Distribution Transformers Short-Term Load Forecasting Models

Styliani I. Kampezidou, Student Member, *IEEE* and Santiago Grijalva, Senior Member, *IEEE*

***Abstract*—The increasing penetration of variable energy sources, plug-in electric vehicles (PEV) and storage in the distribution system generates the need for load forecasting at a more granular level than at the substations. Load forecasting at the distribution transformer level can provide an estimate of how the load is distributed in the distribution network rather than a substation total estimate. It opens possibilities for various applications and use cases including better demand-response management, resource scheduling, enhanced losses minimization, and more accurate management of distribution transformer loading. In order to achieve higher accuracy of transformer loading forecasting, distribution transformer meters or equivalent aggregation of the transformer served load based on smart meters is required. This paper presents two linear models for computationally efficient short-term load forecasting on distribution transformers. The first model uses an aggregate load and the second uses an average load approach. The average load method exhibits prediction accuracy superior to the aggregate load approach. The proposed models are tested on smart grid community data provided by Pecan Street.**

***Index Terms*—Short-Term Load Forecasting, Distribution Transformer Load Forecasting, Linear Regression Models.**

I. Introduction

LOAD forecasting is a critical function and activity in power system operations and planning for both transmission and distribution. Utilities need accurate load forecasting models for a variety of business processes including generation and transmission planning, distribution planning, energy procurement, real-time operation and dispatch, demand side management (DSM) and demand response, and financial planning [10]. From regulated utilities to deregulated environments such as [4], load prediction models are utilized every day.

The increasing deployment of distributed energy resources, electric vehicles and demand response programs pose important challenges for distribution system operations. Reliable operation of emerging distribution systems requires more advanced load forecasting. In particular, load forecasting in smaller scales, such as distribution transformers, microgrids and the customer becomes more highly desirable. Deployment of advanced metering infrastructure (AMI) has resulted in opportunities to collect load data from individual homes and commercial buildings, and to determine aggregates of those load quantities at the distribution transformer level, and feeder level, for a variety of applications.

In this paper we propose two computationally efficient models for distribution load forecasting and develop a comparison of them to illustrate their advantages. The rest of this paper is structured as follows. Section II presents classification of methods for load forecasting and the state of the art in linear methods. In section III, we propose our disaggregated approach. In section IV, analysis and results are illustrated while in section V a conclusion is provided.

II. Background

A. Classification of Load Forecasting Methods

Load forecasting methods can be grouped in various ways. According to the prediction horizon, [10] proposes four categories:

1) Very-short-term (1 day ahead, update: few minutes),
2) Short-term (2 weeks ahead, update every 1 day),
3) Medium-term (3 years ahead, update every 1 month)
4) Long-term (30 years ahead, update every 1 year).

The load forecasting models can also be classified according to the type of load they predict: base load models or peak load models. Finally, various human activities, calendar events, meteorological data, spatial information or price variations [13] can modify the purpose of the model and classify it accordingly.

Regarding the mathematical model use in load forecasting, the problem can be classified in two ways: 1) statistical methods, such as regression analysis [14] and time series analysis [15] and 2) artificial intelligence methods, such as artificial neural networks (ANN) [2, 16] , fuzzy logic [11], and support vector machines (SVM) [3,1717]. Combinations of those methods have also been proposed [12].

B. Aggregate and Average Load Forecasting

Aggregate load and average load forecasting methods have both been proposed in the past using one of the two approaches mentioned before. See [7], [8] and [25]. The aggregation of a daily load of a number of users has the advantage of smoothing the spikes in demand caused by the consumption profile of different devices, as explained in [5]. An example of such a device is the PEV (random-high spikes).

C. Advantages of Linear Modelling Techniques

Multiple regression analysis has several advantages including modelling of different parameters that affect the load, i.e. holidays can be modeled as binary variables, different days of the week can have a different error variance assigned (heteroskedasticity) and reducing the effect of outliers, especially in the case of large load forecast errors, as mentioned in [14]. In this paper we use linear regression models because they are the simplest to develop and the easiest to fully interpret. However, any attempt of predicting a single user's load with a linear regression model (i.e. using panel data structures in Mixed Linear Models proposed in statistics publications like [22]) would lack accuracy because the model would be highly driven by the spikes in a single

978-1-5090-4168-8/16/$31.00 ©2016 IEEE

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on November 04,2022 at 21:43:48 UTC from IEEE Xplore. Restrictions apply.

user's consumption rather than the baseline load. Thus, we present an aggregate and an average linear load forecasting model applied to distribution transformers.

