.50%]	
10	[−11.76%; 8.89%]	
11	[−11.87%; 8.78%]	WNN (LSTM & Gauss2)
12	[−12.17%; 9.11%]	
13	[−12.75%; 9.61%]	
14	[−13.05%; 9.53%]	
15	[−13.39%; 9.88%]	
16	[−13.58%; 9.86%]	
17	[−13.82%; 9.73%]	
18	[−14.05%; 10.00%]	
19	[−14.10%; 9.85%]	
20	[−14.55%; 10.16%]	
21	[−14.98%; 10.65%]	
22	[−15.49%; 10.99%]	WNN (LSTM & Gauss2)
23	[−15.73%; 11.14%]	
24	[−15.67%; 10.99%]	
25	[−16.02%; 11.31%]	
26	[−16.36%; 11.89%]	
27	[−16.55%; 12.00%]	
28	[−16.72%; 12.12%]	
29	[−16.70%; 12.22%]	
30	[−16.78%; 12.22%]	

Notes: MLP: multilayer perceptron; LSTM: long short-term memory; WDNN: wavelet decomposition neural network; WNN: wavelet neural network.

Table 6
Ranking of the most successful network topologies for a predictive period of one day (pred. = 1 day).

Network topology	[%] — Share of data sets best predicted (pred. = 1 day)
MLP	40%
WNN (MLP & Gauss1)	35%
LSTM	10%
WNN (MLP & Gauss2)	5%
WDNN (Level 1)	5%
WNN (LSTM & Gauss2)	5%

Notes: MLP: multilayer perceptron; LSTM: long short-term memory; WDNN: wavelet decomposition neural network; WNN: wavelet neural network.

[9] Describes the number of predictions meeting the target accuracy interval (in line with Gupta & Lam, 1996).

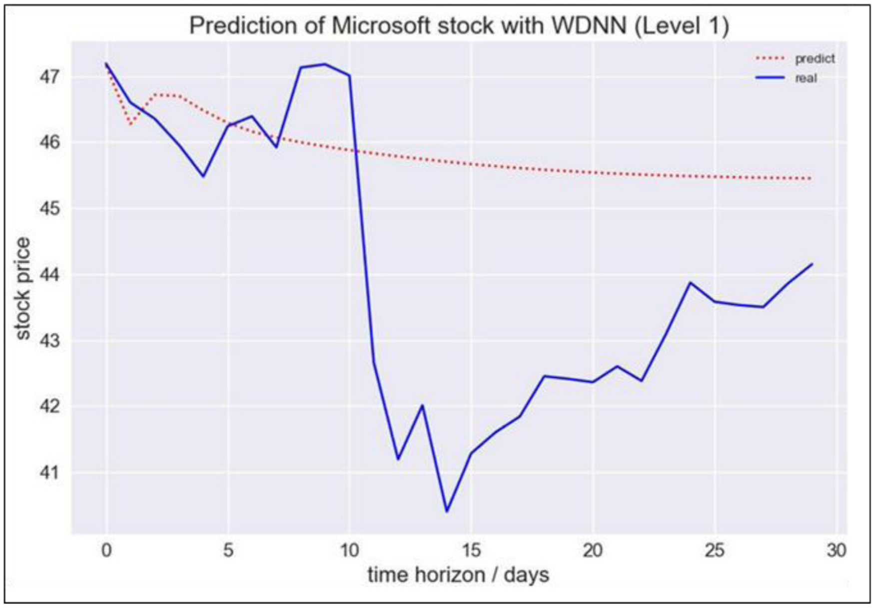

Fig. 4. Example of 30 day prediction with experimental wavelet decomposition neural network (Level 1) based on the Microsoft stock data set. Notes: WDNN: wavelet decomposition neural network.

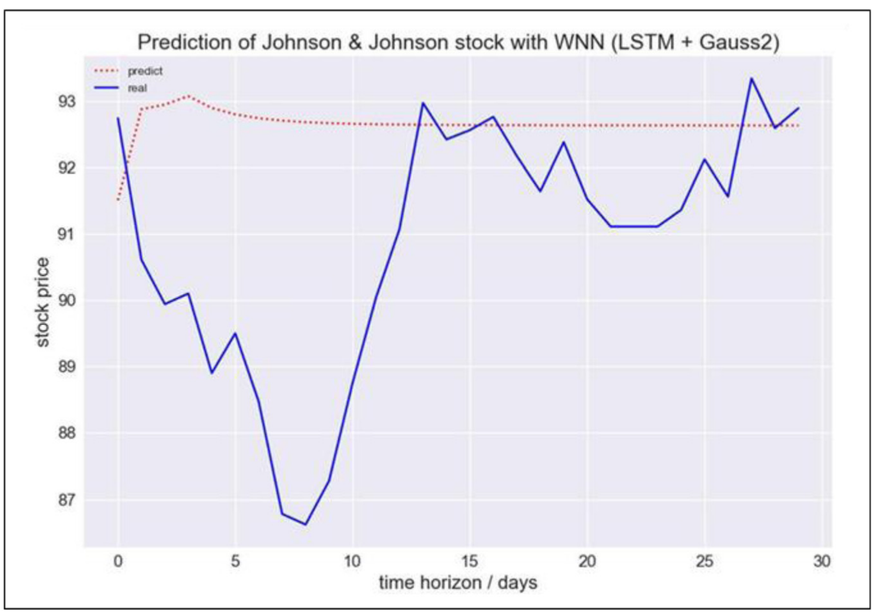

Fig. 5. Example of 30 day prediction with experimental wavelet neural network (LSTM & Gauss2) based on the Johnson & Johnson stock data set. Notes: WNN: wavelet neural network.

