





Article

# Short-Term Natural Gas and Carbon Price Forecasting Using Artificial Neural Networks

Laura Böhm <sup>†</sup>, Sebastian Kolb <sup>\*,†</sup> , Thomas Plankenbühler , Jonas Miederer, Simon Markthaler   
and Jürgen Karl 

Chair of Energy Process Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU),  
Fürther Straße 244f, 90429 Nürnberg, Germany; thomas.plankenbuehler@fau.de (T.P.);  
jonas.miederer@fau.de (J.M.); simon.markthaler@fau.de (S.M.); juergen.karl@fau.de (J.K.)

\* Correspondence: sebastian.kolb@fau.de; Tel.: +49-911-5302-99028; Fax: +49-911-5302-99030

<sup>†</sup> These authors contributed equally to this work.

**Abstract:** Methods of computational intelligence show a high potential for short-term price forecasting of the energy market as they offer the possibility to cope with the complexity, multi-parameter dependency, and non-linearity of pricing mechanisms. While there is a large number of publications applying neural networks to the prediction of electricity prices, the analysis of natural gas and carbon prices remains scarce. Identifying a best practice from the literature, this study presents an iterative approach to optimize both the input values and network configuration of neural networks. We apply the approach to the natural gas and carbon market, sequentially testing autoregressive and exogenous explanatory variables as well as different neural network architectures. We subsequently discuss the influence of architectural properties, input parameters, data preparation, and the models' resilience to singular events. Results show that the selection of appropriate lags of gas and carbon prices to account for autoregressive properties of the respective time series leads to a high degree of forecasting accuracy. Additionally, including ambient temperature data can slightly reduce errors of natural gas price forecasting whereas carbon price predictions benefit from electricity prices as a further explanatory input. The best configurations presented in this contribution achieve a root mean square error (RMSE) of 0.64 EUR/MWh (natural gas prices) corresponding to a normalized RMSE of 0.037 and 0.33 EUR/t<sub>CO2</sub> (carbon prices) corresponding to a normalized RMSE of 0.023.

**Keywords:** neural network; natural gas price; carbon price; multi-layer perceptron; price prediction



**Citation:** Böhm, L.; Kolb, S.; Plankenbühler, T.; Miederer, J.; Markthaler, S.; Karl, J. Short-Term Natural Gas and Carbon Price Forecasting Using Artificial Neural Networks. *Energies* **2023**, *16*, 6643. <https://doi.org/10.3390/en16186643>

Academic Editor: Rocío Pérez de Prado

Received: 14 July 2023

Revised: 15 August 2023

Accepted: 16 August 2023

Published: 15 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Despite the ongoing displacement of conventional and nuclear energy carriers by renewable sources, natural gas and natural gas substitutes are still seen as a central element of many countries' future energy systems. Comparing 36 scenarios on the transformation of the German energy sector, the meta-analysis of Scharf et al. [1] illustrates that the demand for natural gas will not decrease considerably until German greenhouse gas reductions compared to 1990 exceed 70%. The authors argue that one reason for the constant high gas consumption in the scenarios is the conventional fuel switch. That is, natural gas replaces energy provision from commodities with a higher carbon intensity such as hard coal or lignite. The main driver for this switch is carbon prices, offering an economic benefit to natural gas due to its (comparably) low CO<sub>2</sub> emissions (the actual emissions of natural gas, especially with regard to its upstream fugitive methane losses, during its production and transportation will, with reference to [2], not be discussed here).

The predicted high consumption of natural gas as well as the impact of carbon prices on energy market demands adequate models for their price prediction. In this context, a multitude of modeling approaches for the prediction of energy prices exists, as Weron [3] highlights and classifies in his extensive review. Multi-agent [4–9] and fundamental [10–12] models replicate the existing plant fleet and market interactions in a high degree of detail

and allow the analysis of long-term horizons. For shorter (e.g., day-to-day) predictions, statistical models (e.g., ARX-type [13–16], GARCH-type [17–19]) or computational intelligence models [20–23] are common. The review of Nowotarski and Weron [24] points out that the majority of recent articles on electricity price forecasting make use of neural networks as they offer the possibility to cope with the complexity, multi-parameter dependency, and non-linearity of pricing mechanisms [3].

While there is a large number of publications applying methods of computational intelligence to the prediction of electricity prices, the analysis of natural gas and carbon prices remains scarce (cf. Čeperić et al. [25]). This contribution aims at transferring the benefits of computational intelligence to those commodities as well. Of the few existing studies, the majority focus on the consideration of autoregressive properties of gas and carbon price (i.e., not exploring the impact of exogenous parameters such as weather data, demand, or cross-correlations with other commodities). In doing so, the studies neglect a vast part of potential explanatory variables. This potentially leads to less accurate and robust results. To address these research gaps, this contribution suggests an approach to systematically vary input parameters and network configurations of neural networks for energy market price predictions. The proposed methodology enables us to find optimized model setups in a reproducible way, delivering competitive results despite its straightforward implementation. This gives the potential to provide accurate and reliable, yet computationally cheap, predictions of gas and carbon prices. Such predictions are crucial for policymakers, energy traders, and investors to make informed decisions, optimize resource allocation, and mitigate financial risks. Using the strengths of artificial neural networks, we introduce a novel methodology that leverages complex patterns and correlations in historical data to achieve precise forecasts.

We show the applicability of the methodology for the prediction of natural gas and CO<sub>2</sub> prices. Gas prices in this context represent the day-ahead contracts of the NetConnect Germany (NCG) market zone traded at the pan-European gas cooperation (PEGAS). CO<sub>2</sub> prices refer to the daily settlement prices of daily future contracts (ECP D0) traded at the international climate exchange (ICE).

Using a large pool of data (2007–2020 for natural gas and 2012–2020 for carbon prices) and the best practice of relevant network architectures from a literature review, we model a generic neural network that sequentially and systematically tests different network configurations and input parameters. This helps us to identify the most appropriate setup and to discuss the impact of different exogenous input parameters on the prediction performance. In doing so, we treat the neural networks as a blackbox (cf. also Zhang et al. [26] and Tzeng and Ma [27]). The proposed approach offers a way to experimentally, practically, and results-driven optimize these networks.

To that end, Section 2 presents a literature review of existing studies on the prediction of gas and CO<sub>2</sub> prices using neural networks. This serves as the basis for all further programming work, helps to identify and pre-evaluate relevant input and network parameters ex-ante and to elaborate a best practice. In Section 3, we describe the suggested sequential approach in general and its application for natural gas and carbon price prediction in particular. This also implies the collection, analysis, and pre-processing of the required data. Section 4 presents the results, evaluates the performance of the different models, and discusses the implications of the outcomes. Finally, we draw a conclusion and derive recommendations for future modeling work from the results in Section 5.

## 2. Literature Review

Artificial neural networks (ANN) are a promising method for short-term commodity price predictions. As Abiodun et al. [28] point out in their comprehensive survey on artificial neural network applications, ANNs exhibit significant advantages compared to other prediction methods (e.g., statistical models). Inter alia, they feature a high degree of flexibility and fault tolerance. Furthermore, they are capable of handling incomplete data, they do not need a testing hypothesis, and they are able to solve non-linear problems.

Abrishami and Varahrami [29] add that ANNs' parsimony and parametrization are superior to traditional approaches. Compared to linear ARIMA models, for example, natural gas price predictions using neural networks showed an improvement in performance of 33% in Salehnia et al. [30] as well as in Siddiqui [31]. More studies showing the outperformance of ARIMA by ANNs can be found in the review of Zhang et al. [26]. Hence, the following literature review identifies and discusses recent approaches to predict gas (Section 2.1) and CO<sub>2</sub> prices (Section 2.2) using neural networks.

### 2.1. Prediction of Gas Prices Using Artificial Neural Networks

With respect to the prediction of gas prices using neural networks, 14 articles serve as the basis for this literature review (Table 1). Most of the published work focuses on the US gas market. In order to design the hereby presented model, we distinguish between three fundamental aspects:

1. Input data selection and preparation;
2. Model architecture and layout;
3. Model parametrization and hyper-parameters, e.g., activation functions, learning rates, etc.

Regarding the input data, the majority of authors choose an autoregressive approach, i.e., relying solely on historic gas price data. Furthermore, historic gas prices are the principal ingredient in any published model. In order to identify the most significant lags (i.e., the amount and specific position of historic time data), some authors like Siddiqui [31] and Maitra [32] apply autocorrelation analyses. However, the sensitivity analyses by both Hosseinipoor et al. [33] and Busse et al. [34] show that the model errors can be further reduced by including additional explanatory data. Busse et al. [34], for example, indicate that weather forecasts for 2–4 days in the future are the most important additional variables. The authors chose this time horizon as they assume weather forecasts to become reliable from about four days beforehand. Unsurprisingly, the level of gas production and local demands show correlations to the gas price as well. Further data used in the literature comprise USD/EUR exchange rates, gas flows through certain international hubs, and domestic holidays. Maitra [32] uses up to 46 additional input parameters in his model. However, some studies conclude that the higher model complexity often leads to only minor improvements, thus favoring the autoregressive approach for the sake of simplicity.

The high focus on autoregression also leads to one principal finding which many authors (e.g., Wang et al. [35]) are also (self-)aware of: the resulting gas price prediction models work well during normal times without significant disruptions. However, black swan events like economic or political crises and wars and similar unforeseen events with global repercussions obviously lead to drastic mispredictions.

**Table 1.** Overview of the reviewed neural networks for gas price predictions including their inputs and network configuration.

Ref.	Year	Inputs	Market	Architecture	Layer/Nodes	Activation Function
[35]	2020	Gas prices	USA	Hybrid LSTM (RNN)	n/a	Tanh, Sigmoid
[31]	2019	Gas prices	USA	ARNN (RNN)	1/13	n/a
[36]	2019	Gas prices	IRN	Hybrid (unspecified)	1/8	Sigmoid
[37]	2019	Gas prices, oil prices, CDD, HDD, drilling activities, gas offer and demand, gas imports, gas storage levels	USA	NAR (RNN)	1/10	n/a
[33]	2016	Gas prices, CDD, HDD, temperature, oil prices, gas demand, storage capacity and operation	USA	NAR (RNN) NARX (RNN)	1/1 1/6	n/a

Table 1. Cont.

Ref.	Year	Inputs	Market	Architecture	Layer/Nodes	Activation Function
[38]	2015	Gas prices	USA	Hybrid MLP (FFNN)	1/3	n/a
[39]	2015	Gas prices	USA	Hybrid MLP (FFNN)	1/n/a	n/a
[32]	2015	Gas prices, 46 others	USA	GRU (RNN)	n/a	Sigmoid, ReLU
[30]	2013	Gas prices	USA	MLP (FFNN)	n/a	n/a
[40]	2012	Gas prices, economic indicators	IRN	Hybrid (FFNN)	n/a	n/a
		Gas prices, EUR/USD exchange rate, gas imports, oil prices, gas storage level, weekend and holiday indicator				
[34]	2012		GER	NARX (RNN)	1/7	Tanh
[41]	2012	Gas prices	EU, USA	MoG NN (FFNN)	n/a	n/a
[29]	2011	Gas prices	USA	Hybrid (FFNN)	2/n/a	n/a
				Hybrid MLP (FFNN)	1/6	Tanh
[42]	2010	Gas prices, components of Wavelet transformation	UK	RBFNN (FFNN)	1/40	RBF

n/a = not available; LSTM = Long short-term memory; RNN = Recurrent neural network; FFNN = Feed-forward neural network; ARNN = Autoregressive neural network; GRU = Gated recurrent unit; ReLU = Rectified linear unit; MLP = Multilayer perceptron; RBFNN = Radial basis function neural network; NARX = Nonlinear autoregressive with exogenous variables; NAR = Nonlinear autoregressive; CDD/HDD = Cooling degree days/heating degree days; Tanh = hyperbolic tangent.

