Benson Mbuya[1,2], Matteo Moncecchi[3], Marco Merlo[3] Thomas Kivevele[1,*]

J. Electrical Systems 15-4 (2019): 593-606

JES Journal of Electrical Systems

Regular paper

Short-term load forecasting in a hybrid microgrid: a case study in Tanzania

Most emerging countries such as Tanzania are promoting rural electrification through installation of microgrids. This paper proposes an approach for short-term day-ahead load forecast in rural hybrid microgrids in emerging countries. Energy4Growing research project by Politecnico di Milano department of energy in collaboration with EKOENERGY (www.ekoenergy.org) implemented in Ngarenanyuki Secondary School (Arusha, Tanzania) innovative control switchboards to form an energy smart-hub. The smart-hub was designed to manage the school's 10kW hybrid micro-grid comprising: PV-inverter, battery storage, micro-hydro system, and genset. Ngarenanyuki school microgrid's data was used for the experimental short-term load forecast in this case study. A short-term load forecast model framework consisting of hybrid feature selection and prediction model was developed using MATLAB environment. Prediction error performance evaluation of the developed model was done by varying input predictors and using the principal subset features to perform supervised training of 20 different conventional prediction models and their hybrid variants. The objective function was feature minimization and error performance optimization. The experimental and comparative day-ahead load forecast analysis performed showed the importance of using different feature selection algorithms and formation of hybrid prediction models approach to optimize overall prediction error performance. The proposed principal k-features subset union approach registered low error performance values than standard feature selection methods when it was used with 'linearSVM' prediction model. Furthermore, a hybrid prediction model formed from the elementwise maximum forecast instances of two regression models ('linearSVM' and 'cubicSVM') yielded better MAE prediction error than the individual regression models fused to form the hybrid.

Keywords: Load forecast; Feature selection; Hybrid micro-grid, Emerging countries.

Article history: Received 22 July 2019, Accepted 17 October 2019

1. Introduction

Electricity in today's world is a necessity, unfortunately most countries in the developing world have limited access to it, this is critical in Africa. Taking Tanzania as an example, in the 2012 census, only 17% of households had access to electricity and only 5.3% rural households had access to electricity where 70% percent live [1]. A plausible and sensible solution has been to promote rural electrification, especially through use of renewable energy since it will take a longer and big funding for the national grid to reach rural settlements, which happen to be scattered [2]. Renewable energy sources have been highly promoted in rural electrification, but are often affected by intermittency, which is in most cases resolved by installing storage systems. It is also common to mix renewable energy and non-renewable energy. However, this requires that an Energy Management System (EMS) optimizes energy generated and stored [3]. One of the vital organs and functions of EMS is Load forecasting.

Load forecast is an important tool for a utility power company decision-making in a decentralized electricity day-ahead power market, demand side management, and unit

* Corresponding author: T. Kivevele, Department of Materials and Energy Science and Engineering, The Nelson Mandela African Institution of Science and Technology, P.O Box 447, Arusha, Tanzania, E-mail: thomas.kivevele@nm-aist.ac.tz

1 Department of Materials and Energy Science and Engineering, The Nelson Mandela African Institution of Science and Technology, P.O Box 296, Arusha, Tanzania

2 Arusha Technical College, P.O Box 447, Arusha, Tanzania

3 Energy department, Politecnico di Milano, via la masa, 20156 Milano, Italy

commitment. There is room for improvements in the accuracy of forecasting models, which are also influenced by geographical location of microgrids [4] [5].Bio-inspired artificial intelligence based forecasting algorithms have been found to have better performance than statistical based models [6] [7]. Ensembled or hybrid forecasting models result in higher accuracy and performance than the individual models forming the hybrid model [8][9][10]. There has been few or no works on EMS load forecasting models studies for emerging countries such as Tanzania. The focus of this study is to propose a short-term load forecast framework that can be used in hybrid microgrids.

2. Brief overview of the Microgrid

Energy4Growing project by Politecnico di Milano department of Energy in collaboration with EKOENERGY (www.ekoenergy.org) and SunEdison implemented at Ngarenanyuki Secondary School (Arusha, Tanzania) an innovative converter and control switchboards designed to manage the school's 10kW hybrid micro-grid comprising: a run-of-river hydropower system (3 kW), backup generator (5 kW), PV-inverter and battery storage [11]. Apart from installing the switchboards, the project involved upgrading the school's PV (by 3 kW) and battery bank by 30X202 Ah/12V lead-acid batteries. The system can operate in Manual/Automatic modes. Automatic mode automatically connects/disconnects sources and loads based on their priority index by the PLC (Programmable Logic Controller). The PLC also serves as a data logger set to 1 second sampling rate. In figure 1, Q1 represents the inverter control board while Q2 denotes PLC switchboard [12] [13]. A detailed description of the microgrid, from the design phase to the deployment and operation is available on Energy4Growing research team Facebook page. [14].