6. Discussion and implications

The empirical results of this study confirm that hybrid models are potentially more advantageous than classical NNs (e.g. Bao et al., 2017; Cristea et al., 2000; Yang & Wang, 2021; Zhang et al., 2001). In addition, regarding the present experiments, the MLP is one of the most reliable predictors, which lends credence to its continued application in research and practice (Guresen et al., 2011; Sezer et al., 2020). To be more detailed, we present empirical implications in Section 6.1, while subsequently elaborating on the performance comparison with buy-and-hold investment strategies in Section 6.2, elucidating ethical considerations in Section 6.3, propose limitations in Section 6.4 and state future avenues of research in 6.5.

6.1. Empirical implications

To the best of our knowledge and following the implications of these results, the premise that simply increasing the amount of data will lead to enhanced predictive results is not confirmed with respect to the experimental setting of this study. Even before the final experiment, the time series for logarithmic returns and trading volumes per day tested as additional inputs led to inadequate predictions of all relevant network topologies. Therefore, we discard the latter approach. Regarding WDNN (Level 5), adding additional time series merely to increase data volume is not favourable. These findings align with the conclusions of Walczak (2001), who states that larger amounts of data

Table 7
Ranking of the most successful network topologies for a predictive period of multiple days (pred. > 1 day).

Network topology	[%] — Share of data sets best predicted (pred. = 1 day)
WNN (LSTM & Gauss1)	40%
LSTM	35%
WDNN (Level 5)	15%
WDNN (Level 1)	10%
WNN (LSTM & Gauss2)	0%

Notes: MLP: multilayer perceptron; LSTM: long short-term memory; WDNN: wavelet decomposition neural network; WNN: wavelet neural network.

Table 8
Precision of different topologies displayed as "hit rate", which describes the number of predictions meeting the target accuracy interval (in line with Gupta & Lam, 1996).

One-day prediction	[%] — Share of forecasts in the interval		
Topology	[−0.1%; 0.1%]	[−1.0%; 1.0%]	[−2.0%; 2.0%]
MLP	5%	36%	66%
WNN (MLP & Gauss1)	3%	31%	54%
WDNN (Level 1)	3%	24%	43%
WNN (MLP & Gauss2)	2%	31%	58%
LSTM	2%	21%	42%
WNN (LSTM & Gauss1)	1%	18%	37%
WNN (LSTM & Gauss2)	1%	21%	39%
WDNN (Level 5)	1%	2%	5%

Notes: MLP: multilayer perceptron; LSTM: long short-term memory; WDNN: wavelet decomposition neural network; WNN: wavelet neural network.

do not consistently produce better forecasting results. Although WDNN (Level 5) achieves good results in a few cases, the time required by the network for training is increased significantly to the point that the respective cost–benefit ratio no longer matches that of the other topologies (e.g. MLP or LSTM).

Further, considering Norvig's remark, namely, that better data is more important than simply more data, we must discuss whether the data that a WDNN additionally receives is really better data. By performing low-pass filtering during decomposition, smoothing[10] (see Fig. 6) of the input data sets occurs, as illustrated in Fig. 5. However, this is a correction of the data by high-frequency price fluctuations. Yet, it contrastingly represents a falsification of the actual true dataset. Thus, the possibility of mutually offsetting effects must be considered. Therefore, we assume that the increase in quantity per se does not ensure better results if the added data do not increase the overall data quality. Furthermore, predictions of MLPs, which receive only a single time series as input, are not significantly worse in comparison with those that process four or more time series. By contrast, the performance in many cases is even better, as can be seen, for example, in the analysis of Alphabet Inc. and Microsoft Corp.

Moreover, we assess whether a larger dataset leads to better long-term predictions. We explicate a former presupposition by the two shorter time series such as Facebook Inc. (since 2012) and Alibaba Group (since 2014). Complementing the latter comparison, the longest data sets, namely Johnson & Johnson and JP Morgan Chase (both since 1980), are noted. Considering the longest prediction periods of 25 to 30 days, Facebook Inc. and Alibaba Group show interval widths in the range of 30% to 40%; hence, they are significantly less accurate predictions than the other datasets. However, smaller intervals only occur for one of the previously mentioned stocks with a long data period, namely, Johnson & Johnson. The inaccuracy of the Johnson & Johnson predictions lies within the range of less than 20%. However, the JP Morgan Chase predictions resemble the short data sets at approximately 30%. Thus, we cannot currently conclude that longer input periods also lead to more accurate long-term forecasts. Simultaneously, regarding the referred data sets, whether greater input with respect to the data period leads to an increase in performance in general is yet to be discussed. We exclude the Apple Inc. data set, which is the least predictable across all NNs. Nevertheless, for the other stocks and indices, no clear relation between the length of a data set and the quality of the prediction is visible. Predominantly, one-day-ahead forecasts range within 6% and 8% fluctuations around the true value.

