Article

Cluster Forecasting of Corruption Using Nonlinear Autoregressive Models with Exogenous Variables (NARX)—An Artificial Neural Network Analysis

SeyedAli Ghahari [1],* , Cesar Queiroz [2] , Samuel Labi [1] and Sue McNeil [3,4]

1 Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA; labi@purdue.edu
2 World Bank, Washington, DC 20433, USA; queiroz.cesar@gmail.com
3 Department of Civil and Environmental Engineering, University of Delaware, Newark, DE 19716, USA;
   smcneil@udel.edu
4 School of Civil and Environmental Engineering, University of New South Wales,
   Sydney, NSW 2052, Australia
* Correspondence: sghahari@purdue.edu

Abstract: Any effort to combat corruption can benefit from an examination of past and projected worldwide trends. In this paper, we forecast the level of corruption in countries by integrating artificial neural network modeling and time series analysis. The data were obtained from 113 countries from 2007 to 2017. The study is carried out at two levels: (a) the global level, where all countries are considered as a monolithic group; and (b) the cluster level, where countries are placed into groups based on their development-related attributes. For each cluster, we use the findings from our previous study on the cluster analysis of global corruption using machine learning methods that identified the four most influential corruption factors, and we use those as independent variables. Then, using the identified influential factors, we forecast the level of corruption in each cluster using nonlinear autoregressive recurrent neural network models with exogenous inputs (NARX), an artificial neural network technique. The NARX models were developed for each cluster, with an objective function in terms of the Corruption Perceptions Index (CPI). For each model, the optimal neural network is determined by fine-tuning the hyperparameters. The analysis was repeated for all countries as a single group. The accuracy of the models is assessed by comparing the mean square errors (MSEs) of the time series models. The results suggest that the NARX artificial neural network technique yields reliable future values of CPI globally or for each cluster of countries. This can assist policymakers and organizations in assessing the expected efficacies of their current or future corruption control policies from a global perspective as well as for groups of countries.

Keywords: policy; corruption; artificial neural networks (ANNs); nonlinear autoregressive exoge-
nous models (NARX)

1. Introduction

Transparency International [1] defines corruption as "the abuse of public power for private benefit (or profit)". Fraudulent practice, according to the World Bank [2] guidelines, is "any act or omission, including a misrepresentation, that knowingly or recklessly misleads, or attempts to mislead, a party to obtain a financial or other benefit or to avoid an obligation"; a collusive practice is "an arrangement between two or more parties designed to achieve an improper purpose, including influencing improperly the actions of another party"; and a corrupt practice is defined as "the offering, giving, receiving or soliciting, directly or indirectly, of anything of value to influence improperly the actions of another party" [2].

Corruption, and fraudulent, collusive, and corrupt practices related to development result in inefficiencies, inequities, and the diversion of resources. Corruption is a multifaceted

phenomenon that ranges from a minor infraction or small act of forbidden compensation to pervasive mass looting by public officials. Hence, it has considerable detrimental effects on sustainable development [3,4]. Sustainable development is defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [5]. In other words, sustainable development is the conservation of resources and the minimization of waste and pollution [6].

The principles of sustainability include enhancing or maximizing the quality and quantity of natural resources through reduction of use, reuse, and recycling, which minimizes the damage to the physical environment [6]. A corrupt society, however, fails to take the constructive steps toward sustainable development, such as (a) avoiding adverse institutional effects; (b) maintaining or enhancing the current and future quality of life; (c) providing flexibility for changes in stakeholder requirements; (d) basing policy and business on values such as fairness, duty, knowledge-based solutions, and efficient production; and sharing responsibility for decision making, planning, and results. Corruption causes short-term economic inefficiency (specifically in the private market), and in the long-term, dynamic inefficiency and instability in economic growth and sustainability.

An accurate picture of how global corruption is evolving is needed to develop effective policies and corruption control measures, not only from a monitoring standpoint, but also from the perspective of being able to assess the long-term effectiveness of programs, policies, and initiatives targeted towards corruption mitigation. The objective of this study is to forecast corruption levels globally as well as in clusters of like countries using artificial neural network (ANN) techniques. Using the findings from our previous study on the cluster analysis of global corruption using machine learning methods that identified the four most influential corruption factors, this work uses data from 113 countries that span the time period from 2007 to 2017. This study considers two levels of analysis. The first is the global level (all countries considered together as a single group). Then, to ensure model flexibility by avoiding making the same predictions for countries that are very dissimilar in terms of development-related attributes, cluster-level analysis was carried out using techniques established in the literature [7]. The four most influential factors of corruption (measured in terms of Corruption Perceptions Index (CPI)) for each cluster identified in the previous study are the independent variables in the model in this paper. The model type used in this study is the nonlinear autoregressive recurrent neural network with exogenous inputs (NARX) technique.

In the next section, we review the literature on related studies in the area of corruption. Data collection and the methodology of the research follow. The results of this research are thoroughly discussed afterward, and the final section presents conclusions and recommendations for future work.

2. Literature Review

Many factors affect the levels of corruption in countries; some exacerbate, and others inhibit corruption. Multiple linear regression models or other similar methods are insufficient to model such complex systems, which exhibit non-linear relationships among crucial attributes and outcomes. Therefore, a method that can handle time series in complex systems is required to model and predict corruption [7]. Artificial neural networks (ANNs), machine learning algorithms, using data that have been processed in prior research using other machine learning techniques, are applied in this study due to their potential for solving problems of this nature [8–11]. ANNs also possess superior predictive accuracy compared to multi-linear regression, support vector machine (SVM), and multivariate adaptive regression splines [12,13]. Of the various well-known ANN approaches and reliable training algorithms, the nonlinear autoregressive recurrent neural network with exogenous inputs (NARX) forecasting method (a feed-backward approach) with the Bayesian regularization training algorithm has been proven to be efficacious in various applications and disciplines [14–17]. In addition, NARX is an effective instrument to forecast time series [18], and in non-linear time series projection, it can utilize its memory capability to

recollect the preceding values of the predicted time series. NARX provides more accurate results compared with other neural network techniques and time series models, such as autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) approaches [19].

The NARX neural network method has been used in various research studies, for example, forecasting heating and cooling electrical loads [20,21], network traffic flows [22], rainfall [23,24], and crop yield and price [25,26]. Peña et al. [24] found that NARX provides significantly more accurate results for rainfall predictions compared with nonlinear regression models and SVM techniques, and Paul and Sinha [25] determined that NARX outperforms ARIMA time series models in forecasting crop yield. NARX has also been applied in macroeconomic modeling. For example, recognizing the episodic and non-linear nature of the gross domestic product (GDP) of a country, researchers have espoused the use of machine learning (ML) techniques such as NARX to improve forecast accuracy of that variable. One example is Cicceri et al. [15], who showed that the great recession in Italy in 2008–2009 could have been forecast by NARX neural network methods [15]. Tang [27] assessed the feasibility of applying NARX for macroeconomic forecasting, national goal setting, and global competitiveness assessment and carried out case studies using data from countries including China, U.S., and Russia, demonstrating the capability of NARX in forecasting macroeconomic indicators. Khan et al. [16] conducted a performance evaluation of NARX in the foreign exchange market [16]. With regard to corruption forecasts, the NARX technique seems to be a promising technique. In this research study, we seek to provide some insights in this regard.

3. Data

There is rather limited data that can be used in studies of this nature. The data were from the following databases [7]: the World Bank Group (WBG) [28], the United Nations Department of Economic and Social Affairs (UNDESA) [29], the United Nations Development Program (UNDP) [30], the World Economic Forum (WEF) [31], and Transparency International (TI) [1] (see Table 1). All data are numerical.

Table 1. Data used for the study and their sources.

	Source (Database)				
	WBG	**UNDESA**	**UNDP**	**WEF**	**TI**
Variables	Gross National Income per Capita (GNI)	E-Governance Index (EGI)	Human Development Index (HDI)	Global Competitiveness Index (GCI): undue influence; public-sector performance; security; transport infrastructure; goods market efficiency; labor market efficiency; financial market development; technological readiness; market size; business sophistication	Corruption Perceptions Index (CPI)
Code	C1	C2	C3	C4–C13	C0

The Gross National Income (GNI) is the dollar value of a country's annual income, and data on GNI are from the World Bank national accounts database [32]. UNDESA publishes the E-Governance Index (EGI) data, which indicate the consistency of being able to supervise all scales and levels of government authority as well as the digital interaction of governments and citizens [29]. According to UNDP, people and their capabilities are the fundamental benchmark to evaluate the development of a country (HDI), and not its economic growth alone [30]. In this study, the Human Development Index (HDI) is taken from the UNDP database [30]. The Global Competitiveness Index (GCI) shows the competitiveness landscape of economies, offering exceptional vision into the contributors to the productivity and prosperity of countries [33]. In this study, we use the following attributes from GCI: undue influence, public sector performance, security, transport infrastructure, goods market efficiency, labor market efficiency, financial market development,

technological readiness, market size, and business sophistication. Finally, the Corruption Perceptions Index (CPI) from TI is a ranking indicator that indicates the perceived levels of public sector corruption [1]; the CPI is our dependent variable.

