Liran KATZIR, Yakir LOEWENSTERN, Doron SHMILOVITZ

Tel Aviv University, Tel Aviv, Israel

doi:10.15199/48.2015.11.31

The Effect of System Characteristics on Very-Short-Term Load Forecasting

***Abstract**. The rise of the Smart Grid and Microgrid concepts require load demand control at short lead times, at a resolution of minutes, leading to the need for Very Short Term Load Forecasting (VSTLF). This study builds upon previous research of load forecast and investigates the relationship between system characteristics and the achievable of VSTLF accuracy. The results presented here are based on study and simulated forecasting of three years' worth of real load data obtained from the New York Independent System Operator (NYISO).*

Streszczenie.** Koncepcje Sieci Inteligentnych oraz MicroSieci wymagają sterowania z krótkim czasem wyprzedzania, rzędu minut, co prowadzi do zapotrzebowania na Bardzo Krótko Terminowe Przewidywanie Obciążenia (ang.: Very Short Term Load Forecasting - VSTLF). Przedstawione badnia są kontynuacją poprzednich nad przewidywaniem obciążenia i dotyczą związku między cechami systemu i osiągalną dokładnością VSTLF. Przedstawione wyniki są oparte na badaniu oraz na modelowaniu trzyletniego przewidywania obciążenia rzeczywistego, na podstawie danych otrzymanych od New York Independent System Operator (NYISO). **Odziaływanie cech systemu na krótkoterminowe przewidywanie obciążenia

Keywords: Load forecasting, load modelling, power systems, Smart Grid.
Słowa kluczowe: Przewidywanie obciążenia, modelowanie obciążenia, system energetyczny, Sieci Inteligentne.

Introduction

With the rise of Artificial Intelligence (AI) and Machine Learning (ML) techniques, Load Forecasting (LF) has become one of the primary fields of research in power systems engineering. As shown in Figure 1, the number of published papers on LF has increased exponentially over the past three decades. A large percentage of the studies have used AI/ML methods, with Artificial Neural Networks (ANN)[1] a particularly popular method. Other ML techniques used for LF include Fuzzy Logic, Support Vector Machines, univariate methods which are commonly used include ARIMA models and exponential Smoothing[2]-[6], and many more as discussed in Hippert's excellent review[1]. The proliferation of research in this area indicates the importance of LF and the evolving needs for accurate forecasts. Generally speaking, three kinds of forecasting lead-times are considered, and by extension, three types of load forecasting are defined: long, medium and short term[7]. Long term (years in advance) forecasting is important for planning investment in infrastructure, modeling pricing policies, required power generation forecasts, maintenance and the like. Short-term (hours in advance, generally up to 24 hours) forecasting is used for prediction of maximum load, load-flow study results, planning load switching and/or shedding and economic dispatch. Short term forecasting is the most common task considered in the literature; do to its criticality for adequate grid operations.

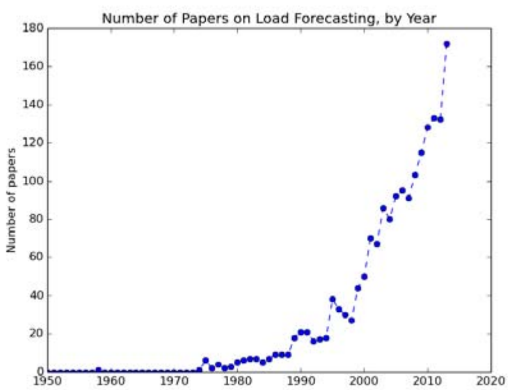

Fig. 1. Number of published papers on Load Forecasting, by year. Data taken from Medline(PubMed)

Medium-term (days/weeks/months in advance) forecasting is also utilized to assist in the above tasks, with the predictions more accurate than long-term forecasts but inferior than the forecasts obtained by short-term forecasting. Some of the factors which enable more accurate forecasts at shorter time horizons include real-time weather forecasts, the naturally slow variations in aggregate load and real-time knowledge of anomalous events which may not have been anticipated in advance. With the rise of the Smart Grid and Microgrid concepts, real-time forecasting at very short lead times are required. This is due to the stochastic nature of renewable energy sources such as photovoltaic (PV) panels and wind farms. For this purpose, Very Short Term Load Forecasting (VSTLF) has been defined and studied. Regardless of the lead-time, the multitude of techniques demonstrate the underlying fact that there is no accepted "best" LF method[8], as evidenced by similar methods obtaining very different results for different systems.