III. Proposed Method

A. *Disaggregated Approach*

The increasing penetration of variable sources (mostly PV) and storage (PEV or batteries for PV) in the distribution system requires distribution transformer load forecasting models to enable energy management [28]. Moreover, congestion issues that might arise in heavily loaded elements or subnetworks can be in general improved with such a forecast model. Power losses that occur by transferring the power away from the generation point for consumption can also be reduced. Overloading of distribution transformers can be avoided. Optimal PEV charging periods could be planned.

Recent literature has made limited attempts to integrate frameworks for more disaggregated load forecasting methods to address some of the previous issues and benefit from the advantages of distributed generation. [26] describes the need for separate short-term forecast models in geographically distributed loads but only focuses on the substation level. It also proposes an optimal region selection technique based on load and weather forecast variation. [27] proposes learning-based distributed load forecasting on subnetworks of the system which are, however, not explicitly defined. In this paper, we fully define a distinct way to forecast disaggregated load. This is performed on distribution transformers, as shown in Fig. 1.

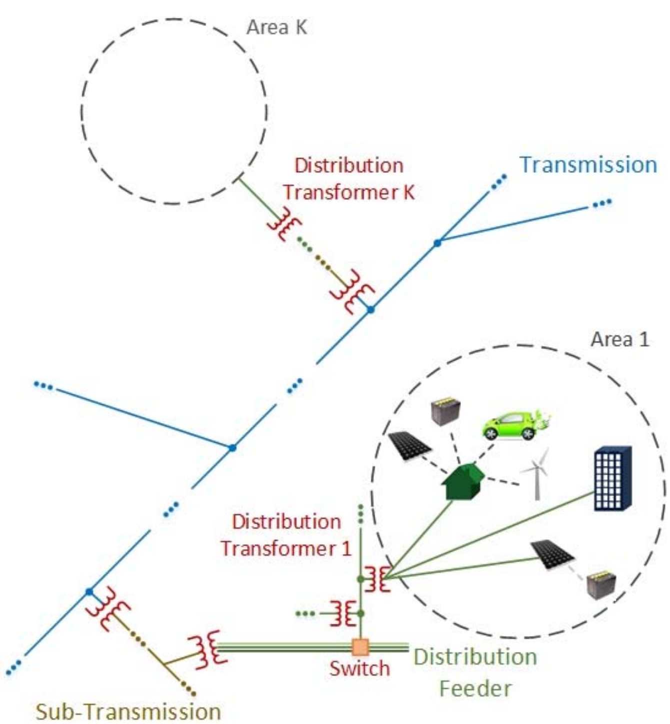

Fig 1. Distribution transformer load forecasting

We use two linear, short-term load forecasting models for the Customer Baseline Load on aggregate and on average distribution transformers data. Only total consumption values were used because most installed smart meters in homes only provide total consumption data, rather than data from individual appliances. Even if there are smart meters installed on appliances, we consciously decide to discard these signals from this study to ensure real-time computational tractability. The authors in [1], propose an approach to model each household device by using the aggregate power signal and Hidden Markov Models (EDHMM-Diff). These models can be used for smart house applications on the house level. However, when forecasting load on a network bus, detailed models of all the appliances of all the houses connected to that node would convert the problem to a prohibitively very complex one. Thus, appliance models are not used here.

This study takes into consideration only weather data and the total houses area. Weather data are easily accessible and very accurate for the day ahead prediction (temperature and humidity). This study does not take into account price data because there is no unified price environment in the USA. [21] proposes a two-step, substation level, linear regression model with adaptation based on weather conditions but does not take into account the total area of the house while ours does. [18] provides a study for the STLF of the peak load based only on weather data. Others, like [20] use ANN to model the nonlinear relationship between the weather variables and the load. However, ANN are difficult to interpret (cannot define confidence intervals) and more complex computationally. In [19] the nonlinearity between the weather variables and the load is assessed with a nonlinear transformation for the purpose of peak STLF. Last, [21] gives a nonparametric approach using probability density functions that does not require weather data.