More discussion on the reasoning behind choosing these attributes can be found in our recent research study on the cluster analysis of corruption level in continents using principal component analysis and machine learning techniques [7]. In this paper, a principal component analysis (PCA), cluster analysis, and a random forest technique are used to determine CPI values for countries. By performing PCA, we were able to deal with the potential correlations between the thirteen attributes (C1 to C13 shown in Table 1), and we were able to condense the original potentially correlated attributes into principal components—with a minimum potential loss of data. Then, we used the top three selected principal components (PC1, PC2, and PC3) to measure the Euclidean distance between the components for each of the 113 countries to form the clusters. We verified the optimum number of clusters using the K-means machine learning (ML) technique, and we categorized the countries into four clusters. Table 2 shows the list of the countries within their corresponding clusters. A description of the data and details on the cluster analysis and ML methods related to the data presented in this study are thoroughly discussed in *Detecting and Measuring Corruption and Inefficiency in Infrastructure Projects Using Machine Learning and Data Analytics* [34].

Table 2. The cluster analysis results.

Cluster	Countries	No. of Countries in Each Continent	
1	Albania, Armenia, Azerbaijan, Bahrain, Barbados, Chile, China, Costa Rica, Cyprus, Czech Republic, Estonia, France, Georgia, Hungary, Iceland, India, Indonesia, Italy, Jordan, Kazakhstan, Korea (Rep.), Latvia, Lithuania, Mauritius, Mexico, Montenegro, Morocco, Oman, Panama, Poland, Portugal, Russia, KSA, Slovakia, Slovenia, South Africa, Spain, Thailand, Turkey, Uruguay, Vietnam	Africa Asia and Oceania Europe North America South America	3 13 19 4 2
2	Algeria, Argentina, Bangladesh, Bolivia, Bosnia-Herzegovina, Brazil, Bulgaria, Croatia, Dominican Republic, Egypt, El Salvador, Greece, Guatemala, Honduras, Iran, Lebanon, Moldova, Nicaragua, Nigeria, Pakistan, Paraguay, Peru, Philippines, Romania, Serbia, Trinidad and Tobago, Tunisia, Ukraine	Africa Asia and Oceania Europe North America South America	4 5 8 6 5
3	Australia, Austria, Belgium, Canada, Denmark, Finland, Germany, Ireland, Israel, Japan, Luxembourg, Malaysia, Netherlands, New Zealand, Norway, Singapore, Sweden, Switzerland, UK, USA	Africa Asia and Oceania Europe North America South America	0 6 11 3 0
4	Benin, Burkina Faso, Burundi, Cameroon, Côte d'Ivoire, Ethiopia, Gabon, Gambia, Ghana, Guyana, Jamaica, Kenya, Lesotho, Libya, Madagascar, Malawi, Mauritania, Mozambique, Namibia, Nepal, Rwanda, Senegal, Uganda, Zambia	Africa Asia and Oceania Europe North America South America	21 1 0 1 1

Furthermore, using a random forest (RF) algorithm, we found the marginal effects of the given variables on the outcome, and we identified and ranked the most important attributes in determining the CPI values using Gini charts [34]. RF generally offers higher precision compared with other machine learning prediction methods [35] and is a robust technique that builds on decision trees to predict models and perform analysis on the behavior of objects. RF considers each object independently and chooses the one with the maximum number of returns as the designated prediction.

In the time series analysis that is presented in this paper, we use the top four important attributes corresponding to each cluster to predict the CPI values. These attributes are shown in Table 3 in descending order of importance/influence. In a complex dynamic system such as the one that is presented in this paper, analyzing fewer variables would be misleading due to the fact that we are dealing with more than 13 variables in reality; however, the variables were narrowed down to four to provide feasible policy suggestions.

Table 3. Attributes corresponding to the world-level and cluster-level, 2007 to 2017.

Level	Top Four Influential Attributes
World	C11—technological readiness, C3—Human Development Index, C2—E-Governance Index, C4—undue influence
Cluster 1	C11—technological readiness, C1—Gross National Income, C6—security, and C4—undue influence
Cluster 2	C3—Human Development Index, C4—undue influence, C2—E-Governance Index, C5—public sector performance
Cluster 3	C5—public sector performance, C9—labor market efficiency C2—E-Governance Index, C11—technological readiness
Cluster 4	C4—undue influence, C5—public sector performance, C6—security, C1—Gross National Income

In this study, four different distance measures were used: median clustering method, Ward's method, the nearest neighbor algorithm, and the average linkage clustering technique. The results are compared using the cophenetic correlation coefficients (CCCs) for verification and comparison purposes and for the purposes of identifying the best clustering method. CCC helped gage the extent to which the dendrogram upholds the pairwise distances between the original unmodeled data points [36]. The maximum value of the CCC helped identify the best method for the hierarchical agglomerative clustering. Table 4, which presents the cophenetic correlation coefficients, suggests that the average linkage method is the best approach for clustering the data in this study. Furthermore, the optimal number of clusters using the K-means clustering method—a machine-learning based clustering technique—was identified as four clusters (Table 5).

Table 4. The cophenetic correlation coefficient values.

Methods	Median Method	Ward's Method	Nearest Neighbor Algorithm	Average Linkage Technique
Values	0.872	0.889	0.907	**0.928**

Table 5. K-means clustering method results.

Machine Learning Trials	Optimum Number of Clusters
4	2
3	3
8	**4**
3	7
2	10
1	12
1	15

4. Methodology

To meet our objective to forecast corruption levels globally as well as in clusters of like countries using artificial neural network (ANN) techniques, we used a NARX neural network and the processed data described in the preceding section. In this Methodology Section, we put this methodology in the context of ANN processes and then describe the NARX methodology as applied to our data in more detail.

4.1. Artificial Neural Network Techniques

An ANN model has an input layer, hidden layers, and an output layer, using neurons to find a pattern within a dataset and expanding the pattern to the other or future events. The model is established using a nonlinear relationship between the input layers and the output layers [37]. ANN accuracy varies with the network structure. Therefore, different training/learning algorithms and changes in the number of hidden layers, neurons, lags, iterations, hyperparameters, etc., can change the output [38]. ANN techniques can be categorized as follows: feed-forward and feed-backward. As shown in Figure 1, each category consists of different training algorithms [39].

Feed-forward NN training algorithms include single-layer perceptron, multi-layer perceptron, and radial-based function network. On the other hand, recurrent or feed-backward NN algorithms include Bayesian regularization NNs, Hopfield networks, competitive networks, art models, and Kohonen's self-organizing map. One well-known ANN approach and reliable training algorithm for nonlinear complex time series analysis is the nonlinear autoregressive with exogenous variables (NARX) NN time series forecasting method (a feed-backward approach) with the Bayesian regularization training algorithm [14–17,19,40].
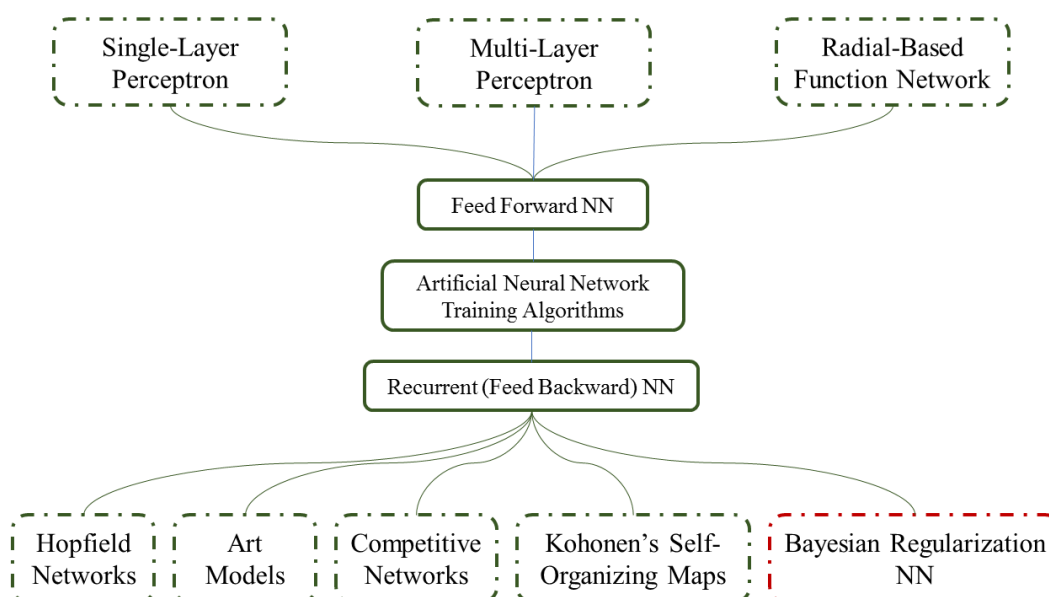

Figure 1. Training algorithm classification for artificial neural networks.

