Journal of Engg. Research **Vol. 6 No. (3) September 2018 pp** *116-135*

Long-term electric load forecast in Kuwaiti and Egyptian power systems

M. F. AlHajri, M. R. AlRashidi and K. M. EL-Naggar*

mfalhajri@yahoo.com; malrash2002@yahoo.com; knaggar60@hotmail.com

Electrical Engineering Technology Department, Colleage of Technological Studies, (PAAET), Kuwait

**Corresponding Author: knaggar60@hotmail.com*

ABSTRACT

This paper presents an efficient methodology for forecasting annual peak demands in electrical power systems. The proposed approach is developed as an accurate alternative forecasting method to other existing methods. The method is based on cuckoo search algorithm. It is used to minimize the error associated with the estimated model parameters in order for the forecasted demands to follow the real load data. Real data from Kuwaiti and Egyptian networks are used to perform this study. Three long-term forecasting models have been used in this research work to measure the robustness of the developed estimation tool. Durbin-Watson statistical test is conducted to validate selected models' adequacy, and model transformation is applied as a remedial measure to ill-conditioned time series data when needed. Forecasting outcomes are reported and compared to those obtained using other forecasting techniques. The performance of the proposed method is examined and evaluated. Results reveal that cuckoo search algorithm has is a promising potential as a viable tool for parameter estimation.

Keywords: estimation; heuristic optimization; long-term forecasting.; power system planning; peak load demand.

INTRODUCTION

Electric load forecasting is a crucial task that has to be done on regular basis by many participants of the electrical power industry as it serves in unit scheduling, system security, market operation, and planning of system reserve. Prior to deregulating different sectors of this industry, load forecasting was an exclusive task assigned to a single entity, i.e.that is, vertically integrated power utility. However, it is no longer the case as different market players (buyers and sellers at generation, transmission, and distribution levels) are incorporating load forecasting information into their trade decisions. Load forecasting can be classified in terms of time duration as short-term (ranging from few minutes to few days), medium-term (ranging from several months to one year), and long-term (ranging from one year to ten years). Each type of load forecasting serves different purposes. For example, long-term load forecasting outcomes target expansion planning, inter-tie tariff setting, and long-term capital investment return problems, while short-term load forecast results are applied in unit commitment, maintenance, and economic dispatch problems. Power system planning and operations highly depend on the expected load profile. However, there is always indecisive nature of forecasting process as there are large numbers of considerable factors such as weather, time, and seasonal events, ..etc that characterize and directly or indirectly affect the underlying forecasting process; most of them are uncertain and uncontrollable.

Forecasting techniques can be generally classified into two main categories: traditional and artificial intelligence based methods. The first category includes regression, statistical, and time series methods, while the second one includes fuzzy based techniques, artificial neural network (ANN), genetic algorithm (GA), particle swarm optimization (PSO), support vector machine (SVM), differential evolution (DE), and expert system. Traditional estimation methods have been proposed and applied to long-term load forecasting to identify model characteristics, including static and dynamic state estimation methods (AlFares and Nazeeruddin 2002; IEEE Power System Engineering Committee 1980; Jabr 2006; Willis and Northcote-Green 1984). Least squares method has been the most favorable technique for many years. However, this popular method has some drawbacks and disadvantages. For instance, when outliers exist in collected data, the estimated parameters may be imprecise. To overcome this problem, a larger set of data collection is needed. A method that has the advantage of being able to detect bad data is presented in reference (Temraz et al., 1998). A dDifferent class of parameter estimators is dynamic filters where data points are used recursively to update the estimates. Such filtering techniques are suitable for control system applications (Cassola and Burlando 2012).

New biologically and naturally inspired algorithms such as ANN and SVM have been used successfully in this area (Kandil et al., 2006; Kermanshahi and Iwamiya 2002; Nie et al., 2012; Wang et al., 2012). Other techniques based on expert systems have been applied to load forecasting problem in recent years (Rahman and Hazim 1996). Both ANN and SVM have a drawback of how to configure the network properly (Nie, Liu, Liu, & Wang 2012; Wang, Li, Niu, & Tan 2012). Naturally inspired algorithms such as GA have been applied to the load forecasting problem (Karabulut et al., 2008). Hybrid techniques that combine more than one method, whether traditional or emergingtechniques , such as fuzzy rules, ANN, GA, and SVM, were also proposed in many references to enhance capabilities of such hybrid tools (Li et al., 2013; Wu 2010). Recently, (Cetinkaya, N. (2013) applied artificial neural-fuzzy inference system to forecast long-term energy consumption and peak load to Turkey network. The presented results were compared with those obtained using the classic artificial neural network method. Takiyar proposed hybrid load forecasting tool based on artificial neural network and PSO for tackling the long-term load forecasting problem. In the aforementioned reference, effects of utilizing renewable energy resources were considered (Takiyar, S. 2015). Khuntia et al. presented a review study for load forecasting techniques developed recently for the mid- and long-term horizons of electrical power systems, and discussions on future research directions in such field were presented (Khuntia, S. R., et -al.). Performance of these techniques highly depends on the algorithm's parameters, and, in some cases, proper tuning of these parameters may be a tedious task due to various parameters involved.

In this paper, a cuckoo search algorithm (CSA) based technique is developed for solving the long-term load forecasting problem. CSA is a recently developed general purpose optimization technique that deals with problems in which a best solution is sought without using derivative information nor doing complex mathematical operations. CSA based approach is utilized to find different forecasting model parameters. Real networks data are used to evaluate the developed algorithm. Results are compared with those of other competing forecasting methods.

MATHEMATICAL FORMULATION

The fForecasting process is a prediction technique that estimates future points based on historical data. Such data is gathered chronologically in time and, accordingly, referred to as time series,; while the corresponding forecasting process that is based on such past data is known as time series method. Regression analysis is commonly used to quantify the forecast relation between time as the dependent variable and the time series dependent variable and often called time series regression (Farnum and Stanton, 1989). The time series regression as a forecasting process utilizes the statistical properties of the data to build an appropriate stochastic model that relates the dependent and the independent variables and to estimate the model parameters usually via minimizing the least error squared. In this paper, the objective function is the sum of squared errors (SSE), and CSA is employed to minimize the nonlinear SSE according to.

$$\underset{\boldsymbol{\beta}}{\text{Minimize}} \quad \left\|\boldsymbol{\varepsilon}_t\right\|^2 = \left\|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\right\|^2 \tag{1}$$

where $\boldsymbol{\varepsilon}_t$ is an m error vector, $\mathbf{X}$ is an overdeterminant $m \times p$ matrix of independent chronicle values, $\mathbf{Y}$ is m vector of the time series variables, and $\boldsymbol{\beta}$ is the p vector of the model unknown parameters.

Long-term p eak demand forecasting

Electric peak load demand forecast depends on many factors like type of load itself, weather condition, as well asand economic and social activities. A robust long-term load forecasting model that utilizes time series regression is needed to represent energy consumption and peak demand that would meet the requirements imposed by utilities and consequently to predict future demand accurately.