Regarding the types and layouts of the networks, the majority of studies favor feed-forward neural networks (FFNN) with one hidden layer. The number of nodes ranges from 3 to 40 (if stated in the article). Some authors like Thakur et al. [39], Jin and Kim [38], Salehnia et al. [30], and Hosseinipoor [33] examine different neuron configurations to check the influences on forecasting accuracy. In such cases, the value listed in Table 1 represents the respective optimum. The simple network layouts are unsurprising given the autoregressive character of most models: as Zhang et al. [26] state, one-layered FFNN typically deliver satisfying accuracies when describing any kind of nonlinear functions. Yet, some authors utilize multilayer perceptrons with two hidden layers (e.g., Abrishami and Varahrami [29]) or recurrent neural networks (RNN) (e.g., Wang et al. [35] using a hybrid model containing a RNN and Busse et al. [34] using a NARX). In contrast to FFNN, the latter contain connections between individual neurons in a backward direction and typically excel in time series predictions. Wang et al. [35] use an LSTM, a special RNN, which tackles the vanishing gradient problem of normal RNN by introducing up to three gates controlling the data flow and a mathematical equivalent of forgetting.

Some studies use (or characterize their models as) hybrid models between neural networks and other approaches. Abrishami and Varahrami [29] do so by coupling the FFNN with a rule-based expert system (RES), containing rules derived from regression for describing correlations between irregular events and price fluctuations. Other authors use nonlinear autoregressive models (NAR, e.g., Hosseinipoor [33] and Busse et al. [34] with exogenous variables (NARX)) or Fuzzy Linear Regression (Azadeh et al. [40]). The observed and reported autoregressive character of the task typically motivates this kind of approach.

Most studies do not explicitly mention many of the values for underlying hyper-parameters, activation functions, or training algorithms. The ones giving insight rely on sigmoid or hyperbolic tangent activation functions. Regarding the optimization, Levenberg–Marquardt is used primarily, followed by Bayesian optimization. Siddiqui [31] justifies the choice of the Bayesian method by the fact that this method has a higher capability to process noisy data whereas Thakur et al. [39] and Su et al. [37] prefer Levenberg–Marquardt because of its fast convergence behavior.

## 2.2. Prediction of CO<sub>2</sub> Prices Using Artificial Neural Networks

As the trading of emission certificates emerged just a few decades ago, the number of existing studies on CO<sub>2</sub> price forecasting is small. Table 2 gives an overview of the reviewed literature, again classifying the approaches by their input parameters, network architectures, and training strategies.

Regarding the choice of input data, the results of Lu et al. [43] show how strongly the forecasting quality of a neural network depends on the data set itself, testing it on eight different Chinese markets. In addition to historical CO<sub>2</sub> prices, other authors like Han et al. [44] and Zhang et al. [45] mainly use oil, gas, and coal prices as well as environmental and economic indicators for price forecasting. Additionally, the results of Yahsi et al. [46] imply that the clean energy index, the German stock index, and coal prices have the greatest influences on CO<sub>2</sub> prices whereas oil, gas, and electricity prices have no significant influence. The results of Han et al. [44] also suggest that CO<sub>2</sub> prices are more sensitive to coal prices, temperature data, and the air quality index than to other input parameters. Furthermore, the authors' review shows that there is a general disagreement in the literature about whether and to what extent temperature data affect CO<sub>2</sub> prices.

Feed-forward networks or a combination of FFNN and other approaches in hybrid models dominate the network architecture used for CO<sub>2</sub> price prediction. One reason for this are the autoregressive components of CO<sub>2</sub> prices. This is also reflected by the choice of the combination of FFNN and regression models in hybrid modeling approaches. For example, Han et al. [38] use a FFNN as part of a hybrid model with a regression approach and Jiang and Wu [47] use a hybrid model consisting of an FFNN and an ARIMA model, among others.

Furthermore, the complicated characteristics of the time series often deal with time series decompositions in the hybrid models. Zhang et al. [45], Zhu [48], and Lu et al. [43] use time series decomposition, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), and empirical mode decomposition (EMD), respectively, in addition to the neural network. The authors justify their choice with better processing of instationary, non-linear, and multifrequent components of the CO<sub>2</sub> price time series.

Atsalakis [49] tests a total of 50 different network structures and concludes that the best network consists of one hidden layer. The author does not specify an optimal number of hidden nodes whereas other authors such as Zhu [48] determine the number of hidden nodes depending on the number of input parameters. Considering the activation function, the sigmoid function is employed by Atsalakis [49] while the radial basis function (RBF) also serves in studies like those of Lu et al. [43] and Tsai and Kuo [50]. The latter justify their choice by saying the RBF deals well with the complex interactions between oil, coal, gas, and CO<sub>2</sub> prices. All authors specifying the training algorithm in their publications make use of backpropagation.

**Table 2.** Overview on the reviewed neural networks for CO<sub>2</sub> price predictions including their inputs and network configuration.

Ref.	Year	Inputs	Market	Architecture	Layer/Nodes	Activation Function
[43]	2020	CO <sub>2</sub> prices	8 different markets in China	Hybrid MLP (FFNN)	1/n/a	RBF
[46]	2019	CO <sub>2</sub> , oil, coal, gas, and electricity prices; DAX; clean energy index	EU-ETS	FFNN	1/n/a	n/a
[44]	2019	CO <sub>2</sub> , oil, and coal prices; leading stock index, air quality index, temperature	Shenzhen	Hybrid (FFNN)	1/p	n/a
[45]	2018	CO <sub>2</sub> prices, Stoxx600, global cal Newcastle index	EU-ETS	Hybrid GNN	1/3	n/a
[51]	2017	CO <sub>2</sub> prices	EU-ETS	FFNN	1/10	n/a



Table 2. Cont.

Ref.	Year	Inputs	Market	Architecture	Layer/Nodes	Activation Function
[49]	2016	CO <sub>2</sub> prices	EU-ETS	Hybrid FFNN	1/n/a	Sigmoid
[52]	2015	CO <sub>2</sub> prices	EU-ETS	MLP (FFNN)	1/7	Purelin
[47]	2015	CO <sub>2</sub> prices	EU-ETS	Hybrid FFNN	n/a	n/a
[50]	2014	CO <sub>2</sub> , oil, coal, and gas prices	EU-ETS	MLP (FFNN)	n/a	RBF
[53]	2013	CO <sub>2</sub> , oil, coal, and gas prices	EU-ETS	Hybrid MLP (FFNN)	1/13	RBF
[54]	2012	CO <sub>2</sub> prices	EU-ETS	Hybrid FFNN	1/2*p + 1	n/a

n/a = not available, MLP = Multilayer perceptron, FFNN = Feed-forward neural network, RBF = radial basis function, BPN = Backpropagation neural network, GNN = Grey neural network, p = number of neurons in input layer, DAX = German share index.

### 3. Methodology and Data

Starting from the insights gathered from the literature in Section 2, the following chapter describes the development and layout of a price prediction modelling and optimization methodology. This comprises a discussion of the general workflow including the suitable evaluation criteria (Section 3.1) as well as the data collection and preparation for both natural gas (Section 3.2) and carbon prices (Section 3.3).

#### 3.1. General Approach for Network Optimisation

The general approach followed in this study starts with the findings from the literature reviews. Based on best practice network configurations (i.e., regarding the network type, number of hidden layers/neurons, activation function, etc.), an own generalized model setup is implemented in Matlab R2020b. The design and quantification of the best practice network configuration is further described in Section 3.2.2 for gas and Section 3.3.2 for carbon price predictions.

Subsequently, autocorrelation analyses, market assessments, and literature reviews help to identify relevant input variables and lags to be tested. After the collection and preparation of the data (i.e., cleaning up missing values and standardization), we develop a heuristic approach depicted in Figure 1 allowing for simple and practicable yet effective model development and optimization. The next step uses the basic model setup in order to identify the ideal selection of price lags as well as additional explanatory data. For this, additional input parameters are introduced in an iterative procedure step by step. Depending on the improvement/deterioration of the model with a new input variable/lag, the respective input is kept or discarded in further iteration steps. This allows us to consider only the input parameters and lags that have enough explanatory power to improve the model's performance. After iterating all inputs/lags and determining the best configuration, the network architecture is optimized in the workflow. This comprises fine-tuning of the most promising setup(s) by varying the activation function and introducing additional layers or modified train-test splits, for instance.

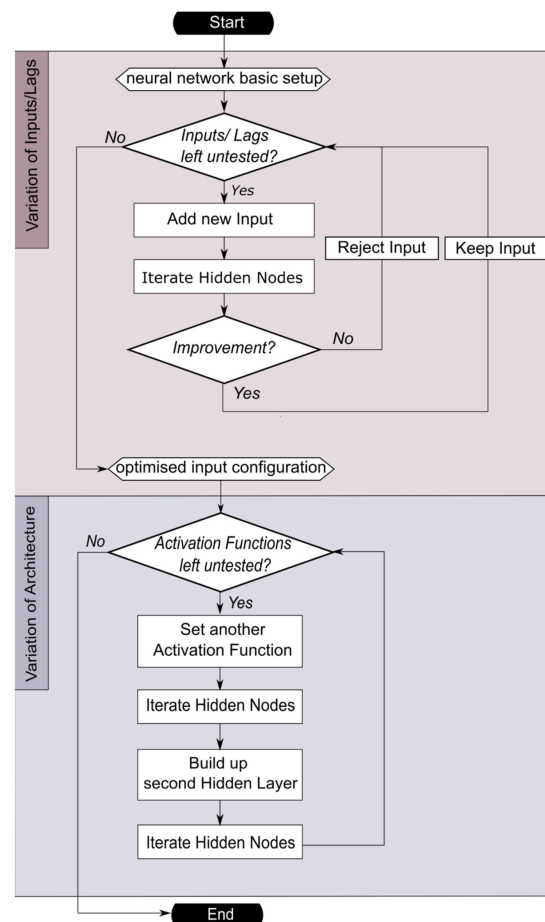
There is a validation with the forecasting standards at each level of the iteration. For this purpose, we firstly use the root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{x}_t - x_t)^2} \quad (1)$$

with  $\hat{x}_t$  being the predicted and  $x_t$  the historic value at time  $t$ . We furthermore apply the coefficient of determination ( $R^2$ ) for model evaluation and comparison

$$R^2 = 1 - \frac{\sum_{t=1}^T (\hat{x}_t - x_t)^2}{\sum_{t=1}^T (x_t - \frac{1}{T} \cdot \sum_{t=1}^T x_t)^2} \quad (2)$$

Tracking the errors of both, the test and training data sets enable the assessment of potential overfitting and underfitting. To validate the generalizability of the neural network, a five-fold validation on a rolling basis is applied. A conventional k-fold cross validation cuts data randomly and generates k subsets. Since time series are ordered, a rolling five-fold validation based on a forward chaining of the five subsets of related data is used instead. The entire time series is cut into five parts of equal length, each of them containing consecutive and connected time series. In the first run, fold 1 is the training fold and fold 2 serves for testing. In the second run, the training fold consists of fold 1 and 2 whereas fold 3 becomes the testing fold. With each new run, the test data move forward by one subset and the training data set increases by the previous test data subset.