4.2. Nonlinear Autoregressive Recurrent Neural Network with Exogenous Inputs (NARX) Models

The nonlinear autoregressive recurrent neural network with exogenous inputs (NARX) technique is a time series modeling technique that relates the current value of a time series to both past values of the time series and the current and past values of the exogenous inputs time series [41]. In fact, this characteristic of NARX, which accepts dynamic variables from different time series sets, makes it superior over other feed-forward backpropagation through-time algorithm (BPTT) neural networks [42,43]. The recurrent NNs, including NARX, are cyclic in nature. Time lag connections, which transfer values between successive activations, form the cycles that include exogenous inputs and endogenous inputs [25]. NARX NN performs this procedure via autonomous learning [19]. The NARX technique

builds complex interconnections among the exogenous variables and ultimately creates a function, and this renders NARX a reliable approach for time series forecast analysis [44,45].

The architecture of the NARX neural network applied to the world-level corruption data is illustrated in Figure 2. Under this architecture, the output is forecast from the past values of CPI as well as the past and present values of the exogenous variables. The NARX technique is defined according to the following equation:

$$y(t + n) = f(y(t + 1), \ldots, y(t + n - 1), x(t - l + 1), \ldots, x(t - l + n - 1)) + \varepsilon \quad (1)$$

where n is the discrete time step, $y(t + n)$ is the predicted value of CPI, $f(.)$ is the neural network mapping function, $y(t + 1), \ldots, y(t + n - 1)$ are the past predicted values for CPI, l is the number of lags, $x(t - l + 1), \ldots, x(t - l + n - 1)$ are the past values for the exogenous variables (including l number of lags), and ε is the error term. The variable $y(t)$ (Figure 2) is defined as follows:

$$y(t) = \sum_i \alpha_i \varnothing \left(\sum_{j=1}^{k} (\beta_{ji} x_{t-j} + \gamma_{ji} y_{t-1}) \right) \quad (2)$$

where $\varnothing$ is the hidden layer activation function, β_{ji} and γ_{ji} are the hidden layer input weights at the neuron j, α_i is the hidden layer output weights, and k is the number of input nodes.

In the NARX technique, a recurrent multi-layer perceptron (RMLP), is utilized to estimate the mapping function of $f(.)$, which consists of input layers, hidden layers, and output layers. RMLP also includes neurons, activation functions, and weights. Within the hidden layer, neural network functions are operated through the interior neurons. The neurons multiply the previous layers' input vectors by the weight vectors, and they provide the scalar output. The connection weights are tuned using the Bayesian regularization algorithm. Afterward, the activation function maps each output layer to generate the neuron output to be forwarded to the next layer. In other words, to compute the output, the weighted sum of the inputs is applied to the activation function. When the generalization improvement (in the training period) ends, and the changes in the mean square error values (MSEs) become stable, the training process automatically stops. MSE is a crucial performance evaluation criterion that assists with determining the optimum initial hyperparameters for the neural network. MSE can be obtained according to Equation (3):

$$MSE = SSE/df \quad (3)$$

where SSE is the sum of square errors and df is the degree of freedom. Obviously, the lowest MSE value for the neural network models leads to the optimum model [46]. After the first model is fitted through the series-parallel architecture, more time steps can be forecast in a closed-loop parallel architecture, where each predicted output (in the previous step) is fed into the model to predict a future output.

In NARX, the number of hidden layers, lags, and neurons, as the main hyperparameters, influence the accuracy of the results. Therefore, we investigate several different numbers of hidden layers, lags, and neurons to find the optimum model. It is ideal to consider all hyperparameters at the same time; however, it increases the process time [47]. Therefore, we optimized the hyperparameters based on their importance level. The variation of the number of hidden layers, lags, and the number of neurons selected in this study were one to seven for the number of hidden layers, one to three for the number of lags, and one to 20 for the number of neurons, respectively. Two other hyperparameters are important: the number of epochs and the learning rate. One epoch is one cycle that the full training dataset performs; an epoch is made up of one or more batches, and in the current panel time series dataset that is used in this study, the countries are considered for the epochs' training dataset batches. The epochs in this study were initially set to 100 and then varied from 100 to 1000. The learning rate is the step size at each iteration while

moving toward a minimum of a loss function and was initially set to 0.1 and then varied from 0.0001 to 0.1.

The NARX structure and methodology recognizes the structure of the data as panel data that includes lagged exogenous variables. For example, the hyperparameter values selected for training the model recognize this structure. This is in fact an important strength of the methodology. We evaluate the precision of the models by comparing the mean square error values (MSEs) accordingly. Our exogenous variables (inputs), as shown in Table 3, include GNI, E-governance index (EGI), human development index (HDI), undue influence, public-sector performance, security, labor market efficiency, and technological readiness. CPI is used as the dependent variable (output). Data from 2007 to 2017 for 113 countries are assembled for all variables, with 70% of the data used for training the model, 15% for validation, and 15% of the data used to test the model. In the next section we present the results of the NARX analysis outcome.
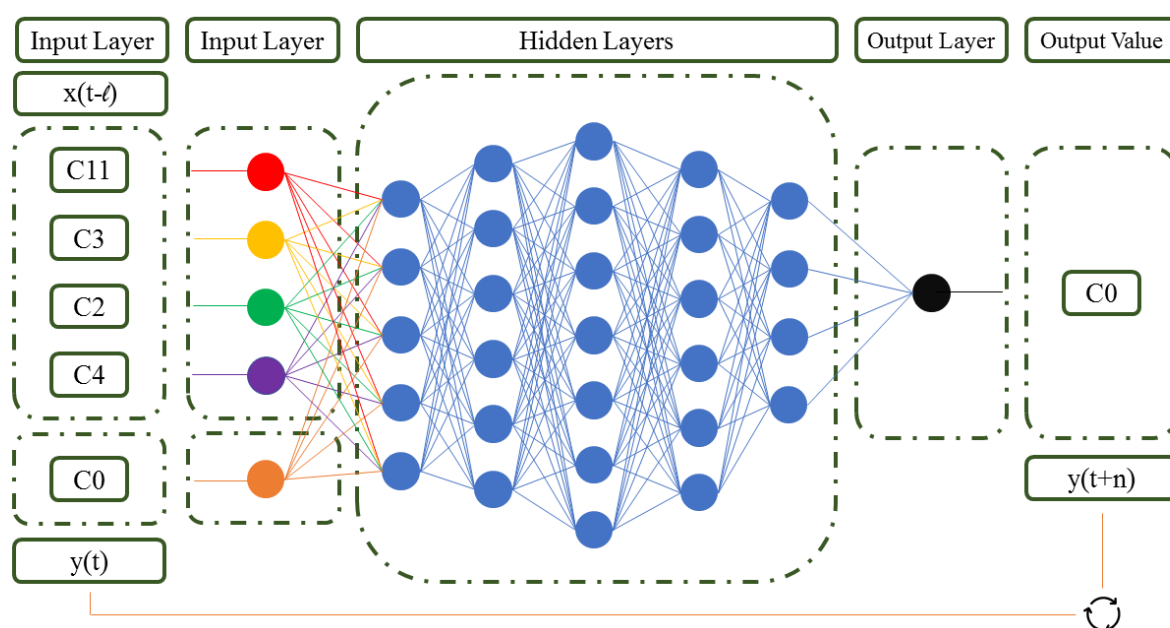

Figure 2. The architecture of the NARX neural network applied to the world-level data.

5. Results and Discussion

Hyperparameters play a critical role in the accuracy of the NARX analysis or any neural network analysis [14]. The hyperparameters in the NARX analysis, which need to be tuned to give models with higher accuracies, are the number of hidden layers, lags, neurons, and epochs as well as the learning rate. In many cases, a higher number of hidden layers causes overfitting in the model and lower prediction accuracy [48,49]. Ideally, all hyperparameters should be optimized in parallel. However, this significantly increases processing time. In this study, we optimized the hyperparameters sequentially based on the importance level. This approach is supported in the literature ([48,49]). First, we investigate different numbers of hidden layers and lags to initiate the neural network analysis, and we choose the least error associated with a hidden layer and a lag. Table 6 presents the errors associated with each hyperparameter (number of lags and number of hidden layers) for the world-level data. Consistent with best practices, the hyperparameters are optimized using the training and validation data and the error values reported for the testing data. The data shows that four hidden layers with one lag gives the least possible error among the other number of hidden layers and lags. The training MSE is calculated as 0.261, the error for the validation phase is 0.180, and the testing error is 0.243. The four hidden layer training MSEs for two lags and three lags are 26.82% and 15.71% higher than that of the four hidden layer MSE for one lag, respectively. When it comes to the testing MSE, two lags and three

lags show 93.49% and 181.61% higher MSE compared to one lag, respectively. While the training MSE for four layers and one lag is 1.91% higher that the training MSE for six layers and one lag, the validation and testing MSEs are 4.76% and 64.68% lower, respectively. This also shows that when the number of hidden layers increases, lower prediction accuracy is obtained.