In most current literature, a LF method is chosen for a forecasting task based on the lead-time and the nature of the past (e.g. historical loads) and future data (e.g. weather forecasts) available for the system being predicted. Some previous work[9] discussed Very Short Term Load Forecasting has studied the power system characteristics and shown that a given VSTLF method's accuracy is extremely variable, based mostly on the statistical characteristics of the system it is applied to. This earlier research concluded that a VSTLF method should be chosen based on system characteristics, the required accuracy and tolerance for model complexity. Furthermore, most research has studied very large grids, which often supply power to an entire country and are not representative of smaller power systems which are of very different statistical characteristics. In this paper, the previous work is built upon and a heuristic lower-bound for VSTLF accuracy is observed when applying forecasting methods to different systems. The results are based on three years' worth of load data collected from the eleven power systems operated by the New York Independent System Operator (NYISO)[10], sampled once every five minutes. The NYISO zone map is shown in Fig. 2. Five VSTLF techniques are compared, and show that three of them converge to solutions which produce small, usually uncorrelated residuals.

The contribution this paper makes to smart-grid research is twofold. STLF is required for demand response in systems powered traditional generation facilities - generators, etc. With increased penetration of renewable energy sources of highly stochastic nature (PV and wind power) into the smart grid, the demand-response problem requires shorter-termed, accurate forecasts. VSTLF will ensure that generation can accurately anticipate load demand and meet it even while relying on these highly stochastic power sources. Additionally, with increased smart-grid research focusing on load-forecasting aggregation [11][12] and demand response[13] at the consumer level, aggregation techniques can be applied to the high voltage system to improve overall demand response capabilities. The VSTLF methods discussed in this work and the system variability investigated can contribute to development of effective aggregation techniques.

Fig. 2. NYISO zone map

Very short term load forecasting

A. Definition and Framework

While no formally accepted definition of VSTLF exists, past studies have used the term to indicate load forecasting from one minute to half an hour[5][14] lead time. Its primary function is providing a generation target for economic dispatch and load frequency control. By accurately forecasting the load at a future point, the grid frequency can be held at its nominal value by preventing under-generation, while holding extra power generation to a minimum. Another use of VSTLF results is in forecasting the values of state variables for quasistatic[15] and dynamic[16] Power System State Estimation schemes(PSSE). In modern power systems, PSSE is general carried out every several minutes, which makes VSTLF a natural candidate for the forecasting step. In this study, we define VSTLF as forecasting the load five minutes ahead of time.

B. System Characteristics

As discussed in previous work[9] and corroborated later in this work in Figure (6), system size and autocorrelation are of critical importance to the quality of forecasts. At opposite sides of the characteristic spectrum are the systems of New York City (NYC) and Millwood. Figures (3) and (4) graphically illustrate the significant differences between the sizes of the two systems and the nature of the load fluctuations observed in each one. To compare the load changes in two systems of very different sizes, Load Percentage Difference is defined as