Exceptions can be determined for three data sets, namely, Amazon Inc. with 10% interval width, Johnson & Johnson with 3%, and EURO STOXX50 with 4%. Some of the longer price time series, such as Microsoft Corp., result in accuracy intervals around 5% to 6%. Data sets that consider 30 to 40 years of data show an average fluctuation of around 7%, while those displaying only less than 20 years of data points also present the same fluctuation. This finding leads to the conclusion that a longer data period is not necessarily advantageous compared to shorter data sets. Moreover, Walczak's (2001) claim that larger data amounts do not enhance NN forecasts is substantiated. Referring to the most similar publication to this present study, namely, the forecasting results of Bao et al. (2017), showing NNs with wavelet pre-processed input to be at an advantage against other networks, for example, LSTM, are proven.

Agreeing with the presupposition of Bao et al. (2017) that NNs with wavelet components can outperform classical topologies, our study reveals that not only intermediate procedures outperform classical topologies; the presented WNNs developed from the generalisation methodology perform at an even higher precision compared to the respective intermediate results. Therefore, more innovative WNNs (as proposed and elaborated on Yang & Wang, 2021) are preferred over WDNNs in the majority of time horizons and data sets. The discussion envisages the attempt to understand and render the functioning of NNs comprehensible by proven methods to reach its limits in many cases. Therefore, we cannot confirm some detailed aspects (e.g. the outperformance of WNN over MLP as stated by Zhang et al., 2001 derived from the literature in this study). Consequently, it is particularly difficult to grasp the performance of NNs from a big picture perspective. The methods in the research area of explainable AI, that is, the investigation of how NNs achieve results as stated in Adadi and Berrada (2018), should begin at this point and gain further importance and necessity in the future, while employing an increasing number of applications of AI.

We evaluate the precision of the NNs under analysis, as far as exact predictions are concerned, as is not completely practical so far. As only a single digit percentage of the prediction results can be found within 0.1% of the exact value, we neglect the execution of these models as a sole prediction tool. As a workaround, we suggest training multiple NNs of a topology on the same data set and then consider and evaluate multiple predictions in the same period, which may optionally average the latter to achieve higher accuracy. This may mitigate the stochastic aspect to which NNs are fundamentally subject, even though this increases the already high expenditure on computational resources even further. Alternatively, the presented networks are suitable as a complement to the tools applied in fund management practices so far. Reducing the black-box problem with the aforementioned explainable AI is expected to help in terms of understanding and acceptance (Remus & O'Connor, 2001). This development can already be seen in Zhou et al. (2019), implementing NNs in terms of predictive algorithms, while Li and Kuo (2008) generally show that wavelet algorithms help to maximise returns on given timescales for financial institutions. Moreover, we state Puchalsky et al. (2018) proposing several optimisation algorithms for WNNs, which can be seen as direct enhancement of our findings in terms algorithmic improvement possibilities. In addition, Kanarachos et al. (2017) applies WNNs to detect successfully real-time anomalies within markets, which our findings are capable of displaying further improvement capabilities due to the comparability character of our study. Regarding the generalisability of our findings, WNNs are also applied in different scientific domains facing the same optimisation problems. For example, Alexandridis and Zapranis (2013, 2014) implement the discussed WNN solutions for financial, chaotic, wind (refer to

[10] That is, denoising the time series (Torrence & Compo, 1998).

Table 9
Best network topology for a one-day predictive period (pred. = 1 day) and for longer periods (pred. > 1 day) for each selected data set within our respective analysis.

Name	Best network (pred. = 1 day)	Best network (pred. > 1 day)
Stocks		
Alibaba Group	MLP	WDNN (Level 5)
L'Oréal S.A.	MLP	WNN (LSTM & Gauss1)
LVMH SE	MLP	LSTM
Allianz SE	WDNN (Level 1)	WNN (LSTM & Gauss1)
Linde plc	WNN (MLP & Gauss1)	WNN (LSTM & Gauss1)
SAP SE	WNN (MLP & Gauss1)	LSTM
Siemens	WNN (MLP & Gauss1)	LSTM
Apple Inc.	MLP	WDNN (Level 5)
Alphabet Inc. C	WNN (MLP & Gauss2)	WNN (LSTM & Gauss1)
Amazon Inc.	WNN (MLP & Gauss1)	WNN (LSTM & Gauss1)
Facebook Inc.	WNN (LSTM & Gauss2)	LSTM
Johnson & Johnson	LSTM	LSTM
Johnson & Johnson	LSTM	LSTM
JP Morgan Chase	LSTM	WDNN (Level 1)
Microsoft Corp.	WNN (MLP & Gauss1)	WNN (LSTM & Gauss1)
Visa Inc.	WNN (MLP & Gauss1)	WNN (LSTM & Gauss1)
Indices		
China Securities Index 300	WNN (MLP & Gauss1)	WDNN (Level 1)
EURO STOXX 50	MLP	WDNN (Level 5)
DAX 30	MLP	WNN (LSTM & Gauss1)
Dow Jones Industrial Average	MLP	LSTM
NASDAQ 100	MLP	LSTM

Notes: MLP: multilayer perceptron; LSTM: long short-term memory; WDNN: wavelet decomposition neural network; WNN: wavelet neural network.