**Figure 1.** Simplified representation of the chosen neural network optimization approach.

Additionally, the normalized *RMSE* (*NRMSE*) serves as a standardized quality parameter:

$$NRMSE = \frac{RMSE}{\bar{x}} \quad (3)$$

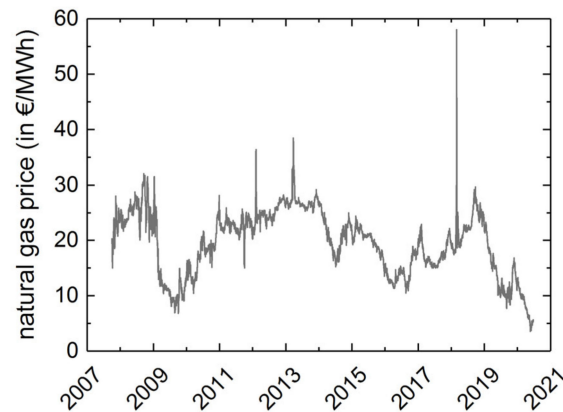
$\bar{x}$  represents the average price of the observed fold. Hence, relating the *RMSE* of a particular fold to the average price within the same fold leads to the *NRMSE* [55]. The smaller the *NRMSE*, the more powerful the model. This relative specification of the error allows an assessment of the results without knowing typical price regions.

For the evaluation, we use the mean value of run 3 to 5 only because the ratio of the training and testing data set approaches the 70:30 ratio often used in the literature [30,33].

### 3.2. Network and Input Preparation for Gas Price Predictions

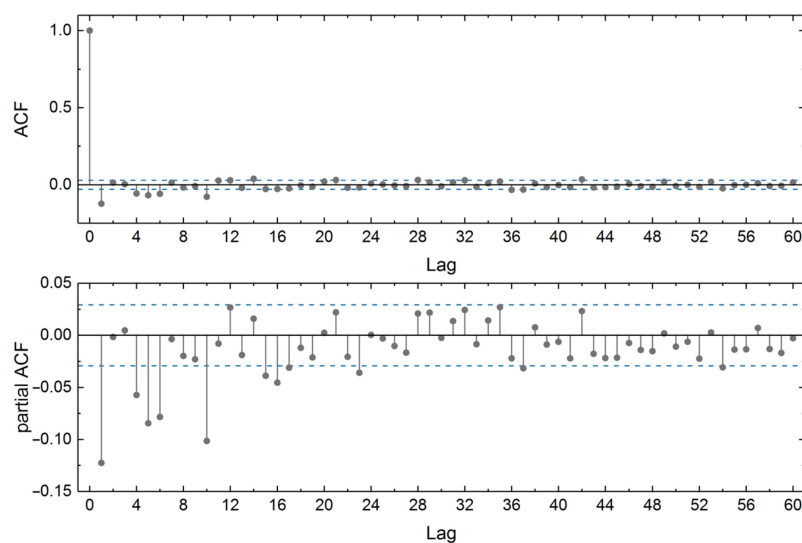
#### 3.2.1. Data Collection and Preparation

As shown in the literature review in Section 2.1, historic gas prices are of paramount importance for gas price predictions. For this cause, we use approximately 13 years of price data from the Norwegian information provider Montel. This database features a higher amount of settlement price data compared to other potential sources like the European gas spot index (EGSI) which only dates back to 2017. The specific commodity used is the PEGAS NCG (natural gas regardless of gas quality L or H) continuous day ahead contract in daily resolution. For Saturdays and Sundays, these contracts are not available, so NCG Saturday/Sunday prices serve to bridge those gaps. Additional missing data points (e.g., due to holidays or recording errors) are approximated by calculating the mean value of the preceding and following values. The deviation between the MONTEL NCG and EGSI prices accounts for 1.39% on average, justifying the use of the NCG continuous day ahead contract. In total, gas market price data deliver 4565 data points, i.e., one data point per day (cf. Figure 2).



**Figure 2.** Natural gas price time series from 1 October 2007 to 23 June 2020 used for modeling.

As shown in the literature review, the selection of lags seems to be the most influential factor for the model's capabilities. Autocorrelation and partial autocorrelation analyses on both the original price data as well as data after a decomposition into seasonal and short-term data and noise help to identify appropriate lags. Figure 3 displays the autocorrelation as well as partial autocorrelation after the elimination of seasonal and short-term trends.



**Figure 3.** Autocorrelation and partial autocorrelation function for the gas price time series from 1 October 2007 to 23 June 2020 (seasonal and short-term trend eliminated); dashed lines represent  $2\sigma$ .



The analysis of the time series shows only a weak seasonal trend (reflected in a similar course of the time series itself and the ACF of original as well as seasonally adjusted time series) while the short-term as well as autoregressive components dominate by far (reflected in a very slowly decreasing autocorrelation of the original time series, cf. also [54,56]). On the one hand, this could be surprising regarding the observable seasonality in gas prices. On the other hand, the literature review already shows that most studies rely on short-term data as well; the seasonal data might be insignificant in comparison to other factors like the influence of storage capacities and external events (regarding the global economy and its outlook, crises, etc.). Yet, the exact selection (and amount of) input data is no trivial task and subject to empiric considerations. A significant (in this case  $2\sigma$ ) correlation of historic data is visible until around 42 days (i.e., 6 weeks prior). Thus, the model will face various combinations of different lags in order to identify an optimum model.

Since several authors like Busse et al. [34] as well as Hosseinipoor [33] report improved predictions when including additional data sources, oil and coal prices, and temperature data as well as gas net flow and gas storage levels are tested in the approach. Table 3 gives an overview on the underlying time spans, data points, and the respective sources.

**Table 3.** Overview on the data (daily settlement prices) used for the gas price model parametrization.

Data	Available Time Span	Data Points	Source
Gas prices	1 October 2007–23 June 2020	4565	Montel [57]
Oil prices	1 October 2007–23 June 2020	3286	Montel [57]
Coal prices	1 October 2007–23 June 2020	3286	Montel [57]
Temperature	1 October 2007–28 June 2020	4650	German meteorological service [58]
Net flow	22 June 2014–23 June 2020	2194	Montel [57]
Gas storage level	10 January 2011–23 June 2020	3262	AGSI+ [59]

Just as for gas prices, missing values are also supplemented with mean values. Due to a lack of recorded data for coal and oil prices on weekends, the values on Saturdays and Sundays equal the ones on Fridays in this analysis. The model uses averaged German daily temperature data, meaning an average value from one city of each federal state of Germany in an hourly resolution. By means of shifting, we simulate an artificial temperature prediction instead of using real (i.e., potentially flawed) prognoses. The net flow data originate from a balance of import and export data at the German hubs with Luxemburg, the Czech Republic, Poland, Austria, France, Switzerland, the Netherlands and, finally, Sweden. Again, the filling of gaps happens by means of average values of previous as well as following values. Due to reasons of data availability, the time series of net flow as well as gas storage exhibit a shorter time span than the other parameters. Hence, in analyses comprising these parameters, all data are adjusted to the size of net flow and gas storage data, respectively.

### 3.2.2. Network Implementation and Variation

The initial model's setup is inspired by the literature review described in Section 2.1. We use an FFNN in an MLP setup with one hidden layer. MLPs are not only used in gas and carbon price predictions, as pointed out in the literature review, but are also the common choice in electricity price forecasting [3]. The reason for this is that they can especially handle the complex non-linear relationships between inputs and outputs intrinsically present in energy price data. MLPs are well-researched and understood and therefore a standard choice. Yet, they require high-quality data which is no issue for the present task due to the availability of multiple years to decades of data. Furthermore, they are prone to overfitting, which can be prevented using the suggested network optimization approach.

A tangens hyperbolicus serves as an activation function for the hidden neurons. It is zero-centered and non-linear, capturing both positive and negative values. However, one needs to carefully take into consideration that the tangens hyperbolicus can lead to the vanishing gradient problem. The impact of the activation function is assessed at a

later stage of this manuscript. Regarding the optimization, the Levenberg–Marquardt algorithm is applied. Being a second-order optimizer, this brings the added value of good convergence yet comes at the cost of higher memory utilization than other algorithms for supervised learning.

The calculation starts with random initialization for weights and biases. To ensure reproducible results, it is necessary to fix the random starting weights for each run. In order to deal with the weekly and monthly periodicity of the gas prices, any of the following model setups contain indicators for weeks and months.

As presented in Section 3.1, the general approach is a sequential one. After identifying an appropriate amount and selection of gas price lags (including the weekly markers for 3–6 weeks prior), additional data are added step-wise. This begins with the introduction of temperature predictions and the examination of its influence. This step includes another evaluation of lags (speaking of lags for temperature data). In case of a significant improvement, the further model development comprises this data set; otherwise, it is rejected. This process repeats for all other input parameters.

Table 4 shows the investigated models and data sets including their nomenclature, i.e., model G1 to G7. This notation finds application in the following sections displaying and discussing the modelling results.

**Table 4.** Input parameter combinations tested for the variation and optimization of the neural network for gas price predictions.

Model Name	Inputs <sup>1</sup>	Lags
Model G1	Historic gas prices	up to 3 weeks back additively + single inputs of 4, 5, and 6 weeks back
Model G2	Future temperature	3, 1, 2, and 4 days ahead as well as a day of price forecast itself additively
Model G3	Historic temperature	single inputs up to 5 days back
Model G4	Historic oil prices	single inputs up to 1 month back
Model G5	Historic coal prices	single inputs up to 1 month back
Model G6	Net flow	single inputs up to 1 month back
Model G7	Gas storage level	single inputs up to 1 month back

<sup>1</sup> In addition to the inputs of the currently best configuration.

After examining and testing all input parameters, the approach optimizes the network architecture. For this, three activation functions serve for iterations with up to 100 neurons and 2 hidden layers. Due to the standardization, the gas prices have positive as well as negative values which is why only activation functions covering positive and negative values provide reasonable results. Therefore, a hyperbolic tangent (tansig), a linear transfer function (purelin), and a symmetric saturating linear transfer function (satlins) find applications.

### 3.3. Network and Input Preparation for CO<sub>2</sub> Price Predictions

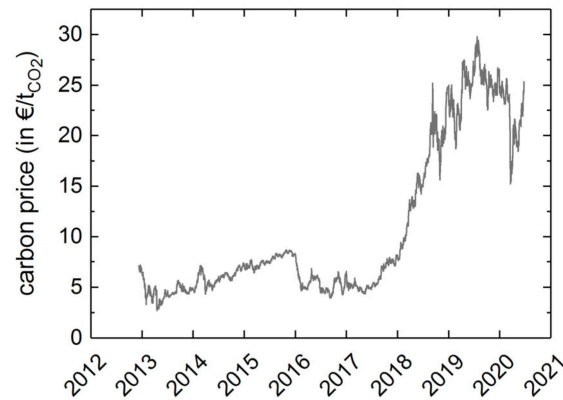
The following chapter explicitly deals with the price trend of the CO<sub>2</sub> time series in the time span of phase 3 of the ETS framework, therefore using data from end of 2012 until the middle of 2020 for price predictions.