(1) $$LPD[n] = \frac{x[n] - x[n-1]}{x[n-1]}$$

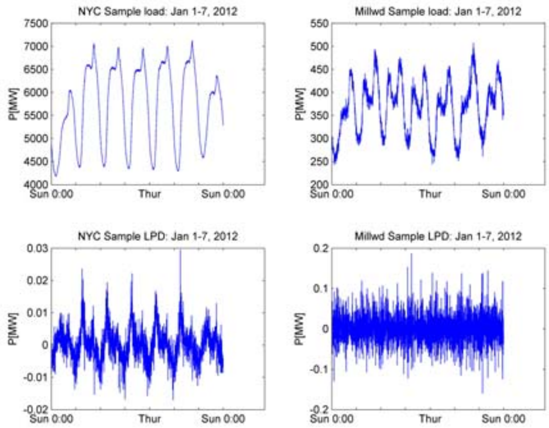

Fig. 3. Sample weekly data for New York City and Millwood

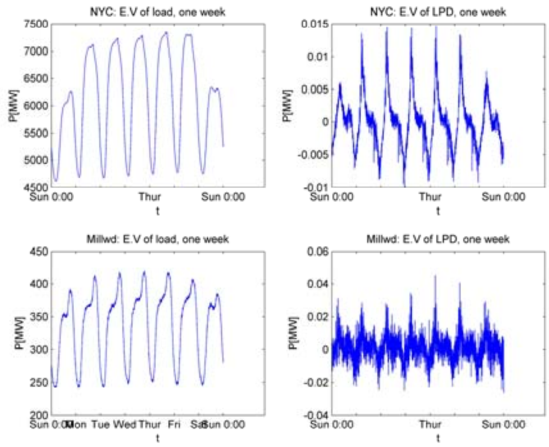

Fig. 4. Average weekly data for New York City and Millwood

Table 1 Size and autocorrelation of the NYISO power systems

	Mean Load[GW]	Range	Load Difference Max. AC(index)	Diff. AC(1)
CAPITL	1.52	2.05	0.22(288)	-0.13
CENTRL	1.97	1.85	-0.20(1)	-0.2
DUNWOD	0.85	2.32	-0.33(1)	-0.33
GENESE	1.29	2.09	0.44(2)	0.23
HUD_VL	1.35	2.33	-0.32(1)	-0.32
LONGIL	3.22	2.51	0.61(288)	0.46
MHK_VL	0.96	1.97	0.23(288)	-0.06
MILLWD	0.38	2.39	-0.44(1)	-0.44
NORTH	0.72	1.19	-0.19(1)	-0.19
NYC	7.37	2.04	0.57(2)	0.43
WEST	1.9	1.78	-0.20(1)	-0.2

The size of each system is indicated by its mean load, and the variability by the range, defined as

(2) $$\text{range} = \frac{\text{95th percentile load}}{\text{5th percentile load}}$$

The table exhibits the normalized lag-1 autocorrelation, as well as the highest (in absolute value) autocorrelation value, with its corresponding lag parenthesized. For instance, in

NYC, the highest autocorrelation factor is 0.57, at lag 2, with an autocorrelation factor of 0.43 at lag 1.

C. Methods and Models

The five methods used for this study are described below.

1) Wiener Process (WP)

Due to the short time interval between load samples and the slow load variations, reasonably accurate forecasts can be obtained by modeling the load signal as a random walk:

(3) $$X[n] = X[n-1] + w[n]$$

with $w[n]$ Gaussian white noise of zero mean and variance σ^2. Due to the noise having zero mean, the forecast load is:

(4) $$\hat{X}[n+1] = X[n]$$

While all other forecasting algorithms presented will improve upon this method, this technique is important to present as a benchmark due to its simplicity; it requires no historical data or parameter estimation.

2) Difference Averaging (DA)

We define the load difference as:

(5) $$d[n] = X[n] - X[n-1]$$

Rearranging terms, the load can be expressed as

(6) $$X[n] = X[n-1] + d[n]$$

Both load and load difference have periodic means[9], so given $X[n-1]$, the load at time n can be estimated as

(7) $$\hat{X}[n] = X[n-1] + \hat{d}[n]$$

After averaging the difference signal to obtain the periodic mean

(8) $$\hat{\eta}_d[n] = E\{d[n]\} = \frac{1}{k}\sum_{i=1}^{k} d[n-iT]$$

The estimated difference signal is applied to Eq. (6).

3) Difference Averaging with MA(1) correction (DAMA1)

We define the error process as:

(9) $$e[n] = X[n] - \hat{X}[n]$$

Examination of the error process obtained using the simple DA model shows significant correlation between adjacent error samples (one lag) and near-zero correlation for longer lags. Therefore, modeling the error as a first order moving-average (MA(1)) process is appropriate, and offers a potential correction factor, which can also be viewed as feedback.

Assuming the error process is of a zero-mean, Gaussian distribution, the optimal estimator for the error at time n is the optimal linear estimator. If we denote, for brevity's sake, $y = e[n]; z = e[n-1]$, the optimal estimator is given by

(10) $$\hat{z} = \frac{\sigma_{zy}}{\sigma_{yy}^2}(y - \eta_y) + \eta_z$$

substituting $\eta_z = \eta_y = 0$ yields

(11) $$\sigma_{zy} = E\{(z-\eta_z)(y-\eta_y)\} = E\{zy\} = R_{ee}[1]$$

and since

(12) $$\sigma_{zz}^2 = \sigma_{yy}^2 = E\{e[n]^2\} = R_{ee}[0]$$

the error process can be optimally estimated as

(13) $$e[n] = \hat{\alpha} \cdot e[n-1], \qquad \hat{\alpha} = \frac{\hat{R}_{ee}[1]}{\hat{R}_{ee}[0]}$$

given α, the error resulting from DA forecasting is estimated as:

(14) $$\hat{e}_{DA}[n] = \hat{\alpha} \cdot e_{DA}[n-1] = \hat{\alpha} \cdot (x[n-1] - \hat{x}[n-1])$$

and the forecast load value is corrected by subtracting the estimated DA error:

(15) $$\begin{aligned} \hat{x}[n] &= x[n-1] + \hat{d}[n] - \hat{e}_{DA}[n] = \\ &= (1-\alpha)x[n-1] + \alpha\hat{x}[n-1] + \hat{d}[n] \end{aligned}$$

this correction improves upon the simple DA for almost all systems, as evidenced by lower MAPE and less correlated residuals.