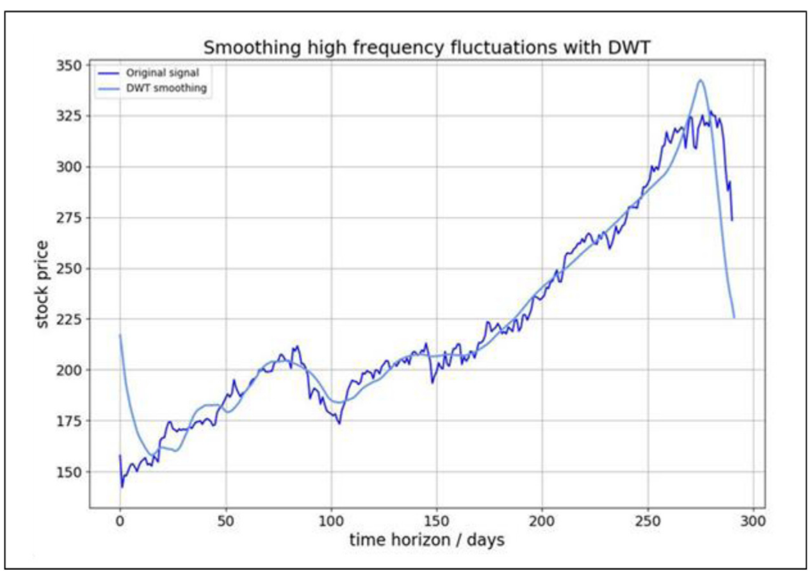

Fig. 6. Example for the pre-processed input sequence into wavelet decomposition neural networks, demonstrating the low-pass smoothing characteristics.

Doucoure et al., 2016; Liu et al., 2013) or breast cancer datasets and can further be applied to image processing, signal denoising, density estimates and time-scale decomposition (see Berghorn, 2015).

6.2. Performance comparison "buy-and-hold"

Especially in the financial domain, the discussion of the necessity of advanced algorithms and neural network solutions in comparison with "buy-and-hold" investment strategies is ongoing (Couillard & Davison, 2005; Vogl, 2021). The reason of the prominence of "buy-and-hold" or passive investment strategies, as introductory mentioned, lays within the still broadly stated efficient market hypothesis (EMH), which in short states that financial market data is obliged to randomness and forecasting attempts, thus, are futile in their nature (Fama, 1965a, 1965b; Fama & French, 1989). The EMH is vastly critiqued by Mandelbrot (1963) and Mandelbrot and Taylor (1967) and faced with the existence of stylised facts (i.e. empirical observations) such as dynamical nonlinearity (see Alexandridis et al., 2017), volatility dynamics (see Adams et al., 2017) or momentum-induced (multi-)fractal trends (e.g. Berghorn, 2015 or Daniel & Moskowitz, 2016). Even if the EMH is questioned further in its entirety due to stated deterministic chaotic characteristics (see Vogl & Rötzel, 2022), the implications and baseline comparisons of respective back-tests (see López de Prado, 2018) still level around beating "buy-and-hold" scenarios, although not all institutional trading facilities propagate such passive investments. Furthermore, following Berghorn (2015) states the theoretical possibility of outperformance, contrasting the EMH. Comparing the performance of NNs or advanced machine learning algorithms with "buy-and-hold" strategies in the literature can be seen, for example, in Nobre and Neves (2019) proposing hybrid AI models for trade signal generation for intra-day trading funds. In addition, Chalvatzis and Hristu-Varsakelis (2020) state NN outperformance over "buy-and-hold" strategies in automated asset trading scenarios, which our results may propose a solid groundwork for further return maximisation. Moreover, Sezer et al. (2017) state outperformance of machine learning algorithms, while Mitra (2009) shows higher risk-adjusted returns by ANNs than via "buy-and-hold", which also is of interest to institutional entities. Finally, referring to the encompassing review of Kumbure et al. (2022), show advanced algorithms to be capable to deal with complex system dynamics such as financial markets and propose detailed comparisons of methodologies and evaluations, while conclusively stating NNs to be capable of further prospering.