Time series data may be relatively stationary when the mean response value is independent of time or it may exhibit a trend pattern over time. Daniels' tTest is a powerful nonparametric test that is utilized to detect the trend in time series data (Daniels 1950). Generally, trends in load data are mainly classified into two categories: linear and curvilinear trends. The former exhibits a steady change in the forecasted response as the time passes,; while the amount of change in the latter trend varies with time. The time series regression that models a linear trend is formulated as shown in . Polynomial, exponential growth and S-shaped patterns are examples of the curvilinear type of trends, and their corresponding regression time series models are, respectively, expressed in equations and.

$$P_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{2}$$

$$P_t = \beta_0 + \sum_{n=1}^{k} \beta_n t^n + \varepsilon_t \tag{3}$$

$$P_t = \beta_0 \cdot \beta_1^t \cdot \varepsilon_t \tag{4}$$

where P_t is the peak load demand as the time series dependent variable at time t; β_0, $\beta_1, \ldots,$ β_k are the model coefficients to be estimated by CSA; k is the model order; and ε_t is the associated error at time t. Furthermore, β_0 represents the country's gross domestic product (GDP), population

(Pop), GDP per capita (GDP/Cap), and electricity utilization by population (EP). Although, there is no unique formulation for the regression intercept coefficient β_0 since it is country as well as residents' behaviour dependant, it can be expressed as follows:

$$\beta_o : f\left(GDP, Pop,\ GDP/Cap, EP\ \right) \tag{5}$$

Underlying trend in the peak load time series is essential in selecting an appropriate model that would fit the exhibited pattern and increase the likelihood of obtaining better forecast. To apply the time series regression as a long-term peak load forecasting procedure, the collected time series peak load data is divided into two sets. The first one is used to regress the peak load against time and consequently estimate the model parameters. The second set is used to validate the prediction strength of the selected model by evaluating the projected loads for the remaining time periods and comparing them with the actual data points.

A discrete set of equations can be obtained using peak power demand (P_t) at each year t via model equations. For *m* years, there will be *m* equations with $p = k + 1$ unknowns, and an overdetermined system $(m > p)$ is created.

In this paper, the model given by equation is used. Therefore, a discrete system of equations can be written in a matrix notation as:

$$\mathbf{P}_t = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \tag{6}$$

Having *n* as the number of years and *m* as the number of independent variables to be estimated,; that is, *m* = 2 for linear model and *m* = 3 for quadratic model and *m*=3 for cubic model,; $\mathbf{P_t}$ is defined as ($n \times 1$) column, $\boldsymbol{\beta}$ as ($m \times 1$) column, and $\mathbf{X}$ is the time series ($n \times m$) matrix that relates $\mathbf{P_t}$ to $\boldsymbol{\beta}$.

Linear, quadratic, and cubic models are investigated in the an endeavor of to estimating estimate the unknown vector $\boldsymbol{\beta}$ that would minimize the error vector $\boldsymbol{\varepsilon}$ of the associated model. Once the problem parameters are estimated, the prediction process is implemented and the future loads can be obtained. Hence, validating the selected model is a crucial procedure in adopting the model to rely on in forecasting the response in future time periods.

Model adequacy

When fitting a regression model, the residuals are assumed to be normally and independently distributed with zero mean and constant variance, that is, $e_t \sim NID(0, \sigma^2)$. Any violation of such assumption would seriously degrade the adequacy of the constructed model. Since the annual peak load demand records in power systems are of time series type of data, the assumption of independent errors is often violated and the residuals are said to be autocorrelated. A fitted model with autocorrelated errors will likely have over- or underestimated standard errors in addition to inflated t-statistics and coefficient of determination R^2. Such violation will also increase the probability of Type I or II depending on the autocorrelation type. Such symptoms lead to unreliable hypothesis testing. Thus, Durbin-Watson (DW) statistical test (Durbin and Watson 1950; Durbin and Watson 1971; Durbin and Wheeler 1951) is utilized in detecting the presence of autocorrelation, that is, testing the lack of independence among the fitted model residuals. The DW test hypotheses and test statistic are formulated as expressed in (7) and (8), respectively.

$$H_o : \rho = 0$$
$$H_1 : \rho > 0 \tag{7}$$

$$d = \frac{\sum_{t=2}^{m} \left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{m} e_t^2} \quad , \; t = 1,\; 2,\ldots,m \tag{8}$$

where H_o, H_1, ρ, and e_t are the DW statistical test null and alternative hypotheses, the autocorrelation parameter, and the fitted model residuals, respectively. For a specific sample size, number of regressors, and type I error rate α, DW test statistic has a tabulated lower and upper values, (d_L and d_U), and the corresponding decision to be drawn depends on whether the test statistic value d is lower than d_L, higher than d_U, or lies in between. Smaller values of d than d_L implies that the null hypothesis should be rejected in favor of the alternative hypothesis, and the fitted model errors are autocorrelated. Higher d values than d_U lead to conclude $H_o : \rho = 0$, i.e.that is, no evidence of autocorrelation in the forecasting model residuals. On the other hand, the test is inconclusive if the d value lies between d_L and d_U.

Eliminating the autocorrelated residuals problem from the fitted model can be achieved by using Cochrane-Orcutt (Cochrane and Orcutt 1949; Montgomery et al. 2008) remedial measure, which is based on employing transformed variables via an estimator r defined in (9).

$$r = \frac{\sum_{t=2}^{m} e_{t-1}\, e_t}{\sum_{t=2}^{m} e_{t-1}^2} \tag{9}$$

By obtaining r initially from fitting the ill-conditioned autocorrelated original model, the response as well asand the k predictors are, respectively, transformed to $y_t^* = y_t - ry_{t-1}$ and $X_t^* = X_t - rX_{t-1} \;\; \forall\; t = 2,\ldots,m$, and, accordingly, the transformed model can be reformulated as expressed in (10) and (11).

$$y_t^* = \beta_0 \left(1 - r\right) + \sum_{i=1}^{k} \beta_i X_{it}^* + \varepsilon_t \tag{10}$$

$$y_t - ry_{t-1} = \beta_0 \left(1 - r\right) + \sum_{i=1}^{k} \beta_i \left(X_{it} - rX_{i(t-1)}\right) + \varepsilon_t \tag{11}$$

After fitting the transformed model, the DW statistic test is employed again for testing autocorrelation. The process is concluded if $d > d_U$, and, consequently, the fitted model is transformed back to its original variables. However, if the autocorrelation between the fitted model errors persists, Cochrane-Orcutt corrective action is repeated with an updated r form from the most recent fitted model errors until the DW statistic test d rejects the alternative hypothesis H_1 expressed in (7).

CUCKOO SEARCH ALGORITHM

Cuckoo search algorithm (CSA) method is a relatively new nature-inspired heuristic global search algorithm that was introduced by Yang and Deb in 2009 to solve a wide spectrum of optimization problems. This class of optimization algorithms is getting added attention in recent years mainly due to their high efficiency in searching the solution hyperspace using high level yet simple strategies. Scientists and researchers are using collective intelligent behaviors of some primitive species to develop smart and effective algorithms that can overcome many of the existing algorithms' shortcomings.

CSA is a population-based gradient-free search procedure that is inspired by cuckoo birds' aggressive parasitic breeding behavior while laying their own eggs combined with an efficient flight strategy. Cuckoo bird species use other bird nests to lay their own eggs and might even mimic the host egg outer shell to increase the probability of being undiscovered. When hatched, the new hatchlings take advantage of the food provided obliviously by the foster host birds, and sometimes the cuckoo deliberately remove some of the host bird eggs to increase their baby's food share. Such parasitic behavior has its own risk the cuckoo birds are willing to take. That is, the host bird might discover the alien egg and consequently either remove them or abandon the nest to build a new one.

