Very Short-Term Electrical Energy Consumption Forecasting of a Household for the Integration of Smart Grids

Kasım Zor, Adana Science and Technology University, Turkey
Oğuzhan Timur, Çukurova University, Turkey
Özgür Çelik, Adana Science and Technology University, Turkey
Hatice Başak Yıldırım, Adana Science and Technology University, Turkey
Ahmet Teke, Çukurova University, Turkey

The European Conference on Sustainability, Energy & the Environment 2018
Official Conference Proceedings

Abstract
The recent integration of smart grid systems to present electric power systems and the increasing penetration of renewable energy sources make electrical energy consumption forecasting not only a prominent subject but also an arduous challenge due to nonlinear and nonstationary characteristics of electric loads which can be affected by seasonal effects, weather conditions, socioeconomic dynamics, and random effects. Very short-term electrical energy consumption forecasting (VSTCF), which includes few minutes to an hour ahead forecasting of electrical energy consumption, ensures monitoring energy consumption, identifying base and peak loads, making feasible decisions for renewable energy investments such as photovoltaic (PV) systems, and improving energy management quality of a household for the smart grid integration. In this paper, for the first time in Turkey, electrical energy consumption data of a household with an averaging period of 10-minute is obtained by an energy logger during a 1-month period in order to perform VSTCF by using several artificial intelligence (AI) techniques including decision trees (DT), genetic algorithm (GA), artificial neural networks (ANN), and support vector machines (SVM) in the literature. After data pre-processing, various AI techniques will be applied to real-time data obtained from a household in the Mediterranean Region of Turkey for the calculation of mean absolute error (MAE) performance metric. Results indicate that gradient boosted decision trees (GBDT) have the best performance in comparison with other techniques for VSTCF.

Keywords: Very-short term, energy forecasting, household, smart grid integration, artificial intelligence, decision trees, genetic algorithm, artificial neural networks, support vector machines, mean absolute error.

Introduction

After the deregulation of electric power system, distributed generation of electricity has become more important due to onsite generation and efficient use of electrical energy in a small environment. Increasing share of renewable energy technologies among today's power plants and current integration of smart grid systems to modern day's electric power systems not only make energy forecasting a popular subject in energetics, but also categorise it as a demanding challenge with highly unpredictability because of the influencing factors such as social activities, climate and seasonal factors.

According to time period, energy forecasting can be classified as shown in Figure 1. VSTCF (or ultra-short term electrical energy consumption forecasting) includes between 1-minute and 1-hour ahead forecasts, while short-term electrical energy consumption forecasting (STCF) contains among 1-hour and 2 weeks ahead forecasts. Medium-term (or mid-term) electrical energy consumption forecasting (MTCF) refers to future predictions from 2 weeks to 3 years and long-term electrical energy consumption forecasting (LTCF) is performed for forecasts from 3 years up to 50 years (Zor et al., 2017b).

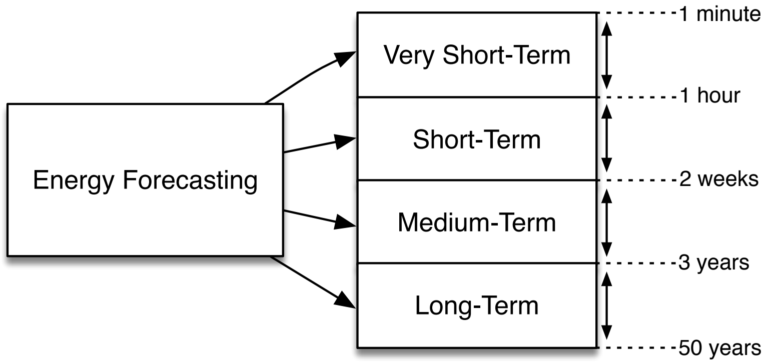

Figure 1: Energy forecasting classification according to time period.

Although several techniques have been developed for use in STCF, the existing literature related to VSTCF is notably numbered. In the literature, VSTCF is commonly employed for smart grid and automated demand response applications. Perpetual developments in advanced metering infrastructure (AMI) system and smart meter provide obtaining electrical energy consumption data from individual households instantly by initialising bi-directional communication between electricity distribution companies and individual households. This results in accelerating personalised auto demand response applications in individual households, which leads to customised contracts and rates, such as a dynamic rate and bi-directional transaction bidding, and causes effective deployment of electricity (Hsiao, 2015).

In this paper, electrical energy consumption data of a household with an averaging period of 10-minute is obtained for the first time in Turkey by an energy logger during a 1-month period in order to perform VSTCF by employing various AI techniques including DT, GA, ANN, and SVM. After the introduction section, the recent literature, material and methods containing household properties, data acquisition, and data set information, evaluation criterion, discussion and results, and conclusions are explained respectively.

Literature Review

At the beginning of the VSTCF literature, Liu et al. made a comparison of VSTCF techniques named as fuzzy logic (FL), neural networks (NN), and auto-regressive model (AR) for an automatic generation control (AGC) system in a multi-area interconnected power system to match area generation to area load, to regulate system frequency and area net interchange to their scheduled values, and to distribute area generation economically among available resources (Liu et al., 1996). Feng et al. proposed a method for VSTCF based on ANN in order to address problems and solutions related to forecasting in a lead time of 10 minutes (Feng et al., 1997).

