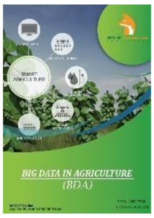
ZIBELINE INTERNATIONAL™ PUBLISHING

ISSN: 2682-7786 (Online)
CODEN: BDAIDR

Big Data in Agriculture (BDA)

DOI: http://doi.org/10.26480/bda.01.2022.17.21



CrossMark

RESEARCH ARTICLE

A COMPARATIVE ANALYSIS OF HOLTS-WINTERS' AND NEURAL NETWORK PREDICTION MODELS ON ANNUAL BLOEMFONTEIN'S PRECIPITATION: RISK AVERSION

Bernard Moeketsi Hlalele*

Department of Business Support Studies, Faculty of Management Sciences, Central University of Technology, Free State, Bloemfontein, 9300, South Africa
****Corresponding Author E-mail: hlalele.moeketsi@gmail.com***

ARTICLE DETAILS

Article History:

Received 12 January 2022
Accepted 15 February 2022
Available online 20 April 2022

ABSTRACT

Neural Networks are a series of algorithms that mimic the human brain's operation to recognise various relationships amongst vast amounts of data. These algorithms are better in predictive analytics than linear regression models due to the use of the hidden layers of neutrons. Failure to accurately predict precipitation pattern may lead major risks agriculture and other heavy water-using economic sectors. The present study's aim was to aid water resources management sector with accurate annual predicted precipitation values for informed decision making. Prior to predictions, data quality controls were conducted, where outliers in the dataset were detected, removed and replaced by expectation maximum algorithm aided by SPSS computer software. Holt-Winter and Neural Network models were comparatively deployed in the predictions of the annual precipitation in the study area. Evident from these results is that NN provided more accurate results in prediction than HW where MSE and MAE for NN were relatively smaller. It can therefore be concluded that NN produces relatively more accurate results than linear models. The study therefore recommends that researchers in meteorology, agrometeorology, water resources management sectors deploy NN for forecasting climate variables for better informed decision making for protection and mitigation of climate-related risks.

KEYWORDS

Holts-Winters, Neural Network, risk, water resources management, Bloemfontein

1. Introduction

Predictive modelling is the method of taking known outcomes and creating a model that can predict values for new occurrences (Haworth & Velliste, 2018). It predicts future events based on historical evidence (Bissantz et al., 2007). Predictive modelling techniques include ANOVA, linear regression (ordinary least squares), logistic regression, ridge regression, time series, decision trees, neural networks, and many others (Stengos & Wu, 2010). Choosing the right predictive modelling methodology at the start of your project will save you a lot of time (Ghosh et al., 2004). Using the incorrect modelling technique will result in inaccurate predictions and residual plots with non-constant variance and/or mean (Kolassa, 2011). Holt-Winters's forecasting is a method for modelling and predicting the behaviour of a time series of values (Woodford, 2018). Holt-Winters is a common time series forecasting technique. It's decades old, but it's still commonly used in many applications, including tracking, where it's used for things like anomaly detection and power planning (Wittenberg et al., 2014). The Holt-Winters forecasting technique is a variant of exponential smoothing that is easy to implement but usually works well in practice and is particularly useful for generating short-term forecasts (Joy et al., 2019; Gregoriou & Kontonikas, 2006). However, artificial neural networks are algorithms that can be used to perform nonlinear statistical modelling, providing a new alternative to logistic regression, the most widely used approach in medicine for constructing predictive models for dichotomous outcomes (Billah et al., 2006). Neural networks have many benefits, including the ability to detect complex nonlinear associations between dependent and independent variables, the ability to detect all possible correlations between predictor variables, and the availability of several training algorithms (Ramesh & Mujumdar, 1996). The water resources management, agriculture and agricultural projects are the most affected by rainfall variability risks. These risks are usually difficult to characterise and predict. The objective of this study is to assist practitioners in these field with the most relevant and accurate tools to predict rainfall patterns in order to effect appropriate risk responses. The next section provides a full account on the methodology used in this study.

2. Methods and Materials

2.1 Quality Control

Outliers were detected by means of SPSS computer software, and all were removed and replaced by Expectation maximum algorithm.