CSA implements the following three rules towards finding the optimal solution: cCuckoo birds lay their eggs in randomly chosen nests; the best nest that carries the best solution is the one to be carried over to the next iteration; and alien eggs will be discovered by host bird with a probability $P_a \in [0,1]$. CSA, as a meta-heuristic algorithm, starts with initial random estimate and searches for the global optimum solution using exploration and exploitation strategies. The eExploration component of CSA explores the search space globally using large jumps to avoid possible stagnation and thus diversifies the solution search space,; while the exploitation process exploits worthy positive feedback information sent from a feasible solution to locally search the respective space intensively.

It was shown in the literature that the movement of animals when searching their landscape generates Levy flights pattern. In CSA endeavor towards carrying out the global search, the algorithm, while mimicking cuckoo's breeding behavior, utilizes Levy flights in generating new point estimate during the solution process. Such combination is effectively utilized in solving constrained nonlinear, mix-integer, nonconvex, and even non-differentiable optimization problems. Levy flights are a modified class of random walks,; in which its step length is not constant but rather are is randomly drawn from Levy distribution. Allowing the search step length to be drawn from such heavy-tailed probability distribution enables the step length to have large jumps and therefore explore the search space efficiently. Compared to other heuristic optimization methods, CSA is easy to implement and has fewer parameters that are needed to be specified and adjusted. They are, namely, the total nests number and the probability P_a. Moreover, it was shown by Yang et al. (Yang et al., 2013; Yang and Deb 2009) that the probability parameter has a moderate effect on the CSA rate of convergence.

CSA formulation

The Levy search length along with an appropriate step size scalar constitutes the next approximated solution point that would be utilized in starting a new CSA iteration. The next new generated solution for m dimension optimization problem is expressed as

$$\boldsymbol{\beta}^{(n+1)} = \boldsymbol{\beta}^{(n)} + \alpha\, \mathbf{S}_{Levy} \delta \boldsymbol{\beta}^{(n)} \mathrm{rand}n(m) \tag{12}$$

where n, $\beta^{(n+1)}$, $\beta^{(n)}$, α, S_{Levy}, $\delta\beta^{(n)}$, and *randn (m)* are, respectively, the iteration index, the new and the old feasible solution vector, the step size, the search step length performed via Levy flights, difference factor vector, and a random vector of m numbers drawn from standard normal distribution with a zero mean and unity standard deviation, $N(0,1)$. The step size α can be represented as a constant positive random number. It determines how far the random walker will travel for a specific number of iterations. If is too small, the search will be intensified and the change will be too small to be significant. On the other hand, if it is too large, the next solution point will be either too far from the old one or might even step out of the problem boundaries;, i.e.that is, this will increase the diversification and reduce the intensification process greatly. It might be a fixed step, i.e.that is, $\alpha \in [1,2]$, or a varying number. (Yang, X.-S. & and Deb, S. (2009) suggested that α=1 for most cases; (Walton et al. (2011) propose that α=0.1 for small scale problems; and (Yang, X.-S. (2013) proposed that, the (α/L) ratio of the step size to the optimization problem scale, can be bounded between 0.01 to and 0.001 for most applications. The step size could be also represented as a function that decreases with increasing the number of CSA iterations (Valian et al., 2011; Walton et al., 2011).

The difference factor$\delta\beta^{(n)}$is the difference between the best global solution so far and the up-to-date solution point. The Levy step length $\mathbf{S}_{Levy}$ is generated randomly according to Mantegna's algorithm (Mantegna 1994) as follows:

$$\mathbf{S}_{Levy} = \frac{u}{|v|^{1/\beta}}, \tag{13}$$

where u is a stochastic variable drawn from a normal distribution given as $\upsilon \sim N(0, \sigma_u^2)$. The corresponding variance σ_u is calculated according to

$$\sigma_u = \left[\frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma((1+\beta)/2)\beta\, 2^{(\beta-1)/2}} \right]^{1/\beta} \tag{14}$$

where Γ(•) denotes the gamma function, β is the Levy flight control parameter or scale factor and is bounded as $\beta \in [1,2]$ (Kaveh, A., Bakhshpoori, T. 2013). This factor is utilized to estimate the step length in Mantegna's algorithm. Yan and Deb advised using β=1.5 in their CSA original standard implementation (Yang, X.-S. & Deb, S. 2010) .

On the other hand, the stochastic number v is drawn from a standard normal distribution, i.e.that is, $\upsilon \sim N(0,1)$. As an illustration, for an m-dimension optimization problem, that is, R^m, the Levy step length for the i^{th} element is mathematically expressed as:

$$\mathbf{S}_{Levy}^{(i)} = \frac{\mathrm{rand}n(m)_i \cdot \sigma_u}{|\mathrm{rand}n(m)_i|^{1/\beta}} \tag{15}$$

RESULTS AND ANALYSIS

In this paper, the CSA technique was utilized as a long-term forecasting tool so as to predict the demand peak load of two middle eastern countries, i.e.that is, Egypt and Kuwait. Each national network data was divided into two sets;, the first one was employed in performing the time series regression by means of CSA, PSO, and least squares (LS) methods and consequently formulating the long-term forecasting regression model, and, afterwards, the second set was used to validate the prediction model. The comparison between all the forecasting models were was performed via the percentage average absolute error (AAE) that is mathematically expressed as follows:

$$\%AAE = \frac{\sum_{j=1}^{m} \left| \frac{P_{Actual}^{(j)} - P_{Calculated}^{(j)}}{P_{Actual}^{(j)}} \right|}{m} \times 100\% = \frac{\sum_{j=1}^{m} AE^{(j)}}{m} \times 100\% \qquad (16)$$

The CSA method was implemented using Matlab® R2012a, and the cuckoo bird in the proposed heuristic method was assumed to randomly lay one egg at a time in a chosen nest. The CSA number of nests is 25, and a probability of 25% was assigned for the host bird to discover the alien egg. The CSA iterative solution method was assigned a 0.01 step size so as to restrict the Levy jumps from leaving the problem search space.

Egyptian power network

The Egyptian demand peak load data was collected over 30 years covering the period from 1977 to 2006 (Ministry Of Electricity & Energy - Egyptian Electricity Authority 2006). The data was divided into regression and prediction sets; the former set $\in [1977,2001]$, while the latter $\in [2002,2006]$. The Daniels' test for stationarity revealed that the null hypothesis H_0 is to be rejected, and therefore there is a trend in the peak load time series data. Accordingly, linear, quadratic, and cubic polynomial regression models were fitted using the proposed CSA method and compared against PSO and LS methods. The DW test was utilized in examining the fitted models' residuals against autocorrelation. The DW test revealed that all the fitted models were ill-conditioned and autocorrelation exists among the residuals. Therefore, the Cochrane-Orcutt remedial measure was applied, and all variables were transformed accordingly. Next, forecasting regression models were fitted to the obtained transformed dependent and predictor variables. Table 1 shows the improvement in the standard error of the peak demand estimate after applying the remedial measure.