B. *Model Implementation*

The methodology proposed to generate the two STLF models and extract the forecast is summarized in Fig. 2 below:

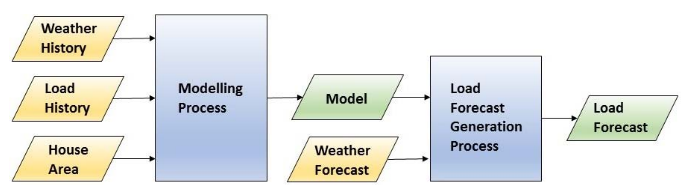

Fig 2. Modelling process and forecast extraction

The modelling process is explained in Fig. 3. Initially all predictors are standardized individually to be fairly treated by the AIC later, during model selection (Akaike's Information Criterion penalizes for large variance and large values). For example, temperature and temperature square are both standardized instead of standardizing the temperature and then squaring it. This standardization technique is proposed in [23] to reduce the multicollinearity between the predictors. In addition, the consumption data was log-transformed to become more linear and comply with the linearity assumption of linear regression theory. Subsampling without replacement is then applied on the standardized raw data to generate the training set. For the aggregate model we have the following: T is the size of the training set (in our case 60 houses), k is the number of random sets generated by the random subsampling

978-1-5090-4168-8/16/$31.00 ©2016 IEEE

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on November 04,2022 at 21:43:48 UTC from IEEE Xplore. Restrictions apply.

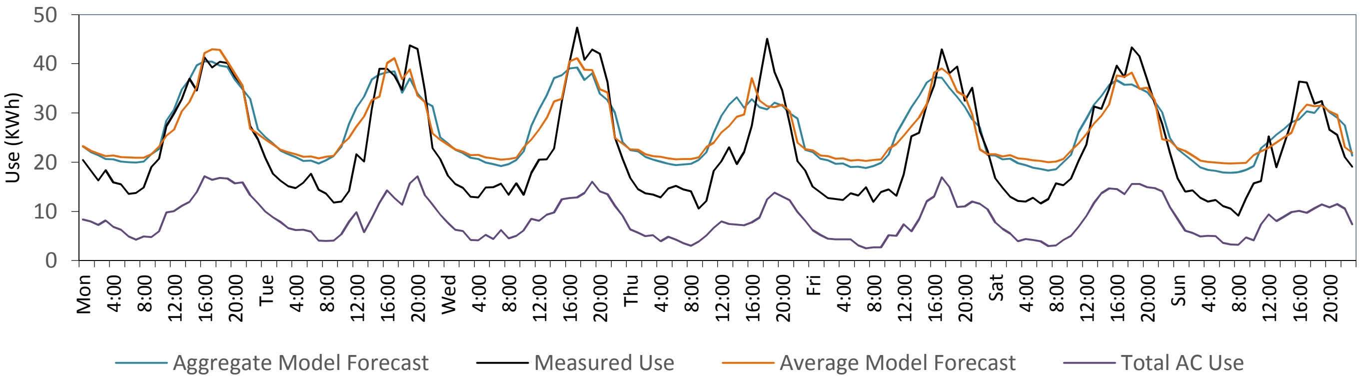

Fig 4. Use forecast of a distribution transformer

(in our case k=50 was a good number to achieve statistical convergence) and n is the number of houses in each randomly generated set (in our case n=12). Each random set of 12 houses of the distribution transformer was used to generate a linear model according to:

$$\widehat{Y^k} = \widehat{\beta_0^k} + \sum_{i=1}^{N} \widehat{\beta_i}^k X_i^k \quad (1)$$

where N: number of predictors used. In the case of the average model: T=99 homes (initial full set), k=100 (number of training nodes randomly generated from the 99 homes) and n=12 (houses in each randomly generated node of the training set). An (1) linear model was used for average modelling.

