Multi-Resolution Forecast Aggregation for Time Series in Agri Datasets*

Fouad Bahrpeyma
Fouad.Bahrpeyma2@mail.dcu.ie

Mark Roantree
Mark.Roantree@dcu.ie

Andrew McCarren
Andrew.McCarren@dcu.ie

Insight Centre for Data Analytics, School of Computing,
Dublin City University, Glasnevin, Dublin 9, Ireland.

Abstract

A wide variety of phenomena are characterized by time dependent dynamics that can be analyzed using time series methods. Various time series analysis techniques have been presented, each addressing certain aspects of the data. In time series analysis, forecasting is a challenging problem when attempting to estimate extended time horizons which effectively encapsulate multi-step-ahead (MSA) predictions. Two original solutions to MSA are the direct and the recursive approaches. Recent studies have mainly focused on combining previous methods as an attempt to overcome the problem of discarding sequential correlation in the direct strategy or accumulation of error in the recursive strategy. This paper introduces a technique known as Multi-Resolution Forecast Aggregation (MRFA) which incorporates an additional concept known as Resolutions of Impact. MRFA is shown to have favourable prediction capabilities in comparison to a number of state of the art methods.

1 Introduction

The usage of analytics and machine learning in the agricultural sector has grown considerably over the past few years. Analytics will play an important role in activities such as optimising farming practices and accelerating crop movements while improving profitability through the development of more robust algorithms for predicting prices across the sector [24]. However, the predictive power of algorithms relies heavily on the data used to learn the models. In many areas, data in the agri or commodity sectors is collected on a weekly or monthly basis. This differs from the majority of data streams which are on a much more frequent basis and thus, models are learnt using significant volumes

*This work is supported by Science Foundation Ireland grant number SFI/12/RC/2289.

of data. For example, in [19], a deep learning approach was used to interpret the interaction between variables using large volumes of agri data streams. This type of approach is not suited when data volumes are low.

In general, predictions are made in short-run or long-run multiple responses. Extending an analysis from a short run, or one-step-ahead prediction (OSAP), to a long run response is known as multi-step ahead prediction (MSAP) and the period applied to the MSAP is known as the Prediction Horizon (PH). The number of steps in a MSAP problem equals the size of the PH, with each step addressing one particular time unit (TU) inside the PH. A TU is the shortest period of time at which a time series is successively sampled, such as daily or weekly price indices. Typically, the TU is referred to as a time stamp or sampling interval, and consists of non-overlapping successive periods of the same length [21].

Time series data quantitative analysis attempts to model sequential patterns over time which possess features such as auto-correlation, conditional heteroskedasticity or non-stationary behaviour. Auto-correlation is the sequential correlation between TUs; conditional heteroskedasticity occurs when the error variance does not remain constant over time, and a time series is stationary if its mean, variance and sequential correlation remain unchanged over time [13]. Additionally, the auto-correlated nature of error terms is a concern when attempting to conduct predictions on such data and has the effect of inflating traditional confidence boundaries [8].

Contribution. In this paper, a novel recursive approach is presented which addresses the shortcomings of the typical recursive strategy by introducing a concept known as the Resolutions of Impact (ROI). The introduction of ROI is an attempt to address the limitations of the sliding window (SW) techniques. SW enables ML algorithms to be applied to time series data [11]. The ROI approach examines how long a time series signal reacts to various local patterns. In our evaluation, a comparative analysis is conducted which compares the forecasting capabilities of our approach with the current state of the art's methods for the Irish Pig Price.

Paper Structure. The remainder of this paper is organized as follows: in §2, we provide a review of other approaches to this problem; in §3, we present our approach which we call the MRFA method, which incorporates a recursive strategy using the Resolutions of Impact; in §4, we discuss the experiments and results of our evaluation; and finally, in §5, we provide our conclusions.