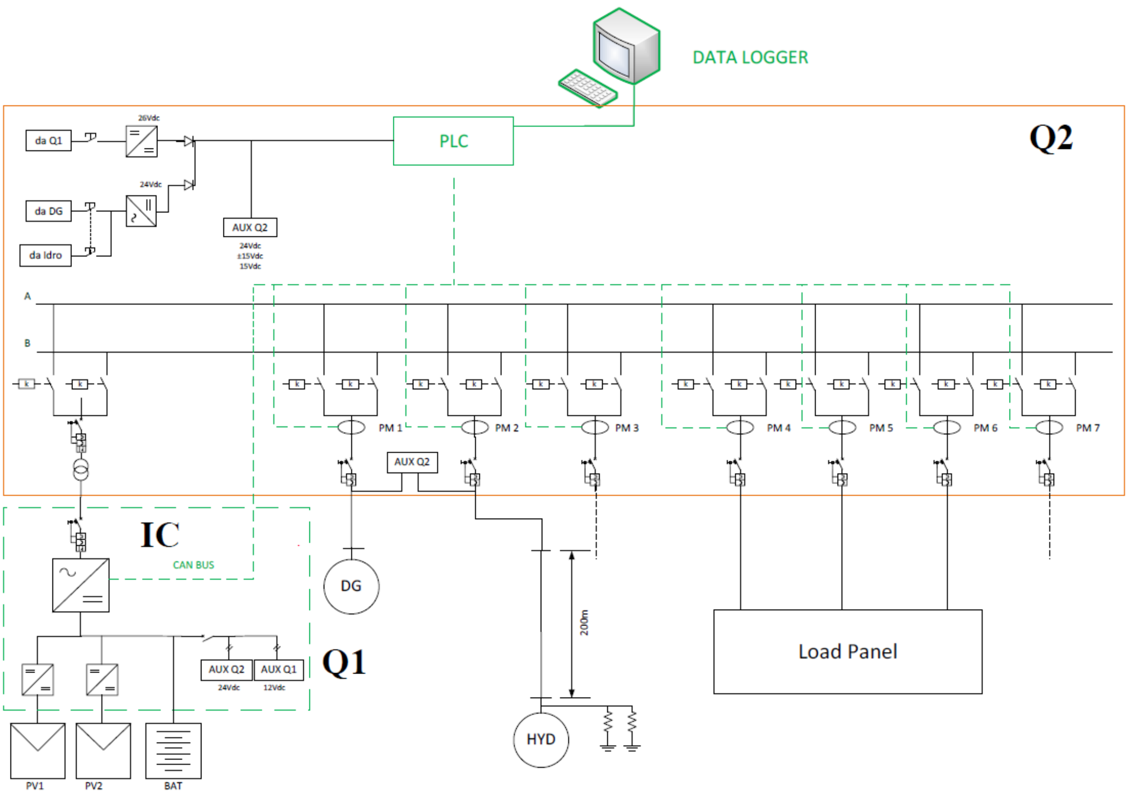

Figure 1. Ngarenanyuki microgrid architecture with PLC based demand-side management and basic battery management and resource management.

Ngarenanyuki microgrid has already been described in literature : Carmeli et al [12] made a comparison between AC and DC bus configurations, control stategies, reliability and efficiency before the implementation of the architecture in figure 1. Mandelli et al [13] reported an overview of typical configuration and economic models of batteries in off-grid system and describes the application of batteries in the Ngarenanyuki school grid experimental project. Carmeli et al[11] performed analysis of the school's actual power supply system prior to deployment of the architecture in figure 1 as well as simulation of operations and dynamics of the architecture. Mandelli et al[15] examined the school's consumption pattern, simulation of the electro-mechanical operation and the power flow from generators and loads. Nyari et al [16] studied consumption of electrical appliances on stand-by and active operation states. A Matlab-based stochastic procedure that allows to generate load profiles of microgrids in small communities was developed in Mandelli et al[17], and validated with Ngarenanyuki microgrid data. Mauri et al [18], developed a neural-fuzzy EMS for Ngarenayuki school grid. This work puts forward a day ahead load forecasting model framework derived from conventional feature selection (FS) and regression algorithms variants.

3. Methodology

This section describes the approach developed in this paper. An overview of the proposed model framework is first given, followed by a description of the individual blocks that compose the model, namely: input data pre-processing; feature selection, prediction models and performance evaluation index. This work proposes an approach for short-term day-ahead load forecast in rural hybrid microgrids of emerging countries. Furthermore, the load forecast is achieved through performing an experimental comparative analysis of feature selection and prediction algorithm variants.

3.1. Proposed load forecasting model framework

Figure 2 shows the proposed model framework adopted in this work. The model was implemented in Matlab environment. Data exploration, cleaning and transformation stages were implemented so as to enhance input data and the model's behaviour.

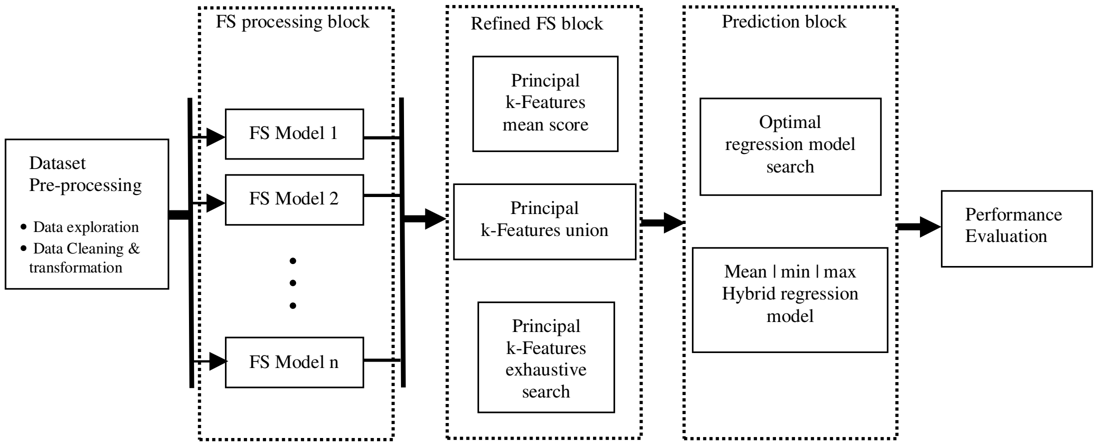

Figure 2. Proposed day ahead load forecasting model framework

Feature selection stage aims at minimizing number of predictors in order to reduce computation complexity without compromising prediction performance. As it will also be shown later, feature selection models differ in their mechanism, thus they yield different results. It is therefore proposed to combine more than one FS model in order to increase performance at the prediction stage of the load forecast procedure. With respect to the proposed model in figure 2, input data is fed to each of the FS models. In this work, 5 FS models were used. Each of the 5 feature selection models ranked the same features with different weight scores. In general, it was observed that features which were voted to be the most important according to one FS model, were also given high importance in the other FS models. To overcome the dilemma of feature selection, three approaches are proposed in selection of a subset of k principal features from a superset of n features through: 1) mean score; 2) subset union; 3) refined exhaustive search based on k-combination, they are described later. For computational time reasons, either the mean score or subset union or both approaches can be used and compared to select features for training and prediction. Refined exhaustive search method should be opted as the last resort, since it is computationally much more intensive than mean score and subset union approaches.