Charytoniuk and Chen presented a novel approach that leads to a better accuracy for VSTCF by the application ANN to model load dynamics (Charytoniuk and Chen, 2000). Shamsollahi et al. developed and implemented an ANN based VSTCF model for the interim electricity market of ISO New England (Shamsollahi et al., 2001). Chen et al. reported upon the implementation and performance analysis of VSTCF in electronic dispatch project in ISO-NE (Chen et al., 2001). Trudnowski et al. described a strategy for developing a very short-term load predictor using slow and fast Kalman estimators (Trudnowski et al., 2001). In 2006, Yang et al. presented an improved fuzzy neural system (FNS) for electric VSTCF problem based on chaotic dynamics reconstruction technique (Yang et al., 2006). James W. Taylor used minute-by-minute British electricity demand observations to evaluate methods for prediction between 10 and 30 minutes ahead (Taylor, 2008). Setiawan et al. performed a new approach for VSTCF by applying support vector regression to predict the load demand every five minutes based on historical data from the Australian electricity operator NEMMCO from 2006 to 2008 (Setiawan et al., 2009). De Andrade and Da Silva tried to achieve a comparative analysis among autoregressive integrated moving average (ARIMA) model, ANN and adaptive neuro-fuzzy inference system (ANFIS) techniques for load demand forecasting in distributed substations of cities located in Sao Paulo state of Brazil (De Andrade and Da Silva, 2009). Guan et al. presented a methodology based on multilevel wavelet neural networks with novel pre-filtering in order to detect and eliminate spikes within load, apply the wavelet technique to decompose the load into several frequency components, perform appropriate transformation on each component, and feed it together with other appropriate input to a separate neural network (Guan et al., 2009).

Koprinska et al. used autocorrelation analysis to extract 6 nested feature sets of previous electricity loads for 5 minute ahead electricity load forecasting (Koprinska et al., 2010). Qingle and Min proposed a novel approach to very short-term load by the application of ANN and rough set (Qingle and Min, 2010). Guan et al. presented a method of multilevel wavelet neural networks trained by hybrid Kalman algorithms (MWNNHK) to forecast next hour's load in five-minute steps and generate a moving prediction every five minutes, around which a good confidence interval (CI) is estimated at the same time (Guan et al., 2010). Cheah et al. used a quarter-hourly ahead load forecasting model employing a multilayer neural network with a backpropagation learning algorithm in NI LabVIEW (Cheah et al., 2011). Kotillova performed 30-minute Australian electricity demand observations to evaluate time series forecasting methods for prediction 30 minutes ahead (Kotillova, 2011). Neusser et al. employed VSTCF for a complete real-time distributed demand side management system in absence of historical data (Neusser et al., 2012). Shankar et al. used Kalman filter prediction recursive algorithms to obtain a bank of hourly predicted load data for 5-minute look ahead forecasting (Shankar et al., 2012). An et al. proposed a method of first treating the data by scale through wavelet analysis and then selecting partially similar day to forecast various loads in different frequencies with more load forecast models for VSTCF under the influence of electric railway (An et al., 2013). Shang presented a number of functional modelling and forecasting methods for predicting very short-term (such as minute-by-minute) electricity demand. The suggested functional methods slice a seasonal univariate time series (TS) into a TS of curves; reduce the dimensionality of curves by applying functional principal component analysis (PCA) before using a univariate TS forecasting method and regression techniques (Shang, 2013). Khan et al. applied a neuro-evolutionary technique known as Cartesian genetic programming evolved recurrent neural network to develop a load forecasting model for very short-term of half an hour (Khan et al., 2013).

Hsiao performed a novel approach to model the load of an individual household based on context information and its daily schedule in Taiwan with a VSTCF horizon of 30 minutes (Hsiao, 2015). Golestaneh et al. proposed a nonparametric approach to generate very short-term predictive densities, i.e., for lead times between a few minutes to one hour ahead, with fast frequency updates especially by relying on an extreme learning machine (ELM) as a fast regression model (Golestaneh et al., 2016). Yoon et al. suggested a VSTCF method based on pattern ratio for an office building in Korea (Yoon et al., 2016). Barbieri et al. presented an overview of the various tools needed to forecast photovoltaic (PV) power within a very short-term horizon (Barbieri et al., 2017). Sepasi et al. employed two parallel-series techniques for load forecasting to optimize the performance of a grid-scale battery energy storage system (BESS) (1 MW, 1.1 kWh) in 15-minute steps within a moving 24-hour window (Sepasi et al., 2017). Lastly, Capuno et al. presented a model for VSTCF based on algebraic prediction (AP) using a modified concept of the Hankel rank of a sequence. Moreover, AP is coupled with support vector regression (SVR) to accommodate weather forecast parameters for improved accuracy of a longer prediction horizon; thus, a hybrid model was also proposed (Capuno et al., 2017).

Material and Methods

The household is located on the second floor in an apartment which is settled in Mahfesığmaz neighbourhood in Çukurova district, Adana, Turkey. Geographical properties of the household are given in Table 1 and household location is shown in Figure 2.

Table 1: Geographical properties of the household

Household Location	Latitude	Longitude	Altitude
Mahfesığmaz, Çukurova	37.042 N	35.314 E	81 m

VSTCF of a household is an arduous challenge because of the fact that electric loads are characterised as nonlinear, and electrical appliances in the household vary due to their operation. For instance, a refrigerator has continuous operation, while a television (TV) and a TV console operate in standby generally. There are also other appliances such as a washing machine, a dishwasher, or a vacuum cleaner which have operation on demand.