4) ARMA modeling using the Box-Jenkins

The Box-Jenkins method[18] is a common method used for time-series analysis utilizing ARMA or ARIMA models to forecast future values based on past values. In this study, ARMA models ar used to forecast the load difference d[n], and it is substitute into Eq. (6) to obtain a load forecast. Since the size and statistical properties of the different power systems vary, the order and estimated parameter values vary accordingly between the systems. For a given ARMA(p,q) model, the load difference is modeled as:

(16) $$d[n] = \sum_{i=1}^{p} a_i \cdot d[n-i] + \sum_{i=1}^{q} b_i \cdot w[n-i] + w[n]$$

When forecasting $d[n]$, the past values $d\{[m] : m < n\}$ are known. The coefficients a_i; b_i are estimated during the model training stage, using Least-Squares techniques, and the past white-noise samples $\{w[m] : m < n\}$ are estimated "online", along with the load differences.

5) Double Exponential Smoothing (DES)

The well-known Holt-Winters double exponential smoothing (DES) method[19][20] has been used extensively for general time-series forecasting problems, including load-driven power system state forecasting[21]. The DES method is well-suited to time-series exhibiting a trend, but no seasonal factors. In this model, the level $S[n]$ and trend $T[n]$ are estimated as

(17) $$\begin{aligned} S[n] &= \alpha X[n] + (1-\alpha)(S[n-1] + t[n-1]) \\ T[n] &= \gamma(S[n] - S[n-1]) + (1-\gamma)T[n-1] \end{aligned}$$

and the next load is forecast to be

(18) $$\hat{X}[n+1] = S[n] + T[n]$$

Using this method, the parameters α and γ are estimated from the training data, and the model is applied to the test data to obtain error values. Incorporation of seasonal factors has been explored for STLF applications[22], and while in this study we experimented with seasonal factors as well, this did not improve results. This is likely due to the fact that the load signal changes slowly and is sampled at high frequency, so enough very recent measurements make "distant" samples which are correlated due to seasonal and periodic factors, redundant.

D. Residual Analysis

We evaluate the accuracy of the forecasting methods by measuring the Mean Absolute Percentage Error (MAPE), the widely accepted measure for forecasting accuracy[1]:

(19) $$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|x_i - \hat{x}_i|}{x_i} \quad \%$$

After calculating the MAPE, the question of potential improvement should be asked. One indication of little room for progress is the convergence of the error values for several different techniques. A more rigorous approach is commonly applied using the Box-Jenkins methodology[18].

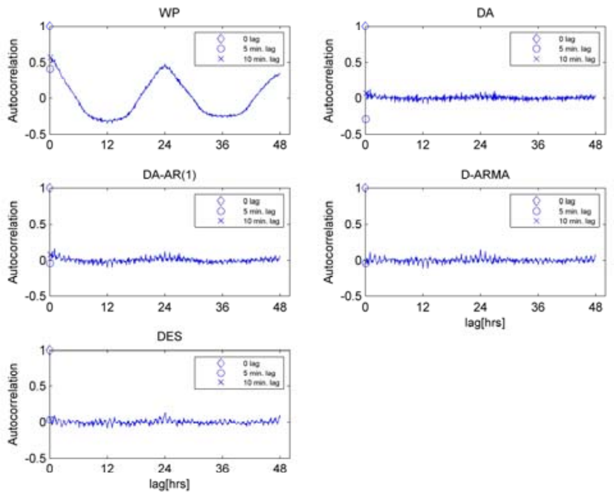

Fig. 5. Autocorrelation of errors obtained by forecasting the NYC system loads during a week in August 2012. The zero, one and two-sample-lag coefficients are marked with a diamond, circle and 'x', respectively. It can be observed that the one-lag correlation evident in the DA plot is nullified in the DAMA1 plot.