The identified k principal features can be trained by 20 different regression models, thereafter, performance efficacy assessed by mean absolute error (MAE) and root mean square error (RMSE) of each regression model compared. All the possible combinations of 2 regression models out of the 20 models can be fused to form a hybrid model based on the mean or min or max values of the best 2 models. The resulting hybrid regression model show improved prediction performance. Ultimately, the choice of which path to follow in the proposed framework depends on the computational resources available as well as the degree of prediction performance desired.

3.2 Input data description

The period of the particular dataset used in this paper is from 15 May 2015 to 7 March 2018. The data sampling was one second, but during pre-processing it was converted to hourly aggregated observations. The microgrid dataset contains a total of 12,912 samples. The training and cross validation dataset used was from May 2015 to January 2018, while test data used was from 1 February 2018 to March 2018. Data used are available for research purposes thanks to the Energy4Growing project [19].

Table 1: Dataset features description

#	Feature / predictor	Feature Description	Evaluation time
1	Month	Month number of the year	day D
2	Day	Day of the month	day D
3	WeekDay	Day of the week	day D
4	Hour	Hour of the day	day D hour h
5	Weekend	Weekend and holiday indicator	day D
6	temp	Ambient temperature (control room temperature)	day D hour h
7	P_DG	Back-up diesel generator power	day D hour h
8	P_HYD	Micro hydro power	day D hour h
9	P_inv	Power from inverter	day D hour h
10	Vdc_bus	PV-inverter DC bus system voltage	day D hour h
11	PPV	PV array output power	day D hour h
12	SOC	Battery bank state of charge	day D hour h
13	AirTemp	Outdoor air temperature from nearby airport	day D hour h
14	atmPressure	Atmospheric pressure	day D hour h
15	RHumidity	Relative humidity	day D hour h
16	WindSpeed	Wind speed	day D hour h
17	DewpointTemp	Dew point temperature	day D hour h
18	T2_temp	Ambient temperature 2 days before	day D-2 hour h
19	T2_Load	Load 2 days before	day D-2 hour h
20	T1_temp	Previous day control room ambient temperature	day D-1 hour h
21	Year	Data log Year	day D
22	T1_Load	Previous day load	day D-1 hour h
23	Load	Current day load	day D hour h

The dataset comprised in total of 23 features from previous day as predictors and next day (24 hours) load profile as the target or response. Prediction of the entire next day is performed at midnight (00hrs) which is the start next day. For each hour *h* of the next day, the forecast is based on the values of the 23 predictors at the hour *h* of the day before. Weather data from a nearby airport was used as part of the 23 features [20]. The weather parameters incorporated in the dataset were outdoor relative humidity, air temperature, wind speed, atmospheric pressure, and dew point temperature.

3.3. Data pre-processing

The first step in this work was to explore the data in order to check for integrity, to spot missing values, and examine the relationship between load and the other features in the dataset. As an example, Figure 3 shows the unstable nature of load consumption of the microgrid for October 2017. Data cleaning stage involved detection and correction of missing values and outlier anomalies. Days in the dataset found with missing values were filled with mean values of adjacent neighbouring records. Days with no logged entries were ignored. For smoothing, the dataset was transformed by retiming from 1 second observation entries into mean hourly observations, and further smoothed to remove outliers. *Sgolay (Savitzky-Golay filter)* algorithm based data smoothing method was opted out of seven other data smoothing methods which are *movmean*, *movmedian*, *lowess*, *rlowess*, *loess*, *rloess*, and *Gaussian* smoothing method available in MATLAB software. *Sgolay* was used because it is effective in preserving higher moment peaks in a signal [21][22]. Data smoothing has the effect of removing noise from data hence improving performance of the prediction model.

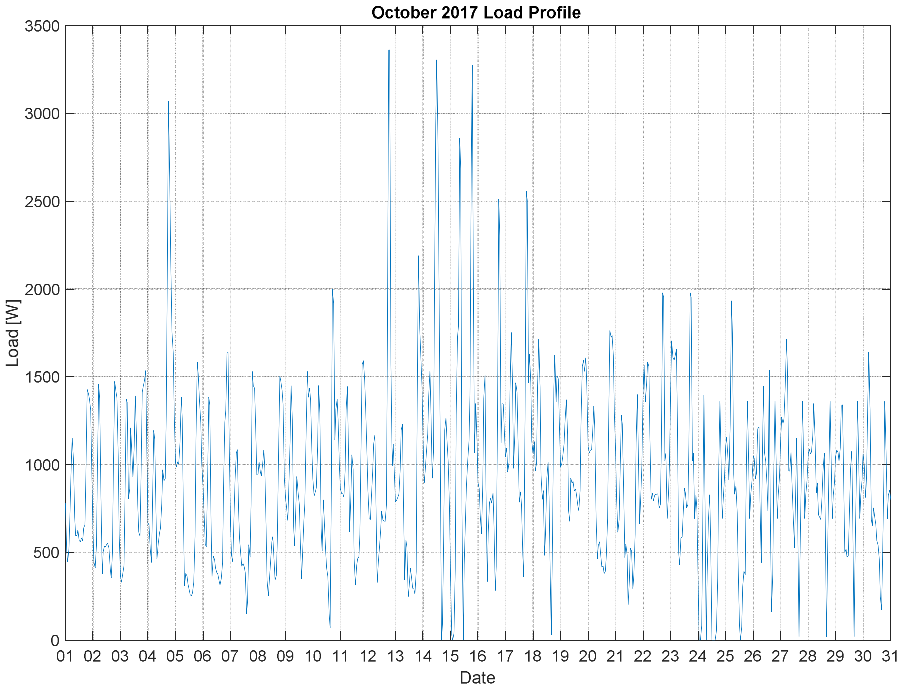

Figure 3. Ngarenanyuki microgrid unstable load profile nature.