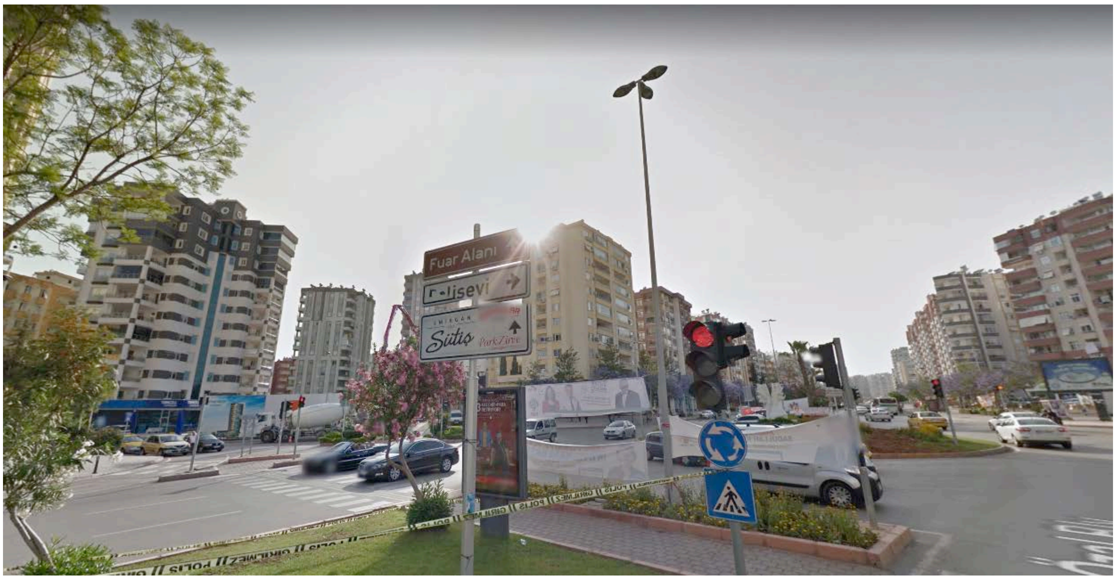

Figure 2: Household location (GoogleMaps, 2015).

Data acquisition stage of electrical energy consumption is performed between May 11 and June 8, 2018 by an energy logger through the distribution panel indicated in Figure 3. Obtained electrical energy consumption data is demonstrated in Figure 4.

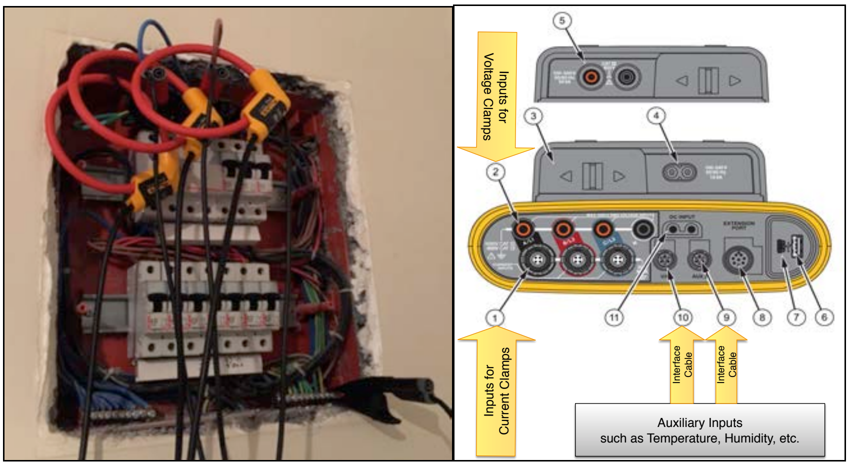

Figure 3: Distribution panel and energy logger connection schematic (Fluke, 2013).

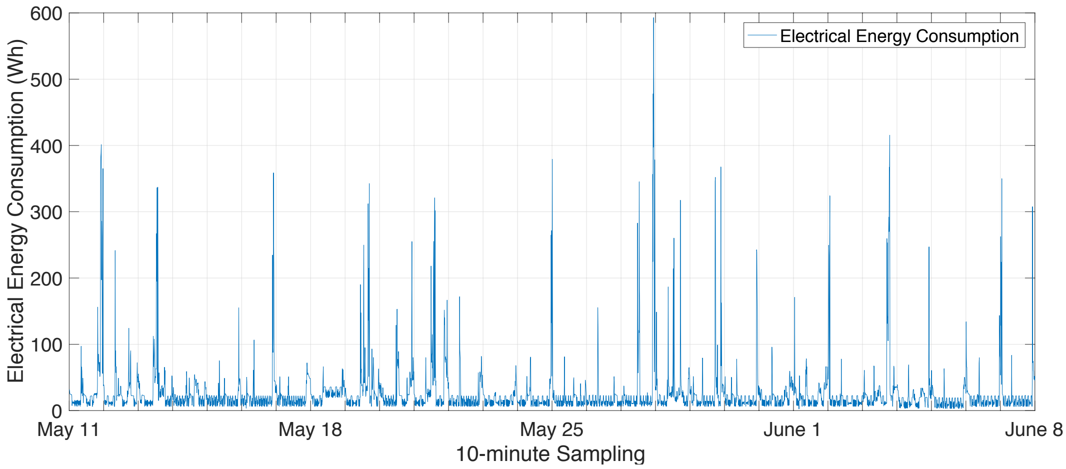

Figure 4: Electrical energy consumption data.

For weather data, MERRA-2 which stands for Modern-Era Retrospective Analysis for Research and Applications – Version 2 data (GMAO, 2015) is utilised. MERRA-2 presents a time series of temperature, relative humidity, pressure, wind speed and direction, rainfall, snowfall, and snow depth with time steps ranging from 1-minute up to 1-month (Gelaro et al., 2017). MERRA-2 data are illustrated in the following figures.