6.3. Ethical considerations

As AI and advanced technology is incorporating more space in our reality, a short debrief on ethical considerations in terms of its implementation is deemed relevant. AI technologies in general represent dual use, namely, the incorporation in peaceful civilian as well as military-driven systems, which is mostly ignored (de Ágreda, 2020). To cope with such a realisation, a ban of research activity is neither favourable nor is it possible to stop respective activities, thus, the need for ethical principles guiding such endeavours in imminent (de Ágreda, 2020). Said principles should follow two characteristics, namely, the algorithm functionality is understood and humans retain enough system control to intervene if required, since nefarious deployment of data is possible and AI is developed at a rapid growth rate (de Ágreda, 2020). Therefore, a vast discussion on ethical frameworks for advanced technology exists (e.g. AI4PEOPLE, EAD2, COMEST or DEEPMIND), revelling around the dimensions of beneficence, human dignity, privacy, human autonomy, fairness and explainability (de Ágreda, 2020; Parasuraman et al., 2000). Building upon these insights, AI provides high benefits for humans in general, yet, can be jeopardised without implementation of formerly denoted ethical codes and security measures (de Ágreda, 2020). Thus, following Yu et al. (2018), AI systems render themselves increasingly ubiquitous, which in the public experience, AI governance incorporating ethical standards become more relevant. Yu et al. (2018), therefore, propose a taxonomy incorporating the dimensions of ethical dilemmas, individuality in ethical decisions, given frameworks and human AI interaction. Moreover, owing to AI structuring, historical data application in social life applications often result in unjust biases and discrimination by machine learning algorithms, raising discussions about fairness and debiasing (Birhane, 2021). Especially, social systems are vulnerable due to the increase in mathematicalisation and formalisation of social issues owed to the advanced propositions of AI systems, leading to operations characterisable as value-free, neutral or even amoral (Birhane, 2021). Further following Birhane (2021) state a historical discussion about the roots of these developments and offers guidance on the "correctness" of biases in terms of their definition.

Due to the lack of a moral agent in machines, Etzione and Etzione (2017) propose reasons for their incorporation in AI systems, yet, stating the risk of the latter drawing on extreme outliers, which would lead to fatal errors, labelling this occurrence as outlier fallacy. Therefore, Banerjee (2020) suggests a computational framework for engineering intelligence to understand the concept of machine consciousness better. Moreover, Gruson et al. (2019) discuss the generating process of AI systems in the dimensions of ethics, legal predicaments, privacy and financing in terms of tailor-made versus off-the-shelf AI solutions, while discussing its implications regarding augmented reality. Gal et al. (2020) show the discourse of AI in people analytics, especially pointing out its impact in management support systems, stressing bias-free and lack of ethics. Building upon the moral agent in AI systems, Nath and Sahu (2020) state its inexistence due to the lack of the answer to the questions of "why being moral matters" in terms of a moral agent to be functional in AI systems. Ashok et al. (2022), therefore, propose generalised digital ethical frameworks for AI and digital technologies.

Referring to the implementation of WNNs and advanced algorithms in financial disciplines or practical processes, the authors state the management decision problem as given. Furthermore, black-box systems, such as NN solutions, render a full "understanding" merely impossible. Autonomous trading systems are prone to outliers as well as possible to investments in morally discussable assets and market regimes. Biases in investment decisions are reflected potentially in AI system once past trading results (originally conducted by a human) is applied for training purposes. Presupposing AI systems to become more autonomous or even sentient in the future, the authors see the potential implementation of investment guidelines (e.g. CFA code of conduct) into such AI systems conducted for trading and forecasting as favourable to prevent misuse or fraud. Nonetheless, a formal framework for AI forecasts in the financial domain is still lacking.

6.4. Limitations and future research

This study has several shortcomings. First, we focus only on a few indices and stocks as a singular asset class and neglect a broader pool of available financial market data (e.g. commodities and bonds). Furthermore, we only analyse daily-frequented data sets and disregard an elaboration on higher frequencies (e.g. intraday data sets). Moreover, we do not discuss in detail the implications stemming from the choice of price versus different returns on the research questions at hand. Furthermore, we note that in this study, we focus mainly on plain or standard NN implementations in terms of topology selection (e.g. LSTM or MLP) and do not discuss or test our wavelet decomposition approach with more complex topologies (as proposed in Yang & Wang, 2021) that are currently discussed in AI research. In addition, we do not elaborate on the potential of customised wavelet functions, which we deem to be interesting in combination with more sophisticated network topologies. One aspect that has hardly been addressed recently is the real-time application and usability of NNs in the field under discussion (refer to Kanarachos et al., 2017). Thus, we note that the training of all implemented NN topologies requires a high amount of time and computational effort (Bao et al., 2017). Further, practical research should focus on improving efficiency and simplifying handling in terms of NNs (Remus & O'Connor, 2001). Making AI accessible to groups without an affinity for computer science should be brought into focus to promote more widespread adoption of its advantageous applicability. In addition, the dissemination and execution of AI methods in critical research fields (e.g. quantitative finance) pose further challenges in risk management and associated legal consequences, which must be addressed (Adadi & Berrada, 2018). In particular, regarding the described fluctuations in the quality of the results of the discussed network topologies, risk management should not be neglected. Regarding methodological restrictions, the application of WNNs is limited to applications of small input dimensions owing to computationally expensiveness if facing high dimensional input vectors, even if capable of handling nonlinear and non-stationary datasets (Alexandridis & Zapranis, 2013). Finally, training of WNNs with backpropagation requires storage and inversion of some matrices, which in case of larger datasets, grow fondly large and, thus, computationally expensive (Alexandridis & Zapranis, 2014).