#### 3.3.1. Data Collection and Preparation

The CO<sub>2</sub> prices used are daily settlement prices of contract type ECP D0 traded on the international climate exchange (ICE). The ICE lists the contract type ECP as a daily future contract (cf. <https://www.theice.com/products/18709519>, accessed on 17 March 2023). Figure 4 shows the development of CO<sub>2</sub> prices over the period under consideration with a price maximum of 29.77 EUR/tCO<sub>2</sub> and a lowest price of 2.70 EUR/tCO<sub>2</sub>.

The low price between the years 2012 and 2017 is due to an oversupply of EUAs. From 2018 onwards, prices rose massively due to governmental decisions to reduce the amount of EUAs in order to make saving CO<sub>2</sub> profitable again. This triggered rapid buying of EUAs by companies. The increased demand led to significantly higher prices. In March

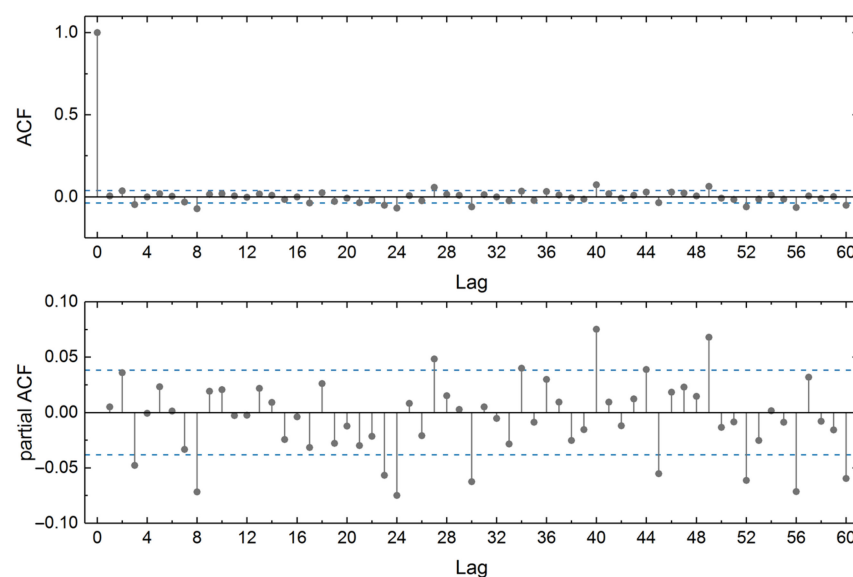
2020, EUA prices fell slightly again. One possible reason could be the shutdown due to the COVID-19 pandemic which led to reduced CO<sub>2</sub> emissions due to a shutdown of the economy and thus smaller demand of CO<sub>2</sub> certificates. Another reason could be imprudent sales of the certificates due to the crisis mood.



**Figure 4.** Carbon price time series from 10 December 2012 to 23 June 2020 used for modelling.

Since CO<sub>2</sub> certificates are only traded during the week, values for the weekend are missing from the entire data set. Therefore, prices of the respective previous Friday serve for weekends. From 26 April 2018 to 8 May 2018 there is a data gap in the time series. As there is no physical explanation for this, this is most likely due to a recording error. CO<sub>2</sub> prices of the contract type MidDec+1 serve as a proxy for the missing values. With a mean deviation 30 days before and after the gap of 0.30%, it can be assumed that the inserted data do not significantly affect the performance of the result.

Figure 5 displays the ACF and PACF of the carbon price time series in order to derive important lags. As mentioned in the beginning as well as is visible in the time series in Figure 4, CO<sub>2</sub> prices show a trend. This is also evident considering the ACF which decreases very slowly (showing an existing trend in the time series as well as autoregressive components). However, the significance of the lags decreases as they lie further back. After trend removal, the course of the ACF shown in Figure 5 changes significantly. In general, the time series shows no correlations after trend adjustment. The CO<sub>2</sub> prices do not show any seasonal dependency which is why there is no further seasonal adjustment.



**Figure 5.** Autocorrelation and partial autocorrelation function for the carbon price time series from 10 December 2012 to 23 June 2020 (short-term trend eliminated); dashed lines represent  $2\sigma$ .

The literature offers only little guidance on the choice of an optimal CO<sub>2</sub> price lag. Therefore, the choice includes the following: lag 1 to lag 21 as well as the lags 23, 24, 27, 30, 40, 45, and 49 as they reach or even exceed the significance level ( $2\sigma$ ) for both the ACF and the PACF.

The literature review shows that CO<sub>2</sub> prices are strongly autoregressive. Nevertheless, many authors like Tsai and Kuo [50] as well as Zhang et al. [45] use exogenous input parameters which is why we also investigate if there is an improvement by introducing further variables. In this case, coal, oil, electricity, and gas prices from Montel as well as temperature data from the German meteorological service serve as additional explanatory variables. Table 5 shows the time span of used CO<sub>2</sub> prices as well as further selected input parameters for the simulation.

**Table 5.** Overview on the data used for the carbon price model parametrization.

Data	Available Time Span	Data Points	Source
CO <sub>2</sub> -prices	10 December 2012–23 June 2020	1936	Montel [57]
Coal prices	10 December 2012–23 June 2020	1936	Montel [57]
Oil prices	10 December 2012–23 June 2020	1936	Montel [57]
Gas prices	10 December 2012–23 June 2020	2743	Montel [57]
Temperature	10 December 2012–23 June 2020	2747	German meteorological service [58]

The hypothesis underlying the choice of parameters is that commodity prices indirectly influence CO<sub>2</sub> prices because they determine the use of various resources for the generation of electricity. Low coal prices, for example, would lead to an increase in (CO<sub>2</sub>-intensive) coal burning which in turn influences demand for CO<sub>2</sub> allowances and thus, prices. The same goes for changes in prices for gas which produces less CO<sub>2</sub> and therefore also affects the demand of EUAs [60].

To detect any correlations between CO<sub>2</sub> prices and the individual input parameters, any prices of weekends are adjusted in the same way as applied for CO<sub>2</sub> prices. This means that for each time series, the value from Friday serves for the following weekend. As with the gas market data processing, the application of a z-transformation ensures the standardized input. Missing values are handled in the same way as described in Section 3.2.1. Also, in cases where the time series of a parameter contains less data points than the previous ones, all the time series are adjusted to the same size.

### 3.3.2. Network Implementation and Variation

Starting with the same basic model setup described in Section 3.2.2, also including indicators for weekly and monthly periodicity, the stepwise introduction of further exogenous variables follows the same heuristic approach as for the gas market (Table 6). These include historic CO<sub>2</sub>, coal, oil, gas, and electricity prices as well as temperature data as a proxy for the overall energy demand. After simulation of different models for the determination of correlations between CO<sub>2</sub> prices and other variables, especially coal, oil, and gas prices, a network optimization follows under the same conditions as presented in Section 3.2.2.

**Table 6.** Input parameter combinations tested for the variation and optimization of the neural network for CO<sub>2</sub> price predictions.

Model Name	Input Parameter	Lags
Model C1	Historic CO <sub>2</sub> prices	up to 3 weeks back additively + single inputs of t-23, t-24, t-27, t-30, t-40, t-45, and t-49
Model C2	Historic coal prices	single inputs up to 1 month back
Model C3	Historic oil prices	single inputs up to 1 month back
Model C4	Historic gas prices	single inputs up to 1 month back
Model C5	Temperature	single inputs up to 1 month back + single inputs of up to 5 days ahead
Model C6	Historic electricity prices	single inputs up to 1 month back

## 4. Results and Discussion

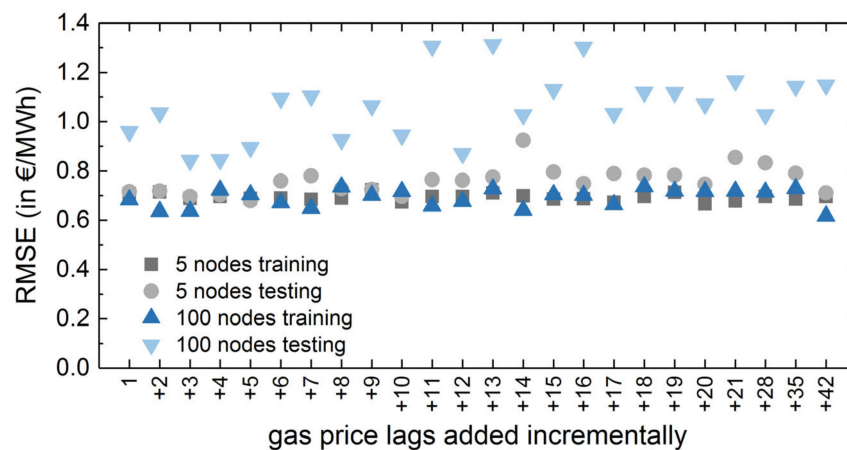
### 4.1. Prediction of Natural Gas Prices

Given that both the literature review and the time series analysis show that gas prices are autoregressive, this section presents the sole use of gas prices and the effects on the neural network independently (Section 4.1.1) before subsequently showing the influences of exogenous variables (Section 4.1.2). Finally, we discuss determinants on the results and performance of the model in order to better understand potential sources of errors (Section 4.1.3) and compare the outcomes with other studies (Section 4.1.4).

#### 4.1.1. Autoregressive Analysis

As part of the stepwise approach, the focus firstly lies on a single input parameter at the beginning of the modelling to check the performance of a purely autoregressive model using the lags provided by the time series analysis presented. Just as in studies of Siddiqui [31] (stating a MSE of 0.026 (USD/mBtu)<sup>2</sup> corresponding to a RMSE of 0.49 EUR/MWh), we obtain results with low RMSEs using only historical gas prices as inputs.

In general, the results show that an increase in the number of gas price lags and/or model complexity (i.e., the number of hidden neurons) does not necessarily lead to better results. More precisely, looking at the performance across all lags and hidden nodes, the results imply that an increase in the number of lags does not influence the forecasting performance significantly. In fact, it is quite the contrary: an increase in the number of hidden neurons not only does not lead to better results but even worsens the prediction beyond a certain point (approximately at lag 20). Figure 6 displays the results of model G1, i.e., using historic gas prices as well as indicators for one week and month, only. It presents the training and testing RMSE as a function of the gas price lags considered in the model for a network setup of 5 hidden nodes and of 100 hidden nodes (which presents the lower and upper boundary of the range of hidden nodes under consideration).



**Figure 6.** Training and testing RMSE of model G1 (historic gas prices only) for 5 and 100 nodes in the hidden layer.

The figure shows that an increase in lags (and hence in the number of inputs) does not show a significant impact on the results of the model but increases the model's complexity. However, an increase in hidden neurons tends to decrease the generalization capabilities. The overfitting, occurring in the case of 100 hidden nodes, is also evident from the larger difference between training RMSE and testing RMSE.