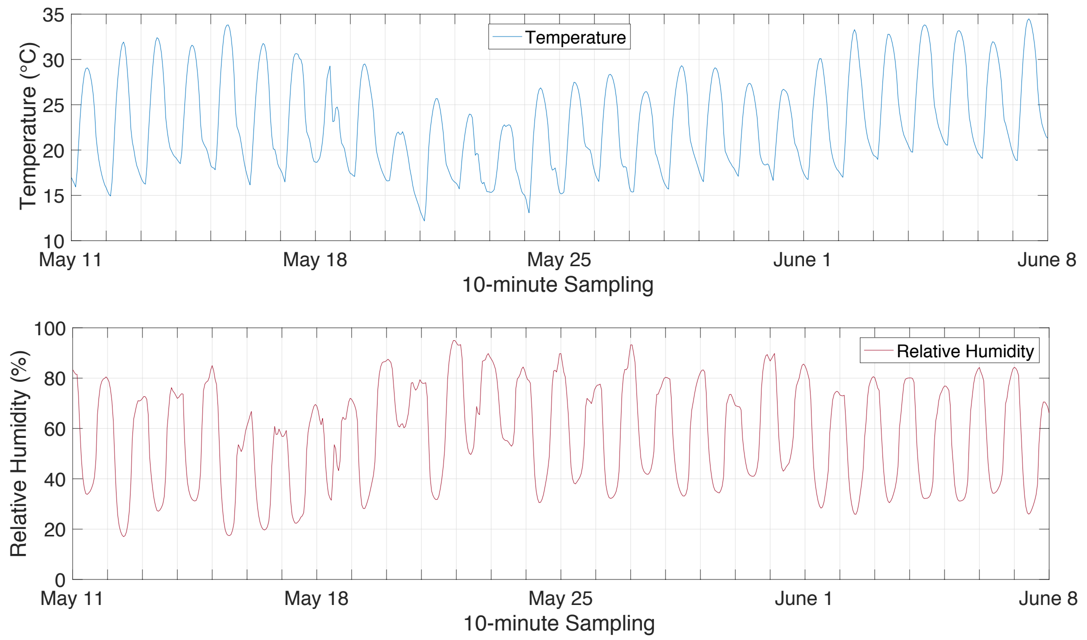

Figure 5: Temperature and relative humidity data.

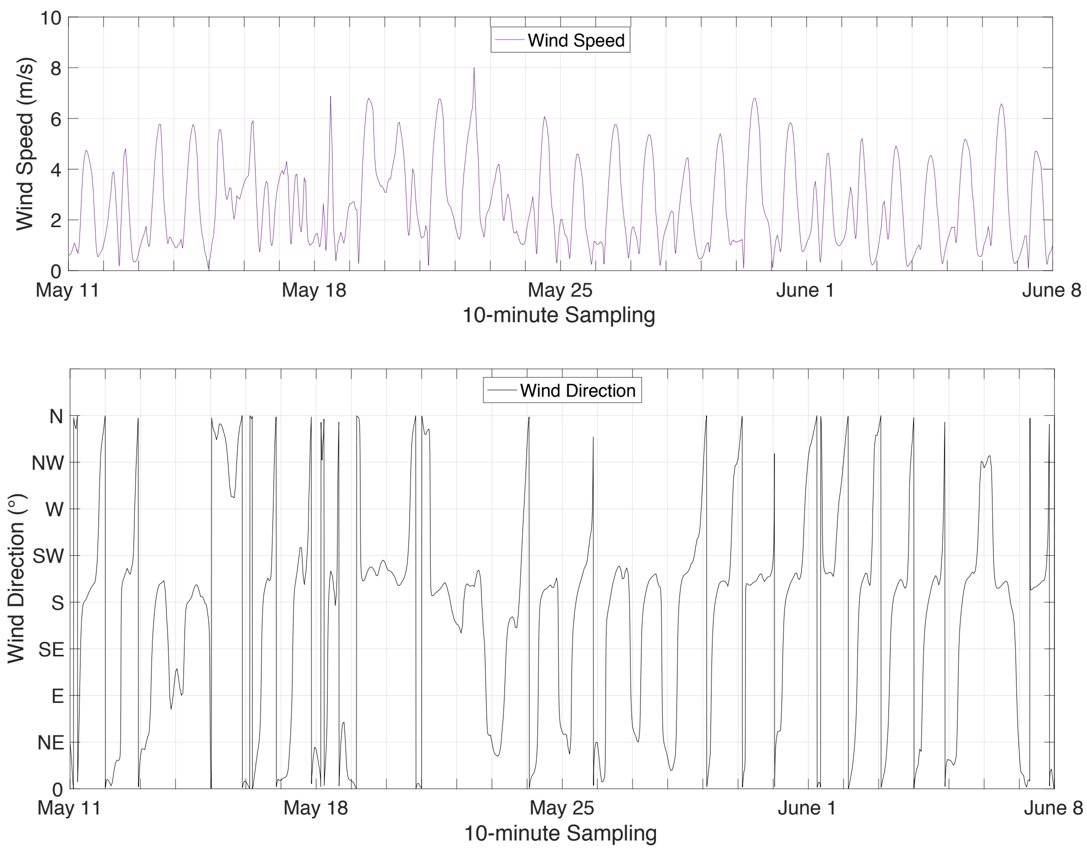

Figure 6: Wind speed and direction data.