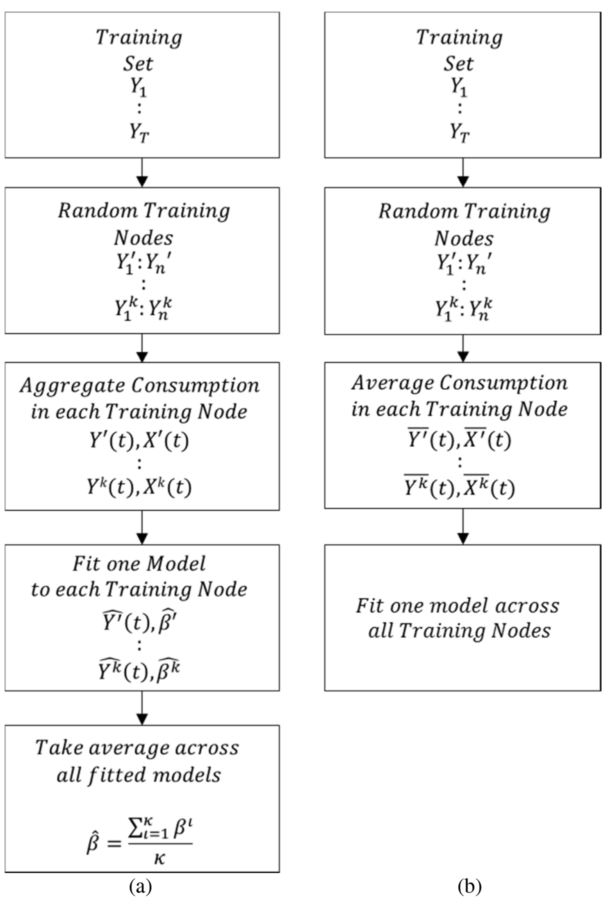

Fig 3. Model Extraction: (a) Aggregate model, (b) Average model

IV. Analysis & Results

A. Explanatory Data Analysis

The concise selection of an initial data set is crucial for the statistical analysis because an excessive number of predictors and data points in linear regression can lead to an overfitted model with poor prediction ability. The initial raw data set contains hourly total consumption data of 99 houses, temperature, and humidity values for the year 2014. In a first explanatory data analysis, different categorical and numerical variables were tested for statistical significance: the total area of all the houses connected to the same distribution transformer, multiple powers of temperature and humidity, past values of temperature and humidity, months of the year and time of the day. In the two different models proposed, some of these predictors were proved significant (using p-value of t-test) and others not. The coefficients for one of the distribution transformers are given in Table I. The time of the day was grouped in four different groups (morning, midday, evening and night) and represented by categorical variables to reduce the number of variables in the problem, as shown in Table I.

B. Variable Selection

1) Aggregate Model

The AIC criterion, provided the significant predictors for each fitted model. We decided to keep the predictors that appeared more than 80% times in the total of 50 fitted models. The coefficient values of the final model are shown in table I.

2) Average Model

AIC was used to choose the categorical variables of the model and LASSO was used to choose the numerical variables of the model. In LASSO method, we calculated the optimal shrinkage parameter by choosing the fraction that allows for the optimal subset through the Cross Validation. The optimal parameter s was automatically chosen by LARS package (R software) to be equal to 1, as an attempt to minimize the least squares. This means that it failed to minimize the least squares before it became s=1 and thus no predictor was excluded from the model. The shrinkage selection procedure is described in literature in [24]. As more variables were added to the model we calculated AIC and multiple R squared until multiple R squared slightly changed. We defined a threshold for the procedure to be terminated. When the next multiple R squared

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on November 04,2022 at 21:43:48 UTC from IEEE Xplore. Restrictions apply.

differed by less than 0.0001 from the previous multiple R squared we terminated the procedure and extracted the corresponding s which was 0.68. No predictor was excluded at this value of s from LASSO. However, to find the coefficients of categorical and numerical variables, linear regression was run because the coefficients calculated from LASSO are not as accurate as the ones from linear regression. The average model coefficients are shown in table I.

C. *Model Validation*

To validate the performance of the models, out-of-sample testing is performed. The results are shown in Fig. 4. We observe that the average model has less yearly MAPE. The errors are shown in Table II.

TABLE I
MODEL COEFFICIENTS FOR ONE DISTRIBUTION TRANSFORMER

Predictor	**Aggregate Model Coefficient Value**	**Average Model Coefficient Value**
Intercept	-7.582	-0.496
Temperature T	2.776	0.195
Lagged T(t-1)	-4.125	-0.385
Lagged T(t-2)	4.859	0.437
T^2	4.641	0.235
T^3	1.941	0.015
Humidity	0.908	0.100
Lagged H (t-1)	-	-0.024
Lagged H (t-2)	-1.108	-0.081
H^2	-0.625	-0.042
Aggregate/Average Area	0.001	0.0004
Midday (11am - 6pm)	2.199	0.120
Evening (7pm - 10pm)	4.601	0.328
Night (11pm - 4am)	0.017	-0.011
January	1.293	0.099
February	0.274	0.013
March	0.245	-0.023
May	0.879	0.115
June	2.746	0.307
July	2.732	0.335
August	3.707	0.407
September	3.131	0.343
October	1.668	0.201
November	0.587	0.036
December	1.777	0.124