2 Related Research

The time series prediction domain has long been influenced by linear statistical approaches, in particular the ARIMA (Autoregressive Integrated Moving Average) family of models (introduced by Box-Jenkins [8]). The AR and the MA models are the key components of ARIMA and, together with an initial differencing phase for non-stationary data, contribute in fitting a linear time series model. Due to its success in analyzing time series data, ARIMA became popular and found an important place in time series prediction applications[2]. ARIMA, however, does not have the ability to model heteroskedasticity [10] and this has led to the development of methods such as the Autoregressive Conditional Heteroskedastic (ARCH) model [12] and Threshold Autoregressive (TAR) model [25]. In fact, such methods have been incorporated with wavelet analysis with varying success on Agri commodity prices [5]. Machine Learning (ML) approaches have also been widely applied to time series prediction [1]. Artificial Neural Networks (ANNs) or NNs are one of the most frequently used ML models which has proven successful in time series prediction problems [14].

In contrast to OSAP methods, where the prediction horizon (PH) consists of only one TU, MSAP is not a straightforward ML problem. For this reason, solutions to MSAP are generally addressed as MSA strategies. Among the existing strategies, the direct approach and the recursive approach are the most prominent. Others either are extensions of them or combine them to make further improvements [7].

The direct and the recursive approaches are the two most distinctive MSA strategies. While the drawback with the direct strategy is that the sequential auto-correlation is discarded, the recursive strategy suffers error accumulation. A recent study highlighted the range of research projects which are addressing these challenges [7]. Some approaches have been made to reformulate either the direct or recursive methods, or overcome the shortcomings of both by combining them. Generally, in the recursive strategy [20], a single model is trained to perform OSAP several times, each time addressing one TU in the PH.

$$y_{t+1} = f(y_t, ..., y_{t-d}) \tag{1}$$

Equation 1 illustrates this approach where d is the number of lags used as inputs to the OSAP model, t is time, and y is the time series.

In the direct strategy [3], for each TU one specific model is trained as shown in Equation 2. Here, k denotes the position of the TU being modeled, and L is the size of the PH.

$$y_{t+k} = f_k(y_t, ..., y_{t-d}); k = 1, 2, ..., L \tag{2}$$

However, the direct strategy discards sequential correlations in a time series which negatively impacts its performance [23].

In [22], a combined MSA strategy called the DirREC strategy is presented, which combines the principles of the direct and the recursive strategies. Like the direct strategy, DirREC predicts each TU in the PH with a different model and, like the recursive strategy, it uses the predictions of previous TUs as additional inputs. DirREC learns k models according to Equation 3.

$$y_{t+k} = f_k(y_{t+k-1}, ..., y_{t-d}) \tag{3}$$

To obtain the predictions, the k learned models are used in Equation 4. The main drawback with DirREC is that as the number of inputs (with prediction error) increases, the complexity of the model increases correspondingly, without input selection [22].

$$\begin{cases} f_k(y_L, ..., y_{L-d}) & \text{if } k = 1 \\ f_k(y_{L+k-1}, ..., y_{L+1}, y_L, ..., y_{L-d}) & \text{if } 1 \leq k \leq L. \end{cases} \tag{4}$$

MIMO (Multiple-Inputs Multiple-Outputs) [6] is another MSA strategy that aims at addressing the conditional interdependence assumption made by the direct strategy, and the accumulation of errors in the recursive strategy. MIMO, like the direct strategy, considers one specific output for each TU but all are the outputs of one model. MIMO returns multiple predictions covering the entire PH by estimating one specific output for each TU as shown in Equation 5.

$$[y_{t+K}, ..., y_{t+1}] = f_k(y_t, ..., y_{t-d}) \tag{5}$$

MIMO incorporates the same model structure for each TU which effectively limits its flexibility, [23]. In [23], the authors combined the direct and MIMO strategies in an attempt to take advantage of the prominent characteristics of both methods (DirMO). DirMO partitions the PH into several intervals of equal lengths (containing s TUs), each modeled using a different MIMO model. Therefore, DirMO predicts PH by training n ($n = PH/s$) MIMO models as illustrated in Equation 6.