Artificial neural networks (ANNs), also known as neural networks (NNs), are computer structures that are loosely inspired by the biological neural networks that make up animal brains (Samuelides & Cessac, 2007). An ANN is built from a network of linked units or nodes known as artificial neurons, which are loosely modelled after the neurons in the human brain. Each relation, like synapses in a biological brain, has the ability to send a signal to other neurons. An artificial neuron that receives a signal, processes it, and can signal neurons to which it is related. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs (Stark & Guzzetti, 2009). The relations are known as edges (Ramesh & Mujumdar, 1996).

Quick Response Code	Access this article online	
	Website: www.bigdatainagriculture.com.my	**DOI:** 10.26480/bda.01.2022.17.21

Neurons and edges usually have a weight that changes as learning progresses (Manatsa et al., 2008). The weight determines the frequency of the signal at a touch. Neurons which have a threshold that allows a signal to be transmitted only if the aggregate signal reaches that threshold. Usually, neurons are arranged into layers. Different layers can apply different transformations to their inputs. Signals pass from the first layer (the input layer) to the final layer (the output layer), probably several times. It is a time-series forecasting approach that uses the Holt-Winters algorithm (Ghosh et al., 2004). Time series forecasting approaches are used to collect and interpret data and figures, as well as classify outcomes, in order to more reliably predict the future using historical data (Snyder et al., 2002). Users can also smooth a time series and use the data to forecast areas of interest using the Holt-Winters forecasting algorithm. Exponential smoothing applies exponentially decreasing weights and values to historical data in order to reduce the significance of the weight for older data (Chinecherem & Kenneth, 2011). In other words, more recent historical data is given more weight than older findings in forecasting. The next section details the discussion and the results of the study.

3. Results and Discussion

Two prediction models were compared in the prediction pf the precipitation values of the study area using annual precipitation data. Figure 1 shows graph of the study area's annual precipitation plotted against time in years where a monotonic decreasing trend pattern is observed as shown by the regression equation in the graph. This prior analysis was undertaken to get an overview of the behavioural pattern of the input variable, precipitation. A further stationarity test for data quality control was employed to minimise effects of spurious results. A Dickey-Fuller test (ADF) was used as a criterion test as shown in table 1. The results of the stationary test revealed the precipitation dataset stationary with a one tailed p-value 0.002< 0.05 specified significance level.