6.5. Concluding remarks

In this study, we demonstrate experimentally that hybrid models (e.g. WNNs and WDNNs) have advantages over classical topologies (e.g. LSTM) regarding financial market predictions. However, adding more data does not necessarily improve prediction performance because the increased data quantity in the given case is accompanied by a loss of data quality and leads to cancellation effects. The implemented wavelet functions may be partially recommended as an alternative to the sigmoid activation function; however, the choice is dependent on the respective data set. Hypotheses that both approaches, namely, intermediate and generalisation, lead to an increase in forecasting performance compared to classical NNs, can be reinforced, apart from the MLP, which is the best predicting NN with regard to one-day-ahead forecasts. Therefore, WNNs are preferable for longer-term forecasts. To apply the presented concepts in economic practice and better assess the risks for investment fund practice, additional tests on further data sets and a significant increase in precision are required, which reflects the state-of-the-art goal within our sampled literature since outperformance of "buy-and-hold" strategies are already partially applicable. Nevertheless, following the conclusions in academic literature, the proof-of-concept concerning NNs with wavelet components is fully substantiated for the presented models. Therefore, we propose our findings as generalised basis and solid groundwork for future studies, as help to select an optimised combination of topology versus wavelet as well as within other optimisation procedures.

CRediT authorship contribution statement

Markus Vogl: Conceptualisation, Methodology, Validation, Investigation, Resources, Writing – review & editing, Visualisation, Supervision. **Peter Gordon Rötzel:** Supervision, Project administration, Funding acquisition. **Stefan Homes:** Software, Formal analysis, Data curation, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.mlwa.2022.100302.

References

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE*, *6*, 52138–52156, In press.

Adams, Z., Füss, R., & Glück, T. (2017). Are correlations constant? Empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance*, *84*, 9–24.

Adya, M., & Collopy, F. (1998). How effective are neural networks at forecasting and prediction? A review and evaluation. *Journal of Forecasting*, *17*, 481–495. http://dx.doi.org/10.1002/(SICI)1099-131X(1998090)17:5/6<481::AID-FOR709>3.0.CO;2-Q.

de Ágreda, Á. G. (2020). Ethics of autonomous weapons systems and its applicability to any AI systems. *Telecommunications Policy*, *44*(6), Article 101953.

Albon, C. (2018). *Machine learning kochbuch. 1. Auflage*. Heidelberg: O'Reilly Media Inc.

Alexandridis, A. K., Kampouridis, M., & Cramer, S. (2017). A comparison of wavelet networks and genetic programming in the context of temperature derivatives. *International Journal of Forecasting*, *33*, 21–47.

Alexandridis, A. K., & Zapranis, A. D. (2013). Wavelet neural networks: A practical guide. *Neural Networks*, *42*, 1–27. http://dx.doi.org/10.1016/j.neunet.2013.01.008.

Alexandridis, A. K., & Zapranis, A. (2014). Wavelet neural networks: With applications in financial engineering, chaos, and classification.

Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *BMJ (Clinical Research Ed.)*, *331*(7521), 903. http://dx.doi.org/10.1136/bmj.331.7521.903.

Anjoy, P., & Paul, R. K. (2017). Comparative performance of wavelet-based neural network approaches. *Neural Computing and Applications*, 1–11.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, *62*, Article 102433.

Banerjee, S. (2020). A framework for designing compassionate and ethical artificial intelligence and artificial conciousness. *Interdisciplinary Description of Complex Systems*, *18*(2-A), 85–95.

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS One*, *12*.

Barlow, H. B. (1983). Vision: A computational investigation into the human representation and processing of visual information: David Marr. San francisco: W. H. Freeman, 1982. pp. xvi + 397. *Journal of Mathematical Psychology*, *27*, 107–110.

Berghorn, W. (2015). Trend momentum. *Quantitative Finance*, *15*, 261–284.

Bernard, C., Mallat, S., & Slotine, J. -J. (1998). Wavelet interpolation networks. In *Proc. of ESANN' 98* (pp. 22–24). Bruges, Belgium.

Billings, Stephen, & Wei, Hua-Liang (2005). A new class of wavelet networks for nonlinear system identification. *IEEE Transactions on Neural Networks*, *16*, 862–874. http://dx.doi.org/10.1109/TNN.2005.849842.

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, *2*(2), Article 100205.

Burton, K. (2016). Inside a moneymaking machine like no other. *Bloomberg Markets*.

Caruana, R., Lawrence, S., & Giles, C. L. (2000). Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In *NIPS*.

Chakrabarty, A., De, A., Gunasekaran, A., & Dubey, R. (2015). Investment horizon heterogeneity and wavelet: Overview and further research directions. *Physica A: Statistical Mechanics and its Applications*, *429*(C), 45–61.

Chalvatzis, C., & Hristu-Varsakelis, D. (2020). High-performance stock index trading via neural networks and trees. *Applied Soft Computing*, *96*, Article 106567.

Chen, C. H. (1994). Neural networks for financial market prediction. In *Proceedings of 1994 IEEE international conference on neural networks* (pp. 1201).

Chen, Y., Yang, B., & Dong, J. (2006). Time-series prediction using a local linear wavelet neural network. *Neurocomputing*, *69*, 449–465.

Couillard, M., & Davison, M. (2005). A comment on measuring the Hurst exponent of financial time series. *Physica A. Statistical Mechanics and its Applications*, *348*, 404–418.

Cristea, P. D., Tuduce, R., & Cristea, A. I. (2000). Time prediction with wavelet neural networks. In *Proceedings of the 5th seminar on neural network applications in electrical engineering (IEEE Cat. No. 00EX287)* (pp. 5–10).