$$\left[\hat{y}_{L+p.s}, ..., \hat{y}_{L+(p-1).s}\right] = \hat{F}_k(y_L, ..., y_{L-d}) \tag{6}$$

where p=1,. . . ,n. However, DirMO discards sequential correlations where the sub-horizons in the PH are modeled via two different MIMO models. Ignoring sequential correlation in the direct and DirMO strategies, the cumulative error in the recursive strategy, the complexity of DirREC, and the inflexibility of MIMO are the reasons that MSAP is still an open challenge. This paper introduces MRFA as an extension of the recursive strategy, addressing the problem of cumulative error by introducing the ROI concept. ROI can be interpreted as an attempt at making use of the actual signal, similar to the direct strategy, by performing Multi Resolution Analysis (MRA) on ROI. MRFA then deals with sequential correlation by following the principles of the recursive strategy.

3 MRFA Methodology

This section introduces MRFA as a novel solution to MSAP that incorporates recursive strategy principles. Incorporated into MSAP is a novel analytical approach known as Resolutions of Impact (ROI), which acts as the core element of MRFA. When implementing MRFA, multiple forecast models are employed to forecast the prediction horizon (PH) at multiple resolutions.

In our approach, the down-sampling operator divides the time series into fixed intervals and returns the mean value of the time series for each interval. The schematic view of the proposed MRFA approach is illustrated in Figure 1.

Based on Figure 1, the value of the signal in the prediction horizon (PH) is forecasted at multiple resolutions, each resolution with a separate model. The final forecasts are then calculated by aggregating weighted averages of the

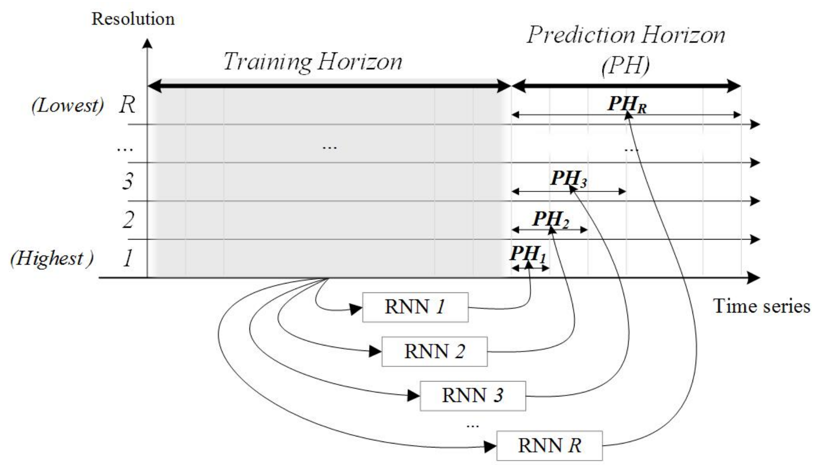


Figure 1: MRFA Approach

different resolutions. The process of MRFA contains three main phases: 1-training the forecast models at multiple resolutions, 2-determining the models' weights, and 3-forecast aggregation.

3.1 Training the forecast models at multiple resolutions

In MRFA, for each resolution $r \in R$ (R is the set of working resolutions), a separate prediction model is employed. In this paper, Recurrent Neural Network (RNN) is used to provide the prediction models. A RNN is a type of ANN that conducts feedback loops in the network architecture as a short term memory, thus providing the facility to model sequentially correlated data [20]. Resolution r defines the forecast target of an RNN model as the mean value of the signal over the next interval of length r.

A RNN is a type of ANN that conducts feedback loops in the network architecture as a short term memory, thus providing the facility to model sequentially correlated data [20].

In MRFA, the resolution determines the forecast target of the corresponding RNN model while, for every RNN, the input is fed from the time series of highest resolution: i.e. the signal itself. An RNN model of resolution r is trained to perform one step prediction in resolution r from the past values of the time series of highest resolution. For multi-steps ahead forecasting, a feedback loop is conducted in the MRFA model, as illustrated in Figure 2.

