TransNav
http://www.transnav.eu
the International Journal on Marine Navigation and Safety of Sea Transportation
Volume 16
Number 1
March 2022
DOI: 10.12716/1001.16.01.16

A Small Wind Turbine Output Model for Spatially Constrained Remote Island Micro-Grids

D. Žigman[1], K. Meštrović[1] & T. Tomiša[2]
[1] *Zagreb University of Applied Sciences, Zagreb, Croatia*
[2] *University of Zagreb, Zagreb, Croatia*

ABSTRACT: Modelling operation of the power supply system for remote island communities is essential for its operation, as well as a survival of a modern society settled in challenging conditions. Micro-grid emerges as a proper solution for a sustainable development of a spatially constrained remote island community, while at the same time reflecting the power requirements of similar maritime subjects, such as large vessels and fleets. Here we present research results in predictive modelling the output of a small wind turbine, as a component of a remote island micro-grid. Based on a month-long experimental data and the machine learning-based predictive model development approach, three candidate models of a small wind turbine output were developed, and assessed on their performance based on an independent set of experimental data. The Random Forest Model out performed competitors (Decision Tree Model and Artificial Neural Network Model), emerging as a candidate methodology for the all-year predictive model development, as a later component of the over-all remote island micro-grid model.

1 INTRODUCTION AND MOTIVATION

Ensuring quality of life through provision of communal infrastructure across the society is one of the cornerstones of modern civilisation. Technology is to devise solutions for robust and uninterrupted telecommunications and logistics services, and water and power provision to remote and/or isolated communities that contribute to societal development. Power supply to remote and isolated communities has become a burning problem, since electricity drives numerous essential technology systems, thus providing fundamentals of developments and survival. Provision of electricity for societies on remote and isolated islands is a category within the general problem, concerned by numerous research teams across the world. Self-sustainable approaches usually take into account the opportunities to generate electricity from resources at hand, and manage the supply through smart micro-grid solutions [1].

We contribute to the problem solution through our research in model development for a spatially constrained remote island micro-grid, that will optimise resource utilisation through advanced compliance of resources availability quality and utilisation costs, and the micro-grid capacity, load, and operational cost [2]. The over-all model development requires evidence-based model development of the micro-grid components, as well as the model of their optimised system integration into an intelligent self-tuning micro-grid [1, 4]. Here we present the results of experimental data-based model development of a small wind turbine output in relation to the essential set of wind energy predictors. Three candidate models were developed using machine learning model development approach [3, 4], and assessed for their performance, thus distilling the optimal small turbine output model as a component of the spatially constrained remote island micro-grid model. Research results presented in this manuscript may serve additional purpose of modelling the alternative power source on vessels, either as a single energy source, or as a component of the vessel's micro-grid.

2 METHODOLOGY

The approach taken utilised experimental data, and followed the essential statistical research principles of experimental model development [3, 4].

2.1 *Over-all concept*

Experimental data observed were analysed statistically, and the machine learning approach [3, 4] was utilised in development of three candidate predictive models of a small wind turbine output based on two wind predictors: wind speed and wind direction. Model performance assessment was conducted for all three models, using a common set of descriptors, including Predicted-vs-Observed (P-O) diagram, and the adjusted R-squared coefficient.

2.2 *Data*

Observations were taken at the experimental research facility Sotavento in Santiago de Compostela, Galicia, Spain (Figure 1), and provided in tabular format on the internet [5]. A small wind turbine used was manufactured as Bornay 1500 Inclin, extended the 1.5 kWh power rating with fibre-glass, and carbon fibre-blades. In this research, we were concerned with the Spring-time period, with the following duration of the experiment selected: 1 May, 2019 – 31 May, 2019. The experimental data set was split between input and output data using variable selection, as follows: (i) inputs (predictors): wind speed [m/s], wind direction [°], (ii) output (target): cumulative generated energy [kWh].

Figure 1. Location of a small wind turbine experimental site

2.3 *Model development methodology*

Machine learning-based approach was utilised in the predictive model development procedure [3, 4]. Three machine learning-based candidate models were developed: (i) decision tree, (ii) random forest, and (iii) artificial neural network with a single hidden layer. The selection of candidate model approaches was taken based on results of the statistical properties of data.

Decision tree [3] is an optimisation-based model development approach that returns a tree-like structured model, comprising the root- (upper), decision- (intermittent), and leaf-nodes (model decisions). The model develops in two essential steps: (i) the feature vector space (X_1, X_2, …, X_p) is divided into non-overlapping regions Ri, and (ii) every new observation of feature vector is assigned to region Ri based on the mean value of the previous (training) observations in the same region R_i. Decision tree is a simple and clear model easily deployed for both the human assessment and as a computer algorithm. Its shortcomings include potential over-fitting (modelling noise rather than a signal) and poor performance with continuous data.