Crowley, P. M. (2007). A guide to wavelets for economists. *Journal of Economic Surveys*, *21*, 207–267. http://dx.doi.org/10.1111/j.1467-6419.2006.00502.x.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 221–247.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Transaction on Information Theory*, *36*, 961–1005.

De Faria, E. L., Albuquerque, M. P., Gonzalez, J. L., Cavalcante, J. T. P., & Albuquerque, M. P. (2009). Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. *Expert Systems with Applications*, *36*(10), 12506–12509.

Dong, B., Li, Z., Rahman, S. M., & Vega, R. E. (2016). A hybrid model approach for forecasting future residential electricity consumption. *Energy and Buildings*, *117*, 341–351.

Doucoure, B., Agbossou, K., & Cardenas, A. (2016). Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. *Renewable Energy*, *92*, 202–211.

Etzione, A., & Etzione, O. (2017). Incorporating ethics into artificial intelligence. *The Journal of Ethics*, *21*, 403–418.

Fama, E. F. (1965a). The behaviour of stock-market prices. *Journal of Business*, *38*(1), 34–105.

Fama, E. F. (1965b). Portfolio analysis in a stable paretian market. *Management Science*, *11*(3), 404–419.

Fama, E., & French, K. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, *25*(1), 23–49.

Gal, U., Jensen, T., & Stein, M. -K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, *30*, Article 100301.

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance*, *51*, 783–810. http://dx.doi.org/10.1111/j.1540-6261.1996.tb02707.x.

Gruson, D., Helleputte, T., Rousseau, P., & Gruson, D. (2019). Data science, artificial intelligence, and machine learning: Opportunities for laboratory medicine and the value of positive regulations. *Clinical Biochemistry*, *69*, 1–7.

Gupta, A., & Lam, M. S. (1996). Estimating missing values using neural networks. *Journal of the Operational Research Society*, *47*(2), 229–238. http://dx.doi.org/10.2307/2584344.

Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, *38*, 10389–10397. http://dx.doi.org/10.1016/j.eswa.2011.02.068.

Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, *24*, 8–12. http://dx.doi.org/10.1109/MIS.2009.36.

Hill, T., O'Connor, M., & Remus, W. (1996). Neural network models for time series forecasts. *Management Science*, *42*(7), 1082–1092, http://www.jstor.org/stable/2634369.

Jordan, M. I., & Mitchell, T. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*, 255–260.

Kaastra, I., & Boyd, M. S. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, *10*, 215–236.

Kanarachos, S., Christopoulos, S. -R. G., & Chroneos, A. (2017). Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform. *Expert Systems with Applications*, *85*, 292–304.

Kumbure, M., Lohrmann, C., Luuka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications*, *197*, Article 116659.

Lahmiri, S. (2014). Wavelet low-and high-frequency components as features for predicting stock prices with backpropagation neural networks. *Journal of King Saud University - Computer and Information Sciences*, *26*(2), 218–227. http://dx.doi.org/10.1016/j.jksuci.2013.12.001.

Lawrence, S., Giles, C. L., & Tsoi, A. C. (1998). *Lessons in neural network training: Overfitting may be harder than expected*. AAAI/IAAI.

Li, S. -T., & Kuo, S. -C. (2008). Knowledge discovery in financial investment for forecasting and trading strategy through wavelet-based SOM networks. *Expert Systems with Applications*, *34*, 935–951.

Lim, T., Loh, W., & Shih, Y. (2004). A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning*, *40*, 203–228.

Liu, H., Tian, H. -q., Pan, D. -f., & Li, Y. -f. (2013). Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks. *Applied Energy*, *107*, 191–208.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLoS One*, *13*.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 674–693.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business*, *36*, 394–419.

Mandelbrot, B. B., & Taylor, H. (1967). On the distribution of stock price differences. *Operations Research*, *15*, 1057–1062.

Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. (2007). *Wavelets and their applications*. Wiley-ISTE Ltd.

Mitra, S. (2009). Optimal combination of trading rules using neural networks. *International Business Research*, *2*(1), 2.

Nath, R., & Sahu, V. (2020). The problem of machine ethics in artificial intelligence. *AI & Society*, *35*, 103–111.

Nguyen, H. H., & Chan, C. W. (2004). Multiple neural networks for a long term time series forecast. *Neural Computing & Applications*, *13*(1), 90–98.

Nobre, J., & Neves, R. (2019). Combining principal component analysis, discrete wavelet transform and xgboost to trade in the financial markets. *Expert Systems with Applications*, *125*, 181–194.

Otuteye, E., & Siddiquee, M. (2019). Underperformance of actively managed portfolios: Some behavioral insights. *Journal of Behavioral Finance*, *21*, 284–300.

Oussar, Y., & Dreyfus, G. (2000). Initialization by selection for wavelet network training. *Neurocomputing*, *34*(1–4), 131–143.

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, *36*, 2–17.

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man & Cybernetics, Part A (Systems & Humans)*, *30*(3), 286–297.

Pati, Y., & Krishnaprasad, P. (1993). Analysis and synthesis of feedforward neural networks using discrete affine wavelet transforms. *IEEE Transactions on Neural Networks*, *4*(1), 73–85.