As shown in Figure 2, the feedback loop feeds the RNN model by a delayed version of the forecasts for further estimates until the entire PH is forecasted. For the time series y_t, the target resolution r is denoted by (PH_r) in Equation 7.

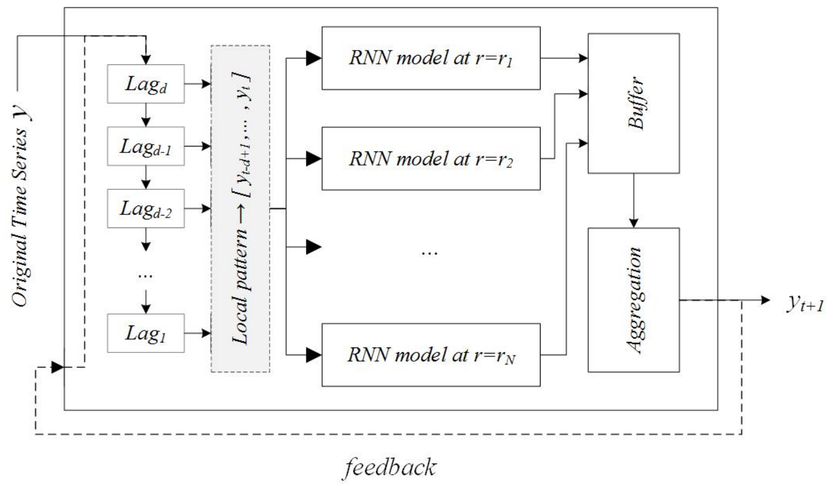


Figure 2: An RNN based forecast model with an output-input feedback loop

$$\underset{\mathbf{r}\in\mathbf{R}}{PH_r(t)} = \frac{1}{r}\sum_{i=1}^{r} y_{t+i} \tag{7}$$

3.2 Determining Weights

MRFA assigns weights to the outputs of the RNN models. A weight reflects the influence of the specific model in the corresponding resolution. Reliability can be determined by the model performance [4], representing the model's accuracy for a single step ahead prediction problem. In [4], a reciprocal value of NMSE (described in section 4.2) is used as model weight. For a single step forecasting model weight at resolution r i.e. W_r is calculated using Equation 8.

$$W_r = \frac{1}{NMSE_r} \tag{8}$$

3.3 Forecast Aggregation

In this phase, the forecasts made by the RNN models of different resolutions are aggregated to provide forecasts for the entire PH. For every TU in the PH, a set of candidate forecasts are introduced by the employed RNN models. At a specific resolution, the final forecast is obtained by aggregating the candidate values introduced by the RNN models at different resolutions, through weighted averaging:

$$\hat{y}(t+1) = \frac{\sum_{r \in R} W_r \sum_{i=0}^{r-1} \hat{y}_r(t+1-i)}{\sum_{r \in R} r.W_r} \tag{9}$$

where $\hat{y}_r(t+1)$ is the forecast made by the RNN model of resolution r. Equation 9 demonstrates how multiple time points contribute to the prediction estimate. The reason for this is that, in MRFA, a RNN of r greater than 1 provides a partial forecast for more than one step. Therefore, as long as r covers the delay i in $\hat{y}_r(t+1-i)$, the corresponding forecast can be used to improve the accuracy.

4 Evaluation

In this section, we describe the experiments and results achieved in our evaluation. We begin with a description of the Irish pig price dataset [17], Figure 3, used in the evaluation; we then describe the metric used for performance evaluation; and finally the experimental results of the study are reported and discussed. No significant Conditional Heteroskedasticity was found in the Irish Pig price dataset using a Ljung-Box Q-test [16].