Random forest utilises the decision-tree concept to form a forest of decisions that eventually yield the random forest decision. The random forest development approach requires the original data set to be split into a number of sub-sets with randomly selected data. Then, decision tree models are developed with every sub-set. Decision, or, estimate, related to new set of observations is performed by all the decision trees, and then integrated using either the democratic procedure (majority/average of votes of separate decision trees) or using weighted approach, favouring influential decision trees. Random forest model encompass variance in data successfully and tackles over-fitting efficiently, but is computationally intensive, and not suitable for real-time predictions.

Artificial neural network mimics a human or animal ones, with artificial neurons being kicked-off by the appropriate input level, and exchanging their outputs with other neurons it is connected with. The artificial neural network (ANN) consists with neuron layers that receive the inputs (input layer), those that reside internally within the network (hidden layers), and the one that provides decision/estimation results (output layer). While theoretically an ANN may consist of many internal layers, a one- or two-hidden layer-architecture may produce optimal results. ANN is suitable for modelling the complex systems, where prediction of behaviour is required without explaining the system.

Model performance assessment was conducted using two essential model performance indicators: (i) Predicted-Observed diagram, (ii) adjusted R-squared coefficient. The P-O diagram is a simple graphical indicator of model's performance, designed as a graphical presentation of observed-predicted pairs. The adjusted R-squared indicator is defined as follows. Let denote observations as $\vec{y}$, and model-derived estimates as $\vec{x}$. The Coefficient of Determination (R-squared) indicator is then defined as given in (1), with the $\bar{y}$ denoting mean of observations.

$$R^2 \equiv 1 - \frac{\sum_{i=1}^{n}\left(y_i - x_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (1)$$

The Coefficient of Determination may mislead mode comparison, so the adjustment to model structure given in (2) is suggested, yielding the adjusted R-squared indicator, with:
s_n … number of observations,
p … number of predictors.

$$adjR^2 = 1 - \left(1 - R^2\right) \cdot \frac{s_n - 1}{s_n - p} \qquad (2)$$

3 RESEARCH RESULTS

This Section contains research results in a form of developed predictive models definition and performance assessment, as described in Section 2. Predictive model development procedure was coded in the open-source R framework for statistical computing, using the R library rattle and associated R libraries.

3.1 *Decision Tree Model (DTM)*

Decision tree model was developed using a common approach. Statistical analysis embedded in the model development procedure found wind direction statistically insignificant, rendering the model based on the wind speed only. DTM is shown in Figure 2.

Validation sub-set of data was used in the DTM assessment procedure, with the results presented in Figure 3. Although performance analysis returned a rather high adjusted R-squared value, the model's P-O diagram returns layered structure with significant variations from the central line. Further to this, DTM model tends to increase inaccuracy for values at the edge of the modelling range.

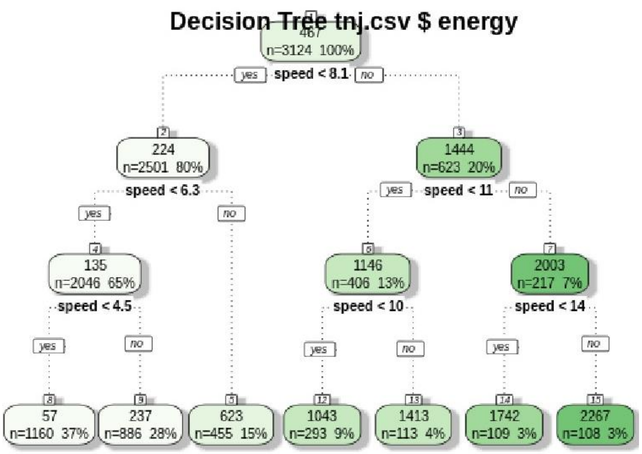

Figure 2. Decision tree model of a small wind-turbine output

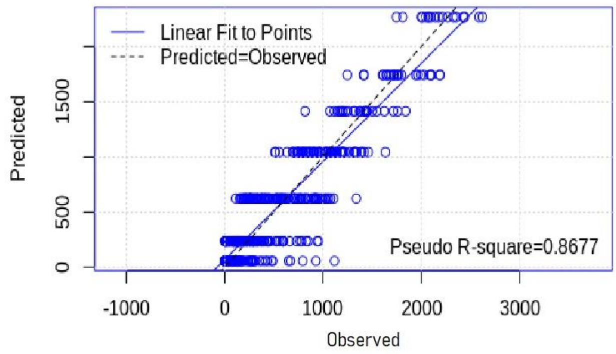

Figure 3. DTM performance assessment

3.2 *Random Forest Model (RFM)*

Model consisted of 500 decision trees build-up on the randmised sub-sets of the original observations. The RFM performance analysis returned a very high adjusted R-squared index value, in justification of the good model fit. Again, RFM was assessed based on a separate validation data, compiled as a sub-set of the original population. P-O diagram (Figure 4) shows that observation-predicted values tie up with the P-O line firmly, with just a handful of exceptions.