Table 6. NARX errors associated with the number of hidden layers and the number of lags (world-level category) (neuron = 1, epochs = 100, and learning rate = 0.1).

Lag	No. of Hidden Layers	Training MSE	Validation MSE	Testing MSE
1	1	0.278	0.175	1.073
2		0.494	0.385	0.253
3		0.315	1.484	0.715
1	2	0.351	0.228	0.561
2		0.375	0.669	0.309
3		0.619	1.164	0.675
1	3	0.267	0.188	0.280
2		0.341	0.485	0.540
3		0.527	0.793	0.466
1	**4**	**0.261**	**0.180**	**0.243**
2		0.331	0.505	0.553
3		0.302	0.735	0.487
1	5	0.379	0.287	0.413
2		0.271	0.374	0.499
3		1.128	0.652	0.425
1	6	0.256	0.189	0.688
2		1.388	0.372	0.375
3		0.488	0.502	0.814
1	7	0.368	0.360	1.507
2		0.327	0.851	0.966
3		0.278	0.175	1.073

Next, to fine-tune another crucial NARX NN hyperparameter, we focus on the number of neurons at each hidden layer. Table 7 presents the errors associated with the number of hidden layers (H) and the number of neurons (N) for the world-level data. The data indicate that four hidden layers with five neurons gives the least possible error among the other number of hidden layers and neurons. The training MSE for H4 | N5 is calculated as 0.236, and the testing error is 0.209. H3, H5, and H6 with five neurons show 2.48%, 4.45%, and 0.84% higher training MSEs compared with H4 | N5, respectively. Likewise, testing MSEs for H4 | N5 are 15.72%, 36.08%, and 53.76% lower than those of the H3, H5, and H6 testing MSEs with five neurons. When the number of neurons exceeds 10 neurons, the errors significantly increase. The H4 | N5 training and testing MSE values are 43.13% and 49.64% lower than those of the training and testing MSEs for H4 with 10 neurons. This also confirms the importance of fine-tuning hyperparameters for the NARX NN.

Table 7. Hyperparameter fine-tuning for the world-level data—NARX errors associated with the number of hidden layers (H3–H6) and number of neurons (N1–N10, N15, and N 20) (lag = 1, epochs = 100, and learning rate = 0.1).

	Training MSE				Testing MSE			
	H3	**H4**	**H5**	**H6**	**H3**	**H4**	**H5**	**H6**
N1	0.267	0.261	0.279	0.256	0.280	0.243	0.413	0.688
N2	0.257	0.253	0.271	0.251	0.265	0.235	0.399	0.612
N3	0.250	0.244	0.265	0.246	0.264	0.228	0.379	0.545
N4	0.242	0.239	0.251	0.238	0.254	0.224	0.352	0.462
N5	0.242	**0.236**	0.247	0.238	0.248	**0.209**	0.327	0.452
N6	0.252	0.249	0.254	0.240	0.264	0.233	0.340	0.480
N7	0.271	0.267	0.286	0.278	0.297	0.249	0.371	0.509
N8	0.295	0.282	0.304	0.315	0.329	0.259	0.400	0.691
N9	0.349	0.334	0.361	0.403	0.374	0.347	0.478	0.922
N10	0.431	0.415	0.435	0.615	0.460	0.401	0.563	1.085
N15	1.034	1.008	1.175	1.456	1.199	0.972	1.140	2.266
N20	1.471	1.456	1.640	2.649	1.656	1.370	1.578	3.646

The final hyperparameter tuning is related to epochs and learning rates. The epochs range from 100 to 1000, and learning rates (LRs) from 0.0001 to 0.1 are investigated to determine the optimum MSE. Table 8 presents the results for the training and testing MSE values associated with different ranges of epochs and LRs for the NARX NN with four hidden layers, five neurons, and one lag. According to the results, the differences between the epochs and LRs are insignificant, demonstrating that the number of hidden layers, neurons, and lags were selected properly. We keep the epochs and the learning rate at 100 epochs and 0.1 learning rate, our starting point, to keep the calculation costs as low as possible.
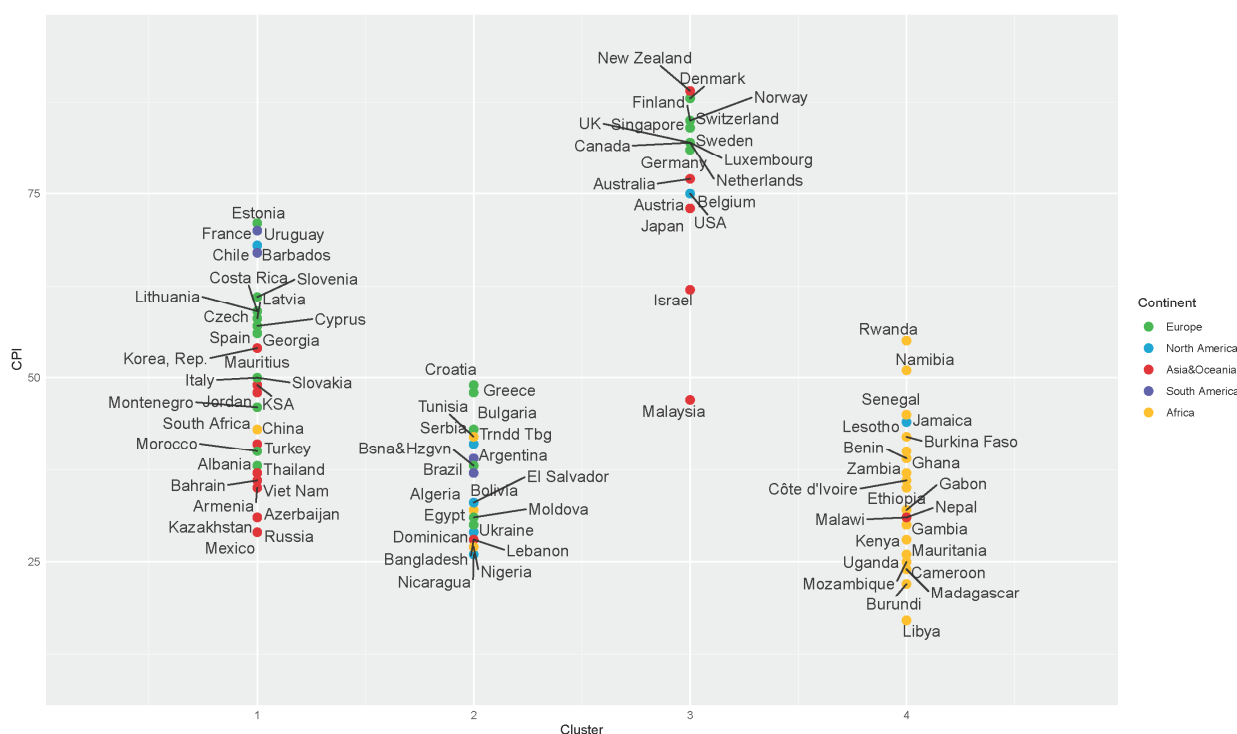
Table 8. Hyperparameter fine-tuning for the world-level data—NARX errors associated with the epochs and learning rates (LR) (lag = 1, hidden layers = 4, and neurons = 5).

	Learning Rate							
Epoch	Training MSE				Testing MSE			
	0.0001	**0.001**	**0.01**	**0.1**	**0.0001**	**0.001**	**0.01**	**0.1**
100	0.240	0.235	0.238	0.236	0.213	0.206	0.211	0.209
200	0.241	0.239	0.240	0.237	0.215	0.207	0.215	0.211
400	0.240	0.234	0.237	0.236	0.213	0.205	0.210	0.209
600	0.250	0.235	0.236	0.236	0.213	0.207	0.211	0.209
800	0.241	0.238	0.238	0.238	0.214	0.209	0.215	0.212
1000	0.241	0.239	0.240	0.240	0.215	0.210	0.216	0.215

The data at each level (world-level data and cluster level data) are distinct; therefore, different errors are likely for each cluster. This means that the number of lags, hidden layers, and neurons can vary for the NARX analysis for each cluster. The analysis of training, validation, and testing MSEs for ranges of the various hyperparameters for each cluster was conducted following the process described for the world-level data. The hyperparameter values resulting in optimal performance and MSEs for the world-level and each cluster are presented in Table 9. According to the results, Cluster 2 and Cluster 3 have a lower MSE value compared with the MSE values for Cluster 1 and Cluster 4, which could be due to fact that the CPI variance among the countries in Cluster 2 and Cluster 3 is significantly less (as shown in Figure 3). Likewise, Cluster 1 and Cluster 4 exhibit higher errors of 0.254 and 0.259 at the testing phase, respectively. Considering the error values, four hidden layers are selected for Cluster 1 and Cluster 4, whereas three hidden layers are found to be optimum for Cluster 2 and Cluster 3.
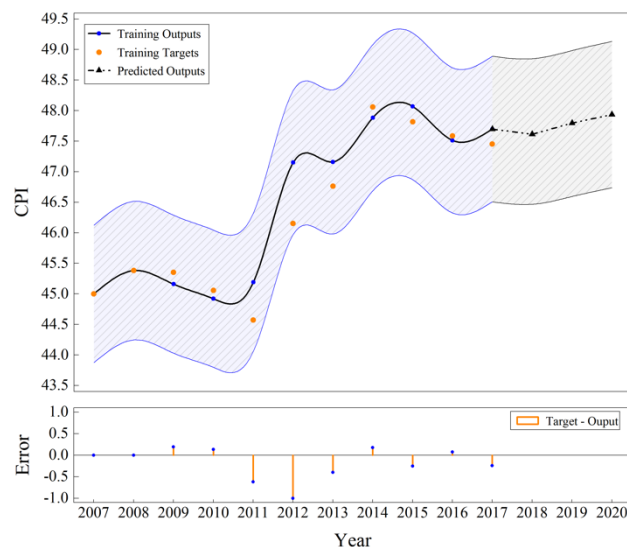
Table 9. NARX model hyperparameters and performance values for the world-level and the cluster-level analysis.