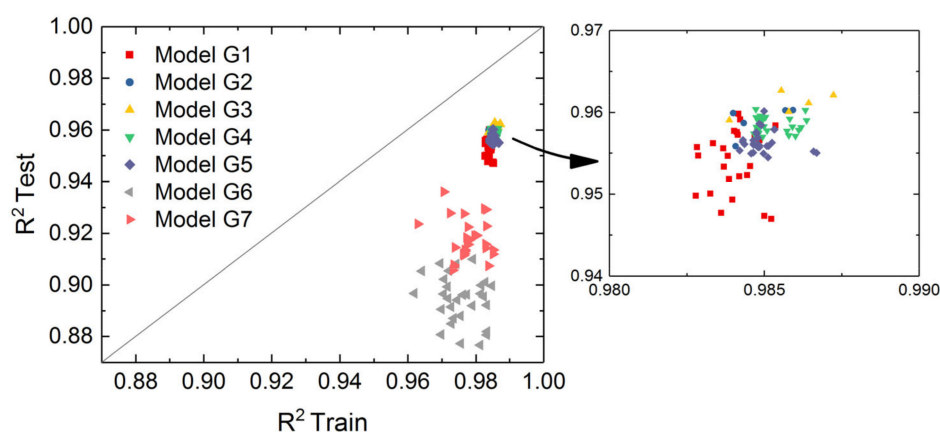
A configuration with 100 hidden nodes tends to show more volatile results, even when working with a more complex input configuration with higher number of lags. The smaller errors in the configuration with fewer hidden nodes imply that even a less complex network structure can appropriately process the amount of input information in the autoregressive case. Hence, the ability to generalize tends to decrease with an increase in hidden neurons using this input configuration. In summary, the minimum testing RMSE of 0.68 EUR/MWh



is achieved at a configuration of five hidden neurons and a gas price input ranging from  $t-1$  to  $t-5$  (i.e., a lag of five days).

#### 4.1.2. Consideration of Exogenous Parameters

The forecasts with historical gas prices already exhibit a high degree of accuracy; however, the results of authors like Jin and Kim [38] as well as Naderi et al. [36] suggest that there is still potential for improvement using further explanatory variables. Subsequently adding exogenous inputs, this section tests if this leads to a further reduction in the forecasting errors. As described in Section 3.2.2, the best configuration of model G1 serves as the basis for the further investigation and addition of input parameters. Figure 7 shows the results of this process. The figure compares the best coefficient of determination for each of the different input parameter combinations (i.e., one point for each input configuration at its best number of hidden neurons). For better visualization, the axis scaling is adapted to the range of 0.87 to 1.0. The figure gives an overview of the prediction qualities of the different model configurations under analysis. The perfect network would exhibit a  $R^2_{\text{Test}}$  and  $R^2_{\text{Train}}$  of 1.



**Figure 7.** Training and testing  $R^2$  for different network configurations of the gas price forecasting model.

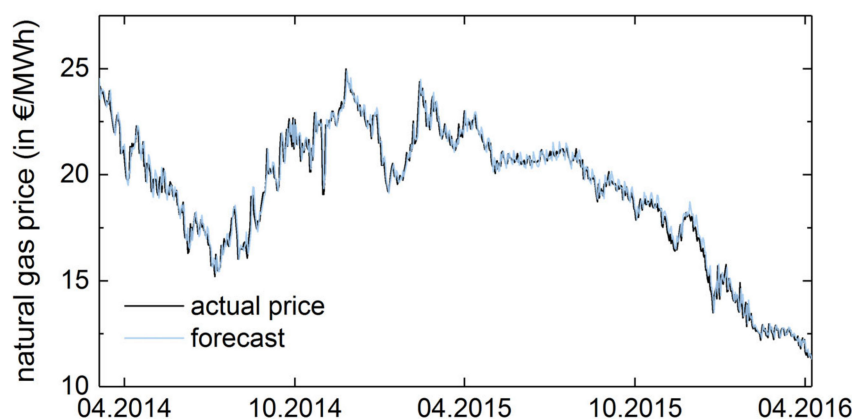
Compared to model G1, models G2 and G3 show a better performance. This implies a positive effect of the use of temperature data on the prediction accuracy of the neural network. Temperature serves as a proxy for the overall gas consumption as gas is a main source of heat generation in Germany. Also, model G4 (adding oil prices) and model G5 (adding coal prices) improve the performance compared to model G1. However, this improvement is still outperformed by model G3. The integration of net-flow data (model G6) further decreases the training errors but increases testing inaccuracies. Similar results apply for the addition of gas storage filling levels (model G7). Hence, this input improves the learning of the neural network but leads to an overfit.

Overfitting not only occurs in the case of too many input parameters but also considering the amount of hidden neurons. All setups with exogenous input parameters consistently exhibit the worst results (both regarding RMSE and  $R^2$ ) when the number of hidden neurons is in the range of 80 to 100 (e.g., a maximum RMSE of 3.36 EUR/MWh at model G6 with 100 hidden nodes). In general, the results of the RMSE analysis for the exogenous variables shows the same trend as the  $R^2$  analysis in Figure 7.

Model G3 shows the best performance of all setups under consideration. Comparing its RMSE (0.64 EUR/MWh), an improvement in 5.88 EUR/MWh compared to model G1 can be observed. As a result, the following optimization of the network architecture applies model G3; however, it does not further improve its prediction capabilities. Using up to 2 hidden layers and up to 100 hidden nodes, tansig as an activation function (best RMSE: 0.64 EUR/MWh as mentioned above) outperforms purelin and satlins (with best RMSEs of 0.72 and 0.69 EUR/MWh, respectively).



In summary, parameter variation shows the following parametrization of the neural network to have the best performance for gas price predictions. Using a feed-forward network with 1 hidden layer containing 7 hidden nodes and a hyperbolic tangent activation function as well as the Levenberg–Marquardt training algorithm. Regarding the choice of input parameters, the following variables (in addition to a weekly and monthly indicator) provide the best forecast. Historical gas prices of up to five days ago and temperature forecasts of up to three days in the future (in the order of three days ahead, one day ahead, and two days ahead) as well as historic temperature from one day before the price prediction. Figure 8 shows the forecasted natural gas prices in validation fold 3 (5 March 2014–8 April 2016) and compares them to the actual historic prices.



**Figure 8.** Visual comparison of historic natural gas prices and prediction results of the optimized network configuration in validation fold 3.

#### 4.1.3. Discussion of Determinants

Gas prices show much lower volatility (mean changes in prices in 24 h in the period under consideration of 1.93%) compared to prices of other energy goods (mean changes of electricity prices in 24 h in 2014–2019 of 41.62%, for example) with a possible volatility of up to 1000% [61]. This facilitates forecasting and to some extent explains the accurate results of the models. Nevertheless, a residual error remains. This asks for a discussion of where this error originates from and how it is influenced. To do so, we distinguish between two dimensions: one is the price range itself coming from the assumption that unusually high or low prices could correlate with higher forecasting errors. The other dimension is the time dependency of the gas economy by means of its seasonality, analyzing to what extent strongly varying demands between winter and summer seasons might influence the forecasting performance. Therefore, an investigation of the forecasting results of model G3 follows to reveal errors depending on the season (divided according to the meteorological seasons) as well as errors depending on the gas price itself.

The results show mean errors of up to 3.50 EUR/MWh in case the gas price was uncommonly high (30–60 EUR/MWh) indicating a strong price dependency in high price ranges. However, this is rather because such price levels are usually reached very erratically and are very rare. Therefore, the error results are instead due to preceding price leaps than by the price itself. Such price leaps and, in extreme cases, black swan events are difficult to predict and therefore increase the error of the neural network. As an example, fold 4 contains the gas price leap in 2018. This leads to an RMSE of 1.00 EUR/MWh in model G3 which is more than twice as high as its RMSE in fold 3 (best possible result with an RMSE of 0.43 EUR/MWh). This result does not only apply for model G3 but also for any iteration step, showing the high impact of such events on the overall error of the model.

With regard to the seasonal analysis, the mean error is invariably higher in colder seasons like winter (December–February) and spring (March–May) than in months with higher temperatures like summer (June–August) and autumn (September–November). In an extreme case (mean error of fold 5), the error in winter is 4.55 times the error in

summer (Indeed, average prices in winter months in this fold were 8.2% higher than in the rest of the year. However, this can only partially explain the significantly higher RMSE). This is in line with the description of Niggemann [62] who states that the demand for gas can fluctuate strongly in the short term due to temperature variations since gas is a much used heating source. This has an impact on spot prices especially in the winter months January to March. To deal with the fluctuations of demand, storage capacities are used. If the storage facilities empty out during the winter and/or delivery failures occur, the gas prices could rise significantly [62].

#### 4.1.4. Comparison with the Literature

A comparison with the existing literature helps to evaluate the results of our model regarding its plausibility and competitiveness despite its simple setup. The final testing RMSE achieved of the model G3 setup is 0.639 EUR/MWh (corresponding to a testing NRMSE of 0.037). Comparing this result with the literature data in Figure 9 shows that this result is well in line with the status-quo of natural gas price prediction. Notably, the RMSEs of all authors are converted from their respective unit (e.g., Mbtu) into EUR/MWh applying the respective conversion factors and exchange rates at the time of their publication. For comparability, Figure 9 shows the performance of neural network individually even if the respective authors use a hybrid model. When necessary, the unit of volume MSCF (million standard cubic feet) is converted to MWh with a calorific value for H-gas of 0.01042 MWh/m<sup>3</sup>.

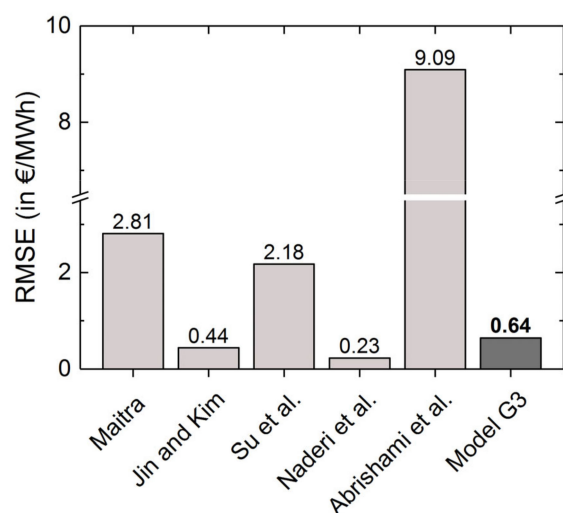


Figure 9. Comparison of the RMSE of model G3 with results from the literature [29,32,36–38].

One of the reasons for the largely varying results in the literature is the prices used. The exceptionally low RMSE of Naderi et al. [36], for example, can partially be explained by the very even prices in a range of approximately 2.5–4 USD/MSCF (corresponding to 5.85–11.69 EUR/MWh). On quite the contrary, Abrishami and Varahrami [29] use more fluctuating gas prices in a range of approximately 3.75–8 USD/MSCF (corresponding to 10.96–28.38 EUR/MWh), also including a price drop between 2008–2010. Furthermore, model G3 outperforms both Maitra [32] and Su et al. [37] which both use various exogenous variables.

