Attempting to Quantify Lightning Flash-Over Threat through the use of a Neural Network.

Oswald van Ginkel
Renier van Rooyen

Eskom Research, Testing and Development
Rosherville, South Africa

***Abstract*—In this paper the possibilities provided by a neural network as a modelling technique to determine the threat lightning storm clusters pose to a high voltage transmission line is investigated. Some background information is provided to support the use of a neural network in this particular context and the methodology used is discussed prior to providing results.**

***Index Terms*—Visualisation, Situational Awareness, Lightning, Neural Networks, Transmission.**

I. Introduction

Lightning is a natural occurrence that has several detrimental effects on an electrical power grid's operation and maintenance. Due to the erratic nature of the phenomena it becomes necessary for a System Operator to be able to anticipate the movement and quantify the danger that any particular lightning storm presents at any given moment.

A flash-over as a result of lightning discharge is very difficult to define, as the event in question involves complex electromagnetics, impedances in the area and gas discharge physics. The nature of the phenomena demands research into new ways of modelling its behaviour.

An artificial neural network is a modelling approach that is known to allow fitting highly non-linear classification or regression problems [1]. Being able to handle noise [2] it could possibly serve as an excellent choice to map lightning characteristics and line design parameters to actual outages. But it is highlighted by various sources that especially the back propagation algorithm requires enough training data to efficiently reach a good generalizable state.

As indicated in [1] it is necessary to standardise the input variables in some way. This process may involve normalisation as well as some of the feature discovery approaches proposed by [3]. In the case of lightning storms, clusters have a derived probability of having a number of strikes above 80kA. This is substantiated by the design methodology stipulated in [4]. This could possibly serve as an additional feature to simplify the solution search space and obtain a more generalizable trained model as was the case with the examples [5] demonstrated.

[3] uses a methodology that uses equal amounts of data where the output needed to indicate an outage or a non-outage greatly enhanced the training result. This ensured a low number of training epochs in order to reach a stable error (typically less than 200 epochs). This enhanced learning capacity after pre-processing the input data agrees with the findings of [3]. In this instance a different form of pre-processing was implemented to prevent biased training (higher probability to output a certain value due to more occurrences in input data).

Since there is no definite known relationship and a great influence of probability in whether a lightning induced outage will occur as a result of a lightning discharge within a storm with certain characteristics [1] supports a neural network as a good approach to modelling this complex relationship. [1] states that a neural network can be construed without prior assumptions on the "functional form of the relationship between predictors and response".

Other interesting and relevant uses of neural networks include the supervised classification of plant communities [5] and the use of a neural network for real-time process control in a water treatment system [6].

II. Methodology

The investigation will consider the following input data:

- A list of outages (time and location)
- Clustered lightning storms indicated to coincide with lines due to occurring in the same columns as the lines (STORM)
- Pre-calculated minimum current required to pose a threat to a transmission line based on available design data (LINE)

It should be noted that at the compilation of this research effort only 1 years' worth of lightning data was used. The model will be tested against its own training data to see if a relationship with the input parameters can be established.

The Methodology is outlined as the following:

Existing logs (Incident Report) from within the Eskom System Operator data warehouse is utilised to extract a list of outages or faults for the period under evaluation.

Data fields utilized from the Incident Reports include:

- "Start time"
- "First Equipment"

Start time is processed by rounding it to the nearest 10 minute interval. This coincides with a chosen storm cluster interval of 10 minutes chosen for its ability to form sensible and relatable clusters.

From the highly detailed lightning feed supplied from the South African Weather Services, lightning clusters are formed and for each cluster the relevant storm characteristics are compiled.

For the experiment these characteristics include:

- A 10 minute time interval field.
- The total number of discharges in the storm
- A list of 10 kA current bins.

After processing outages that correlate with lightning data two sets of data emerge: storm outages and non-storm outages.

To be able to efficiently train a model that should link storm characteristics to an outage and ensure that biasing doesn't occur to either outages or non-outages it is necessary to expose the model to approximately equal amounts of outage and non-outage cases.

Non-storm related outages greatly outnumber storm related outages by the order of 110:1 (for 11000 storm clusters over lines during a year timespan). An equal balance between outage and non-outage numbers in samples is ensured by duplicating actual storm outages in the training data.

III. Results and Discussion

The results from training an artificial neural network based classifier model based on FANN [7] led to the following results:

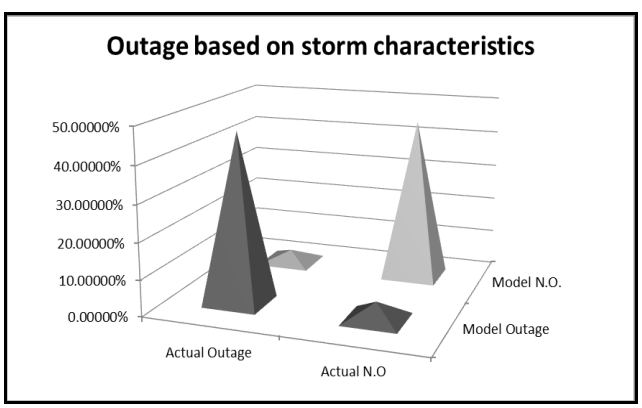

Figure 1 Confusion matrix of predicted vs actual outages without tower design threshold