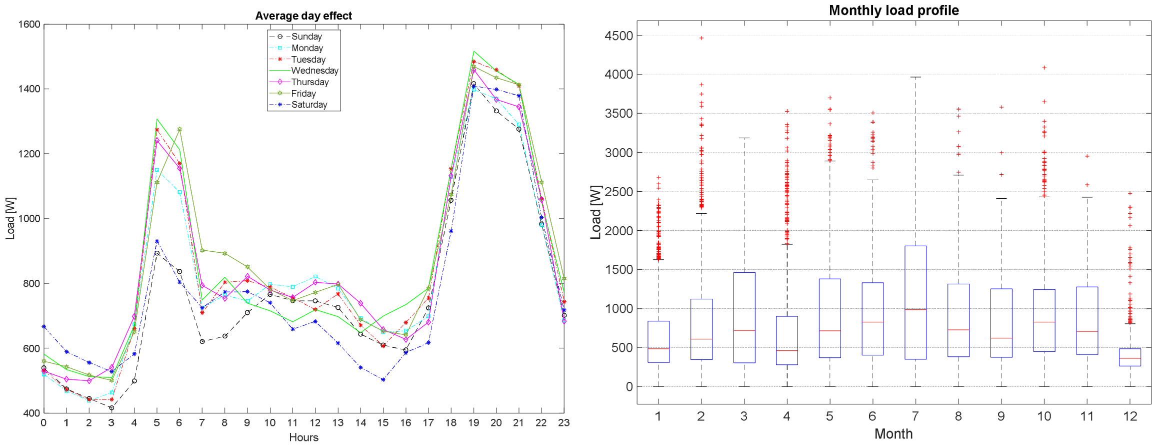

Figure 4 (left) Week day average load, (right) box plots of load values for each month.

Figure 4 (left) is the average hourly load profile line plot for each single day of the week. Power consumption peak hours are observed to be around 05:00 and 18:00 - 21:00 hours. Figure 4 (right) shows hourly load consumption distribution for each month. The central line mark on each box plot is the median value, and the dotted line whiskers are the extreme data points while outliers are plotted using the '+' symbols. The consumption pattern of the box plot in figure 5 (right) is linked to the academic calendar of the school. It had a population of about 500 students and staffs in the period covered by the dataset. Resident students are of four classes. Long School holiday breaks are in June whereby 2 classes of students break for four weeks while the other classes break for one week. All students break for two weeks in April and September. They all break for four weeks in December.

3.5. Feature analysis

Feature selection is a dimensionality reduction technique that ranks and selects an influential subset of the possible predictors or features with the best predictive power of a prediction model. Feature selection is application-oriented [23]. Studies show that the right combination of features is important as the individual features included in prediction model [24]. The 5 feature selection methods used in this work are *Random forest*; *Relieff*; *Ensemble regression tree*; *Compact regression tree*; *Neighborhood component analysis (NCA)*

Random forest (RF) algorithm is a conventional approach in embedded feature selection. In this paper a random forest of 200 bagged ensemble regression trees was grown and used to estimate unbiased feature importance. *Relieff* algorithm works by favouring features that give different values to neighbours of dissimilar response weights while punishing features that give dissimilar weights to neighbours of the same response values [25]. Matlab ReliefF function used in this work was configured to 10 nearest neighbours and regression method for computing weights. The predictor Importance Matlab function was used with *ensemble regression tree* and *compact regression tree* to compute estimates of feature importance. The larger the estimate value the more important the feature. *Neighborhood component analysis* (NCA) is an embedded feature selection method. The *fsrnca* Matlab function was used NCA in this work.

Each of the 5 feature selection models ranked the same features with different importance and weight scores. In order to increase prediction performance, 3 global ranking approaches derived from the 5 FS models are proposed in selection of a subset of k principal features from a superset of N features through: 1) mean score; 2) subset union; 3) refined exhaustive search based on k-combination. A description of the 3 approaches is given in subsequent subsections.

3.5.1 Principal k-features mean score

One approach was to find arbitrary k number of principal features by finding the overall final ranking of the individual features taken as an average score from all the FS methods used. If the total number of features is N, then weight rank will be N, N-1, N-2, …1. Then, the mean score for each feature was evaluated using equation 1:

$$\overline{W_f} = \frac{1}{n}\sum_{i=1}^{n} W_{fi} \tag{1}$$

where n is the total number of FS models used, W_{fi} is the feature weight score.

In this work, for each FS method the most important feature was assigned a weight score of N=23 while the least important feature was assigned a weight score of 1. For example, hour of the day had an estimated importance weight score values of 4.36, 0.02, 9.6, 29.4, 2.3, and 0.27 computed in random forest, relieff, ensemble regression tree, compact regression tree, and NCA features selection respectively. Corresponding values assigned are 21, 20, 15, 20, 7, 19. Resulting in a mean weight score of 17, thus making hour of the day the fourth most important feature. The final weight score of the hour of the day becomes 20 out of 23. Thus, the final rank of a feature is a mean of the individual weighted votes a feature scores from each of the 5 FS algorithms. The overall top 5 most important features from the principal k-features mean score approach were found to be 'Load', 'T1_Load', 'T2_Load', 'Hour', and 'Day'. 'Load' being the first in importance and 'Day' being the fifth in importance.

3.5.2 Principal k-features union

A second approach proposed in this work is to create a features subset comprised of the set union between top k-features from each FS model without redundancy. In this work, principal features were obtained from a union of top 5 most important features from each FS model without redundancy in features selected. The resulting subset had 9 out of the 23 superset features, as follows: 'Load', 'T1_Load', 'Hour', 'T2_Load', 'Day', 'P_DG', 'Month', 'Vdc_bus', 'P_inv'.

3.5.3 Refined k-features exhaustive search

Features recommended by the principal k-features mean score and union can further be minimized to find the best k-features by performing a mathematical k-combination of features given in equation 2. This brute-force approach was refined in this work by only selecting without repetition 5 combination of features that included the 'Load' feature which prior to this step was voted to be the most important feature by principal k-features mean score and union approaches. Computing all the 5 features subset from the 23 features set, resulted into 33,649 combinations, then afterwards choosing only combinations with 'Load' feature inclusive reduced the number of combinations down to 7,315. Taking only combinations which include both 'Load' and 'T1_Load' (previous day load) gives 1,330 combinations. This refined number of combinations can reasonably be run through the 20 conventional load forecasting models used in this work in order to find a good enough regression model.

$$nCk = \frac{n!}{k!(n-k)!} \tag{2}$$

3.6. Forecasting models

Principal k-features from the features selection stage of the proposed framework are used to train and validate 20 different regression models. The regression models used are classified into 5 categories namely: regression tree; Neural network; Gaussian process regression (GPR); Support Vector Machine (SVM); and linear regression. An overview of the forecast models is given below.