## 4.2. Prediction of Carbon Prices

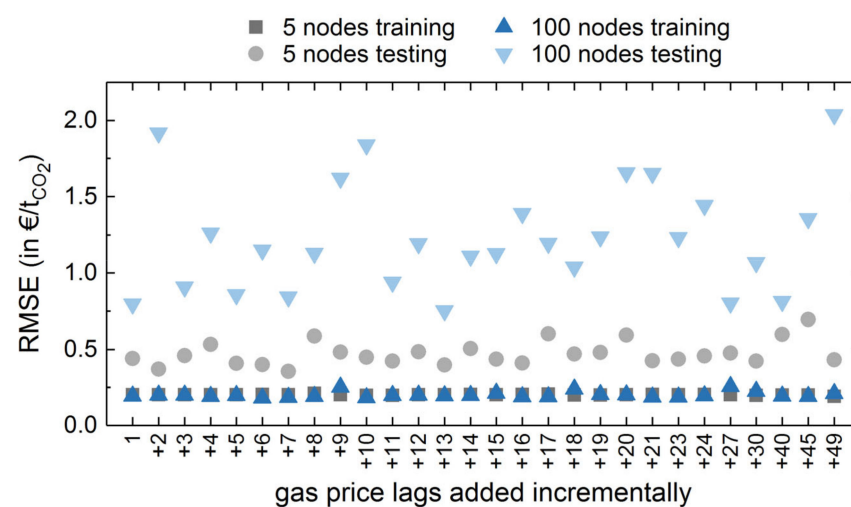
### 4.2.1. Autoregressive Analysis

Following the same approach as for the gas price prediction results, the first neural network model setup for the prediction of carbon prices follows a purely autoregressive approach. Comparing the errors of the neural networks using solely autoregressive data shows the smallest training RMSE of 0.165 EUR/t<sub>CO2</sub>, the smallest testing RMSE of

1.020 EUR/t<sub>CO2</sub>, and the highest testing RMSE of 6.220 EUR/t<sub>CO2</sub>. Considering the individual folds of the k-fold validation, it is noticeable that, without exception, fold 4 delivers poor forecasting results. Taking the configuration of lag t-7 with five hidden nodes as an example, fold 4 has (with a RMSE of 3.350 EUR/t<sub>CO2</sub>) a 22 times higher RMSE compared to fold 3 (with a RMSE of 0.150 EUR/t<sub>CO2</sub>).

As mentioned before, there is a strong increase in CO<sub>2</sub> prices in 2018. This increase is difficult to predict by the neural network as it exhibits price movements and ranges formerly unknown to the network. This time range is part of the testing data set of fold 4 and, hence, explains the weak prediction performance of this fold. To take into account the particularity of this fold, any of the following analyses are conducted once considering fold 4 and once neglecting fold 4. In general, both cases show similar trends; however, we concentrate on the simulation omitting fold 4 in further discussions to ensure comparability with the gas price prediction. Thus, the enormous, and unique in the considered time span, price increase in the time series is not taken into account.

Figure 10 shows the training as well as testing the RMSE of the CO<sub>2</sub> price prediction in dependency of lags for 5 hidden neurons and 100 hidden neurons. In general, the results indicate that the influence of the lag is not significant since the results of both model setups with few as well as with many lags do not differ strongly from each other. In contrast to that, the increase in the number of hidden neurons shows worse results, indicating that the neural network overfits in these setups. Therefore, a simple network setup consisting of 7 lags and 5 hidden neurons with an RMSE of 0.355 EUR/t<sub>CO2</sub> is considered to be the best regarding the autoregressive analysis. Overall, the neural network provides very accurate price predictions in this configuration, confirming the strong autoregressive character of CO<sub>2</sub> prices.



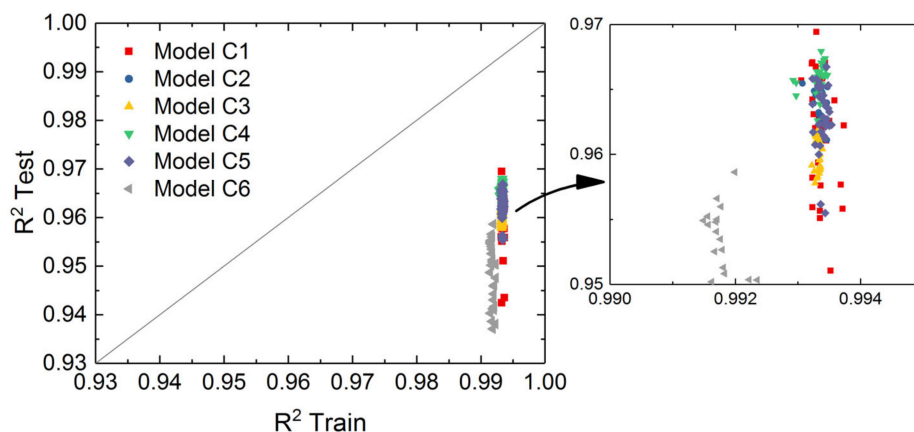
**Figure 10.** Training and testing the RMSE of model C1 (carbon prices only) for 5 and 100 nodes in the hidden layer.

#### 4.2.2. Consideration of Exogenous Parameters

Comparing the RMSE results reveals that most of the further explanatory parameters slightly worsen the results of the carbon price prediction. Contrary to the analysis of a Chinese market by Zhao et al. [63] who state a strong influence of coal prices on CO<sub>2</sub> prices, even the introduction of coal prices does not outperform model C1. However, model C6 (adding electricity prices) improves the RMSE by 1.52%. The improvement by integration of electricity prices could be explained by an indirect influence of coal prices since electricity prices are coupled with electricity production which in turn is coupled with coal consumption and thus also coal prices.

Figure 11 compares the best coefficient of determination for each of the different input parameter combinations. In fact, model C1 (i.e., using historic CO<sub>2</sub> prices only) outperforms

all complex network configurations during training considering  $R^2$ . Generally speaking, the variation of input parameters shows only little effect on the training accuracies of the network. However, it has a higher impact on testing accuracies, as Figure 11 shows.

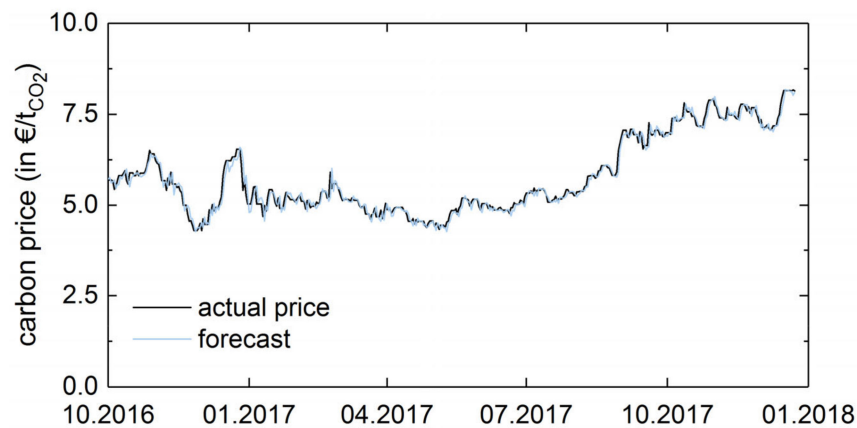


**Figure 11.** Training and testing  $R^2$  for different network configurations of the carbon price forecasting model.

The introduction of exogenous variables seems to have only little influence on the model performance. Even though there is no improvement in model C6 regarding its  $R^2$ , there is a slight reduction in the RMSE. Hence, model C6 serves as a setup for the following network variation. The analysis of network optimization demonstrates that the linear activation function improves the RMSE (0.325 EUR/ $t_{CO_2}$ ) of model C6 by 7.05% and the  $R^2$  (0.96) by 0.61%. Adding a second hidden layer, the RMSE (0.325 EUR/ $t_{CO_2}$ ) decreases again by 0.09%. Since this improvement is negligible but results in a more complex network, a second layer is disregarded.

Using purelin as an activation function, the number of hidden nodes does not show a high sensitivity since the lowest RMSE of 0.325 EUR/ $t_{CO_2}$  (lag 50) and the highest RMSE of 0.335 EUR/ $t_{CO_2}$  (lag 70) diverge by only 2.86%. This fact also supports the thesis of strong autoregressivity [64,65]. The results are consistent with the study of Zhu [48] using an ARIMA model for price prediction which performs (with an RMSE of 0.300 EUR/ $t_{CO_2}$ ) similarly to the neural network with an RMSE of 0.299 EUR/ $t_{CO_2}$  in their study.

To sum up, the parameter variation exhibits the following parametrization of the neural network to have the best performance for  $CO_2$  price prediction: using a feed-forward neural network with 1 hidden layer containing 50 hidden nodes and a linear activation function as well as the Levenberg–Marquardt training algorithm. With regard to the input parameters, the following variables provide the best forecast (in addition to a weekly and monthly indicator): historical carbon prices of up to seven days ago. Further exogenous variables lead to contrary results, showing an improvement in the RMSE by the introduction of electricity prices with a time lag of seven days whereas  $R^2$  worsens by 1.11% in this input configuration. Figure 12 shows the forecasted carbon prices in validation fold 3 (7 October 2016–29 December 2017) and compares them to the actual historic prices.



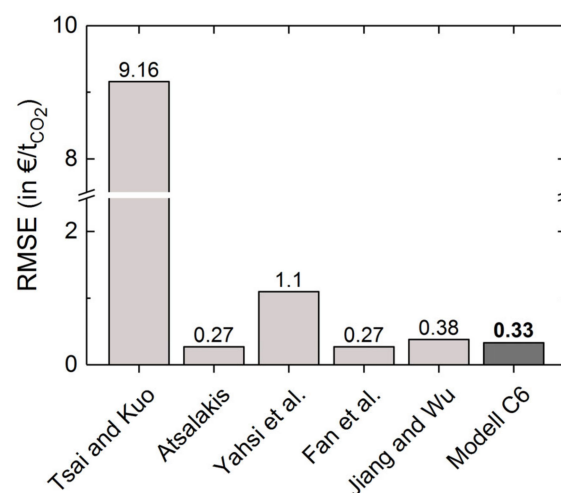
**Figure 12.** Visual comparison of historic carbon prices and prediction results of the optimized network configuration in validation fold 3.

#### 4.2.3. Discussion of Determinants

The CO<sub>2</sub> price time series itself predominantly affects the prediction errors as it contains a strong increase in prices in 2018 due to political regulations described in Section 3.3.1. Comparing both analyses (once considering fold 4 and once neglecting fold 4) for the network setup of CO<sub>2</sub> prices of up to seven days back and five hidden neurons, considering fold 4 leads to a RMSE of 1.354 EUR/tCO<sub>2</sub>, which is 3.8 times higher than when neglecting fold 4 (RMSE: 0.355 EUR/tCO<sub>2</sub>). This is in line with the analysis of eight different markets in China by Lu et al. [43] who identify the data set itself as the main determinant, showing results which differ up to 92% from another.

#### 4.2.4. Comparison with the Literature

The comparison with forecasting results in [46,47,49,50,52] shows that the obtained testing RMSE of 0.33 EUR/tCO<sub>2</sub> of the network configuration in model C6 (corresponding to a testing NRMSE of 0.023) is in line with the status-quo of CO<sub>2</sub> price prediction (Figure 13). Atsalakis [49], for example, states an RMSE of 0.27 EUR/tCO<sub>2</sub>, Zhu [48] an RMSE of 0.3 EUR/tCO<sub>2</sub>, and Fan et al. [52] an RMSE of 0.27 EUR/tCO<sub>2</sub>. The three studies have in common the fact that they achieve these very good results just by using historical CO<sub>2</sub> prices as the input without any exogenous variables. Nevertheless, Zhang et al. [45] also achieve very good results with an average RMSE of about 0.12 EUR/tCO<sub>2</sub> by introducing an economic index and a coal index. In contrast to this, the neural network of Tsai and Kuo [50] show comparatively poor RMSEs which could be a consequence of the very small data set under consideration.



**Figure 13.** Comparison of the RMSE of model C6 with results from the literature [46,47,49,50,52].

#### 4.3. Comparison of the Prediction Performance of Gas and Carbon Prices

For a final comparison of the performance of the gas and carbon price predictions, Table 7 sums up the prediction accuracies of the models. This serves as a quantitative compilation of the graphical results as presented in Figures 7 and 11. The table compares the best performing and worst performing model for each of the network configurations (i.e., G1–G7 and C1–C6).