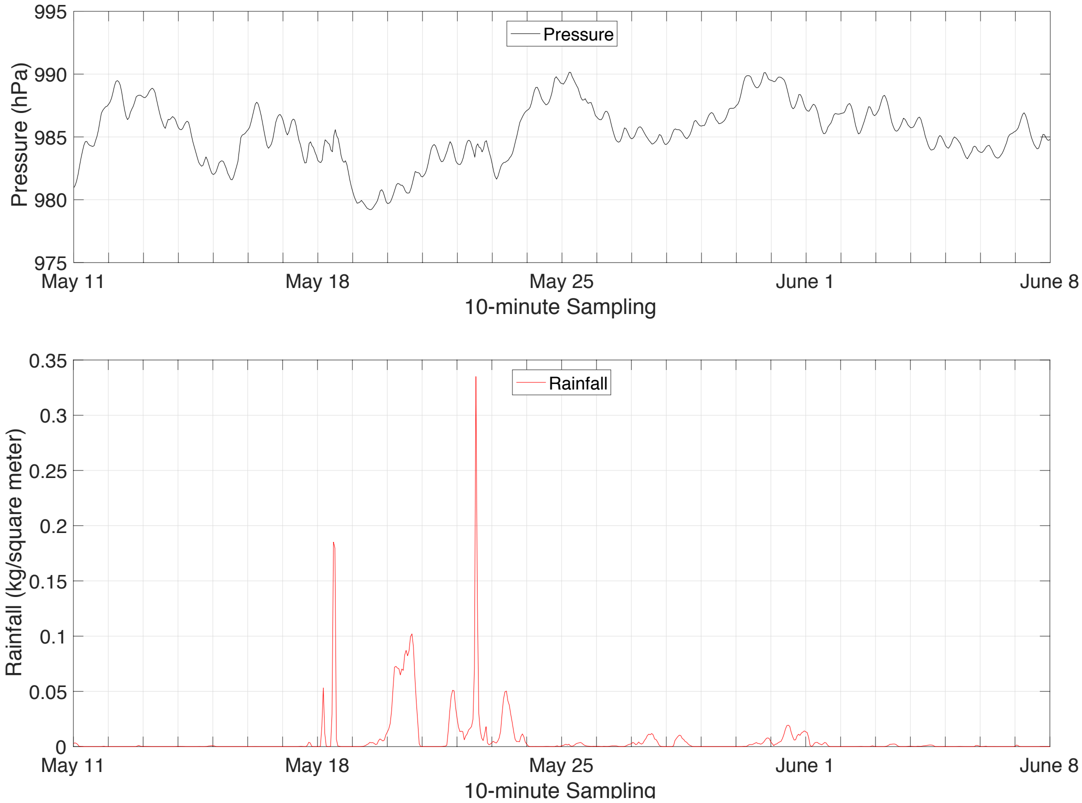

Figure 7: Pressure and rainfall data.

Data set consists of three type of input variables which are electrical, calendar, and weather inputs. Electrical variables are previous day (PrevD), previous hour (PrevH), and previous 10-minute (PrevS). Calendar inputs are day of week (DoW), hour of day (HoD), and sample of hour (SoH). Weather variables are temperature (Temp), relative humidity (RH), wind speed (WindS), wind direction (WindD), pressure (Pres), and rainfall (Rain). The data set is constituted of 4,032 rows and 13 columns (12 input and 1 target). Demonstration of system inputs and target is given in Figure 8. View of the data set in MATLAB environment is shown in Figure 9 (MATLAB, 2017).

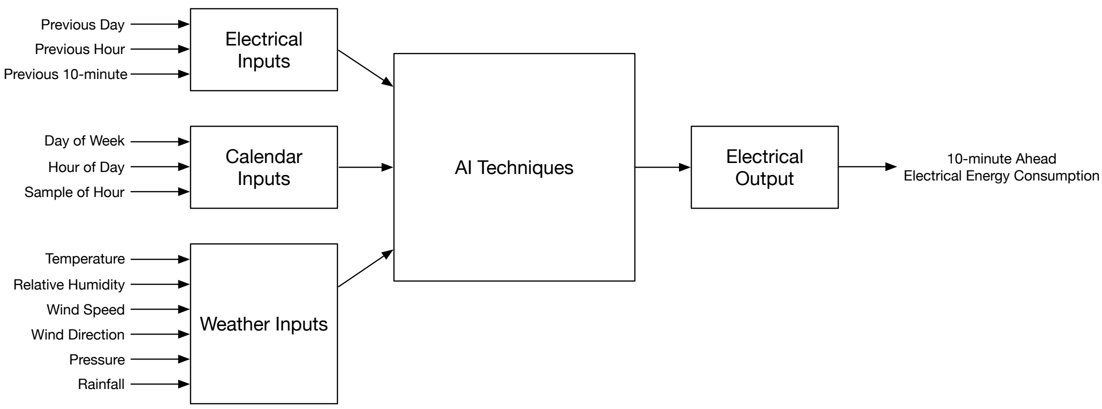

Figure 8: Illustration of system inputs and target.