Table 1. Comparison of the standard error of the response of all formulations.

	Standard Error of Peak Demand (MW)	
	Original Model	**Transformed Model**
Linear Model	382.769	221.546
Quadratic Model	349.894	211.189
Cubic Model	225.253	155.851

The estimated parameters of the Egyptian system forecasting models obtained by CSA, SA, GA, PSO, and LS methods are shown in Table 2. Figures 1‒3 shows the Egyptian linear, quadratic, and cubic forecasting models, respectively, for CSA, SA, GA, PSO, and LS methods; while Figure 4 shows the proposed CSA three forecasting models. The resultant residuals of the CSA models showed a constant variance pattern. All the models' obtained residuals were tested for normality using Anderling-Darling (AD) test, and with 5% significance level, the results showed strong evidence upon the null hypothesis of the residuals being from normal population that should not be rejected. The AD p-values for the CSA linear, quadratic, and cubic models were respectively 0.106, 0.172, and 0.368.

Comparing the CSA results with those of the other heuristic and deterministic technique', all the CSA forecasting models had the least %AAE for both the full and the projected data as illustrated in Table 3. Since the CSA based cubic model had the lowest %AAE, it was selected to be the model that would be employed in forecasting the Egyptian demand peak as follows:

$$P_t = 9.6129 + 794.5553\,t - 32.023\,t^2 + 0.8011\,t^3 \tag{17}$$

Table 2. Estimated parameters of the Egyptian forecasting models based on CSA, SA, GA, PSO, and LS methods.

	Parameters / Methodology	β_o	β_1	β_2	β_3
Linear Model	CSA	-697.2713	549.7600		
	SA	-791.1288	536.0003		
	GA	-379.3177	489.8789		
	PSO	-781.3230	537.8410		
	LS	-753.3839	505.6083		
Quadratic Model	CSA	4116.6650	33.1742	12.5353	
	SA	4657.7549	-11.7621	12.9596	
	GA	4452.8946	8.9139	12.0647	
	PSO	4255.4930	38.4820	11.8800	
	LS	4422.6751	13.4742	11.9174	
Cubic Model	CSA	9.6129	794.5553	-32.0230	0.8011
	SA	23.2833	799.3654	-33.8472	0.8508
	GA	18.0427	783.1851	-32.3231	0.8164
	PSO	312.1100	1009.7600	-46.1200	1.0309
	LS	10.3724	951.4688	-51.3314	1.3188

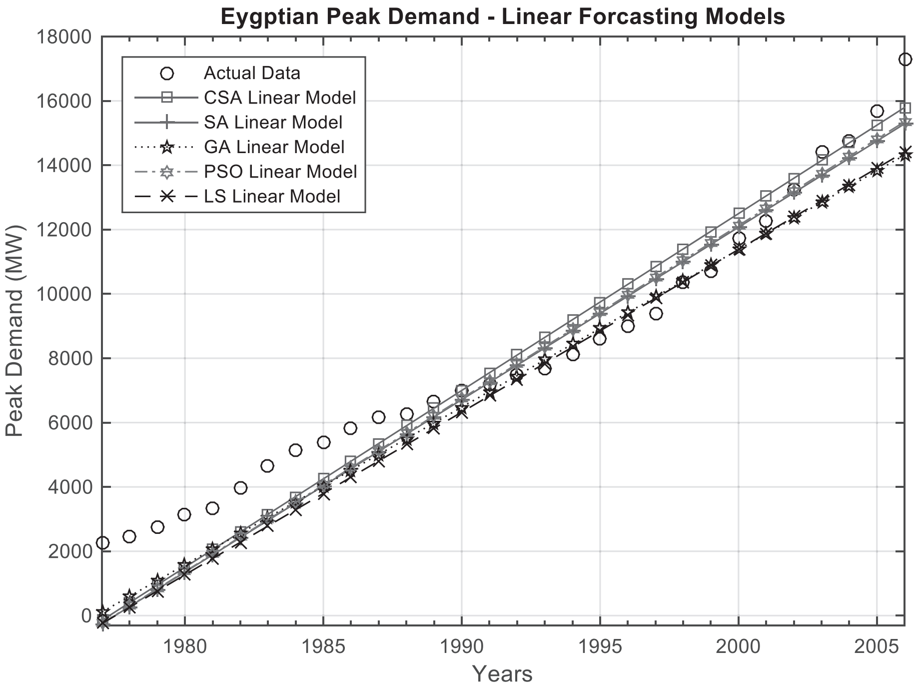

Fig. 1. Egyptian Peak Demand - Linear Forecasting Models.

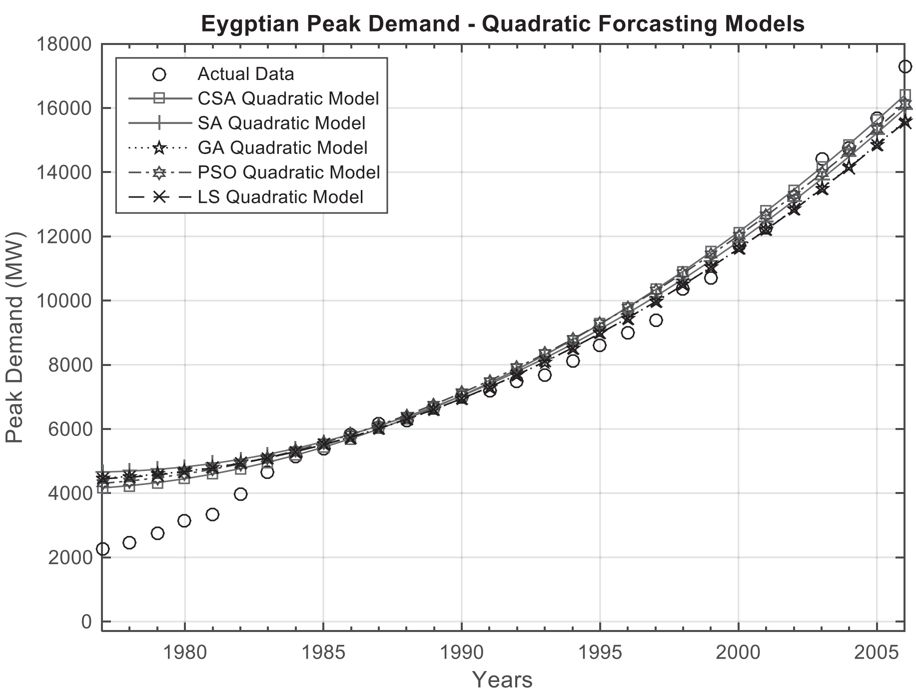

Fig. 2. Egyptian Peak Demand - Quadratic Forecasting Models.