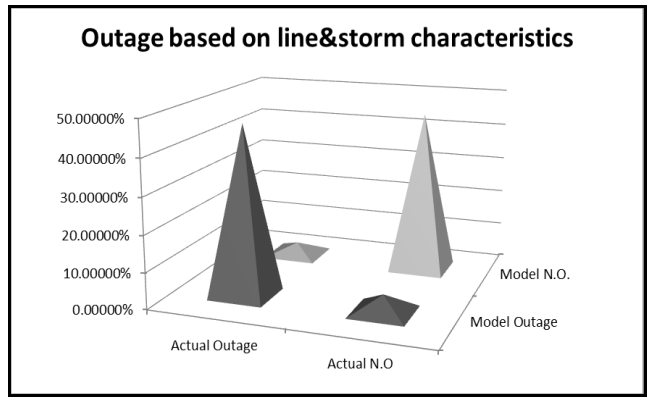

Figure 2 Confusion matrix of predicted vs actual outages with tower design threshold

In the figure provided, it can be seen that the trained model was able to correctly classify most of the storms that would cause outages based on the storm characteristics included in the training of the model. Accurate predictions based on non-outage instances were also mostly correctly predicted. In total the models accuracy can be aggregated to 91.7%.

An additional input to the classifier was evaluated namely the pre-calculated design current for specific towers based on the dominant type of towers within a transmission line assembly. With this additional parameter, the model did not perform significantly different from the model that did not have this design parameter added (90.8% correct classification of outages and non-outages).

To be able to ensure that such a model will effectively function in most cases as a filter for relevant alerts it would be required to train it on larger lightning data sets provided that accurate outage and fault records are available.

Figure 3 represents the sensitivity analysis conducted on the training parameters used in the neural network. From the figure the impact of both the design values of the lines and the lower current and higher current strokes can be seen as separated into each bin. The first two parameters correspond with stroke count and design values respectively. The other parameters indicate the kA bins divisions as previously noted.

It should be noted for clarity sake that the results presented within the figures are as a result of testing the model on the dataset it was trained on. The result seems to offer a promising result in terms of generalizability of a future model.

Further tests using a larger lightning dataset is required in order to make conclusive remarks on the viability of this method and modelling technique.

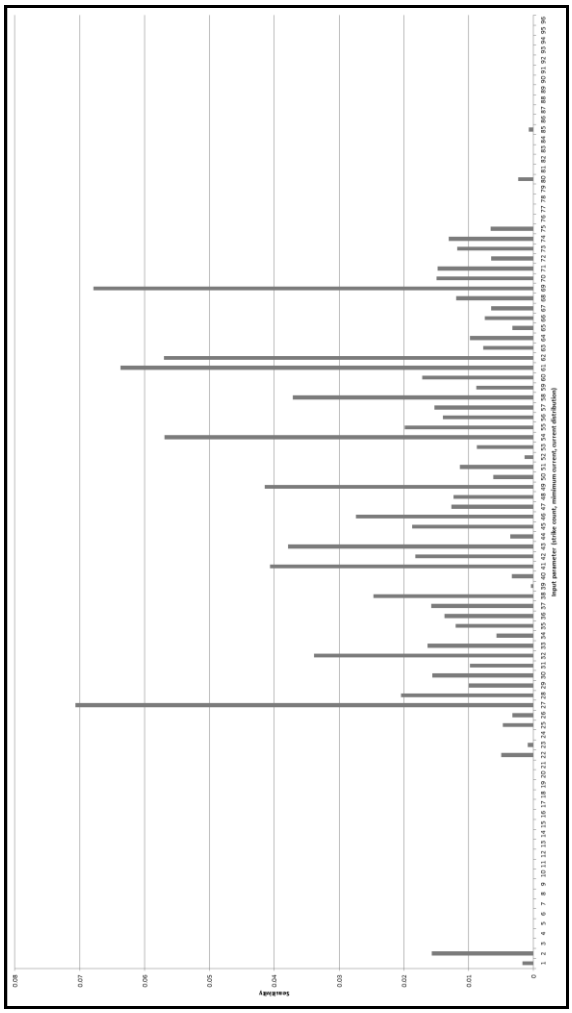

Figure 3 Neural network parameter sensitivity analysis

IV. Conclusion

The research presented in the paper definitively indicates a relationship between the characteristics within a lightning storm and the outages caused as a result thereof through the results of the model that was trained using a neural network.

The results of the training base case create compelling results which demand further investigation.

The overall performance of 91.3% accuracy indicates a viable opportunity to limit alarms raised by lightning discharges' intersection with high voltage transmission lines, thus improving the overall situational awareness of a Transmission System Operator.

Further training and tests are required in order to properly deduce the applicability to the Eskom System operator and the power industry at large.

V. Bibliography

[1] H. S. Stern, *American Statistical Association and the American Society for Quality Control Technometrics*, vol. 38, no. 3, 1996.

[2] D. Rumelhart, J. McClelland and P. R. Group, Parallel distributed processing: explorations the microstructure of cognition vol 1: Foundations, vol. 1, Cambridge: The MIT Press, 1986.

[3] S. Piramuthu, H. Ragavan and M. J. ShawSource, "Management Science," [Online]. Available: http://www.jstor.org/stable/2634678. [Accessed 02 04 2014].

[4] J. Anderson, "Lightning Performance of Transmission Lines," in *Transmission Line Reference Book - 345kV and Above*, Palo Alto, California, Electric Power Research Institute, 1982, pp. 545-596.

[5] L. Černá and M. Chytrý, "Supervised classification of plant communities," [Online]. Available: http://www.jstor.org/stable/409662. [Accessed April 2014].

[6] Q. J. Zhang, R. Shariff, D. W. Smith and S. S. A. Cudrak, "Artificial neural network real-time process control system for small utilities," [Online]. Available: http://www.jstor.org/stable/41313325. [Accessed April 2014].

[7] S. Nissen, "Implementation of a Fast Artificial Neural Network library (FANN)," 2003.