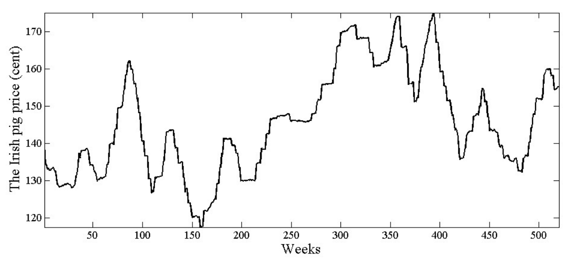


Figure 3: The Irish pig price

4.1 Performance metric

The mean squared error (MSE) is the most popular performance metric for evaluating a forecast model [26]. However, since MSE expresses prediction accuracy within the range of the time series, it cannot detect the performance of the prediction algorithm independent of the dataset. Therefore, the normalized MSE (NMSE) [4] shown in Equation 10 is used in our evaluation.

$$NMSE = \frac{1}{n} \sum_{i=1}^{n} (\frac{T_i - Y_i}{Max(T) - Min(T)})^2, \tag{10}$$

4.2 Experimental Results

The performance of MRFA is highly dependent on the performance of its RNN models. Experimental analyses demonstrated that significant degradation in prediction accuracy occurred when the RNNs in MRFA had more than 12 hidden neurons. In this study, the same RNN configuration was employed in each experiment cycle. MRFA was studied for resolutions ranging from 1 to 7 in size, meaning 7 RNNs were implemented in every experiment. This study also provides a sensitivity analysis on the significant levels of the parameters reported in MRFA analysis. Two parameters have been identified as major contributory factors to the performance of the RNN models: (1) the number of lags being fed to the RNNs as the input, i.e. delays, and (2) the RNN's degree of recurrence that determines the highest number of epochs that the neurons in the RNN's hidden layer remember their own previous outputs. The ranges of parameters for which MRFA was tested is outlined in Table 1.

Table 1: The MRFA's sensitivity parameters

Factors	Range
Delays	[7 8 9 10 11 12 13]
Recurrence delays	[1 2 3 4 5 6 7]

Figure 4 illustrates the impact of growth in PH on MRFA's accuracy.

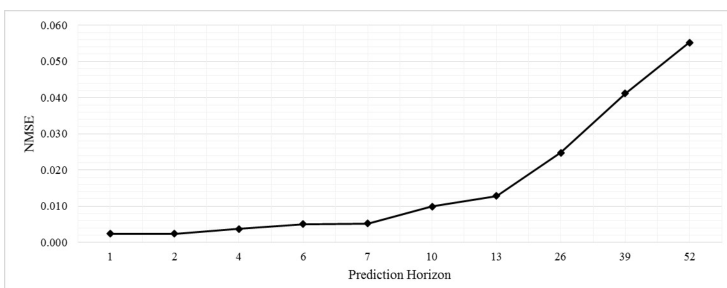


Figure 4: Accuracy of the proposed MRFA with respect to prediction horizon

In order to understand the range of the potential estimates of the RNN model weights, a bootstrap analysis was conducted on the initial estimates, as illustrated in Figure 5. Figure 5 shows how, as PH increases, the error variance increases correspondingly and reaches the highest value at PH=39.

The sensitivity of MRFA on the number of delays and the number of recurrent delays is illustrated in Figure 6. Figure 6a demonstrates the most accurate MRFA model occurs with 7 delays. Figure 6b also reveals a strong relation between the MRFA response to changes in the number of recurrent delays.

In this paper, six state of the art methods have been chosen for comparative analysis: ARIMA [8], Feedforward-SW [14], RNN [15], the direct strategy