Category	Lag	No. of Hidden Layers	No. of Neurons	Training MSE	Validation MSE	Testing MSE
World-level	1	4	5	0.236	0.161	0.209
Cluster 1		4	6	0.324	0.267	0.254
Cluster 2		3	6	0.280	0.189	0.210
Cluster 3		3	5	0.208	0.140	0.150
Cluster 4		4	6	0.350	0.294	0.259

Figure 3. Position of countries in each cluster considering CPI values [7].

Finally, the results of the NARX analysis using the world-level data and the clusters are discussed in this section of this manuscript. The NARX ANN time series response for the world-level data is presented in Figure 4. Figure 4a indicates the training, target, and predicted outputs and the corresponding errors (target output—training output) with a 97.5% confidence band. This figure also presents the predicted values for the 2017–2020 period. Figure 4b presents the optimum ultimate epoch that is selected for obtaining the optimum results regarding the world-level data. The best training performance is identified as occurring at 0.10020 MSE and at the epoch 298, with no observable overfitting. This means that after the initial training of the first neural network model, it retrained the network for 298 epochs until it reached a near-zero change in MSE. According to the results, the highest difference between the training target and training outputs is calculated for 2012, with a value of −0.999, and the second-highest error is achieved for 2011, with a value of −0.619, due to the significant change in the average CPI values from 2011 to 2013. The results show that the predicted CPI values for 2018, 2019, and 2020 (shown in black triangles connected with a dashed line) are comparatively close to the real values reported by Transparency International for those specific years [50]. The CPI actual and forecast values are presented in Table 8, showing generally insignificant error between the two; the differences between the CPI forecasts and actual values in 2017, 2018, 2019, and 2020 are

calculated as 0.25, 0.04, −0.07, and −0.08, respectively. Figure 4a. also indicates that the overall CPI value of the world is increasing. Although a 0.18% decrease in CPI value is seen from 2007 to 2010, the general trend is positive, with a 6.71% increase in the CPI value from 2010 to 2020.
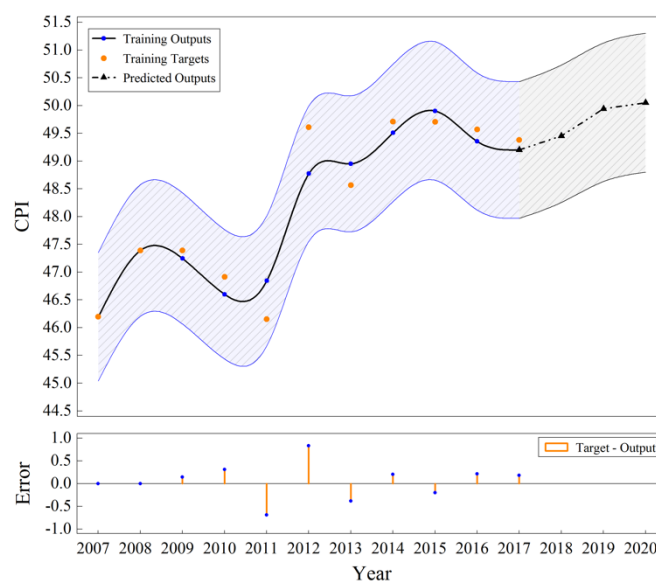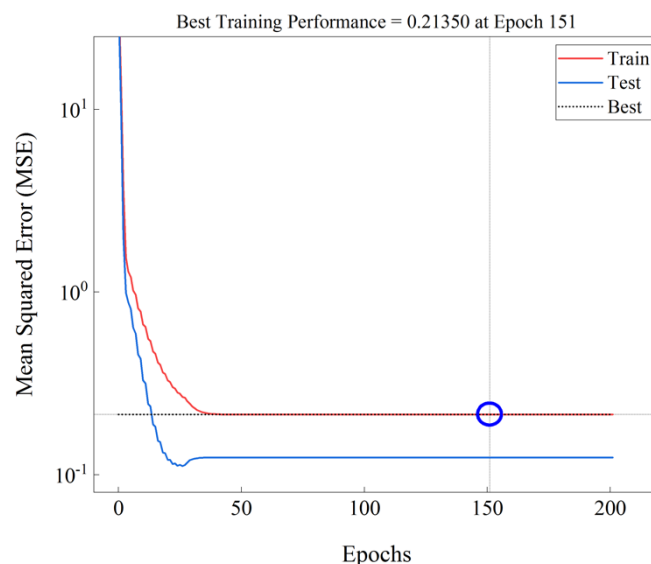

(**a**)

(**b**)

Figure 4. NARX ANN time series response for the world-level data. (**a**) Training, target, and predicted output results and errors; (**b**) epoch and learning rate.

Figure 5 presents the NARX ANN time series response for Cluster 1. Figure 5a illustrates training, target, and predicted outputs and the corresponding errors (target output—training output) with a 97.5% confidence band. In addition, this figure indicates the predicted values for the 2017–2020 period. Figure 5b illustrates the optimum epoch chosen for calculating the optimum results for Cluster 1; the best training performance is set at 0.2135 MSE and epoch 151, with no observable overfitting. This denotes the fact that after the initial training of the first neural network model, it retrained the network for 151 epochs until it reached a near-zero change in MSE. Based on the results, the maximum difference between the training target and training outputs is in 2013, with a value of

0.835, and the second-highest error is in 2011, with a value of −0.690. This could be due to the considerable change in the average CPI values from 2011 to 2013 for this cluster. The predicted CPI value results for 2017, 2018, 2019, and 2020 are close to the real values reported by Transparency International for those specific years [50]. Table 8 presents the actual and forecast CPI values. The results show a 0.01, 0.26, 0.21, and 0.39 difference between the forecast and actual CPI values in 2017, 2018, 2019, and 2020, respectively. Furthermore, Figure 5a shows that the overall CPI value for Cluster 1 is increasing. Despite a 1.67% decrease in CPI value from 2008 to 2010, the general trend is positive, with a 7.41% increase in the CPI value from 2010 to 2020.
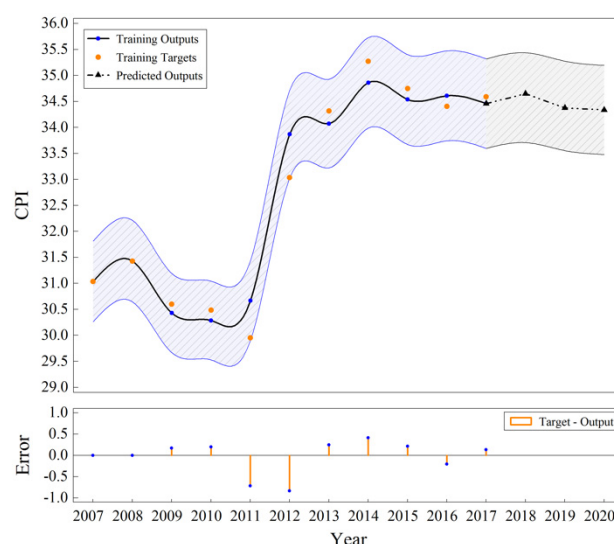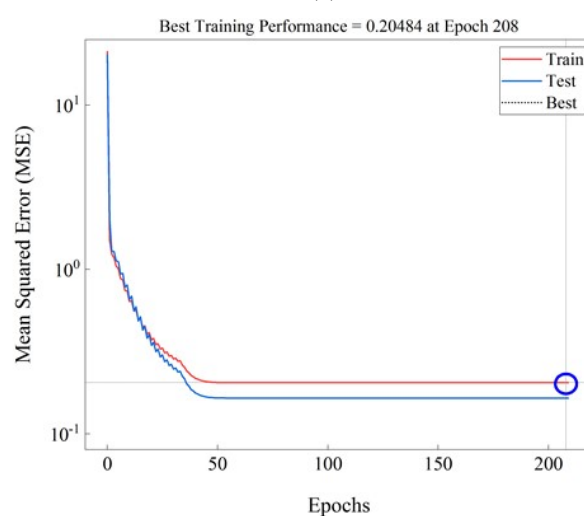

(**a**)

(**b**)

Figure 5. NARX ANN time series response for Cluster 1. (**a**) Training, target, and predicted output results and errors; (**b**) epoch and learning rate.