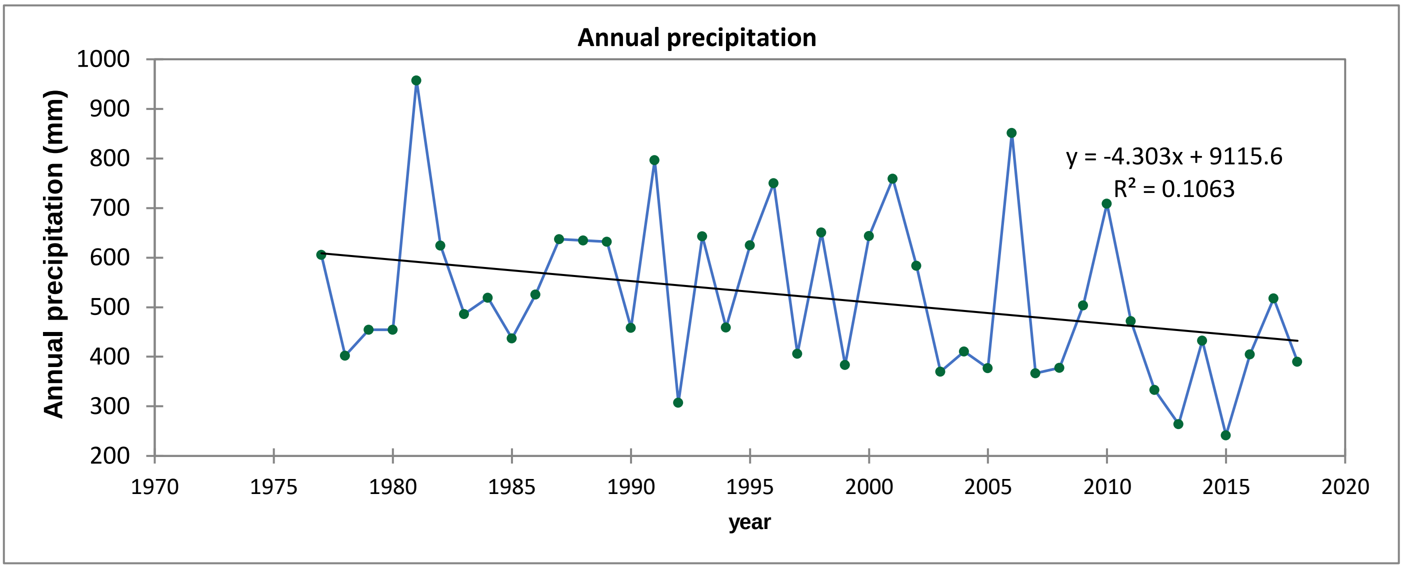
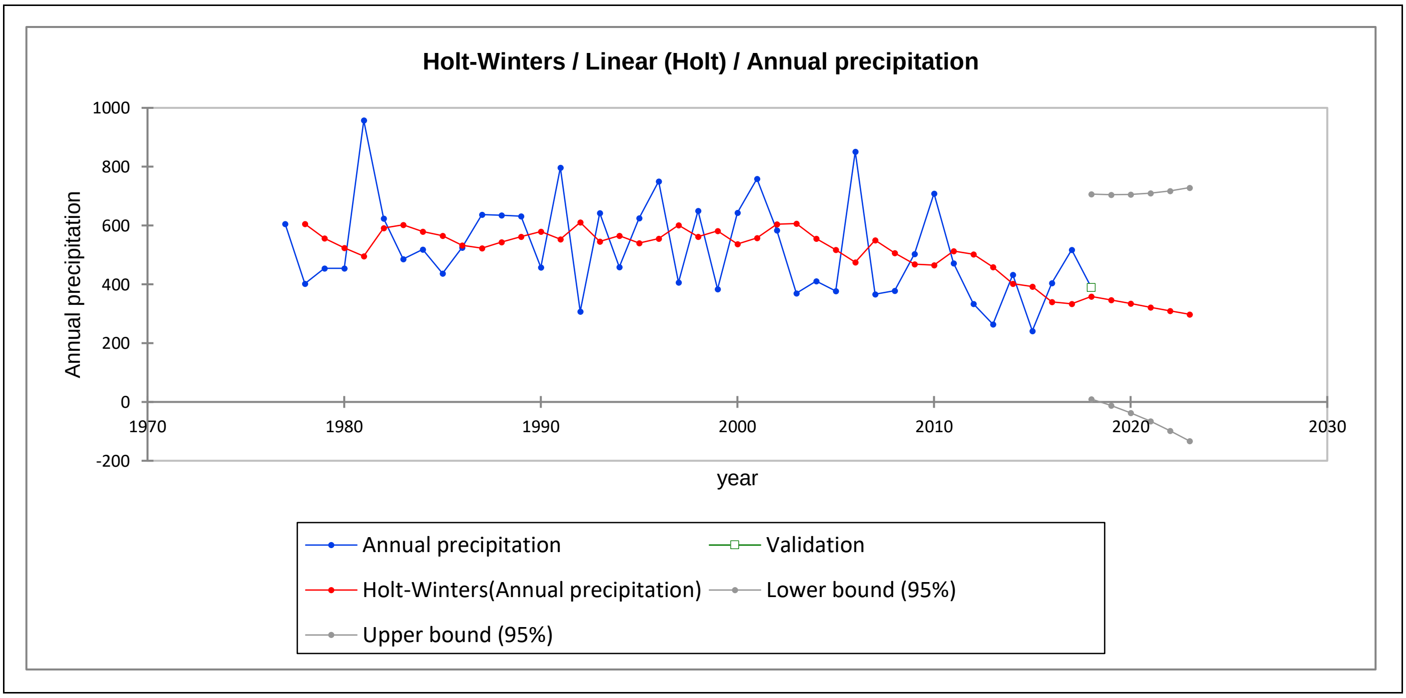


Figure 1: Study area's annual precipitation plot

Table 1: Dickey-Fuller test (ADF (stationary) / Annual precipitation):	
Tau (Observed value)	-4.848
Tau (Critical value)	-3.491
p-value (one-tailed)	0.002
alpha	0.05

After stationarity test, the annual precipitation dataset was subjected to the Holt-Winter linear model for prediction and the prediction results were plotted as shown in figure 2. This model smoothed out the variations in the data and thereby predicted the next three years annual precipitation as indicated in table 4. The goodness of fit statistics for this model is displayed in table 2. This model resulted in the mean squared error (MSE) and the mean absolute error (MAE) of 971.579 and 31.170, respectively. The residual graph was also plotted on the same figure 2 as indicated.