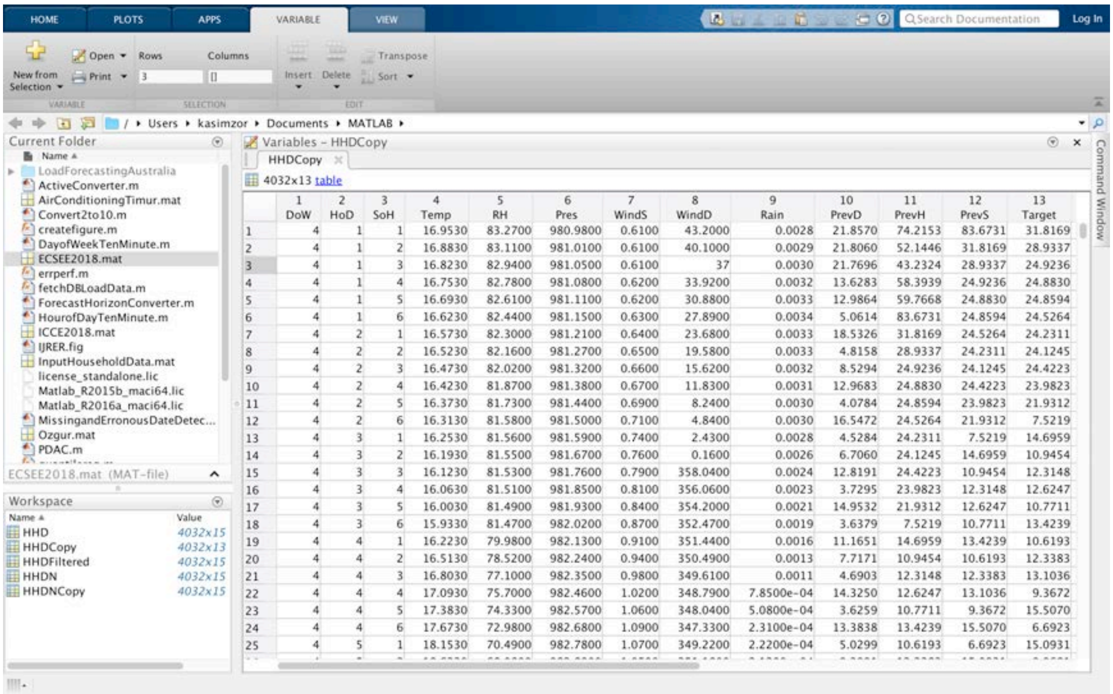

Figure 9: Data set.

Normalisation process is generally employed to eliminate the units of different data types in the data set and compare performances of diversified data columns as well. In order to reach a data distribution between 0 and 1 for each column, the following formula can be applied for $y_{min} = 0$ and $y_{max} = 1$

$$x_{norm} = (y_{max} - y_{min}) \times \left[\frac{x - x_{min}}{x_{max} - x_{min}} \right] + y_{min} \qquad (1)$$

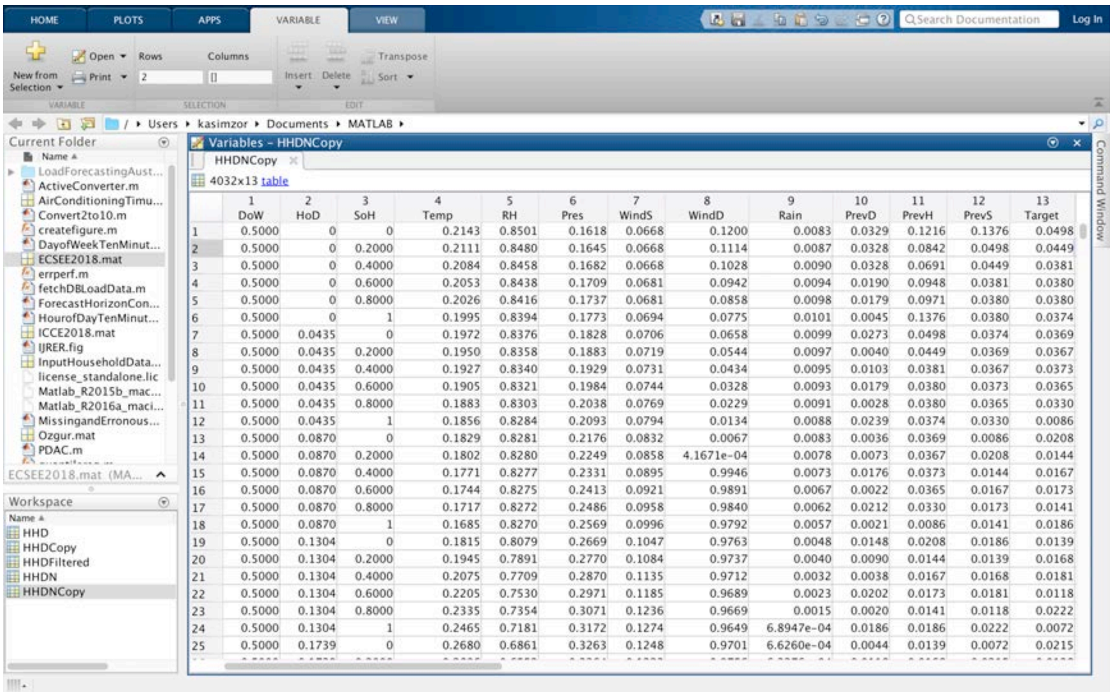

Figure 10: Normalised data set.

where x is an input vector, x_{min} and x_{max} represent minimum and maximum values of the x, y_{min} and y_{max} correspond to boundaries for distribution, and x_{norm} is the normalised version of the vector x (Çelik and Teke, 2017). Normalised data set is illustrated in Figure 10.

In the scope of this paper, DT, GA, ANN, and SVM are employed as AI techniques. Firstly, GBDT technique in DT literature is used for the forecasting process. Huber's quantile cut-off is performed as a loss function which is a hybrid of ordinary least squares (OLS) and least absolute deviation (LAD). For GBDT, number of maximum trees is 400, maximum splitting levels is 5, variable weights are chosen as equal, minimum size node to split is 10. In GBDT, random sampling is utilised for tree validation and tree pruning criterion is selected as MAE.

Secondly, gene expression programming (GEP) technique which performs a genotype/phenotype GA is employed for the prediction of VSTCF of the household. GEP is linear, ramified, and faster than old GAs and applies symbolic regression. 10-fold cross validation (CV) is chosen for the validation.