The NARX ANN time series analysis results for Cluster 2 are presented in Figure 6. Figure 6a shows training, target, and predicted outputs and the corresponding errors (target output—training output) with a 97.5% confidence band. Furthermore, this figure shows

the predicted values for the 2017–2020 period. Figure 6b denotes the optimum ultimate epoch selected in this analysis in order to obtain the optimum results. The best training performance is caught at epoch 248 and 0.20484 MSE, with no observable overfitting. This shows that after the initial training of the first neural network model, it retrained the network for 248 epochs until it reached a near-zero change in MSE. Results show that the highest training target and training output difference is in 2012 and 2011, with values of -0.833 and -0.714, respectively. This could be due to the significant change in the average CPI values from 2011 to 2013. The results indicate that the predicted CPI values for 2017 to 2020 are comparatively close to the real values reported by Transparency International [50]. The values presented in Table 8 indicate a minor error between the real CPI values and the predicted CPI values. Differences between the predicted and real CPI values for 2017 to 2020 are calculated as 0.23, -0.01, -0.35, and -0.21, respectively. Moreover, Figure 6a shows an overall increase in the CPI values in this cluster. Although a 3.64% decrease in the CPI value is seen from 2008 to 2010, the general trend is upward, with a 13.37% increase in the CPI value from 2010 to 2020.
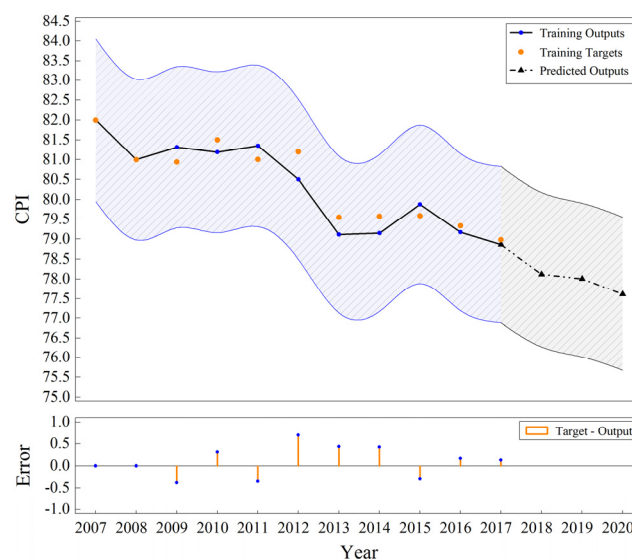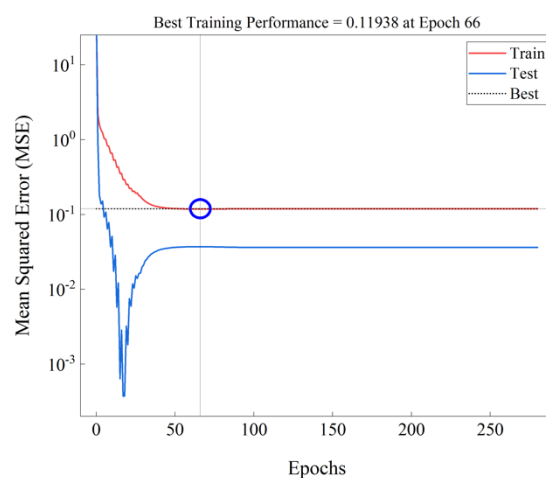

(a)

(b)

Figure 6. NARX ANN time series response for Cluster 2. (**a**) Training, target, and predicted output results and errors; (**b**) epoch and learning rate.

The NARX ANN time series response for the third cluster is illustrated in Figure 7. Figure 7a shows training, target, and predicted outputs and the corresponding errors (target output—training output) with a 97.5% confidence band. This figure also illustrates the predicted values for 2017, 2018, 2019, and 2020. Figure 7b shows the optimum epoch chosen for calculating the optimum results for Cluster 2; the best training performance is obtained at 0.11938 MSE and epoch 66, with no observable overfitting. This indicates that after the initial training of the first neural network model, it retrained the network for 66 epochs until it reached a near zero change in MSE. Based on the results, the difference between the training target and training outputs is at its maximum value of 0.696 in 2012, and the second-highest error at 0.436 is in 2013, which could be due to the significant change in the average CPI values in 2012 and 2013 for this cluster. The predicted CPI value results for 2017 to 2020 are comparatively close to the real values reported by the Transparency International for those specific years [50]. The actual and forecast CPI values are illustrated in Table 8. The results show differences of -0.63, -0.32, -0.35, and -0.16 between the forecast and actual CPI values in 2017 to 2020, respectively. Figure 7a also illustrates that the overall CPI value for Cluster 3 is decreasing. The general CPI trend in this cluster is negative, with a 5.35% decrease in the CPI value from 2007 to 2020.
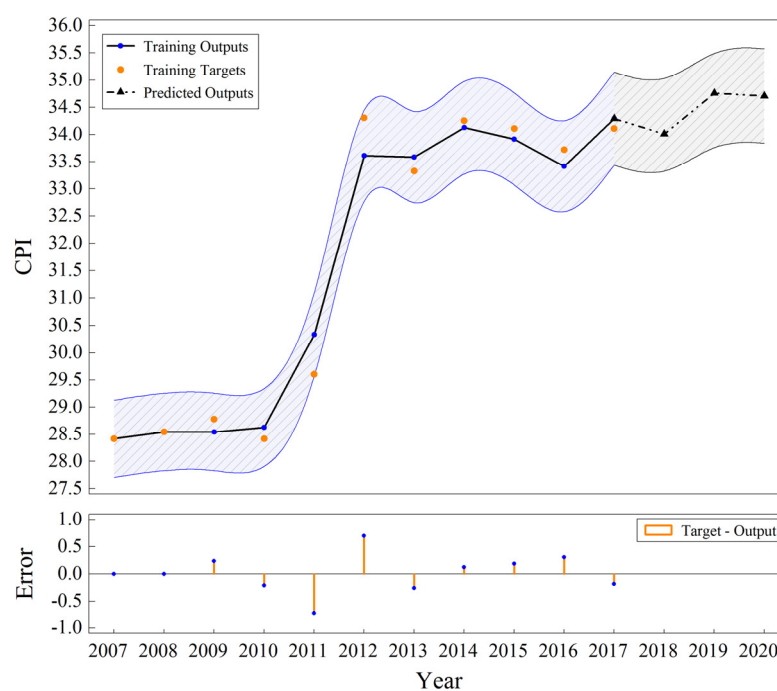

(**a**)

(**b**)

Figure 7. NARX ANN time series response for Cluster 3. (**a**) Training, target, and predicted output results and errors; (**b**) epoch and learning rate.

The NARX ANN time series analysis results for Cluster 4 are presented in Figure 8. Figure 8a shows training, target, and predicted outputs and the corresponding errors (target output—training output) with a 97.5% confidence band. This figure also illustrates the predicted values for 2017–2020. Figure 8b shows the optimum ultimate epoch selected in this analysis to obtain the optimum results; the best training performance is caught at epoch 143 and 0.25328 MSE, with no observable overfitting. This means that after the initial training of the first neural network model, it retrained the network for 143 epochs until it reached a near-zero change in MSE. Results indicate that the highest training target and training output differences are in 2011 and 2012, with values of −0.733 and 0.696, respectively; this could be due to the significant change in the average CPI values from 2010 to 2012 in this cluster. The results denote that the predicted CPI values for 2017–2020 are comparatively close to the real values reported by Transparency International [50]. The values are presented in Table 10, indicating a minor difference between the real CPI values and the predicted CPI values. The differences between the predicted and real CPI values for 2017–2020 are calculated as 0.18, 0.17, −0.17, and 0.08, respectively. Furthermore, Figure 8a indicates an overall increase in the CPI values. The general trend in this cluster is upward, with a 21.25% increase in the CPI value from 2010 to 2020.