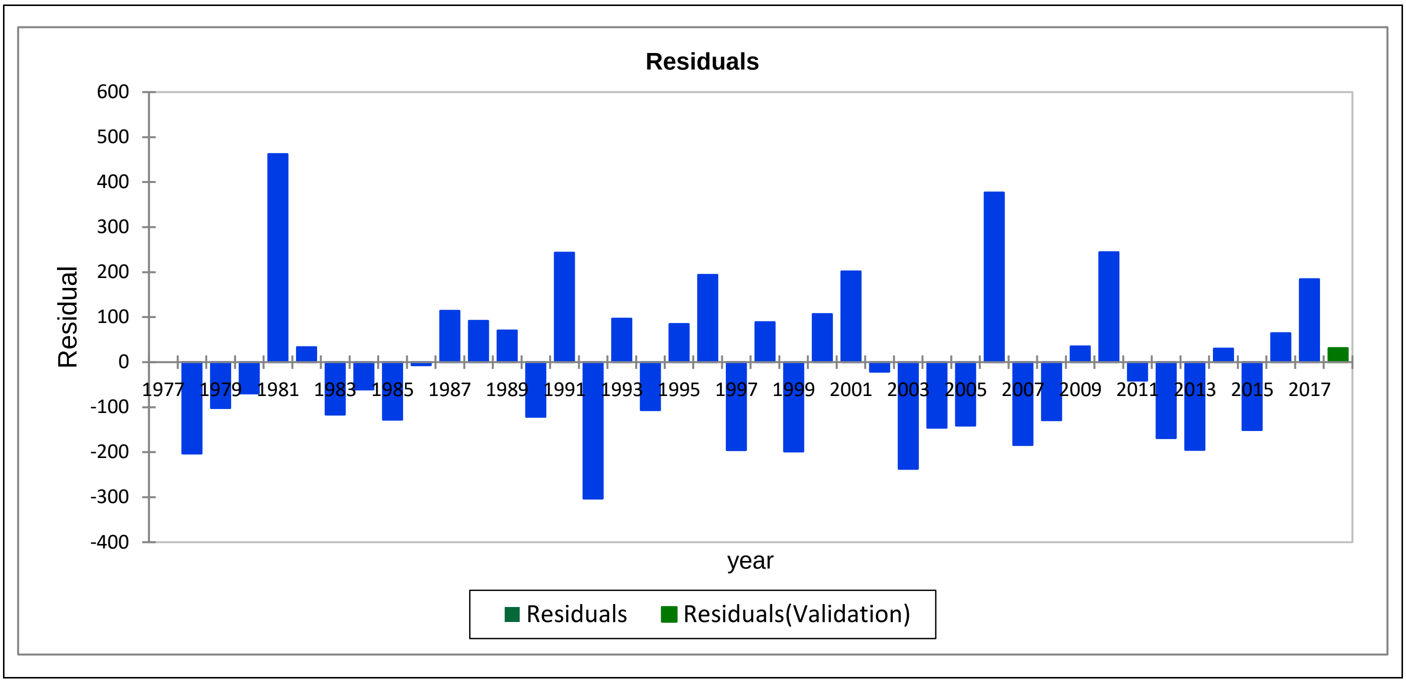


Figure 1: Study area's annual precipitation plot

Table 2: Goodness of fit statistics: Holt-Winters (Annual precipitation):		
Statistic	Data	Validation
Observations	41	1
DF	37	
SSE	1138816.955	971.579
MSE	30778.837	971.579
RMSE	175.439	
MAPE	29.608	7.999
MPE	-9.983	7.999
MAE	142.061	31.170

Figure 3 shows the graph of annual forecasted precipitation by neural network algorithm. Both the linear and neural network models were tested with annual precipitation times series ending in 2020. The neural network model consisted of 12 neutrons both in the input and hidden layers with 1 output layer neutron as shown in table 3. This model resulted in the mean squared error (MSE) and the mean absolute error (MAE) of 0.098497 and 0.157819, respectively.