Thirdly, multilayer perceptron (MLP) neural networks, radial basis function (RBF) networks, generalised regression neural networks (GRNN), and grouping method of data handling (GMDH) type neural networks are investigated as ANN techniques. For MLP neural networks, a topology with 1 hidden layer is used and a search is conducted to find the optimal size of the hidden layer from 2 to 25 neurons. Logistic sigmoid and linear activation functions are utilised for the hidden and output layer respectively. For RBF networks, Gaussian function is used. Minimum and maximum values of r and λ are 0.01 and 573.301, and 0.012 and 9.984 sequentially. Regularisation λ for final weights is $1\times e^{-10}$ after 4 iterations. For GRNN, 4 layers are constituted as input, hidden (kernel), pattern, and decision layer. Gaussian function is employed as a kernel function for the hidden layer. 2 neurons are utilised on the topology for denominator and numerator summation units. Conjugate gradient algorithm is selected for the optimisation of σ values. For GMDH type neural networks, a topology consisting of independent variables, GMDH network, and dependent variables are built. Number of neurons per layer is fixed to 20 and quadratic polynomial with two variables is used. For all ANN techniques, 10-fold CV is selected for the validation.

Lastly, ε-Support vector regression (ε-SVR) is performed for SVM technique. Gaussian RBF type kernel function is used for ε-SVR. Grid and pattern search for optimal values is applied and the search criterion is minimising total error. ε, C, γ, and P parameter values belonging to ε-SVR are 0.001, 1521.702, 0122, and 0.484 respectively. Number of support vectors used for the prediction is 3,797. 10-fold CV is chosen for the validation of ε-SVR.

Evaluation Criterion

MAE is frequently used for evaluating point load forecasts (Xie and Hong, 2017), it calculates the average absolute forecast error of n times forecast results

$$MAE = \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{n} \tag{2}$$

where y_i represents actual or measured output, $\hat{y}_i$ shows predicted output and n indicates the number of observations (Zor et al., 2017a).

Discussion and Results

For both training and validation, MAE performance metric results according to the performed analyses in order to apply VSTCF for the household are demonstrated in Table 2.

Table 2: MAE performance metric results

		MAE (Wh)	
Performed Technique	**Model**	**Training**	**Validation**
Decision Trees	GBDT	9.38*2nd	10.56*1st
Genetic Algorithm	GEP	12.28	12.51
Artificial Neural Networks	MLP	12.68	13.44
	RBF	11.21	14.25
	GRNN	8.62*1st	12.27*3rd
	GMDH	12.05	12.90
Support Vector Machines	ε-SVR	11.27*4th	11.72*2nd
	Average	11.07	12.52

As stated in Table 2, GBDT model has the best performance for household VSTCF problem. It is also considered that GRNN and ε-SVR models accomplished better estimation in comparison with GEP, MLP, RBF, and GMDH models as well.

Conclusions

In this paper, VSTCF for smart grid integration of households is investigated. Data acquisition period occupies an interval between May 11 and June 8, 2018. For electrical data acquisition a three-phase energy logger is used as a data acquisition terminal of the household. Weather data are obtained from MERRA-2.

The data set contains 4,032 samples with 12 inputs and 1 target. Normalisation of data is realised. Several AI techniques including DT, GA, ANN, and SVM are implemented and achieved results are evaluated according to MAE performance metric as an evaluation criterion.

Consequently, results show that GBDT illustrated an excellent performance in applying VSTCF for the household while GRNN and ε-SVR performed good and reasonable performances respectively.

Acknowledgement
The authors would like to acknowledge the Scientific Research Project Unit of Çukurova University owing to financial support for the individual research project named as “Very Short-Term Forecasting of a Household Electrical Energy Consumption” and numbered as “FBA-2017-9344”.

References

An, R., Du, X., Mi, X., Ren, Y., & Wang, J. (2013). Research on forecast of ultra-short-term load under the influence of electric railway. *International Conference on Education Technology and Information System*, 74-80, doi: 10.2991/icetis-13.2013.18.

Barbieri, F., Rajakaruna, S., & Ghosh, A. (2017). Very short-term photovoltaic power forecasting with cloud modeling: a review. *Renewable and Sustainable Energy Reviews*, 75, 242-263.

Capuno, M., Kim, J., & Song, H. (2017). Very short-term load forecasting using hybrid algebraic prediction and support vector regression. *Hindawi Mathematical Problems in Engineering*, 2017, 8298531, 1-9, doi: 10.1155/2017/8298531.

Cheah, P. H., Gooi, H. B., & Soo, F. L. (2011). Quarter-hour-ahead load forecasting for microgrid energy management system. *IEEE Trondheim PowerTech*, doi: 10.1109/PTC.2011.6337027.

Charytoniuk, W. & Chen, M. (2000). Very short-term load forecasting using artificial neural networks. *IEEE Transactions on Power Systems*, 15, 1, 263-268.

Çelik, Ö. & Teke, A. (2017) A hybrid MPPT method for grid connected photovoltaic systems under rapidly changing atmospheric conditions. *Electric Power Systems Research*, 152, 194 - 210, doi: https://doi.org/10.1016/j.epsr.2017.07.011.

De Andrade, L. C. M. & Da Silva, I. N. (2009). Very short-term load forecasting based on ARIMA model and intelligent systems. *15th International Conference on Intelligent System Applications to Power Systems*, doi: 10.1109/ISAP.2009.5352829.

Feng, W., Keng, Y. E., Bo, D., & Xiao, L. (1997). Artificial neural network for ultra short-term load forecasting. *The Proceedings of IFAC Control of Power Systems and Power Plant*s, 30, 17, 541-544.

Fluke. (2013). Fluke 1730 three-phase energy logger user's manual, http://www.produktinfo.conrad.com/datenblaetter/775000-799999/792321-an-01-tr-FLUKE_1730_EU_ENERGY_LOGGER.pdf.

Gelaro et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, 30(14):5419–5454, 2017. doi: 10.1175/JCLI-D-16-0758.1.