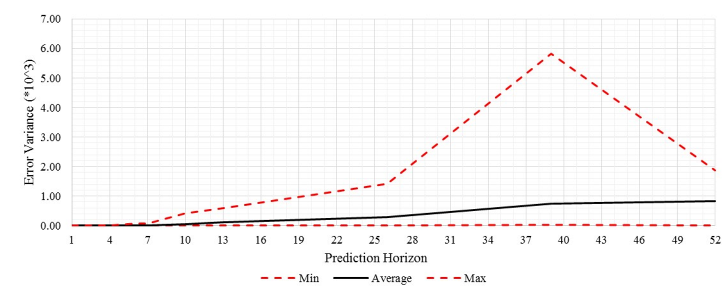
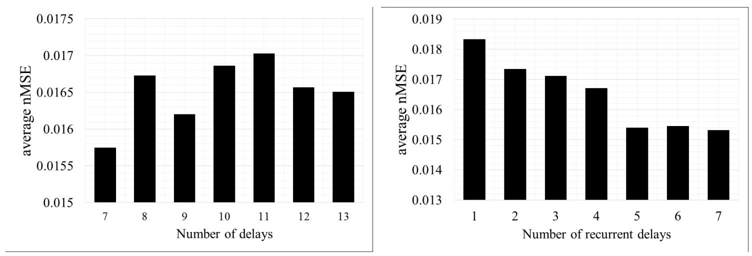


Figure 5: Error variance of the proposed MRFA model



(a) Sensitivity on the number of delays

(b) Sensitivity on the number of recurrent delays

Figure 6: Sensitivity analysis on the MRFA's parameters.

[3], MIMO [6], and ARIMA-NN [18]. A sensitivity analysis on the feedforward model suggests a combination of 12 hidden neurons and a SW of size 10. The sensitivity analysis also suggests that the RNN model exhibits minimum error variance when characterized by 10 input delays, 6 recurrence delays, and 12 hidden neurons. The ARIMA parameters of the ARIMA and the ARIMA-NN models are also identified by analyzing Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, introducing ARIMA(p=1,d=1,q=1) as the appropriate model; where p is the order of the AR component, q is the order of MA model, and d is the differencing order. Further details about the ARIMA parameters is provided in [9]. The performance of the NN model for modeling residuals was affected by the size of the SW and the number of hidden neurons. Sensitivity analysis on these factors demonstrates that the minimum error variance is reached when the SW is of size 13 and the hidden layer contains 10 neurons. Performance comparisons between MRFA, ARIMA, feedforward-SW, RNN, the direct strategy, MIMO and ARIMA-NN with respect to growth in PH are illustrated in Figure 7.

Figure 7 reports the superiority of MRFA over ARIMA, ARIMA-NN, RNN, and feedforward-SW at every PH. The comparison also reveals that, as PH increases, the increase in the prediction error for MRFA occurs more slowly than ARIMA, ARIMA-NN, RNN, the direct strategy, MIMO, and feedforward-

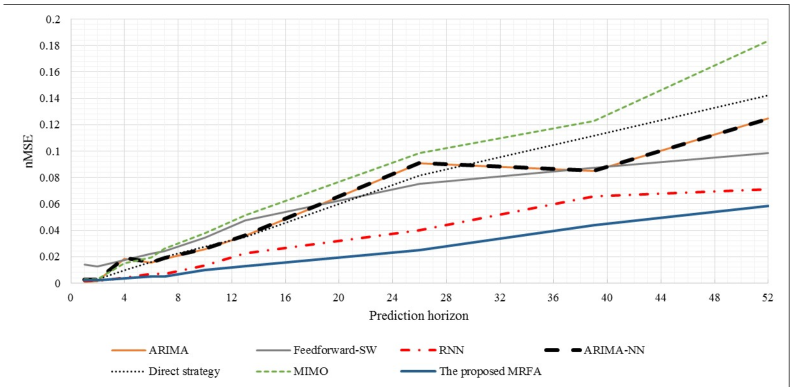


Figure 7: PH comparison between MRFA and ARIMA, feedforward-SW, RNN, the direct strategy, MIMO, and ARIMA-NN

SW.

5 Conclusions

In this paper, a novel RNN based time series prediction approach is presented which improves the prediction accuracy when MSAP is desirable. Our method makes use of multi-resolution analysis on the prediction horizon by incorporating Resolutions Of Impact from local patterns. As our research is based in the agri domain, our evaluation used the Irish pig price dataset. The experimental results demonstrates that the proposed MRFA method outperforms ARIMA, feedforward-SW, RNNs, the direct strategy, MIMO, and ARIMA-NN in forecasting the Irish pig price.