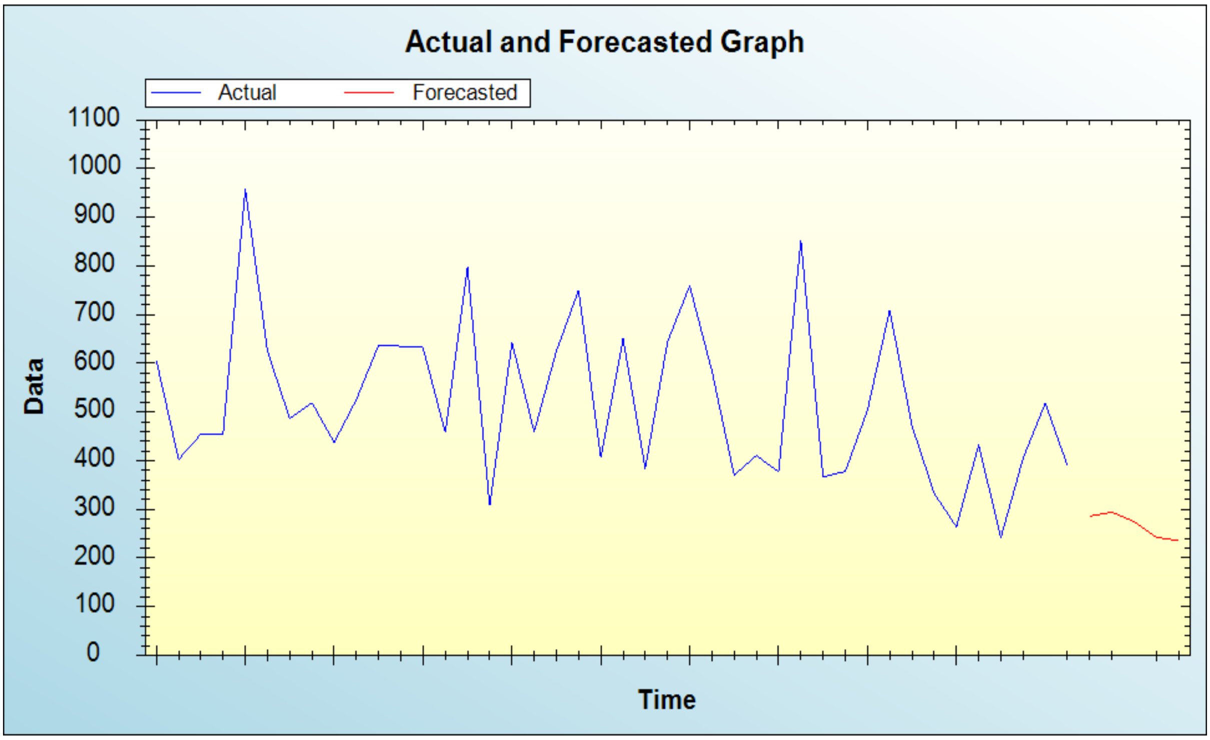


Figure 3: Neural Network's plot showing forecasted values of precipitation

Table 3: Neural Network Model summary	
Variable/Parameter	Annual Precipitation
Network Archiecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Function	**Bipolar Sigmoid Function**
Learning Rate	0.05
Momentum	0.5
Criteria	
Error	0.000009
MSE	0.098497
MAE	0.157819

Table 4 provides the results of the models' predicted annual precipitation. Based on these results, it can be noted that Neural Network model (NN) outsmarted the linear Holt-Winter's model (HW) during the training phase for 2019 and 2020 dataset. The predicted values from NN were closer to the actual values for 2019 and 2020 than those predicted by HW. A further verification test was conducted using a non-parametric mood's test to check if indeed the two predicted set of values from the two models were the same. Table 5 provided evidence that the two predicted sets were indeed different with a statistically significant p-value of 0.002 less than the specified significance level of 5%.

Table 4: Predicted/Forecasted Annual Precipitation Values			
Year	Actual values	Holt-Winters (Annual Precipitation)	Neural Network Predicted Values
2019	285.6660	346.392	286.2347
2020	294.3265	334.293	293.8356
2021	N/A	322.195	274.6935
2022	N/A	310.097	242.3684
2023	N/A	297.998	235.7423

Table 5: Mood Test	
U	10.000
Critical value	3.841
DF	1.000
p-value	0.002
alpha	0.05
An approximation has been used to compute the p-value.	

Test interpretation:

H0: The medians of H and NN are equal.