GMAO. (2015). Merra-2 tavg1 2d slv nx: 2d, 1-hourly, time-averaged, single-level, assimilation, single-level diagnostics v5.12.4. Greenbelt, MD, USA, *Goddard Earth Sciences Data and Information Services Centre (GES DISC)*, Accessed [June 5, 2018], 2015. doi: 10.5067/VJAFPLI1CSIV.

Golestaneh, F., Pinson, P., & Gooi, H. B. (2016). Very short-term nonparametric probabilistic forecasting of renewable energy generation – with application to solar energy. *IEEE Transactions on Power Systems*, 31, 5, 3850-3863.

GoogleMaps. (2015). Street view, https://www.google.com/maps.

Guan, C., Luh, P. B., Coolbeth, M. A., Zhao, Y., Michel, L. D., Chen, Y., Manville, C. J., Friedland, P. B., & Rourke, S. J. (2009). Very short-term load forecasting: Multilevel wavelet neural networks with data pre-filtering. *IEEE Power & Energy Society General Meeting*, doi: 10.1109/PES.2009.5275296.

Guan, C., Luh, P. B., Michel, L. D., Coolbeth, M. A., & Friedland, P. B. (2010). Hybrid Kalman algorithms for very short-term load forecasting and confidence interval estimation. *IEEE Power and Energy Society General Meeting*, DOI: 10.1109/PES.2010.590077.

Hsiao, Y. (2015). Household electricity demand forecast based on context information and user daily schedule analysis from meter data. *IEEE Transactions on Industrial Informatics*, 11, 1, 33- 43.

Khan, G. M., Zafari, F., & Mahmud, S. A. (2013). Very short term load forecasting using Cartesian genetic programming evolved recurrent neural networks (CGPRNN). *12th International Conference on Machine Learning and Applications*, 152-155, doi: 10.1109/ICMLA.2013.181.

Koprinska, I., Sood, R., & Agelidis, V. G. (2010). Variable selection or five-minute ahead electricity load forecasting. *International Conference on Pattern Recognition*, 2901-2904, doi: 10.1109/ICPR.2010.711.

Kotillova, A. (2011). Very short-term load forecasting using exponential smoothing and ARIMA models. *Journal of Information, Control and Management Systems*, 9, 2, 85-92.

Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. L., & Naccarino, J. (1996). Comparison of very short-term load forecasting techniques. *IEEE Transactions on Power Systems*, 11, 2, 877-882.

MATLAB. (2017). R2017a. *The MathWorks Inc.*, Natick, Massachusetts.

Neusser, L., Canha, L. N., Abaide, A., & Finger, M. (2012). Very short-term load forecast for demand side management in absence of historical data. *Renewable Energies and Power Quality Journal*, 1, 10, 787-790.

Qingle, P. & Min, Z. (2010). Very short-term load forecasting based on neural network and rough set. *International Conference on Intelligent Computation Technology and Automation*, 1132-1135.

Sepasi, S., Reihani, E., Howlader, A. M., Roose, L. R., & Matsuura, M. M. (2017). Very short-term load forecasting of a distribution system with high PV penetration. *Renewable Energy*, 106, 142-148.

Setiawan, A., Koprinska, I., & Agelidis, V. G. (2009). Very short-term electricity load demand forecasting using support vector regression. *International Joint Conference on Neural Networks*, doi: 10.1109/IJCNN.2009.5179063.

Shamsollahi, P., Cheung, K. W., Chen, Q., & Germain, E. H. (2001). A neural network based very short term load forecaster for the interim ISO New England electricity market system. *22nd IEEE Power Engineering Society International Conference on Power Industry Computer Applications*, 217-222.

Shang, H. L. (2013). Functional time series approach for forecasting very short-term electricity demand. *Journal of Applied Statistics*, 40, 1, 152-168.

Shankar, R., Chatterjee, K., & Chatterjee, T. K. (2012). A very short-term load forecasting using Kalman filter for load frequency control with economic load dispatch. *Journal of Engineering Science and Technology Review*, 5, 1, 97-103.

Taylor, J. W. (2008). An evaluation of methods for very short-term load forecasting using minute-by-minute British data. *International Journal of Forecasting*, 24, 4, 645-658.

Trudnowski, D. J., McReynolds, W. L., & Johnson, J. M. (2001). Real-time very short-term load prediction for power-system automatic generation control. *IEEE Transactions on Control Systems Technology*, 9, 2, 254-260.

Xie, J. & Hong, T. (2017). Variable selection methods for probabilistic load forecasting: Empirical evidence from seven states of the united states. *IEEE Transactions on Smart Grid*, doi: 10.1109/TSG.2017.2702751.

Yang, H. Y., Ye, H., Wang, G., Khan, J., & Hu, T. (2006). Fuzzy neural very-short-term load forecasting based on chaotic dynamics reconstruction. *Chaos, Solitons and Fractals*, 29, 462- 469.

Yoon, A., Moon, H., & Moon, S. (2016). Very short-term load forecasting based on a pattern ratio in an office building. *International Journal of Smart Grid and Clean Energy*, 5, 2, 94-99.

Zor, K., Timur, O., Çelik, Ö., Yıldırım, H. B., & Teke, A. (2017a). Interpretation of error calculation methods in the context of energy forecasting. *12th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES)*, 0722, 1–9. https://www.researchgate.net/publication/320310871_Interpretation_of_Error_Calculation_Methods_in_the_Context_of_Energy_Forecasting.

Zor, K., Timur, O., & Teke, A. (2017b). A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting. *6th International Youth Conference on Energy*, doi: 10.1109/IYCE.2017.8003734.

Contact email: kzor@adanabtu.edu.tr