Bull. Aust. Math. Soc. **107** (2023), 346–348
doi:10.1017/S0004972723000023

FUZZY APPROACHES AGAINST OUTLIERS AND APPLICATIONS IN WIND ENERGY

SRINIVAS CHAKRAVARTY

(Received 22 December 2022; first published online 3 February 2023)

2020 *Mathematics subject classification*: primary 62P12; secondary 62A86.

Keywords and phrases: deep neural networks, fuzzy regression functions, noise cluster, outlier, possibilistic clustering, robustness, support vector machines, wind energy.

The world-wide rampant and unfettered energy-reliance on fossil fuels has been recognised as the major cause of greenhouse gases in the environment, and of the resultant global warming and climate change as evidenced by extreme weather events, increasing ocean temperatures and rising sea levels, threatening vulnerable populations. World bodies such as the United Nations, responsible countries and global organisations have been pushing for alternative forms of renewable and clean energy. Given the insatiable demand for energy in the world, these circumstances have generated significant interest and research into alternative sources of renewable and clean energy, such as wind energy and solar energy. As part of the mitigation strategies for climate change, wind energy is being harnessed rapidly worldwide. While hydroelectricity has been dominant over several decades, electricity generated by wind has become a strong contender in recent years. Wind speed is a significant criterion for electricity generation by wind turbines. Consequently, it is important to determine the distribution of commercially viable wind speed and its expected values, to assess the wind energy potential in any selected region.

In common with other natural phenomena, since wind speed is not constant, outliers arise, leading to issues of stability around the generation of wind energy. Outliers in wind speed observations lead to difficulties in proposing accurate predictors for machine learning (ML) studies. Machine learning techniques in general lead to misleading conclusions when data contain outliers. Modelling techniques relying on traditional approaches can give rise to debatable results if such outliers are not suitably

Thesis submitted to RMIT University in June 2022; degree approved on 16 September 2022; supervisors Haydar Demirhan and Furkan Baser.

addressed, since outliers can potentially lead to inaccurate predictions. Depending on the context in which data are collected, the traditional approaches of eliminating such outliers altogether or replacing them with pre-defined values, could lead to the loss of genuine information. Techniques robust to outliers are needed to model wind speeds since outliers are invariably present in wind data. Robust ML techniques assist in predicting a dependent feature based on regression, through an analysis of the associated independent features under conditions such as the presence of outliers. Although there is considerable literature around the use of ML methods for the detection of outliers in various contexts, the number of studies that are focused on the impact of outliers on ML methods for both regression and classification problems is quite limited.

The thrust of this thesis therefore, is to develop an algorithm which can yield reliable estimates of a dependent variable based on a set of independent variables in any given data contaminated with outlier observations. This objective has been achieved in three stages.

Chapter 2 addresses the part of the research study which uses synthetic data along with synthetically induced outliers and carries out a range of Monte Carlo (MC) experiments using various contemporary ML techniques including artificial neural networks (ANNs) and support vector machines (SVMs), along with the newly proposed technique of fuzzy regression function with a noise cluster (FRFN). Experiments are subsequently carried out on publicly available benchmark datasets with injected outliers to determine the statistically significant best performing methods.

Realising from the first stage that FRFN is a robust ML technique with potential, Chapter 3 builds on it to develop the modified fuzzy regression with a noise cluster (MFRFN). Further MC experiments are carried out using the top performing ML techniques along with MFRFN to predict the dependent variables in a large set of publicly available benchmark datasets, based on the corresponding sets of independent variables, in the presence of synthetically injected outliers. The MFRFN framework proves to be the statistically significant top performing technique.

Chapter 4 details the research study carried out with daily wind data containing natural outliers over 39 years, obtained from National Aeronautics and Space Administration (NASA) for 12,336 world-wide locations. The MFRFN method, used along with deep neural networks (DNNs) and SVMs, proves to be the best performing technique. The study then uses the best performing implementations of MFRFN to produce wind speed potential maps of the world, using which wind speed potential can be obtained for feasibility assessment of wind power at a location, by specifying the geographical coordinates of locations on the globe. The wind speed potentials are verified against ground level data from 572 locations in 15 regions of the world.

It can be concluded that the variable agnostic MFRFN framework developed in this study can be successfully generalised to predict dependent variables in large datasets by using the corresponding dependent variables in a regression scenario.

Some of this research has been published in [1–3].

https://doi.org/10.1017/S0004972723000023 Published online by Cambridge University Press

References

[1] S. Chakravarty, H. Demirhan and F. Baser, 'Fuzzy regression functions with a noise cluster and the impact of outliers on mainstream machine learning methods in the regression setting', *Appl. Soft Comput.* **96** (2020), Article no. 106535.

[2] S. Chakravarty, H. Demirhan and F. Baser, 'Modified fuzzy regression functions with a noise cluster against outlier contamination', *Expert Syst. Appl.* **205** (2022), Article no. 117717.

[3] S. Chakravarty, H. Demirhan and F. Baser, 'Robust wind speed estimation with modified fuzzy regression functions with a noise cluster', *Energy Conversion Manag.* **266** (2022), Article no. 115815.

SRINIVAS CHAKRAVARTY, School of Science,
RMIT University, Melbourne, Victoria 3001, Australia
e-mail: srinivas.chakravarty@acm.org

https://doi.org/10.1017/S0004972723000023 Published online by Cambridge University Press