A steep increase in CPI between 2011 and 2012 for all clusters except Cluster 3 (countries with higher CPI values) is observable. The criteria for CPI calculation changes every year, and this significant change might mean that there was a change in the method of calculation for CPI for the countries with lower CPI. However, further investigation is needed to pinpoint the changes in the method of calculation.
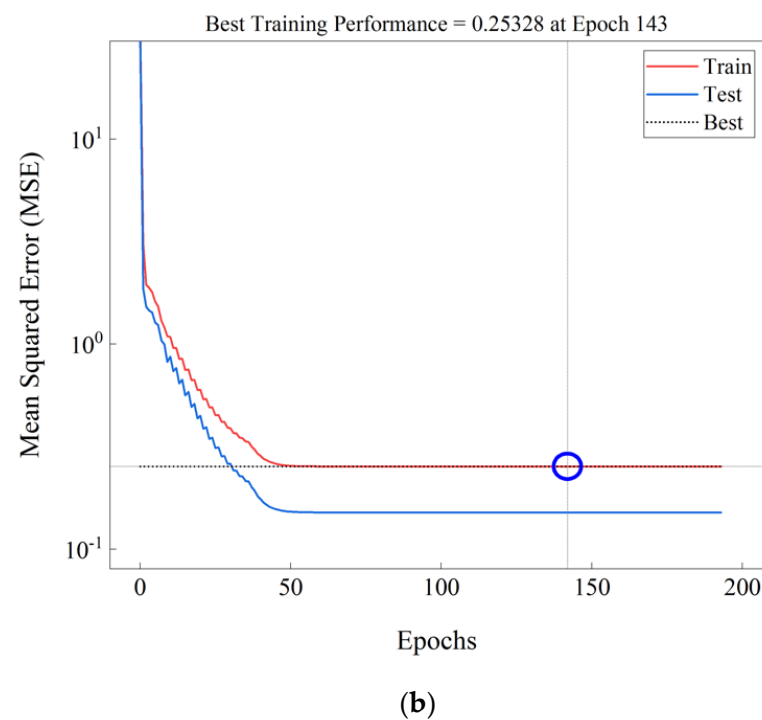

(**a**)

Figure 8. *Cont.*

Figure 8. NARX ANN time series response for Cluster 4. (**a**) Training, target, and predicted output results and errors; (**b**) epoch and learning rate.

Table 10. CPI actual and forecast values.

Year	Category	Actual CPI	CPI Forecast	Error (Forecast—Actual)
	World-level	47.70	47.45	0.25
	Cluster 1	49.39	49.38	0.01
2017	Cluster 2	34.82	34.59	0.23
	Cluster 3	78.35	78.98	−0.63
	Cluster 4	34.29	34.11	0.18
	World-level	47.65	47.61	0.04
	Cluster 1	49.71	49.45	0.26
2018	Cluster 2	34.64	34.65	−0.01
	Cluster 3	77.80	78.12	−0.32
	Cluster 4	34.17	34.00	0.17
	World-level	47.73	47.80	−0.07
	Cluster 1	50.15	49.94	0.21
2019	Cluster 2	34.11	34.37	−0.26
	Cluster 3	77.65	78.00	−0.35
	Cluster 4	34.58	34.75	−0.17
	World-level	47.86	47.94	−0.08
	Cluster 1	50.44	50.05	0.39
2020	Cluster 2	34.14	34.35	−0.21
	Cluster 3	77.45	77.61	−0.16
	Cluster 4	34.79	34.71	0.08

6. Concluding Remarks

Artificial neural networks (ANNs) are effective tools for non-linear mapping of multiple variables on one or more outputs. In this study, we use a well-known neural network method, the nonlinear autoregressive recurrent neural network with exogenous inputs (NARX), to model and forecast corruption in countries. The analysis was carried out using the data on 113 countries from 2007 to 2017. The development-related attributes that have significant influence on the levels of corruption in countries, as measured by the CPI, were identified from the literature. We split the countries into four clusters based on their development-related attributes and developed corruption forecasting models for each cluster. NARX NN training was performed on 70% of the data, 15% of the data were used for validation, and the rest of the data were used for testing the output.

Any reliable neural network model needs precise hyperparameter fine-tuning before training. The variations of the number of hidden layers, lags, and neurons were selected as 1–7, 1–3, and 1–20, respectively. Considering MSE as a baseline for the hyperparameter tuning process showed that one lag, four hidden layers, and five neurons would give an optimum NARX model for forecasting CPI values for the world-level data. For Cluster 1 and Cluster 4, the number of hidden layers was found to be four, versus three for Cluster 2 and Cluster 3. At the same time, the number of neurons for Cluster 1, Cluster 2, and Cluster 4 were chosen to be six, versus five for Cluster 3. Epochs and learning rates were found to have no significant influence on the initial hyperparameter MSE values for the NARX models. It was observed that when the number of neurons and hidden layers increased, a comparatively lower prediction accuracy was obtained, due to the models' overfitting.

As expected, the NARX NN prediction models showed different results for the world-level data analysis and the cluster-level data analysis. For the world-level data, it was found that there is a general uptrend in the value of CPI, showing a 6.71% increase in CPI from 2010 to the predicted value of CPI in 2020. Cluster 1, Cluster 2, and Cluster 4 showed the same uptrend, with a 7.41%, 13.37%, and 21.25% increase in CPI from 2010, respectively, despite having a comparatively minor downtrend in CPI from 2007 to 2010. However, Cluster 3—despite containing a majority of developed countries—showed a 5.35% decrease in CPI from 2007. For countries within the clusters introduced in this paper, the study results can be valuable to policymakers, governments, and NGOs as they continue to assess the efficacy of their current or prospective future corruption-mitigation policies, programs, and initiatives.

In this paper, the lack of adequate data on development-related attributes was one of the main limitations. In future studies, access to other data will be helpful to develop more confident conclusions. Another limitation was the reliance on only one attribute (CPI) as the indicator of corruption. Suggested directions for future research include (a) a detailed investigation of the causes of the uptrend and downtrend momentum in CPI values in each cluster; (b) adequate and explicit assessment of corruption-mitigation initiatives implemented in countries in each cluster, identifying solutions that have worked as well as those that failed, and an overall assessment of the extent to which these solutions succeeded or failed. Furthermore, future studies could investigate project-level data (instead of the country-level data in this study). In this regard, researchers could examine the effect of corruption on infrastructure delivery quality, time delay, and cost overruns and thereby measure, for example, the portion of overrun cost that could be attributed to corruption and the portion that could be attributed to inefficiency. As suggested by a reviewer, the methodology could also be used to explore the impact of corruption on sustainability indices.

Author Contributions: Conceptualization, S.G.; Data curation, S.G. and C.Q.; Formal analysis, S.G., C.Q. and S.M.; Funding acquisition, S.G.; Investigation, S.G. and S.L.; Methodology, S.G., C.Q. and S.M.; Resources, S.G., C.Q. and S.L.; Software, S.G.; Supervision, S.G., C.Q., S.L. and S.M.; Validation, S.G.; Visualization, S.G.; Writing—original draft, S.G.; Writing—review and editing, S.G., C.Q., S.L. and S.M. All authors have read and agreed to the published version of the manuscript.