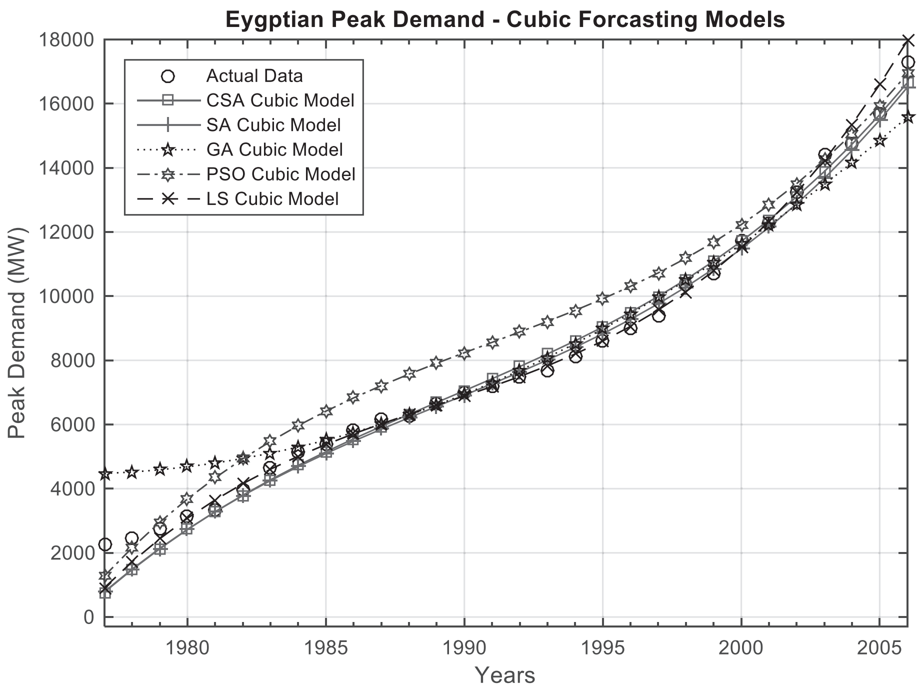

Fig. 3. Egyptian Peak Demand - Cubic Forecasting Models.

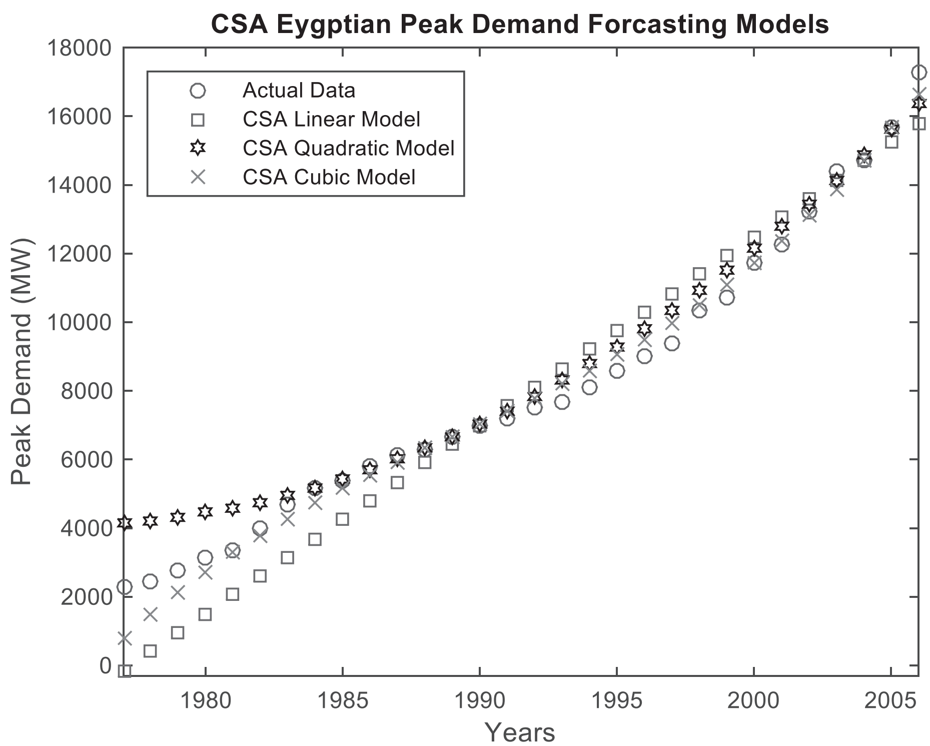

Fig. 4. CSA Egyptian Peak Demand Forecasting

Table 3. Linear, quadratic, and cubic forecasting models for the Egyptian network.

		Linear Model					Quadratic Model					Cubic Model				
	Actual P_D	**CSA**	**SA**	**GA**	**PSO**	**LES**	**CSA**	**SA**	**GA**	**PSO**	**LES**	**CSA**	**SA**	**GA**	**PSO**	**LES**
Year	**(MW)**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**	**% AE**
1977	2284	106.458	111.170	95.159	110.660	110.848	82.241	103.982	95.879	88.523	94.749	66.158	65.427	66.299	44.099	60.078
1978	2449	83.575	88.531	75.482	87.980	89.472	72.852	91.346	84.524	78.848	83.638	39.688	39.019	40.316	11.989	29.827
1979	2742	65.280	70.209	60.236	69.650	72.158	57.878	72.834	67.331	63.306	66.680	22.440	21.965	23.460	7.737	11.072
1980	3161	52.491	57.201	50.010	56.658	59.853	40.776	52.422	48.105	45.508	47.651	13.738	13.520	15.031	16.394	2.583
1981	3350	38.760	43.616	38.207	43.048	47.025	37.192	46.953	43.256	41.639	42.925	2.031	2.081	3.644	29.456	8.934
1982	3981	34.657	39.089	35.696	38.565	42.721	19.743	26.946	24.107	23.438	23.902	4.618	4.930	6.308	23.914	4.399
1983	4672	32.555	36.625	34.721	36.139	40.371	6.231	11.525	9.299	9.311	9.181	8.451	8.987	10.177	17.169	1.375
1984	5158	28.251	32.205	31.374	31.729	36.187	0.510	4.557	2.682	3.212	2.621	8.361	9.120	10.182	15.671	2.828
1985	5361	20.713	24.774	24.835	24.282	29.172	1.298	4.488	2.786	3.789	2.765	3.922	4.940	5.919	19.674	0.302
1986	5803	17.278	21.267	22.118	20.781	25.854	1.742	0.570	0.939	0.436	0.928	4.291	5.515	6.359	17.673	1.590
1987	6152	13.035	17.021	18.574	16.532	21.842	2.498	0.902	2.296	0.581	2.261	3.426	4.858	5.582	17.215	2.133
1988	6279	6.038	10.163	12.419	9.655	15.370	0.650	1.653	0.290	2.373	0.342	0.609	1.070	1.697	20.550	0.576
1989	6664	3.217	7.310	10.127	6.804	12.672	0.036	0.466	0.845	1.493	0.782	0.345	1.501	2.004	18.691	0.932
1990	7004	0.066	4.157	7.496	3.649	9.693	0.486	0.416	0.880	1.695	0.812	0.730	1.273	1.664	17.618	1.644
1991	7215	4.631	0.470	3.411	0.988	5.326	3.045	2.526	1.194	4.029	1.264	2.931	0.755	0.463	18.652	0.432
1992	7503	7.942	3.757	0.590	4.280	2.221	4.711	3.788	2.414	5.458	2.480	4.037	1.732	1.535	18.406	0.108
1993	7657	12.951	8.670	3.809	9.207	2.416	8.441	7.132	5.670	8.959	5.731	7.068	4.616	4.501	20.335	2.258
1994	8108	13.449	9.236	4.076	9.766	2.955	8.229	6.623	5.110	8.501	5.161	6.169	3.688	3.648	17.871	1.095
1995	8609	13.232	9.105	3.710	9.625	2.836	7.704	5.851	4.282	7.740	4.319	5.013	2.541	2.560	15.217	0.065
1996	9017	14.206	10.113	4.450	10.630	3.790	8.620	6.536	4.880	8.430	4.903	5.360	2.893	2.955	14.299	0.452
1997	9369	15.783	11.697	5.755	12.214	5.288	10.379	8.079	6.315	9.966	6.321	6.651	4.198	4.289	14.475	2.120
1998	10346	10.163	6.330	0.503	6.816	0.232	5.486	3.146	1.376	4.891	1.364	1.690	0.578	0.479	8.076	1.981
1999	10694	11.719	7.882	1.813	8.369	1.698	7.638	5.132	3.237	6.837	3.206	3.714	1.503	1.601	9.237	0.860
2000	11726	6.575	2.958	2.970	3.419	2.940	3.472	0.974	0.937	2.524	0.985	0.153	2.157	2.081	4.316	1.841
2001	12276	6.278	2.712	3.326	3.166	3.170	4.110	1.527	0.487	2.986	0.555	0.817	1.058	1.014	4.583	0.371
2002	13226	2.801	0.613	6.566	0.177	6.302	1.717	0.857	2.915	0.460	3.001	0.948	2.623	2.623	2.128	0.016
2003	14401	1.769	5.000	10.788	4.587	10.437	1.739	4.258	6.335	3.097	6.435	3.577	5.026	5.082	1.084	1.134
2004	14735	0.265	3.516	9.486	3.100	9.035	0.938	1.671	3.894	0.598	4.017	0.013	1.287	1.412	2.187	4.231
2005	15678	2.757	5.901	11.805	5.498	11.282	0.364	2.949	5.231	2.012	5.371	0.125	1.207	1.408	1.737	5.868
2006	17300	8.696	11.625	17.242	11.249	16.677	5.239	7.696	9.951	6.925	10.101	3.727	4.552	4.827	2.135	3.840
Total % AAE		21.505	22.097	20.225	22.348	23.328	13.532	16.260	14.915	14.919	14.815	7.693	7.487	7.971	14.420	5.165
%AAE for Projected 2002-2006		**3.258**	5.331	11.178	4.922	10.747	**1.999**	3.486	5.665	2.618	5.785	**1.678**	2.939	3.071	1.854	3.018