TABLE II
FORECAST ERRORS FOR THE SAME DISTRIBUTION TRANSFORMER

Forecast Model	MAPE	RMSE
Aggregate	0.31	0.52
Average	0.22	0.51

The total area of the house, is a statistically significant predictor mainly because it is proportional to the house volume and thus proportional to the air-conditioning consumption levels. Air-conditioning is one of the devices that form the CBL profile as shown in Fig 4.

All the statistical assumptions of linear regression theory are satisfied, verifying the correctness of the modelling process. The predictors of are statistically independent. The log-transformation has cleared most of the nonlinearity between the consumption and the predictors (especially temperature). The residuals follow a normal distribution and their variance is constant. The previous assumptions are validated through residual analysis in Fig. 5 and Fig. 6.

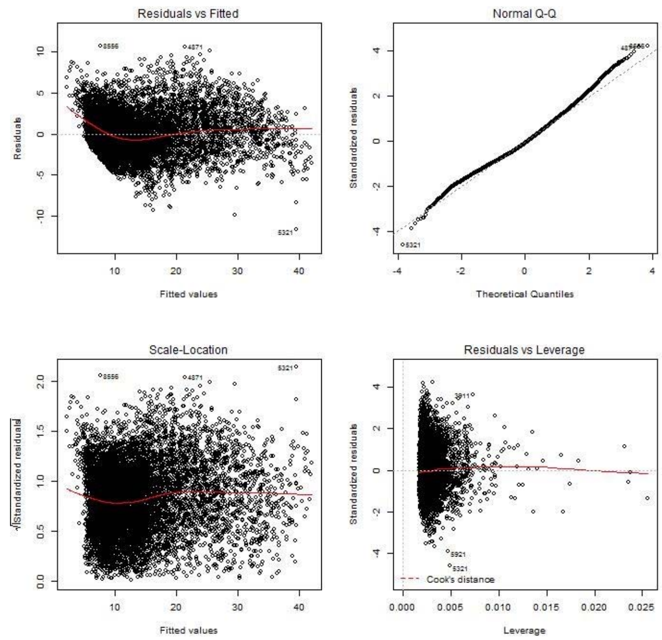

Fig 5. Aggregate model residual analysis

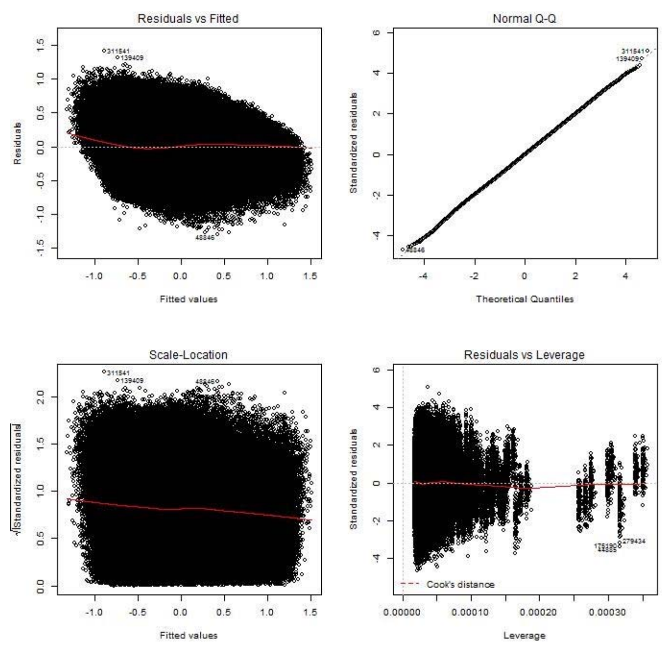

Fig 6. Average model residual analysis

V. CONCLUSION

The distribution transformer load forecast in this paper is performed based on aggregate or average data. A safe comparison of the two methods is offered because the two methods were tested on the same data set and training nodes. For comparison purposes, the output of the average model was scaled (multiplied by the number of customers connected to this distribution transformer).