Linear regression is a linear approach to prediction modelling. The Matlab implementation used in this work used *least-squares*, *robust*, and *stepwise* fitting methods. Matlab implementation used in this work for SVM analysis is the linear epsilon-insensitive SVM (ε-SVM) regression.

For GPR this work used the *fitrgp* Matlab function to train the dataset. The kernel function options used in this work were: 'exponential' for exponential kernel, abbreviated eGPR; 'squaredexponential' for squared exponential kernel, abbreviated seGPR; 'matern52' for matern kernel with parameter 5/2, abbreviated MaternGPR; and 'rationalquadratic' for rational quadratic kernel, abbreviated rqGPR. This work used a two-layer feed-forward neural network consisting of 10 hidden layers and linear output neurons for regression. The network was trained with Levenberg-Marquardt backpropagation algorithm (trainlm).

Regression trees in this work were implemented with the *fitrtree* Matlab function. Variations of the regression trees used were ‘FineTree’ with ‘MinLeafSize’ of 4; ‘MediumTree’ with ‘MinLeafSize’ of 12; ‘CoarseTree’ with ‘MinLeafSize’ of 36. This work used *fitrensemble* Matlab function with the input method ‘bag’ for bootstrap aggregation (bagging) forming a deep ‘BaggedTree’ prediction model. A shallow ‘BoostedTree’ was formed by employing the method ‘LSBoost’ (Least-Squares Boosting) on the fitrensemble. ‘BoostedTree’ model fits to minimize mean-squared error.

3.7. Model evaluation indexes

To assess the model performance in forecasting, mean absolute error (MAE) and root mean square error (RMSE) was used, equation 3 and 4, respectively. They are given by

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \tag{4}$$

where n is the number of power consumption data points, y_j is the observed power consumption value and $\hat{y}_j$ is the predicted power consumption value.

4. Results and discussions

This section shows the results of the comparative analysis on the impact of day-ahead (24hrs) load profile forecasting with respect to variations in input features and prediction models.

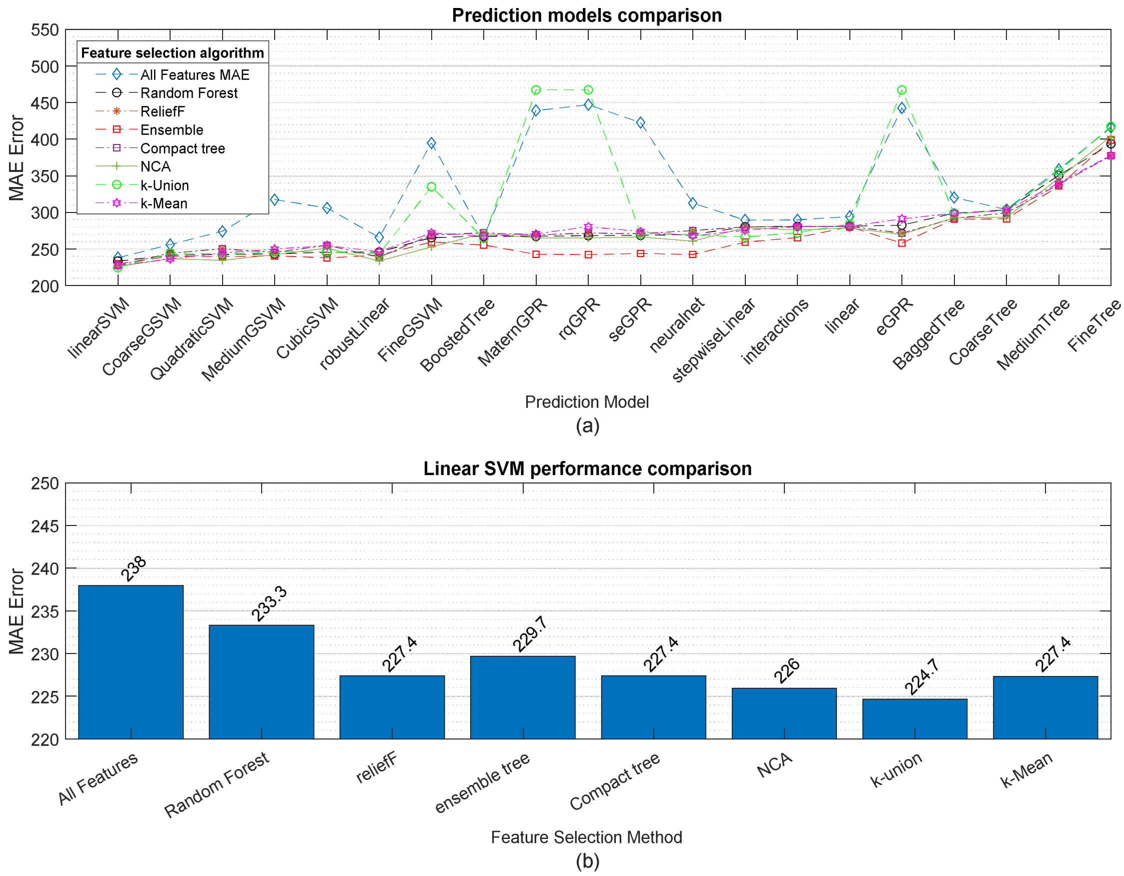

Figure 5: (a) Prediction models evaluation; (b) Linear SVM performance comparison on each of the 8 feature selection methods.

This results in different forecasting strategies namely: 1) load forecast using all the 23 features; 2) load forecast using features from standard FS such as Random Forest, ReliefF algorithm, ensemble regression, compact tree, NCA, 3) proposed FS methods (principal k-features union, principal k-features mean score, refined exhaustive features search); and lastly 4) hybrid load forecast models formed by fusion of two standard regression models through elementwise mean, max, min of the two models outputs.