Kuwaiti power network

Daniels' test for stationarity was performed upon Kuwait power time series data, and the test resulted in accepting the alternative hypothesis H_1, i.e.that is, leading to the conclude conclusion that there is a trend in the peak load time series data. Kuwait power data was not transformed since the DW tests for all the models turned to be insignificant based on 5% p-value level. As with the Egyptian case, the response variable was regressed against the predictor variables to obtain the linear, quadratic, and cubic models using CSA, SA, GA, PSO, and LS methods. The Kuwaiti CSA forecasting models parameters are shown in Table 4, and Figures 57- shows the Kuwaiti linear, quadratic, and cubic forecasting models, respectively, for CSA, SA, GA, PSO, and LS methods; while Figure 8 shows the proposed CSA three forecasting models. The residuals' variance of the fitted models were was homoscedastic, and the AD test for normality showed that their population follow the normal distribution pattern. The CSA corresponding p-values for AD tests at 5% significance level for the three prediction models are 0.142, 0.113, and 0.0620.

The CSA models showed better performance in terms of %AAR as shown in Table 5. It is evident from the corresponding Table that the difference between CSA quadratic and cubic models is insignificant, and to maintain parsimony principle, the simplest model was preferred to represent the Kuwait demand peak as expressed in Eq. (18). It is worth mentioning that the proposed CSA prediction model provided better demand peak estimation for the years 2009 to 2012 than the official prediction of the Kuwaiti Ministry of Electricity and Water (MEW 2011)(Ministry of Electricity and Water 2011) as illustrated in Table 6.

$$P_t = 4116.6650 + 12.5353\,t + 33.1742\,t^2$$

$$P_t = 3745.0122 + 213.7682t + 8.12426t^2 \quad (18)$$

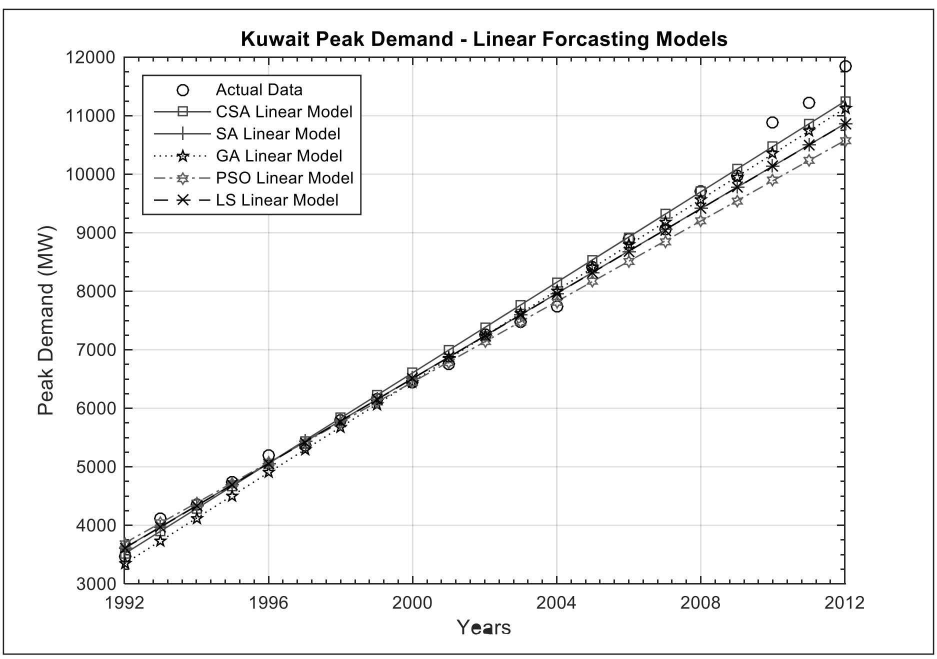

Fig. 5. Kuwait Peak Demand - Linear Forecasting Models.

Table 4. Estimated parameters of the Kuwaiti forecasting models based on CSA, SA, GA, PSO, and LS methods.