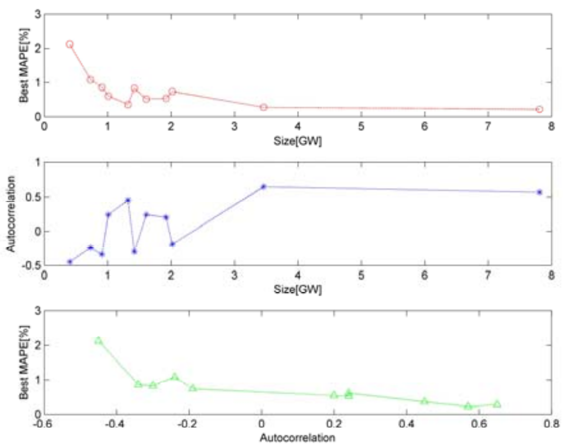

Fig. 6. Top to Bottom: MAPE vs. system size; autocorrelation vs. system size; MAPE vs. autocorrelation

Table 2 MAPE[%] FOR FORECASTING A SAMPLE WEEK IN AUGUST 2012

	WP	DA	DAMA1	D-ARMA	DES
CAPITL	0.67	0.57	0.52	0.55	0.55
CENTRL	0.83	0.76	0.74	0.76	0.77
DUNWOD	1.05	0.99	0.86	0.89	0.88
GENESE	0.52	0.37	0.37	0.37	0.36
HUD_VL	1.02	0.95	0.83	0.85	0.85
LONGIL	0.49	0.29	0.28	0.29	0.29
MHK_VL	0.72	0.64	0.61	0.64	0.65
MILLWD	2.59	2.54	2.14	2.12	2.12
NORTH	1.17	1.17	1.08	1.11	1.13
NYC	0.35	0.23	0.22	0.23	0.23
WEST	0.65	0.59	0.54	0.55	0.56

Results

Table (II) summarizes the performance of all five methods for a sample week in August 2012. It can be seen that the simple Weiner Process forecaster is easily the worst, with Difference Averaging method performing better, with improvement varying greatly among the systems. The latter three techniques are always better than the DA technique, again, with improvement depending on the system in question. The superior techniques exhibit nearly identical performance, for all systems.

We examined the residuals vectors of all results obtained in this study, and found that the errors produced by the latter three techniques are uncorrelated for almost all systems at all times. Figure (5) displays sample error vector autocorrelation values obtained when forecasting NYC (a large system exhibiting high load difference correlations).

Conclusions

We point out a few conclusions from the results shown above:

- When examining the relationships that exist between system size, autocorrelation and forecasting MAPE, it becomes clear, as shown in Figure (6), that large systems tend to be of significant correlation between load changes and tend to be easier to predict. In other words, accurate forecasts are indicative of large system size or correlation, and often both.
- Besides similar MAPE values, the three superior techniques produce error vectors usually devoid of significant autocorrelation. While this is not conclusive proof of optimal performance, given that the three techniques have fundamentally different approaches, it is a strong indicator that further improvement is unlikely in the case where the errors are uncorrelated.

Based on these conclusions we suggest several directions for further research:

- Recent research has focused on load aggregation techniques at the residential level[11][12]. Rudimentary experimentation not detailed in this paper has shown promising results in attempts to apply similar techniques at the high-voltage level.
- Introduction of seasonal factors has been shown to improve forecasting when using ARMA models for VSTLF[17]. Although we experimented with seasonal factors using double exponential smoothing and found them to be of negligible value, we did not apply them to the ARMA technique. Several systems exhibited maximal autocorrelation at a 288-sample (one day) lag, as shown in table (I). This seems to indicate that a seasonal factor of period 288 may be beneficial to the ARMA model.
- As discussed above, MAPE is the widely accepted error measure for LF studies, despite its deficiencies in accurately modeling the costs of poor forecasts or the costs of implementing complicated methods. More sophisticated load forecasting cost functions would undoubtedly improve the framework for tailor-fitting a LF technique to a given application.

Acknowegments

This work was partially supported by Israel Ministry of Energy and National Infrastructure through grant no. 212-11-0033-*9320-*

REFERENCES

[1] H.S. Hippert, C.E. Pedreira and R.C. Souza, "Neural Networks for Short- Term Load Forecasting: A Review and Evaluation", IEEE Transactions on Power Systems, Vol. 16, No. 1, pp. 44-53, February 2001.