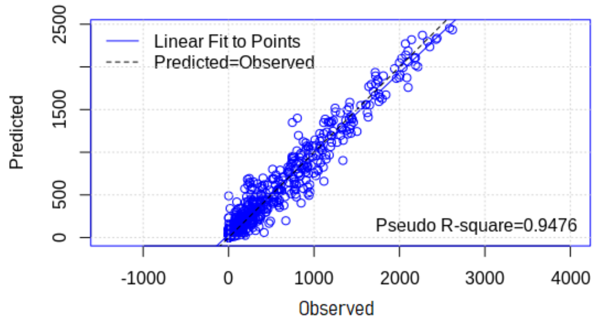

Figure 4. RFM performance assessment

3.3 *Single Hidden Layer Neural Network Model (SHLNNM)*

The SHLNNM was developed using the training set of experimental data, with definition given in Table 1.

Table 1. Definition of the SHLNNM

Weights for node h1 (hidden layer)			
b → h1	i1 → h1	i2 → h1	
120.80	-352.85	306.06	
Weights for node o (output layer)			
b → o	h1 → o	i1 → o	i2 → o
-768.10	254.34	166.94	-0.10

The SHLNNM parameters (Table 1) are described as follows:
b … denotes the bias associated with the node
h1 … marks the hidden layer node 1
i1 … marks the input node 1 (or: the input variable, or predictor, 1)
i2 … marks the input node 2 (or: the input variable, or predictor, 2)
o … marks the output node

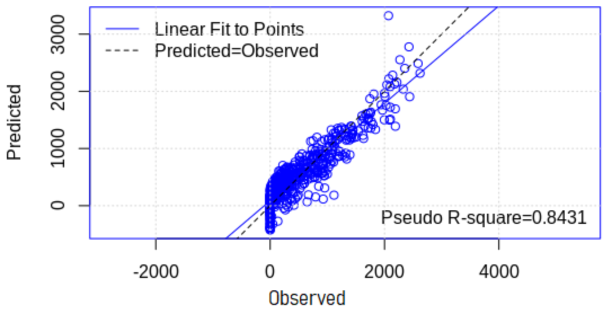

Figure 5. SHLNNM performance assessment

The SHLNNM performance assessment revealed considerable adjusted R-square value, but the model is outperformed by the other candidates. Additionally, the P-O diagram (Figure 5) shows a remarkable deviations from the linear fit for excessive (very small and very high) values.

4 DISCUSSION AND CONCLUSION

The research presented aims at development of a small wind turbine model, based on experimental data collected during May 2019 in near coastal are of northern Spain, and deployment of machine learning techniques, related to statistical properties of the experimental data set. The model is set to become a component of the over-all micro-grid model supposed to be deployed in small remote and isolated island communities, or on vessels or fleets.

Three candidate machine learning-based models of a small wind turbine output were developed using the R framework for statistical computing: (i) Decision Tree Model (DTM), (ii) Random Forest Model (RFM), and (iii) Artificial Neural Network Model. While all three models extends high goodness-of-fit, their response stability over the range of observation values varies significantly. The RFM performance extends far the best adjusted R-squared, with a a linearity of prediction across the range of observation, rendering it as the most suitable model of a small wind turbine output.

Research will continue with the model development that will encompass variance of wind scenarios throughout the year, followed by its integration within the over-all spatially constrained remote island micro-grid model.

REFERENCE

1. Acevedo, M.F.: Introduction to Renewable Power Systems and the Environment with R. CRC Press (2018). https://doi.org/10.1201/b21919.
2. Department of Energy: Small Wind Guidebook, https://windexchange.energy.gov/small-wind-guidebook.
3. Efron, B., Hastie, T.: Computer Age Statistical Inference: Algorithms, Evidence and Data Science. Cambridge University Press (2016).
4. Kuhn, M., Johnson, K.: Feature Engineering and Selection: A Practical Approach for Predictive Models. Chapman and Hall/CRC (2019).
5. Sotavento: Experimental Ecological Park Sotavento Real Time Data Archive, http://www.sotaventogalicia.com/en/technical-area/real-time-data/historical/, last accessed 2021/03/30.