**Table 7.** Comparison of the prediction accuracies of the different models.

Model Name	Best Model		Worst Model	
	R <sup>2</sup> Train	R <sup>2</sup> Test	R <sup>2</sup> Train	R <sup>2</sup> Test
Gas price prediction				
Model G1	0.984	0.960	0.977	0.757
Model G2	0.986	0.960	0.985	0.748
Model G3	0.986	0.965	0.987	0.908
Model G4	0.985	0.960	0.985	0.838
Model G5	0.985	0.960	0.988	0.740
Model G6	0.979	0.910	0.916	0.603
Model G7	0.971	0.936	0.971	0.403
Carbon price prediction				
Model C1	0.993	0.969	0.993	−0.845
Model C2	0.993	0.967	0.993	−2.692
Model C3	0.993	0.962	0.994	−0.002
Model C4	0.993	0.968	0.995	0.171
Model C5	0.993	0.967	0.994	−0.333
Model C6	0.992	0.959	0.990	−1.771

## 5. Conclusions

This paper presents an approach to systematically parametrize and optimize the inputs and network configuration of neural network for short-term price predictions and applied it to natural gas and carbon prices. The simulations carried out show that both natural gas and carbon prices can be forecasted with simple multi-layer perceptrons with a high level of accuracy. Recent successful applications of artificial neural networks for quantum computing (cf. [66–69]) further affirm the potential of such approaches for future robust and fast price predictions.

Results display that the selection of appropriate lags of gas and carbon prices account for the autoregressive properties of the time series which lead to a high degree of forecasting accuracy. Additionally, including temperature data can reduce errors of natural gas forecasting whereas carbon price predictions benefit from electricity prices as a further explanatory input.

The best configurations presented in this contribution are as follows. Using a feed-forward network with one hidden layer containing seven hidden nodes and a hyperbolic tangent activation function as well as the Levenberg–Marquardt training algorithm shows the best results for gas price predictions. Regarding the choice of input parameters, weekly and monthly indicators in combination with historical gas prices of up to five days ago and temperature forecasts of up to three days in the future as well as historic temperature from one day before provide the best forecast. This combination achieves a root mean square error (RMSE) of 0.64 EUR/MWh and an accuracy of XY in the testing data set.

For the prediction of carbon prices, using a feed-forward neural network with 1 hidden layer containing 50 hidden nodes and a linear activation function as well as the Levenberg–Marquardt training algorithm performs best. A weekly and monthly indicator as well as historical carbon prices of up to seven days ago are the best set of input parameters in our study. This leads to an RMSE of 0.33 EUR/t<sub>CO2</sub> and an accuracy of X during testing.



The proposed optimization approach helps to identify the most appropriate setup and to discuss the impact of different exogenous input parameters on the prediction performance. Treating the neural networks as a blackbox (cf. also Zhang et al. [26] and Tzeng and Ma [27]), the proposed approach offers a way to experimentally, practically, and results-driven optimize these networks. In light of the results of this study, we draw a two-dimensional recommendation. Firstly, due to the strong autoregressive character of carbon and natural gas price predictions, simple one-layered FFNN typically delivers satisfying accuracies. Secondly, and directly linked to the first recommendation, the autoregressive character demands the thorough preprocessing of natural gas or carbon price data to identify potentially interesting lags. This goes hand in hand with a critical analysis of the time series included in training and testing: as our results for carbon price forecasts showed, the price increase in 2018 drastically deteriorated the quality of forecasting, which, when not considered in the interpretation, could potentially lead to the selection of suboptimal parameters and network configurations. In a similar manner, the recent developments and massive price increases on especially natural gas but also carbon markets cannot be represented by the current parametrization of the networks as prices are on an historically high and unseen level. This demand further work, namely that including the impact of extreme price gradients as well as differences in overall price levels.

Yet, it has to be noted that the chosen approach for input variable selection certainly affects the results and optimal network configuration. We carefully selected the input variables by an extensive literature review as well as auto correlations and cross correlations. In doing so, we preselected potential variables as well as the order they are added to the model. The variables were then added to the model in the order of their explanatory power for the respective target. While such ex-ante selections of explanatory variables and their order of being added to the model are common (cf., as an example, [70] using a sensitivity score to evaluate the order of variables) they can also lead to suboptimal results as they do not analyze the entire set of possible input variable combinations. However, they show a path towards a systematical yet computational cheap variable selection. Still, further research should focus on the impact of the order in which the variables are added to the models.

Since there are also other influential factors on gas and carbon prices relying on unquantifiable data or events (e.g., in politics, due to pandemics, or due to crises), further increasing the predicative accuracy while maintaining a robust model is a challenging task. Parameters such as strategic behavior or political and geopolitical factors are particularly of the utmost importance for the stock prices but are difficult to be quantified or linearly separated for use in FFNN. Yet, future work on the forecasting of natural gas and carbon prices should consider additional explanatory variables such as the demand for natural gas, environmental indicators such as the clean energy index or the air quality index especially for carbon prices, and economic indicators (such as the German DAX index). In addition, the difference in using predicted values for temperatures and not the actual values (as we applied) could be subject to further analysis. Also, the integration and analysis of a holiday indicator would be of interest (yet difficult to define for the German case due to its federalist structure).

**Author Contributions:** Conceptualization, S.K. and T.P.; methodology, L.B. and S.K.; software, L.B. and S.K.; validation, L.B., S.K., J.M. and S.M.; data curation, L.B. and S.K.; writing—original draft, L.B., S.K., T.P., J.M., S.M. and J.K.; writing—review & editing, L.B., S.K., T.P., J.M., S.M. and J.K.; visualization, L.B. and S.K.; supervision, T.P.; project administration, J.K.; funding acquisition, T.P. and J.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work carried out was funded by the European Union through the Research Fund for Coal and Steel (RFCS) within the project entitled “i3upgrade: Integrated and intelligent upgrade of carbon sources through hydrogen addition for the steel industry”, Grant Agreement No. 800659. This paper reflects only the author’s view and the European Commission is not responsible for any use that may be made of the information contained therein.

**Data Availability Statement:** Restrictions apply to the availability of these data. Data was inter alia obtained from Montel with the permission of Montel.

**Acknowledgments:** The authors would like to thank Montel for providing the data as well as the helpful service.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

### Latin Letters

$NRMSE$	normalized root mean square error
$R^2$	coefficient of determination
$T$	number of time steps $t$ in time series
$\hat{x}_t$	historic value of variable $x$ at time $t$
$x_t$	predicted value of variable $x$ at time $t$

### Abbreviations

ACF	autocorrelation function
AGSI+	aggregated gas storage inventory
ANN	artificial neural network
ARIMA	auto-regressive integrated moving average
ARX	autoregressive model with exogenous variables
BPNN	backpropagation neural network
CEEMDAN	complete ensemble empirical mode decomposition with adaptive noise
DAX	German stock index
ECP D0	daily settlement prices of daily future carbon certificate contracts
EGSI	European gas spot index
EMD	empirical mode decomposition
ETS	EU emissions trading system
EUA	European union allowance
FFNN	feed-forward neural network
GARCH	generalized autoregressive conditional heteroscedasticity
ICE	international climate exchange
LSTM	long short term memory
MLP	multilayer perceptron
MSE	mean square error
NAR	nonlinear autoregressive model
NARX	nonlinear autoregressive model with exogenous variables
NCG	NetConnect Germany market zone
NRMSE	normalized root mean square error
PACF	partial autocorrelation function
PEGAS	pan-European gas cooperation
purelin	linear transfer function
RBF	radial basis function
RES	rule-based expert system
RMSE	root mean square error
RNN	recurrent neural network
satlins	symmetric saturating linear transfer function
tansig	hyperbolic tangent transfer function

## References

1. Scharf, H.; Arnold, F.; Lencz, D. Future natural gas consumption in the context of decarbonization—A meta-analysis of scenarios modeling the German energy system. *Energy Strat. Rev.* **2021**, *33*, 100591. [[CrossRef](#)]
2. Kolb, S.; Plankenbühler, T.; Hofmann, K.; Bergerson, J.; Karl, J. Life cycle greenhouse gas emissions of renewable gas technologies: A comparative review. *Renew. Sustain. Energy Rev.* **2021**, *146*, 111147. [[CrossRef](#)]
3. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecast.* **2014**, *30*, 1030–1081. [[CrossRef](#)]