Figure 5 (a) shows load forecast evaluation using the test dataset on the 20 conventional prediction algorithms. Each of the 8 feature selection methods in Figure 5 (a) was applied on all the 20 conventional prediction models. The principal k-features union approach performed best with 'linearSVM' prediction model, however it performed poorly in the case of: 'FineGSVM'; 'MaternGPR'; 'rqGPR'; and 'eGPR' models. The principal k-features mean score approach performed relatively good across all the 20 prediction models.

Linear SVM prediction algorithm was identified as the best prediction model. For prediction examples, figure 6 shows individual next day load forecast plots for February 3rd to February 6th 2018 using the 'LinearSVM' prediction model. The forecast was done with 95% confidence band. The prediction model is used with the features selected from the principal k-features union approach. Prediction accuracy varies from one day to the next.

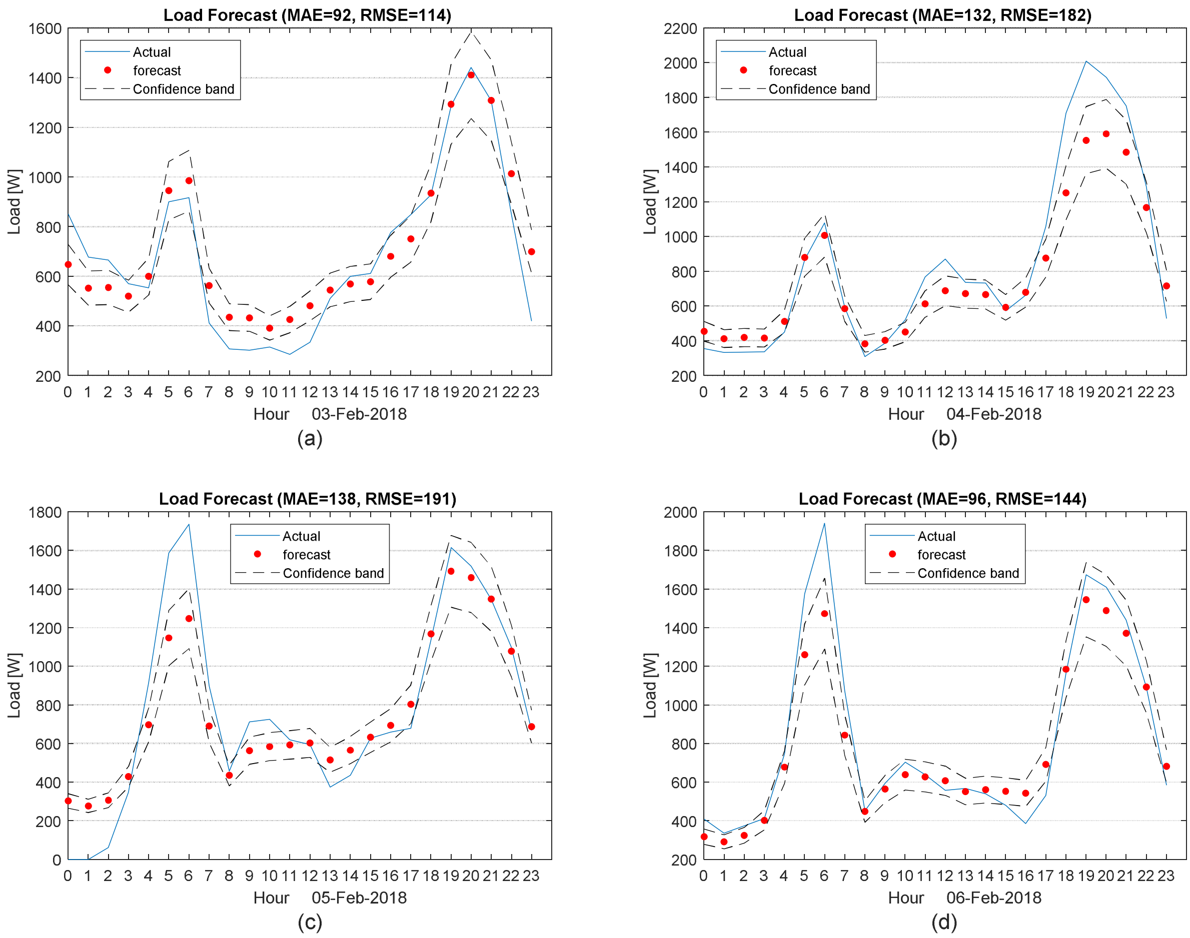

Figure 6: Linear SVM Forecast model performance.

Table 2: Refined exhaustive search load forecast results

	Feature subset	RMSE	MAE
1	'Load,SOC,Vdc_bus,T1_temp,Weekend'	283.3	125.8
2	'Load,RHumidity,SOC,Vdc_bus,temp'	289.6	127.9
3	'Load,SOC,Vdc_bus,T1_temp,DewpointTemp'	288.5	129.1
4	'Load,SOC,Vdc_bus,T2_temp,temp'	289.8	131.5
5	'Load,Vdc_bus,DewpointTemp,atmPressure,Weekend'	296.3	132.3
6	'Load,RHumidity,SOC,Vdc_bus,Weekend'	295.1	132.6
7	'Load,SOC,Vdc_bus,T1_temp,P_HYD'	292.4	133.5
8	'Load,SOC,Vdc_bus,P_HYD,P_DG'	284.7	133.7
9	'Load,SOC,Vdc_bus,temp,P_HYD'	300.2	133.9
10	'Load,SOC,Vdc_bus,atmPressure,Weekend'	295.4	134.0
11	'Load,RHumidity,SOC,Vdc_bus,T1_temp'	300.1	134.2
12	'Load,SOC,Vdc_bus,P_HYD,Weekend'	294.8	134.4
13	'Load,RHumidity,SOC,Vdc_bus,P_HYD'	302.6	134.6
14	'Load,SOC,Vdc_bus,DewpointTemp,Weekend'	290.0	135.4
15	'Load,RHumidity,Vdc_bus,T1_temp,Weekend'	302.4	135.6
16	'Load,Vdc_bus,T1_temp,DewpointTemp,Weekend'	305.6	136.2
17	'Load,SOC,Vdc_bus,WeekDay,Weekend'	298.6	136.7
18	'Load,SOC,Vdc_bus,T1_temp,P_DG'	294.5	136.7
19	'Load,SOC,Vdc_bus,WeekDay,atmPressure'	300.8	137.0
20	'Load,SOC,Vdc_bus,DewpointTemp,P_HYD'	302.1	137.0