Peng, Y., Albuquerque, P. H. M., Kimura, H., & Saavedra, C. A. P. B. (2021). Feature selection and deep neural networks for stock price direction forecasting using technical analysis indicators. *Machine Learning with Applications*, *5*, Article 100060.

Petneházi, G. (2021). Quantile convolutional neural networks for value at risk forecasting. *Machine Learning with Applications*, *6*, Article 100096.

López de Prado, M. (2018). *Advances in financial machine learning*. Hoboken: John Wiley & Sons Inc.

Puchalsky, W., Ribeiro, G., Veiga, C. da., Freire, R., & dos Santos Coelho, L. (2018). Agribusiness time series forecasting using wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack price and perishable products demand. *International Journal of Production Economics*, *203*, 174–189.

Rashid, T. (2016). *Neuronale netze selbst programmieren. 1. Auflage*. Heidelberg: O'Reilly Media Inc.

Refinitiv Limited (2019). Thomson reuters equity indices – Corporate action methodology. (pp. 3–6).

Remus, W., & O'Connor, M. (2001). *Neural networks for time-series forecasting. Principles of forecasting*. New York: Springer Science and Business Media.

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. ArXiv, arxiv:abs/1911.13288.

Sezer, O., Ozbayoglu, M., & Dogdu, E. (2017). A deep neural-network based stock trading system based on evolutionary optimized technical analysis parameters. *Procedia Computer Science*, *114*, 473–480.

Statista Research Department 2019. Größte Börsen der Welt nach dem Handelsvolumen mit Aktien im Jahr 2018. (Accessed 9 March 2020).

STOXX 2021. EURO STOXX 50® INDEX 2021. (Accessed 3 March 2021).

Taspinar, A. (2018). A guide for using the wavelet transform in machine learning. Accessible via https://ataspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning/. (Accessed 9 March 2020).

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*, 61–78.

Tsaih, R., Hsu, Y., & Lai, C. C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system. *Decision Support Systems*, *23*, 161–174.

Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., & Iosifidis, A. (2017). Forecasting stock prices from the limit order book using convolutional neural networks. In *2017 IEEE 19th conference on business informatics*: *Vol. 01*, (pp. 7–12).

Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems with Applications*, *17*, 51–70. http://dx.doi.org/10.1016/S0957-4174(99)00016-0.

Vogl, M. 2021. Frontiers of Quantitative Financial Modelling: A Literature Review on the Evolution in Financial and Risk Modelling after the Financial Crisis (2008-2019). SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3764570. Submitted for publication.

Vogl, M., & Rötzel, P. G. (2022). Chaoticity versus stochasticity in financial markets: Are daily S&P 500 return dynamics chaotic? *Communications in Nonlinear Science and Numerical Simulation*, *108*, Article 106218.

Vui, C. S., Gan, K. S., On, C. K., Alfred, R., & Anthony, P. (2013). A review of stock market prediction with Artificial neural network, ANN. In *2013 IEEE international conference on control system, computing and engineering* (pp. 477-482).

Walczak, S. (2001). An empirical analysis of data requirements for financial forecasting with neural networks. *Journal of Management Information Systems*, *17*, 203–222.

Williams, J. R., & Amaratunga, K. (1997). A discrete wavelet transform without edge effects using wavelet extrapolation. *Journal of Fourier Analysis and Applications*, *3*, 435–449.

Yang, Y., & Wang, J. (2021). Forecasting wavelet neural hybrid network with financial ensemble empirical mode decomposition and MCID evaluation. *Expert Systems with Applications*, *166*, Article 114097.

Yu, H., Shen, Z., Miao, C., Leung, C., Lesser, V., & Yang, Q. (2018). Building ethics into artificial intelligence. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence* (pp. 5527–5533).

Zapranis, A., & Alexandridis, A. (2008). Modelling temperature time dependent speed of mean reversion in context of wheather derivative pricing. *Applied Mathematical Finance*, *15*(4), 355–386.

Zapranis, A., & Alexandridis, A. (2009). Weather derivatives pricing: Modelling the seasonal residuals variance of an Ornstein–Uhlenbeck temperature process with neural networks. *Neurocomputing*, *73*, 37–48.

Zhang, Q. (1997). Using wavelet network in nonparametric estimation. *IEEE Transactions on Neural Networks*, *8*(2), 227–236.

Zhang, Q., & Benveniste, A. (1992). Wavelet networks. *IEEE Transactions on Neural Networks*, *3*(6), 889–898. http://dx.doi.org/10.1109/72.165591.

Zhang, B., Coggins, R. J., Jabri, M. A., Dersch, D. R., & Flower, B. (2001). Multiresolution forecasting for futures trading using wavelet decompositions. *IEEE Transactions on Neural Networks*, *12*(4), 765–775.

Zhang, J., Walter, G. G., Miao, Y., & Lee, W. N. W. (1995). Wavelet neural networks for function learning. *IEEE Transactions on Signal Processing*, *43*(6), 1485–1497. http://dx.doi.org/10.1109/78.388860.

Zhou, F., Zhou, H. -m., Yang, Z., & Yang, L. (2019). EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction. *Expert Systems with Applications*, *115*, 136–151.