4. Kolb, S.; Plankenbühler, T.; Frank, J.; Dettelbacher, J.; Ludwig, R.; Karl, J.; Dillig, M. Scenarios for the integration of renewable gases into the German natural gas market—A simulation-based optimisation approach. *Renew. Sustain. Energy Rev.* **2021**, *139*, 110696. [CrossRef]
5. Kolb, S.; Plankenbühler, T.; Pfaffenberger, F.; Vrzal, J.; Kalu, A.; Holtz-Bacha, C.; Ludwig, R.; Karl, J.; Dillig, M. Scenario-based analysis for the integration of renewable gases into the German gas market. In Proceedings of the 27th European Biomass Conference and Exhibition, EUBCE 2019, Lisbon, Portugal, 27–31 May 2019; pp. 1863–1868. [CrossRef]
6. Abada, I.; Gabriel, S.; Briat, V.; Massol, O. A Generalized Nash–Cournot Model for the Northwestern European Natural Gas Markets with a Fuel Substitution Demand Function: The GaMMES Model. *Netw. Spat. Econ.* **2013**, *13*, 1–42. [CrossRef]
7. Boots, M.G.; Rijkers, F.A.; Hobbs, B.F. Trading in the Downstream European Gas Market: A Successive Oligopoly Approach. *Energy J.* **2004**, *25*, 73–102. [CrossRef]
8. Egging, R. Benders Decomposition for multi-stage stochastic mixed complementarity problems—Applied to a global natural gas market model. *Eur. J. Oper. Res.* **2013**, *226*, 341–353. [CrossRef]
9. Rubin, O.D.; Babcock, B.A. The impact of expansion of wind power capacity and pricing methods on the efficiency of deregulated electricity markets. *Energy* **2013**, *59*, 676–688. [CrossRef]
10. Kolb, S.; Dillig, M.; Plankenbühler, T.; Karl, J. The impact of renewables on electricity prices in Germany—An update for the years 2014–2018. *Renew. Sustain. Energy Rev.* **2020**, *134*, 110307. [CrossRef]
11. Kanamura, T.; Ōhashi, K. A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation. *Energy Econ.* **2007**, *29*, 1010–1032. [CrossRef]
12. Coulon, M.; Howison, S. Stochastic behavior of the electricity bid stack: From fundamental drivers to power prices. *J. Energy Mark.* **2009**, *2*, 29–69. [CrossRef]
13. Nick, S.; Thoenes, S. What drives natural gas prices?—A structural VAR approach. *Energy Econ.* **2014**, *45*, 517–527. [CrossRef]
14. Nogales, F.; Contreras, J.; Conejo, A.; Espinola, R. Forecasting next-day electricity prices by time series models. *IEEE Trans. Power Syst.* **2002**, *17*, 342–348. [CrossRef]
15. Kristiansen, T. Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy* **2012**, *49*, 328–332. [CrossRef]
16. Zareipour, H.; Canizares, C.; Bhattacharya, K.; Thomson, J. Application of Public-Domain Market Information to Forecast Ontario’s Wholesale Electricity Prices. *IEEE Trans. Power Syst.* **2006**, *21*, 1707–1717. [CrossRef]
17. Garcia, R.; Contreras, J.; Vanakkeren, M.; Garcia, J. A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE Trans. Power Syst.* **2005**, *20*, 867–874. [CrossRef]
18. Ketterer, J.C. The impact of wind power generation on the electricity price in Germany. *Energy Econ.* **2014**, *44*, 270–280. [CrossRef]
19. Zeitlberger, A.C.M.; Brauneis, A. Modeling carbon spot and futures price returns with GARCH and Markov switching GARCH models: Evidence from the first commitment period (2008–2012). *Central Eur. J. Oper. Res.* **2016**, *24*, 149–176. [CrossRef]
20. Lago, J.; De Ridder, F.; De Schutter, B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Appl. Energy* **2018**, *221*, 386–405. [CrossRef]
21. Reddy, S.S.; Jung, C.-M.; Seog, K.J. Day-ahead electricity price forecasting using back propagation neural networks and weighted least square technique. *Front. Energy* **2016**, *10*, 105–113. [CrossRef]
22. Singhal, D.; Swarup, K. Electricity price forecasting using artificial neural networks. *Int. J. Electr. Power Energy Syst.* **2011**, *33*, 550–555. [CrossRef]
23. Pindoriya, N.M.; Singh, S.N.; Singh, S.K. An Adaptive Wavelet Neural Network-Based Energy Price Forecasting in Electricity Markets. *IEEE Trans. Power Syst.* **2008**, *23*, 1423–1432. [CrossRef]
24. Nowotarski, J.; Weron, R. Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1548–1568. [CrossRef]
25. Čeperić, E.; Žiković, S.; Čeperić, V. Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. *Energy* **2017**, *140*, 893–900. [CrossRef]
26. Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast.* **1998**, *14*, 35–62. [CrossRef]
27. Tzeng, F.-Y.; Ma, K.-L. Opening the Black Box—Data Driven Visualization of Neural Network. In Proceedings of the IEEE Visualization Conference, Minneapolis, MN, USA, 23–28 October 2005.
28. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [CrossRef] [PubMed]
29. Abrishami, H.; Varahrami, V. Different methods for gas price forecasting. *Cuad. Econ.* **2011**, *34*, 137–144. [CrossRef]
30. Salehnia, N.; Falahi, M.A.; Seifi, A.; Adeli, M.H.M. Forecasting natural gas spot prices with nonlinear modeling using Gamma test analysis. *J. Nat. Gas Sci. Eng.* **2013**, *14*, 238–249. [CrossRef]
31. Siddiqui, A.W. Predicting Natural Gas Spot Prices Using Artificial Neural Network. In Proceedings of the 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 1–3 May 2019; Volume 2019, pp. 4–9. [CrossRef]
32. Maitra, S. Natural Gas Spot Price Prediction Using Artificial Neural Network 2015. Available online: <https://towardsdatascience.com/natural-gas-spot-price-prediction-using-artificial-neural-network-56da369b2346> (accessed on 6 May 2020).
33. Hosseinipoor, S. Forecasting Natural Gas Prices in the United States Using Artificial Neural Networks. Master’s Thesis, University of Oklahoma, Norman, OK, USA, 2016.

34. Busse, S.; Helmholz, P.; Weinmann, M. Forecasting day ahead spot price movements of natural gas—An analysis of potential influence factors on basis of a NARX neural network. In *Multikonferenz Wirtschaftsinformatik*; Gito Verlag: Berlin, Germany, 2012; pp. 1395–1406.
35. Wang, J.; Lei, C.; Guo, M. Daily natural gas price forecasting by a weighted hybrid data-driven model. *J. Pet. Sci. Eng.* **2020**, *192*, 107240. [CrossRef]
36. Naderi, M.; Khamehchi, E.; Karimi, B. Novel statistical forecasting models for crude oil price, gas price, and interest rate based on meta-heuristic bat algorithm. *J. Pet. Sci. Eng.* **2019**, *172*, 13–22. [CrossRef]
37. Su, M.; Zhang, Z.; Zhu, Y.; Zha, D.; Wen, W. Data Driven Natural Gas Spot Price Prediction Models Using Machine Learning Methods. *Energies* **2019**, *12*, 1094. [CrossRef]
38. Jin, J.; Kim, J. Forecasting Natural Gas Prices Using Wavelets, Time Series, and Artificial Neural Networks. *PLoS ONE* **2015**, *10*, e0142064. [CrossRef] [PubMed]
39. Thakur, A.; Kumar, S.; Tiwari, A. Hybrid model of gas price prediction using moving average and neural network. In Proceedings of the 2015 1st International Conference on Next Generation Computing Technologies (NGCT), Dehradun, India, 4–5 September 2015; pp. 735–737. [CrossRef]
40. Azadeh, A.; Sheikhalishahi, M.; Shahmiri, S. A hybrid neuro-fuzzy simulation approach for improvement of natural gas price forecasting in industrial sectors with vague indicators. *Int. J. Adv. Manuf. Technol.* **2012**, *62*, 15–33. [CrossRef]
41. Panella, M.; Barcellona, F.; D’Ecclesia, R.L. Forecasting Energy Commodity Prices Using Neural Networks. *Adv. Decis. Sci.* **2012**, *2012*, 289810. [CrossRef]
42. Nguyen, H.T.; Nabney, I.T. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. *Energy* **2010**, *35*, 3674–3685. [CrossRef]
43. Lu, H.; Ma, X.; Huang, K.; Azimi, M. Carbon trading volume and price forecasting in China using multiple machine learning models. *J. Clean. Prod.* **2020**, *249*, 119386. [CrossRef]
44. Han, M.; Ding, L.; Zhao, X.; Kang, W. Forecasting carbon prices in the Shenzhen market, China: The role of mixed-frequency factors. *Energy* **2019**, *171*, 69–76. [CrossRef]
45. Zhang, J.; Li, D.; Hao, Y.; Tan, Z. A hybrid model using signal processing technology, econometric models and neural network for carbon spot price forecasting. *J. Clean. Prod.* **2018**, *204*, 958–964. [CrossRef]
46. Yahşi, M.; Çanakoğlu, E.; Ağralı, S. Carbon price forecasting models based on big data analytics. *Carbon Manag.* **2019**, *10*, 175–187. [CrossRef]
47. Jiang, L.; Wu, P. International carbon market price forecasting using an integration model based on SVR. In Proceedings of the 2015 International Conference on Engineering Management, Engineering Education and Information Technology, Guangzhou, China, 24–25 October 2015. [CrossRef]
48. Zhu, B. A Novel Multiscale Ensemble Carbon Price Prediction Model Integrating Empirical Mode Decomposition, Genetic Algorithm and Artificial Neural Network. *Energies* **2012**, *5*, 355–370. [CrossRef]
49. Atsalakis, G.S. Using computational intelligence to forecast carbon prices. *Appl. Soft Comput. J.* **2016**, *43*, 107–116. [CrossRef]
50. Tsai, M.T.; Kuo, Y.T. Application of Radial Basis Function Neural Network for Carbon Price Forecasting. *Appl. Mech. Mater.* **2014**, *590*, 683–687. [CrossRef]
51. Zhiyuan, L.; Zongdi, S. The Carbon Trading Price and Trading Volume Forecast in Shanghai City by BP Neural Network. *Int. J. Econ. Manag. Eng.* **2017**, *11*, 598–604.
52. Fan, X.; Li, S.; Tian, L. Chaotic characteristic identification for carbon price and an multi-layer perceptron network prediction model. *Expert Syst. Appl.* **2015**, *42*, 3945–3952. [CrossRef]
53. Tsai, M.-T.; Kuo, Y.-T. A Forecasting System of Carbon Price in the Carbon Trading Markets Using Artificial Neural Network. *Int. J. Environ. Sci. Dev.* **2013**, *4*, 163–167. [CrossRef]
54. MathWorks. Implement Box-Jenkins Model Selection and Estimation Using Econometric. MathWorks Help n.d. Available online: <https://de.mathworks.com/help/econ/box-jenkins-model-selection-using-econometric-modeler.html> (accessed on 25 June 2020).
55. Otto, S.A. How to Normalize the RMSE 2019. Available online: <https://www.marinedatascience.co/blog/2019/01/07/normalizing-the-rmse/> (accessed on 15 November 2021).
56. Kneip, A. *Zeitreihenanalyse*; University of Bonn: Bonn, Germany, 2010.
57. Montel 2020:2020. Available online: <https://www.montelnews.com/de/> (accessed on 22 May 2020).
58. Deutscher Wetterdienst, n.d. Available online: [https://www.dwd.de/DE/Home/home\\_node.html](https://www.dwd.de/DE/Home/home_node.html) (accessed on 17 March 2023).
59. Aggregated Gas Storage Inventory n.d. Available online: <https://agsi.gie.eu/#/> (accessed on 3 September 2020).
60. Wagner, M.W. CO<sub>2</sub>-Emissionszertifikate—Preismodellierung und Derivatebewertung. Ph.D. Thesis, Universität Karlsruhe, Karlsruhe, Germany, 2007. [CrossRef]
61. Frauendorfer, K.; Kiske, K. Die Spotvolatilität der Strompreise an der EEX. *EMW—Energ. Markt Wettbew.* **2010**, *1*, 52–56.
62. Niggemann, M. *Steuerung von Gaspreisrisiken*; Springer: Wiesbaden, Germany, 2013. [CrossRef]
63. Zhao, X.; Zou, Y.; Yin, J.; Fan, X. Cointegration relationship between carbon price and its factors: Evidence from structural breaks analysis. *Energy Procedia* **2017**, *142*, 2503–2510. [CrossRef]
64. AI United. Aktivierungsfunktionen, ihre Arten und Verwendungsmöglichkeiten 2019. Available online: <http://www.ai-united.de/aktivierungsfunktionen-ihre-arten-und-verwendungsmoeglichkeiten/> (accessed on 23 August 2020).

65. AI United. Die Aktivationsfunktion des Neuronennetzes 2019. Available online: <http://www.ai-united.de/die-aktivationsfunktion-des-neuronennetzes-die-sigmoidfunktion-die-lineare-die-gestufte-aktivationsfunktion-relu-der-hyperbeltangens/> (accessed on 16 September 2020).
66. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202. [[CrossRef](#)]
67. Liu, Z.-P.; Zhou, M.-G.; Liu, W.-B.; Li, C.-L.; Gu, J.; Yin, H.-L.; Chen, Z.-B. Automated machine learning for secure key rate in discrete-modulated continuous-variable quantum key distribution. *Opt. Express* **2022**, *30*, 15024–15036. [[CrossRef](#)]
68. Wang, W.; Lo, H.-K. Machine learning for optimal parameter prediction in quantum key distribution. *Phys. Rev. A* **2019**, *100*, 062334. [[CrossRef](#)]
69. Zhou, M.-G.; Liu, Z.-P.; Yin, H.-L.; Li, C.-L.; Xu, T.-K.; Chen, Z.-B. Quantum Neural Network for Quantum Neural Computing. *Research* **2023**, *6*, 0134. [[CrossRef](#)]
70. Ntokas, K.F.F.; Odry, J.; Boucher, M.-A.; Garnaud, C. Investigating ANN architectures and training to estimate snow water equivalent from snow depth. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 3017–3040. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.