[2] J.W. Taylor, L.M. de Menezes and P.E. McSharry, ”A comparison of univariate methods for forecasting electricity demand up to a day ahead”, International Journal of Forecasting 22, pp. 1-16
[3] J. Contreras, R. Espinola, F.J. Nogales, A.J. Conejo, ”Models to Predict Next-Day Electricity Prices” IEEE Transactions on Power Systems, vol. 18, no. 3, pp. 1014-1020, August 2003.
[4] D.G. Infield and D.C. Hill, ”Optimal Smoothing for Trend Removal in Short Term Electricity Demand Forecasting”, IEEE Transactions on Power Systems, Vol. 13, No. 3, pp. 1115-1120, August 1998.
[5] K. Liu, S. Subbarayan, R.R. Shoults, M.T. Manry, C. Kwan, F.L. Lewis and J. Naccarino, ”Comparison of very short-term load forecasting techniques,” IEEE Transactions on Power Systems, vol.11, no.2, pp.877,882, May 1996
[6] W.C. Hong, ”Electric load forecasting by support vector model”, Applied Mathematical Modeling, Vol. 33, pp. 24442454, 2009.
[7] E.A. Feinberg, and D. Genethliou, ”Load Forecasting”, Applied Mathematics for Power Systems, pp. 269-285, Springer, 2005
[8] A. Khosravi and Saeid Nahavandi, ”Load Forecasting Using Interval Type-2 Fuzzy Logic Systems: Optimal Type Reduction”, IEEE Transactions on Industrial Informatics, vol. 10., no. 2, pp. 1055-1063, May 2014
[9] Y. Loewenstern, L. Katzir and D. Shmilovitz, ”Statistical Analysis of Power Systems and Application to Load Forecasting”, Proceeding of the IEEE 28-th Convention of Electrical and Electronics Engineers in Israel, Eilat, 2014
[10] www.nyiso.com
[11] C.E Borges, Y.K. Penya, I. Fernandez, ”Evaluating Combined Load Forecasting in Large Power Systems and Smart Grids”, IEEE Transactions on Industrial Informatics, vol. 9., no. 3, pp. 1570-1577, August 2013
[12] W. Labeeuw, G. Deconinck ”Residential Electrical Load Model Based on Mixture Model Clustering and Markov Models”, IEEE Transactions on Industrial Informatics, vol. 9., no. 3, pp. 1561-1569, August 2013
[13] A. Safdarian, M. Fotuhi-Firuzabad, and M. Lehtonen, ”A Distributed Algorithm for Managing Residential Demand Response in Smart Grids”, IEEE Transactions on Industrial Informatics, vol. 10., no. 4, pp. 2385- 2393, November 2014
[14] W. Charytoniuk and M.S. Chen, ”Very Short-Term Load Forecasting Using Artificial Neural Networks”, IEEE Transactions on Power Systems, vol. 15, no. 1, pp. 263-268, February 2000.
[15] E.A. Blood, M.D. Ilic, B.H. Krogh, ”A Kalman Filter Approach to Quasi-Static State Estimation in Electric Power Systems,” 38th North American Power Symposium, NAPS 2006, pp. 417-422, September 2006
[16] A.K. Sinha, ”Short term load forecasting using artificial neural networks,” Proceedings of IEEE International Conference on Industrial Technology, vol.1, pp.548-553, January 2000
[17] J.W. Taylor and P.E. McSharry, ”Short-Term Load Forecasting Methods: An Evaluation Based on European Data”, IEEE Transactions on Power Systems, vol. 22, no. 4, pp. 2213-2219, November 2007.
[18] G.E.P. Box, G.M. Jenkins and G.C. Reinsel, ”Time Series Analysis” (3rd edition), Prentice-Hall, Inc., 1994.
[19] P.R. Winters, ”Forecasting sales by exponentially weighted moving averages”, Mngt Sci no. 6, pp 324-342, 1960.
[20] C.C. Holt, ”Forecasting seasonals and trends by exponentially weighted moving averages”, International Journal of Forecasting, Volume 20, Issue 1, March 2004, pp. 5-10
[21] A.M. Leite da Silva, M.B. Do Coutto Filho and J.F.de Queiroz, ”State forecasting in electric power systems”, IEE Proceedings C on Generation, Transmission and Distribution, vol.130, no.5, pp.237,244, September 1983
[22] J.W. Taylor, ”Short-term electricity demand forecasting using double seasonal exponential smoothing”, Journal of the Operational Research Society, no. 54, pp. 799-805, 2003.

***Authors**: Liran Katzir, Tel Aviv University, Tel Aviv, Israel, E-mail: lirankat@post.tau.ac.il; Yakir Loewenstern, Tel Aviv University, Tel Aviv, Israel, E-mail: yakirloe@mail.tau.ac.il; Dr. Doron Shmilovitz, Professor, Tel Aviv University, Tel Aviv, Israel, E-mail: shmilo@post.tau.ac.il.*