References

[1] Ahmed, N. K., Atiya, A. F., Gayar, N. E., and El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, 29(5-6):594–621.

[2] Atsalakis, G. S. and Valavanis, K. P. (2009). Surveying stock market forecasting techniques–part ii: Soft computing methods. *Expert Systems with Applications*, 36(3):5932–5941.

[3] Aweya, J. (1999). *Sensitivity Methods for Congestion Control in Computer Networks*. PhD thesis, Ottawa, Ont., Canada, Canada. AAINQ48085.

[4] Bahrpeyma, F., Zakerolhoseini, A., and Haghighi, H. (2015). Using ids fitted q to develop a real-time adaptive controller for dynamic resource provisioning in cloud's virtualized environment. *Applied Soft Computing*, 26:285–298.

[5] Bailey, K., Roantree, M., Crane, M., and McCarren, A. (2017). Data mining in agri warehouses using modwt wavelet analysis. In *International Conference on Information and Software Technologies*, pages 241–253. Springer.

[6] Bontempi, G. (2008). Long term time series prediction with multi-input multi-output local learning.

[7] Bontempi, G., Taieb, S. B., and Le Borgne, Y.-A. (2013). *Machine learning strategies for time series forecasting*, pages 62–77. Springer.

[8] Box, G. E. and Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*. Holden-Day.

[9] Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley and Sons.

[10] De Gooijer, J. G. and Hyndman, R. J. (2006). 25 years of time series forecasting. *International journal of forecasting*, 22(3):443–473.

[11] Dietterich, T. (2002). Machine learning for sequential data: A review. *Structural, syntactic, and statistical pattern recognition*, pages 227–246.

[12] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

[13] Goh, C. and Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism management*, 23(5):499–510.

[14] Ho, S., Xie, M., and Goh, T. (2002). A comparative study of neural network and box-jenkins arima modeling in time series prediction. *Computers and Industrial Engineering*, 42(2):371–375.

[15] Lasheras, F. S., de Cos Juez, F. J., Sánchez, A. S., Krzemień, A., and Fernández, P. R. (2015). Forecasting the comex copper spot price by means of neural networks and arima models. *Resources Policy*, 45:37–43.

[16] Ljung, G. M. and Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

[17] McCarren, A., McCarthy, S., Sullivan, C. O., and Roantree, M. (2017). Anomaly detection in agri warehouse construction. In *Proceedings of the Australasian Computer Science Week Multiconference*, page 17. ACM.

[18] Moeeni, H. and Bonakdari, H. (2016). Forecasting monthly inflow with extreme seasonal variation using the hybrid sarima-ann model. *Stochastic Environmental Research and Risk Assessment*, pages 1–14.

[19] O'Donoghue, J., Roantree, M., and McCarren, A. (2017). Detecting feature interactions in agricultural trade data using a deep neural network. In *International Conference on Big Data Analytics and Knowledge Discovery*, pages 449–458. Springer.

[20] Parlos, A. G., Rais, O. T., and Atiya, A. F. (2000). Multi-step-ahead prediction using dynamic recurrent neural networks. *Neural networks*, 13(7):765–786.

[21] Pollock, D. S. G., Green, R. C., and Nguyen, T. (1999). *Handbook of time series analysis, signal processing, and dynamics*. Academic Press.

[22] Sorjamaa, A. and Lendasse, A. (2006). Time series prediction using dirrec strategy. In *Artificial Neural Networks, 2006. ESANN 2006. The 14th European Symposium on*, pages 143–148, Bruges, Belgium. d-side publications.

[23] Taieb, S. B., Bontempi, G., Sorjamaa, A., and Lendasse, A. (2009). Long-term prediction of time series by combining direct and mimo strategies. In *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pages 3054–3061. IEEE.

[24] Thomasson A., S. G. and A, B. (2016). How would google farm? pages 50–55.

[25] Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 245–292.

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