Ha: Medians of Holts and NN are not equal

As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

4. Conclusion and Recommendations

In conclusion, several prediction models exist in literature that are used particularly to forecast random variables, many of which result in huge errors thereby provide misleading insight about the phenomenon in question. The two linear Holt-Winter (HW) and Neural Network (NN) models were compared in this study to assist decision makers across several water consuming sectors, particularly agriculture and water resources management. Evident from these results is that NN provided more accurate results in prediction than HW where MSE and MAE for NN were relatively smaller. It can therefore be concluded that NN produce relatively more accurate results than linear models. The study therefore recommends that researchers in meteorology, agrometeorology, water resources management sectors deploy NN for forecasting climate variables for better informed decision making for protection and mitigation of climate-related risks.

References

Billah, B., King, M. L., Snyder, R. D., Koehler, A. B. 2006. Exponential smoothing model selection for forecasting. International Journal of Forecasting, 22(2), pp. 239–247. DOI: https://doi.org/10.1016/j.ijforecast.2005.08.002

Bissantz, N., Dümbgen, L., Holzmann, H., Munk, A. 2007. Non-parametric confidence bands in deconvolution density estimation. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 69(3), pp. 483–506. DOI: https://doi.org/10.1111/j.1467-9868.2007.599.x

Chinecherem, O., Kenneth, C. 2011. Neural Networks in Healthcare. Neural Networks in Healthcare, 15(10), pp. 61–68. DOI: https://doi.org/10.4018/978-1-59140-848-2

Ghosh, M., Maiti, T., Kim, D., Chakraborty, S., Tewari, A. 2004. Hierarchical Bayesian neural networks: An application to a prostate cancer study. Journal of the American Statistical Association, 99(467), pp. 601–608. DOI: https://doi.org/10.1198/016214504000000665

Gregoriou, A., Kontonikas, A. 2006. Inflation targeting and the stationarity of inflation: New results from an estar unit root test. Bulletin of Economic Research, 58(4), pp. 309–322. DOI: https://doi.org/10.1111/j.0307-3378.2006.00246.x

Haworth, G. M., Velliste, M. 2018. Chess Endgames and Neural Networks. ICGA Journal, 21(4), pp. 211–227. DOI: https://doi.org/10.3233/icg-1998-21402

Joy, J., Kanga, S., Singh, S. K. 2019. Kerala flood 2018: Flood mapping by participatory GIS approach, Meloor Panchayat. International Journal on Emerging Technologies, 10(1), pp. 197–205.

Kolassa, S. 2011. Combining exponential smoothing forecasts using Akaike weights. International Journal of Forecasting, 27(2), pp. 238–251. DOI: https://doi.org/10.1016/j.ijforecast.2010.04.006

Manatsa, D., Chingombe, W., Matarira, C. H. 2008. The impact of the positive Indian Ocean dipole on Zimbabwe droughts Tropical climate is understood to be dominated by. International Journal of Climatology, 2029(March 2008), pp. 2011–2029. DOI: https://doi.org/10.1002/joc

Ramesh, T. S. V., Mujumdar, P. P. 1996. Rainfall forecasting using neural networks. Stochastic Hydraulics. Proc. Symposium, Mackay, Australia, 1996, pp. 325–332.

Samuelides, M., Cessac, B. 2007. Random recurrent neural networks dynamics. European Physical Journal: Special Topics, 142(1), pp. 89–122. DOI: https://doi.org/10.1140/epjst/e2007-00059-1

Snyder, R. D., Koehler, A. B., Ord, J. K. 2002. Forecasting for inventory control with exponential smoothing. International Journal of Forecasting, 18(1), pp. 5–18. DOI: https://doi.org/10.1016/S0169-2070(01)00109-1

Stark, C. P., Guzzetti, F. 2009. Landslide rupture and the probability distribution of mobilized debris volumes. Journal of Geophysical Research: Earth Surface, 114(2), pp. 1–16. DOI: https://doi.org/10.1029/2008JF001008

Stengos, T., Wu, X. 2010. Information-Theoretic Distribution Test with Application to Normality. Econometric Reviews, 29(3), pp. 307–329. DOI: https://doi.org/10.1080/07474930903451565

Wittenberg, A. T., Rosati, A., Delworth, T. L., Vecchi, G. A., Zeng, F. 2014. ENSO modulation: Is it decadally predictable? Journal of Climate, 27(7), pp. 2667–2681. DOI: https://doi.org/10.1175/JCLI-D-13-00577.1

Woodford, C. 2018. How neural networks work. 5(01), pp. 2016–2018. DOI: https://doi.org/10.26562/IJIRAE.2018.DCAE10085