Model	Methodology \ Parameters	β_o	β_1	β_2	β_3
Linear Model	CSA	3126.7700	386.6000		
	SA	3250.0866	362.48642		
	GA	2955.2102	388.9471		
	PSO	3351.3861	344.0350		
	LS	3244.7500	362.6029		
Quadratic Model	CSA	3745.0122	213.7682	8.1426	
	SA	3299.0348	342.53758	1.220205	
	GA	3566.1609	308.2908	1.9983116	
	PSO	3207.6570	390.0020	-2.8310	
	LS	3313.5893	339.6565	1.3498	
Cubic Model	CSA	3089.8570	469.7721	-17.0361	0.6938
	SA	3097.0367	467.53138	-16.780721	0.7101937
	GA	3088.3317	479.49777	-18.63837	0.7840087
	PSO	3091.8570	450.1110	-16.0361	0.7938
	LS	3095.7143	473.5893	-17.7621	0.7495

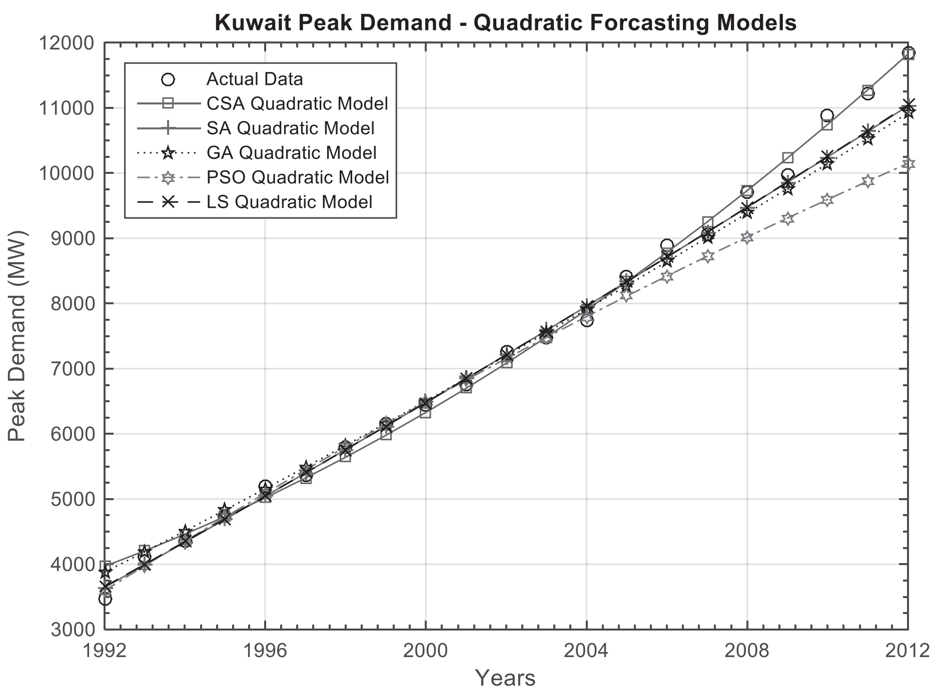

Fig. 6. Kuwait Peak Demand - Quadratic Forecasting Models.

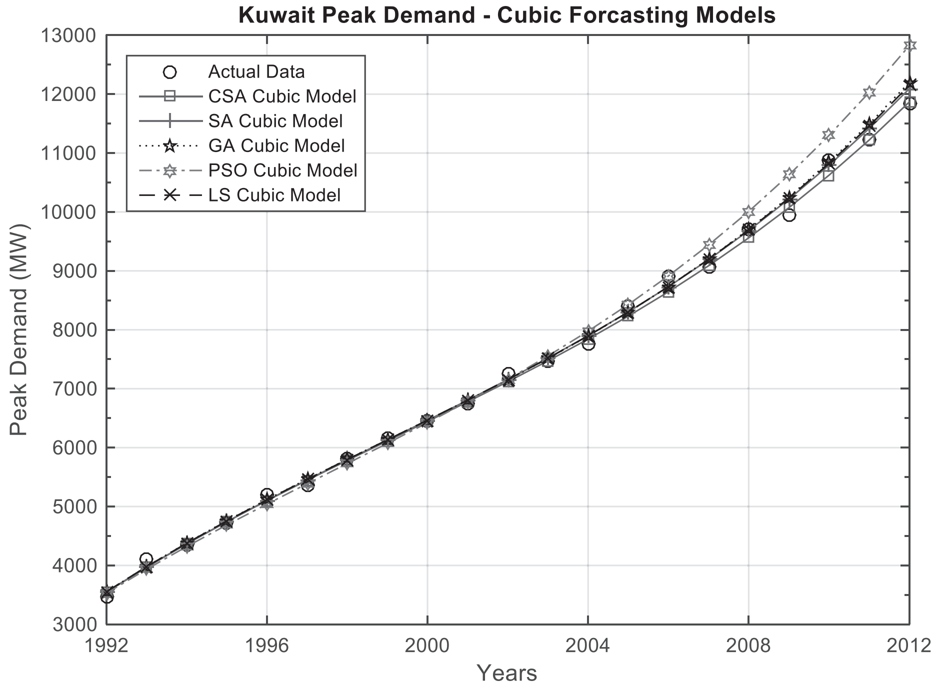

Fig. 7. Kuwait Peak Demand - Cubic Forecasting Models.

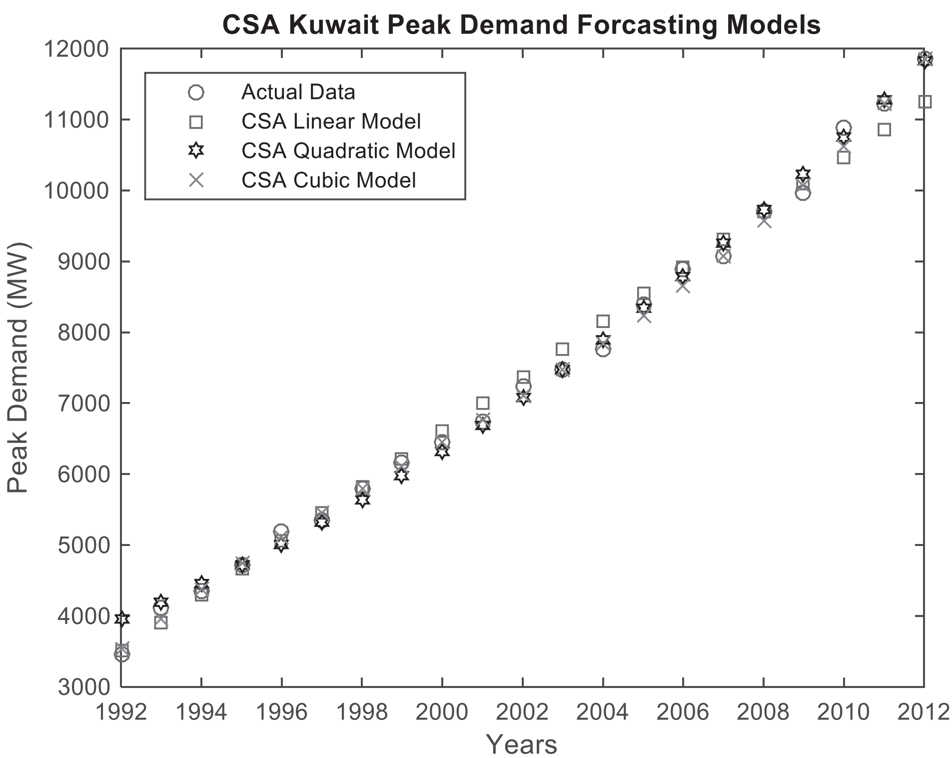

Fig. 8. Kuwait Peak Demand Forecasting Models.

Table 5. Linear, quadratic, and cubic forecasting models for the Kuwaiti network.