The size of the prediction error is justified by the choice of a linear model representation and by the lack of weather independent variables except for the total house area (for simplicity and computational tractability). The CBL is affected by the air-conditioning devices but not only. The

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on November 04,2022 at 21:43:48 UTC from IEEE Xplore. Restrictions apply.

choice of the predictors is able to capture the load profile of other devices as well. Special events or holidays were not taken into consideration which also contributed towards an imperfect prediction. The only term which partially captures other factors, is the intercept. However, the forecast errors are low. Other models in literature with similar errors have been proposed for substation level forecast which has smoother load profiles and thus more accurate prediction.

The linear models proposed, satisfy linear regression theory assumptions and can be used to predict the CBL. On top of these, other models, mentioned in the introduction, can be used to predict the load spikes and profile individualities. These peak values would then be added to the output of the models proposed.

VI. References

[1] Z. Guo, Z. J. Wang, and A. Kashani, "Home Appliance Load Modelling From Aggregated Smart Meter Data", *IEEE Transactions on Power Systems*, vol. *30,* no. *1,* pp. *254-262,* Jan., 2015.

[2] D. C. Park, M. A. El-Sharkawi, R.J. Marks II, L.E. Atlas and M.J. Damborg,"Electric Load Forecasting Using An Artificial Neural Network", *IEEE Transactions on Power Systems*, vol. *6,* no. *2,* pp*. 442-449,* May, 1991.

[3] E. Ceperic, V. Ceperic, and A. Baric, "A Strategy for Short-Term Load Forecasting by Support Vector Regression Machines", *IEEE Transactions on Power Systems*, vol. *28,* no. *4*, pp. *4356-4364*, Nov., 2013.

[4] T. Hubert and S. Grijalva (2011, January): Realizing Smart Grid Benefits Requires Energy Optimization Algorithms at Residential Level. Presented at Innovative Smart Grid Technologies (ISGT).

[5] R. A. Sevlian and R. Rajagopal (2014, July): A Model for the Effect of Aggregation on Short Term Load Forecasting. Presented at Power and Energy Society General Meeting.

[6] F. L. Quilumba, W. Lee, H. Huang, D. Y. Wang and R. L. Szabados, "Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities", *IEEE Transactions on Smart Grid, vol. 6, no. 2, pp. 911-918, Mar. 2015.*

[7] *R. M. Salgado, R. Ballini and T. Ohishi (2009, March): An Aggregate Model Applied to the Short-Term Bus Load Forecasting Problem. Presented at Power Systems Conference and Exposition.*

[8] A. Jain and B. Satish (2009, July): Short Term Load Forecasting by Clustering Technique based on Daily Average and Peak Loads. Presented at Power and Energy Society General Meeting.

[9] J. Xie, T. Hong and J. Stroud, "Long-Term Load Forecasting with Consideration of Residential Customer Attrition", *IEEE Transactions on Smart Grid*, issue *99,* Jan., 2015.

[10] T. Hong, "Short Term Electric Load Forecasting", Ph.D. dissertation, EE, NCSU, Raleigh, NC, 2010.

[11] D. K. Ranaweera, N. F. Hubele, and G. G. Nikolic, "Weather Sensitive Method for Short Term Load Forecasting in Electric Power Utility of Serbia", *IEEE Transactions on Power Systems*, vol. *18*, pp. *1581-1586*, 2003.

[12] R. Gao, L. Zhang and X. Liu (2012, July): "A Novel Hybrid Method for Short Term Load Forecasting Using Fuzzy Logic and Particle Swarm Optimization". Presented at Power System Technology, 2010 International Conference on.

[13] A. Khotanzad, R. Afkhami-Rohani and, D. Maratukulam, "ANNSTLF-Artificial Neural Network Short-Term Load Forecaster Generation Three", *IEEE Transactions on Power Systems,* vol. *13*, pp. *1413-1422*, 1998.

[14] A. D. Papalexopoulos and T. C. Hesterberg, "A Regression-based Approach to Short-Term System Load Forecasting", *IEEE Transactions on Power Systems*, vol. *5*, pp. *1535-1547*, 1990.

[15] M. T. Hagan and S. M. Behr, "The Time Series Approach to Short Term Load Forecasting", *IEEE Transactions on Power Systems*, vol. *2*, pp. *785-791*, 1987.

[16] H. S. Hippert, C. E. Pedeira, and R. C. Souza, "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation", *IEEE Transactions on Power Systems*, vol. *16*, pp. *44-55*, 2001.