Table 2 shows top 20 next day load forecast results obtained using the refined exhaustive search and Linear SVM prediction model. The refined exhaustive search involved choosing 5 features out of all 23 features and using the chosen 5 features on the 'LinearSVM' prediction model. In other words, referring to equation (2), a 5 features subset of a 23 features set was used in predicting day ahead load forecast using 'LinearSVM' prediction model. This resulted in a total of 5989 different combinations of 5 features which at least include the 'Load' variable but exclude the 'Year' variable. The latter was correctly omitted since it has no meaningful influence in day to day short term load forecast because it remains constant from one day to the next.

Figure 7 (a) shows the MAE and RMSE prediction result evaluation when a subset of top 5 most important features selected from mean score votes of the 5 FS conventional models was trained and tested on each of the 20 regression models. Figure 7 (b) shows a general improvement in MAE error performance when hybrid models are formed by fusion of two regression models through elementwise mean, max, min of the two conventional regression models outputs. The resulting hybrid models formed from the maximum forecast instances of two regression models exhibited the lowest MAE prediction error.

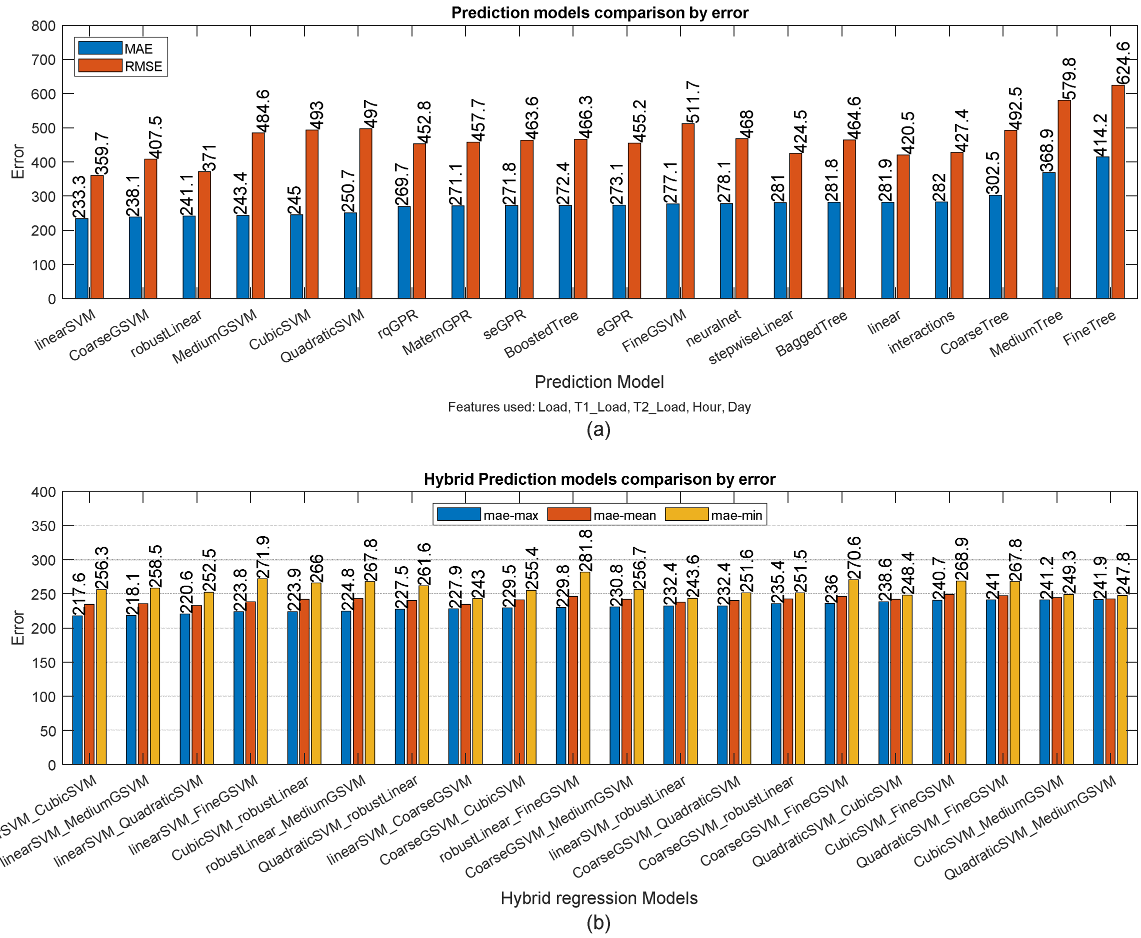

Figure 7: (a) Conventional prediction models evaluation; (b) Hybrid regression models.

Based on the microgrid dataset used in this work, both the k-features mean score and subset union approaches registered lowest error values when they were used with 'linearSVM' prediction model. The principal k-features union approach model registering MAE error of 224.7 while principal k-features mean score approach registered 227.4. The refined exhaustive search used together with 'linearSVM' prediction model registered the lowest MAE error of 125.8, however, it was computational more intensive. Furthermore, a hybrid prediction model formed from the elementwise maximum forecast instances of two regression models yielded better MAE prediction error than the individual regression models fused to form the hybrid. In this case study, the overall best prediction model was found to be the hybrid regression model formed from 'linearSVM' and 'cubicSVM' regression models. Therefore, given a different microgrid it is recommended to find the best features using the proposed principal k-features union approach and in turn use form a hybrid regression model based on top two performing conventional prediction models.

5. Conclusion

In this paper, a one day ahead load forecasting model framework has been proposed. It has been shown that the blending of conventional feature selection methods gives more reliable global subset principal features that improve prediction model performance. Furthermore, blending two conventional regression models forms hybrid regression models with improved prediction performance. Three approaches were proposed for supervised selection of a subset of principal k-features from a superset of N features through: 1) mean score; 2) subset union; 3) refined exhaustive search based on k-combination. For computational time reasons, the mean score and subset union approach can be applied and the best of the two chosen after evaluating their performance on prediction models. If further prediction performance is desired, then refined exhaustive principal k-features search can be applied although it is more resource intensive.