References

1. TI. Corruption Perception Index. 2017. Available online: https://www.transparency.org/news/feature/corruption_perceptions_index_2017 (accessed on 4 June 2018).
2. Integrity Vice Presidency. *Fraud and Corruption Awareness Handbook: How It Works and What to Look For*; World Bank: Washington, DC, USA, 2009.
3. Tabish, S.; Jha, K.N. The Impact of Anti-Corruption Strategies on Corruption Free Performance in Public Construction Projects. *Constr. Manag. Econ.* **2012**, *30*, 21–35. [CrossRef]
4. Loosemore, M.; Lim, B. Inter-Organizational Unfairness in the Construction Industry. *Constr. Manag. Econ.* **2015**, *33*, 310–326. [CrossRef]
5. ASCE. *Policy Statement 418—The Role of the Civil Engineer in Sustainable Development*; American Society of Civil Engineers: Rihcmond, VA, USA, 2010.
6. Brundtland, G.H. *Report of the World Commission on Environment and Development: "Our Common Future"*; United Nations: New York, NY, USA; Oxford University Press: Oxford, UK, 1987.
7. Ghahari, S.A.; Queiroz, C.; Labi, S.; McNeil, S. Impact of E-Governance on National Corruption Indexes: New Evidence Using Panel Vector Auto Regression Analysis. *Preprints* **2021**. [CrossRef]
8. Woldemariam, W.; Murillo-Hoyos, J.; Labi, S. Estimating Annual Maintenance Expenditures for Infrastructure: Artificial Neural Network Approach. *J. Infrastruct. Syst.* **2016**, *22*, 04015025. [CrossRef]
9. López-Iturriaga, F.J.; Sanz, I.P. Predicting Public Corruption with Neural Networks: An Analysis of Spanish Provinces. *Soc. Indic. Res.* **2018**, *140*, 975–998. [CrossRef]
10. Khalil, A.J.; Barhoom, A.M.; Abu-Nasser, B.S.; Musleh, M.M.; Abu-Naser, S.S. Energy Efficiency Prediction Using Artificial Neural Network: 2019. Available online: https://core.ac.uk/download/pdf/237182408.pdf (accessed on 20 December 2020).
11. Lima, M.S.M.; Delen, D. Predicting and Explaining Corruption across Countries: A Machine Learning Approach. *Gov. Inf. Q.* **2020**, *37*, 101407. [CrossRef]
12. Ekonomou, L. Greek Long-Term Energy Consumption Prediction Using Artificial Neural Networks. *Energy* **2010**, *35*, 512–517. [CrossRef]
13. Yin, Z.; Jia, B.; Wu, S.; Dai, J.; Tang, D. Comprehensive Forecast of Urban Water-Energy Demand Based on a Neural Network Model. *Water* **2018**, *10*, 385. [CrossRef]
14. Al-Sbou, Y.A.; Alawasa, K.M. Nonlinear Autoregressive Recurrent Neural Network Model for Solar Radiation Prediction. *Int. J. Appl. Eng. Res.* **2017**, *12*, 4518–4527.
15. Cicceri, G.; Inserra, G.; Limosani, M. A Machine Learning Approach to Forecast Economic Recessions—An Italian Case Study. *Mathematics* **2020**, *8*, 241. [CrossRef]
16. Khan, Z.; Pathak, D.K.; Pandey, A.; Kumar, S. Performance Evaluation of Nonlinear Auto-Regressive with Exogenous Input (Narx) in the Foreign Exchange Market. In Proceedings of the 10th IRF International Conference, Chennai, India, 8 June 2014.
17. Kayri, M. Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. *Math. Comput. Appl.* **2016**, *21*, 20. [CrossRef]
18. Chen, S.; Billings, S.; Grant, P. Non-Linear System Identification Using Neural Networks. *Int. J. Control* **1990**, *51*, 1191–1214. [CrossRef]
19. Yu, X.; Chen, Z.; Qi, L. Comparative Study of Sarima and Narx Models in Predicting the Incidence of Schistosomiasis in China. *Math. Biosci. Eng. MBE* **2019**, *16*, 2266–2276. [CrossRef]
20. Powell, K.M.; Sriprasad, A.; Cole, W.J.; Edgar, T.F. Heating, Cooling, and Electrical Load Forecasting for a Large-Scale District Energy System. *Energy* **2014**, *74*, 877–885. [CrossRef]
21. Buitrago, J.; Asfour, S. Short-Term Forecasting of Electric Loads Using Nonlinear Autoregressive Artificial Neural Networks with Exogenous Vector Inputs. *Energies* **2017**, *10*, 40. [CrossRef]
22. Alfred, R. Performance of Modeling Time Series Using Nonlinear Autoregressive with Exogenous Input (Narx) in the Network Traffic Forecasting. In Proceedings of the 2015 International Conference on Science in Information Technology (ICSITech), Yogyakarta, Indonesia, 27–28 October 2015.
23. Benevides, P.; Catalao, J.; Nico, G. Neural Network Approach to Forecast Hourly Intense Rainfall Using Gnss Precipitable Water Vapor and Meteorological Sensors. *Remote Sens.* **2019**, *11*, 966. [CrossRef]

24. Peña, M.; Vázquez-Patiño, A.; Zhiña, D.; Montenegro, M.; Avilés, A. Improved Rainfall Prediction through Nonlinear Autoregressive Network with Exogenous Variables: A Case Study in Andes High Mountain Region. *Adv. Meteorol.* **2020**, *2020*, 1828319. [CrossRef]

25. Paul, R.K.; Sinha, K. Forecasting Crop Yield: Arimax and Narx Model. *RASHI* **2016**, *1*, 77–85.

26. Khamis, A.; Abdullah, S. Forecasting Wheat Price Using Backpropagation and Narx Neural Network. *Int. J. Eng. Sci.* **2014**, *3*, 19–26.

27. Tang, L. Application of Nonlinear Autoregressive with Exogenous Input (Narx) Neural Network in Macroeconomic Forecasting, National Goal Setting and Global Competitiveness Assessment (15 May 2020). Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3601778 (accessed on 15 December 2020).

28. WBG. Worldwide Governance Indicators. 2017. Available online: https://datacatalog.worldbank.org/dataset/worldwide-governance-indicators (accessed on 4 June 2018).

29. UNDESA. E-Government Development Index. 2017. Available online: https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2018 (accessed on 3 June 2018).

30. UNDP. Human Development Reports. 2017. Available online: http://hdr.undp.org/en/content/human-development-index-hdi (accessed on 4 June 2018).

31. WEF. The Global Competitiveness Report. 2017. Available online: http://www3.weforum.org/docs/GCR2017-2018/05FullReport/TheGlobalCompetitivenessReport2017%E2%80%932018.pdf (accessed on 3 June 2018).

32. World Bank. GNI Per Capita. 2017. Available online: https://data.worldbank.org/indicator/NY.GNP.PCAP.CD (accessed on 3 June 2018).

33. WEF. The Global Competitiveness Report 2018. Paper Presented at the World Economic Forum. 2018. Available online: https://www3.weforum.org/docs/GCR2018/05FullReport/TheGlobalCompetitivenessReport2018.pdf (accessed on 5 July 2019).

34. Ghahari, S.A. Detecting and Measuring Corruption and Inefficiency in Infrastructure Projects Using Machine Learning and Data Analytics. Ph.D. Thesis, Purdue University, West Lafayette, IN, USA, 2021; pp. 1–274.

35. Bosso, M.; Vasconcelos, K.L.; Ho, L.L.; Bernucci, L.L. Use of Regression Trees to Predict Overweight Trucks from Historical Weigh-in-Motion Data. *J. Traffic Transp. Eng. (Engl. Ed.)* **2019**. [CrossRef]

36. Shoba, D.; Vijayan, R.; Robin, S.; Manivannan, N.; Iyanar, K.; Arunachalam, P.; Nadarajan, N.; Pillai, M.A.; Geetha, S. Assessment of Genetic Diversity in Aromatic Rice (*Oryza sativa* L.) Germplasm Using Pca and Cluster Analysis. *Electron. J. Plant Breed.* **2019**, *10*, 1095–1104. [CrossRef]

37. Muyeen, S.; Hasanien, H.M.; Al-Durra, A. Transient Stability Enhancement of Wind Farms Connected to a Multi-Machine Power System by Using an Adaptive Ann-Controlled Smes. *Energy Convers. Manag.* **2014**, *78*, 412–420. [CrossRef]

38. Beyca, O.F.; Ervural, B.C.; Tatoglu, E.; Ozuyar, P.G.; Zaim, S. Using Machine Learning Tools for Forecasting Natural Gas Consumption in the Province of Istanbul. *Energy Econ.* **2019**, *80*, 937–949. [CrossRef]

39. Poznyak, T.; Oria, J.I.C.; Poznyak, A. *Ozonation and Biodegradation in Environmental Engineering: Dynamic Neural Network Approach*; Elsevier: Amsterdam, The Netherlands, 2018.

40. Murat, Y.S.; Ceylan, H. Use of Artificial Neural Networks for Transport Energy Demand Modeling. *Energy Policy* **2006**, *34*, 3165–3172. [CrossRef]

41. Taqvi, S.A.; Tufa, L.D.; Zabiri, H.; Maulud, A.S.; Uddin, F. Fault Detection in Distillation Column Using Narx Neural Network. *Neural Comput. Appl.* **2020**, *32*, 3503–3519. [CrossRef]

42. Jaeger, H. *Tutorial on Training Recurrent Neural Networks, Covering BPPT, RTRL, EKF and the "Echo State Network" Approach*; GMD-Forschungszentrum Informationstechnik Bonn: Bremen, Germany, 2002; Volume 5.

43. Diaconescu, E. The Use of Narx Neural Networks to Predict Chaotic Time Series. *Wseas Trans. Comput. Res.* **2008**, *3*, 182–191.

44. Ruiz, L.G.B.; Cuéllar, M.P.; Calvo-Flores, M.D.; Jiménez, M.D.C.P. An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. *Energies* **2016**, *9*, 684. [CrossRef]

45. Boussaada, Z.; Curea, O.; Remaci, A.; Camblong, H.; Mrabet Bellaaj, N. A Nonlinear Autoregressive Exogenous (Narx) Neural Network Model for the Prediction of the Daily Direct Solar Radiation. *Energies* **2018**, *11*, 620. [CrossRef]

46. Hagan, M.T.; Demuth, H.B.; Beale, M. *Neural Network Design*; PWS Publishing Co.: Boston, PA, USA, 1997.

47. Yu, T.; Zhu, H. Hyper-Parameter Optimization: A Review of Algorithms and Applications. *arXiv* **2020**, arXiv:2003.05689.

48. Liu, H.; Kim, H. Ecological Footprint, Foreign Direct Investment, and Gross Domestic Production: Evidence of Belt & Road Initiative Countries. *Sustainability* **2018**, *10*, 3527.

49. Kim, J.-H.; Seong, N.-C.; Choi, W. Cooling Load Forecasting Via Predictive Optimization of a Nonlinear Autoregressive Exogenous (Narx) Neural Network Model. *Sustainability* **2019**, *11*, 6535. [CrossRef]

50. TI. Corruption Perception Index. 2020. Available online: https://www.transparency.org/en/cpi/2020/index/nzl (accessed on 29 January 2021).