[17] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load Forecasting using Support Vector Machines: a Study on EUNITE Competition 2001", *IEEE Transactions on Power Systems*, vol. *19*, pp. *1821-1830*, 2004.

[18] O. Hyde and P. F. Hodnett, "An Adaptable Automated Procedure for Short-term Electricity Load Forecasting", *IEEE Transactions on Power Systems*, vol. *12*, pp. *84-94*, 1997.

[19] T. Haida and S. Muto, "Regresion Based Peak Load Forecasting Using a Transformation Technique", *IEEE Transactions on Power Systems*, vol. *9*, pp. *1788-1794*, 1994.

[20] K. Methaprayoon, W.-J. Lee, S. Rasmiddatta, J. R. Liao, and R. J. Ross, "Multistage Artificial Neural Network Short-Term Load Forecasting Engine with Front-End Weather Forecast", *IEEE Transactions on Power Systems*, vol. *43*, no. *6*, pp. *1410-1416*, 2007.

[21] W. Charytoniuk, M. S. Chen, P. Kotas, and P. Van Olinda, "Demand Forecasting n Power Distribution Systems Using Nonparametric Probability Density Estimation", *IEEE Transactions on Power Systems*, vol. *14*, pp. *1200-1206*, 1999.

[22] B. Winter (2014, May): A Very Basic Tutorial for Performing Linear Mixed Effects Analyses (Tutorial 2). University of California. Merced, CA, USA. [Online]. Available: http://www.bodowinter.com/tutorial/bw_LME_tutorial2.pdf

[23] D.-S. Kim and S. Dong-Ku: A Standardization Technique to Reduce the Problem of Multicollinearity in Polynomial Regression Analysis. Hanyang University, Seoul, Korea. [Online]. Available: https://www.stat.fi/isi99/proceedings/arkisto/varasto/kim_0574.pdf

[24] R. Tibshrirani, "Regression Shrinkage and Selection via the LASSO", *Journal of the Royal Society, Series B (Methodological)*, vol. *58*, issue *1,* pp. *267-288*, 1996.

[25] S. A. Azad, A. S. Ali, and P. Wolfs (2014, November): Daily Average Load Forecasting Using Dynamic Linear Regression. Presented at Computer Science and Engineerings, 2014 Asia-Pacific World Congress on.

[26] S. Fan, Y.-K. Wu, W.-J. Lee and C.-Y. Lee (2011, July): Comparative Study on Load Forecasting Technologies for Different Geographical Distributed Loads. Presented at Power and Energy Society General Meeting.

[27] K. Kalbat and A. Tajer (2013, Decemeber): Learning-based Distributed Load Forecasting in Energy Grids. Presented at Signal and Information Processing (GlobalSIP), IEEE Global Conference on.

[28] K. Kouzelis, B. Bak-Jensen, P. Mahat and J. R. Pillai (2014, October): A Simplified Short Term Load Forecasting Method Based on Sequential Patterns. Presented at Innovative Smart Grid Technologies Europe (ISGT Europe).

VII. Biographies

Styliani I. Kampezidou was born in Athens, Greece. She received a diploma in Electrical and Computer Engineering from the University of Patras, Greece in 2013. She is currently pursuing her PhD at the School of Electrical and Computer Engineering in Georgia Institute of Technology, Atlanta, GA, USA. Her research interests include power systems automation, operation, control and optimization.

Santiago Grijalva is Georgia Power Distinguished Professor of Electrical and Computer Engineering and Director of the Advanced Computational Electricity Systems (ACES) Laboratory at The Georgia Institute of Technology. His research interest is on decentralized power system control, power system informatics and economics, and future sustainable energy systems. He is the principal investigators for various research projects under Department of Energy, ARPA-E, EPRI, PSERC and other several industry and Government sponsors. From 2002 to 2009 he was with PowerWorld Corporation as a software architect and consultant. From 2013 to 2014 he was on assignment to the National Renewable Energy Laboratory (NREL) as founding Director of the Power System Engineering Center (PSEC). Dr. Grijalva's graduate degrees in Electrical and Computer Engineering, M.Sc. (99), Ph.D. (02) are from the University of Illinois at Urbana-Champaign.

978-1-5090-4168-8/16/$31.00 ©2016 IEEE

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on November 04,2022 at 21:43:48 UTC from IEEE Xplore. Restrictions apply.