The selected principal k-features were trained on 20 different conventional regression models and their prediction performance efficacy evaluated. The hybrid regression model formed from fusion of the best 2 models ('linearSVM' and 'cubicSVM') out of the 20 conventional models showed improved prediction performance than the individual regression models (MAE reduced by 5.4%). Ultimately, the choice of which path to follow in the proposed framework depends on the computational resources available as well as the degree of prediction performance desired. Other bio-inspired FS methods and prediction models could be explored to study their performance in load forecasting.

Acknowledgment

Authors gratefully acknowledge the support of The Nelson Mandela African Institution of Science and Technology, Politecnico di Milano department of energy, Arusha Technical College, Dar es salaam Institute of Technology, Ngarenanyuki secondary school, MCM energy lab - spin off Politecnico di Milano, OIKOS, ELVI Group, Phoenix contact, and Energy Team©.

References

[1] National Bureau of Statistics, "2012 Population and housing census; Population Distribution by Adminstrative Areas," *Natl. Bur. Stat.*, 2013.

[2] International Energy Agency, "World energy outlook," *Paris: International Energy Agency*. 2014.

[3] I. D. Serna-Suarez, G. Ordonez-Plata, and G. Carrillo-Caicedo, "Microgrid's Energy Management Systems : A Survey Microgrid's Energy Management Systems : A Survey," no. MAY, pp. 1–6, 2015.

[4] F. Martínez-Álvarez, A. Troncoso, G. Asencio-Cortés, and J. Riquelme, "A Survey on Data Mining Techniques Applied to Electricity-Related Time Series Forecasting," *Energies*, vol. 8, no. 12, pp. 13162–13193, 2015.

[5] P. Su *et al.*, "Recent Trends in Load Forecasting Technology for the Operation Optimization of Distributed Energy System," *Energies*, vol. 10, no. 9, p. 1303, 2017.

[6] M. Alkhathami, "Introduction to Electric Load Forecasting Methods," vol. 2, no. 1, pp. 1–12, 2015.

[7] H. Shi, M. Xu, and R. Li, "Deep Learning for Household Load Forecasting – A Novel Pooling Deep RNN," *IEEE Trans. Smart Grid*, vol. 3053, no. c, pp. 1–1, 2017.

[8] S. Takiyar and V. Singh, "Trend analysis and evolution of Short Term Load Forecasting Techniques," *2015 4th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir.*, pp. 1–6, 2015.

[9] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and DSM in smart grid: A review," *Renew. Sustain. Energy Rev.*, vol. 54, pp. 1311–1322, 2016.

[10] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network," *Energies*, 2018.

[11] M. S. Carmeli *et al.*, "Hybrid micro-grid experimental application in Tanzania," *Clean Electrical Power (ICCEP), 2015 International Conference on*. pp. 534–541, 2015.

[12] M. S. Carmeli *et al.*, "Hybrid distributed generation system for a rural village in Africa," *3rd Renew. Power Gener. Conf. (RPG 2014)*, pp. 8.41-8.41, 2014.

[13] S. Mandelli, M. Molinas, E. Park, M. Leonardi, E. Colombo, and M. Merlo, "The Role of Storage in Emerging Country Scenarios," *Energy Procedia*, vol. 73, pp. 112–123, 2015.

[14] E4G, "Energy4Growing," 2019. [Online]. Available: https://web.facebook.com/energy4growing2014/. [Accessed: 14-Jun-2019].

[15] S. Mandelli, M. Merlo, E. Tedeschi, and M. Molinas, "Electro-Mechanical Model for Understanding the Operation and Dynamic Behavior of a Micro-Grid : a case study in Tanzania," 2015.

[16] E. A. Nyari, M. Castro-sitiriche, and S. E. Park, "An Experimental Study of Electrical Appliances Consumption Using Panel Meter Data," pp. 132–144, 2017.

[17] S. Mandelli, C. Brivio, M. Moncecchi, F. Riva, G. Bonamini, and M. Merlo, "Novel LoadProGen procedure for micro-grid design in emerging country scenarios: application to energy storage sizing," *Energy Procedia*, vol. 135, no. February, pp. 367–378, 2017.

[18] M. Mauri, M. S. Carmeli, M. Merlo, C. Brivio, and B. Mbuya, "Neural Network based Load Forecasting and Fuzzy Logic EMS for Ngarenanyuki School microgrid," *Int. Symp. Power Electron. Electr. Drives Autom. Motion*, 2016.

[19] E4G, "Ngarenanyuki Microgrid data," 2018. [Online]. Available: www.e4g.polimi.it.

[20] Raspisaniye Pogodi Ltd, "Weather archive in Kilimanjaro (airport), METAR." [Online]. Available: http://rp5.md/Weather_archive_in_Kilimanjaro_(airport),_METAR. [Accessed: 23-Apr-2019].

[21] M. de Oliveira, N. Araujo, R. da Silva, T. da Silva, and J. Epaarachchi, "Use of savitzky--golay filter for performances improvement of SHM systems based on neural networks and distributed PZT sensors," *Sensors*, vol. 18, no. 1, p. 152, 2018.

[22] H. Kaneko, T. Matsumoto, S. Ootakara, and K. Funatsu, “Practical Use of Savitzky-Golay Filtering-Based Ensemble Online SVR,” *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 371–376, 2016.

[23] A. Jovic, K. Brkic, and N. Bogunovic, “A review of feature selection methods with applications,” no. May, pp. 25–29, 2015.

[24] J. Li, K. Cheng, S. Wang, and F. Morstatter, “Feature Selection : A Data Perspective,” vol. 50, no. 6, 2017.

[25] I. The MathWorks, “Rank importance of predictors using ReliefF or RReliefF algorithm - MATLAB relieff.” [Online]. Available: https://www.mathworks.com/help/stats/relieff.html. [Accessed: 16-Apr-2019].