		Linear Model					Quadratic Model					Cubic Model				
	Actual P_D	CSA	SA	GA	PSO	LES	CSA	SA	GA	PSO	LES	CSA	SA	GA	PSO	LES
Year	(MW)	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE	% AE
1992	3460	1.542	4.410	3.348	5.262	4.259	14.651	5.283	12.036	3.897	5.624	2.407	2.558	2.600	1.928	2.667
1993	4120	5.341	3.518	9.391	3.386	3.642	2.066	3.180	1.717	3.487	2.954	3.718	3.625	3.421	4.508	3.450
1994	4350	1.458	0.286	5.240	2.228	0.401	2.520	0.284	3.656	0.050	0.122	0.335	0.409	0.695	0.706	0.617
1995	4730	1.201	0.634	4.630	1.149	0.737	0.008	0.873	2.142	0.161	0.765	0.228	0.299	0.598	0.919	0.504
1996	5200	2.697	2.644	5.770	0.227	2.735	3.511	3.034	0.816	2.175	2.969	1.932	1.847	1.580	3.063	1.667
1997	5360	1.611	1.213	1.327	0.574	1.126	0.732	0.712	2.385	1.600	0.749	1.587	1.707	1.934	0.498	1.860
1998	5800	0.568	0.216	2.106	1.264	0.294	2.752	0.748	0.381	0.018	0.736	0.320	0.153	0.008	1.222	0.042
1999	6160	0.967	0.163	1.513	1.882	0.234	2.982	0.691	0.006	0.219	0.694	0.764	0.537	0.440	1.415	0.464
2000	6450	2.421	0.968	0.089	2.457	0.902	1.884	0.476	0.816	0.595	0.462	0.098	0.206	0.243	0.425	0.243
2001	6750	3.597	1.851	1.403	2.978	1.789	0.786	1.429	1.465	1.105	1.409	0.411	0.807	0.792	0.491	0.813
2002	7250	1.784	0.173	0.226	3.360	0.229	2.321	0.488	0.701	1.309	0.508	1.801	1.315	1.364	1.252	1.330
2003	7480	3.823	1.603	1.906	3.826	1.551	0.037	1.406	0.981	0.000	1.388	0.096	0.511	0.443	1.012	0.483
2004	7750	5.194	2.741	3.374	4.242	2.691	1.937	2.687	2.086	0.635	2.674	1.188	1.926	1.863	2.931	1.899
2005	8400	1.657	0.894	0.006	4.420	0.938	0.789	0.789	1.501	3.419	0.794	2.007	1.164	1.196	0.340	1.176
2006	8900	0.290	2.389	1.243	4.650	2.429	1.308	2.116	2.920	5.385	2.111	2.866	1.896	1.874	0.163	1.881
2007	9070	2.672	0.222	1.195	5.032	0.260	1.983	0.243	0.657	3.826	0.261	0.185	1.330	1.435	4.077	1.389
2008	9710	0.114	3.065	1.470	5.139	3.100	0.230	2.422	3.351	7.111	2.391	1.533	0.261	0.050	3.082	0.147
2009	9960	1.261	1.859	0.038	5.437	1.892	2.721	1.003	1.979	6.522	0.956	1.128	2.585	2.940	6.703	2.775
2010	10890	3.837	6.912	5.003	5.364	6.940	1.321	5.898	6.840	11.885	5.837	2.440	0.886	0.387	3.761	0.620
2011	11220	3.220	6.419	4.330	5.585	6.446	0.512	5.188	6.138	11.985	5.110	0.011	1.755	2.451	7.220	2.126
2012	11850	5.102	8.335	6.134	5.648	8.359	0.211	6.916	7.835	14.353	6.822	0.147	2.042	2.945	8.216	2.523
Total % AAE		**2.398**	2.405	2.845	3.529	2.426	**2.155**	2.184	2.877	3.797	2.159	**1.200**	1.325	1.393	2.568	1.366
%AAE for Projected 2008-2012		**2.707**	5.318	3.395	5.435	5.347	**0.999**	4.286	5.229	10.371	4.223	**1.052**	1.506	1.755	5.796	1.638

Table 6. Comparison between CSA and MEW forecasts.

Year	Actual Demand Peak (MW)	MEW Forecast (MW)	CSA Forecast		
			Linear	Quadratic	Cubic
2009	9960	11385	10085.57	10231.05	10072.3
2010	10890	12520	10472.17	10746.1	10624.27
2011	11220	14255	10858.77	11277.43	11221.26
2012	11850	15395	11245.37	11825.04	11867.43

CONCLUSIONS

A new application of CSA is presented in this paper for long-term load forecasting in power systems. The problem is formulated, and different models are presented. The goal is to minimize the error associated with long-term load modeling process. Real data from Kuwaiti and Egyptian networks are used to evaluate the CSA outcomes. Three forecasting models are used in this study, and the results showed that CSA performed well on all cases. Extensive statistical tests have been applied to test the error auto-correlation. Test results revealed that the resultant error in the Egyptian network case are is highly correlated and thus, remedial measures were taken. A comparison among between CSA, SA, GA, PSO, and LS is presented. This clearly states that the developed algorithm performed well in estimating load forecast model characteristics. In future work, other factors that might impact annual load forecast such as oil prices and social welfare may be incorporated in forecast modeling.

REFERENCES

AlFares, H.K. & Nazeeruddin, M. 2002. Electric load forecasting: literature survey and classification of methods. International Journal of Systems Science, 33, (1) 23–24.

Cassola, F. & Burlando, M. 2012. Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. Applied Energy, 99, (0) 154–166 available from: http://www.sciencedirect.com/science/article/pii/S0306261912002747

Cetinkaya, N. 2013. Long-term Electrical load forecasting based on economic and demographic data for Turkey. Proceedings of IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI), Budapest.

Cochrane, D. & Orcutt, G.H. 1949. Application of least squares regression to relationships containing auto-correlated error terms. Journal f the American Statistical Association, 44, 32–61.

Daniels, H.E. 1950. Rank Correlation and Population Models. Journal of Royal Statistical Society, B, (12) 171–181.

Durbin, J. & Watson, G.S. 1950. Testing for serial correlation in least squares regression I. Biometrika, 37, 409–428.

Durbin, J. & Watson, G.S. 1971. Testing for serial correlation in least squares regression III. Biometrika, 58, 1–19.

Durbin, J. & Wheeler, J.E. 1951. Testing for serial correlation in least squares regression II. Biometrika, 38, 159–178.

Farnum, N.R. & Stanton, L.W. 1989. Quantitative Forecasting Methods PWS-KENT Publishing Company.

IEEE Power System Engineering Committee 1980. Load Forecast Bibliography Phase I. IEEE Transactions on Power Apparatus and Systems, PAS-99, (1) 53–58.

Jabr, R.A. 2006. Power system state estimation using an iteratively reweighted least squares method for sequential L1-regression. International Journal of Electrical Power & Energy Systems, 28, (2) 86–92 available from: http://www.sciencedirect.com/science/article/pii/S014206150500133X

Kaveh, A. &, Bakhshpoori, T. 2013. Optimum designe of space trusess using Cucko search algorithm with leacy flights. Iranian journal of science & technolog., Transactions of civil engineering, 37, (C1), 1–15.

Kandil, N., Wamkeue, R., Saad, M., & Georges, S. 2006. An efficient approach for short term load forecasting using artificial neural networks. International Journal of Electrical Power & Energy Systems, 28, (8) 525–530 available from: http://www.sciencedirect.com/science/article/pii/S0142061506000676

Karabulut, K., Alkan, A., & Yilmaz, A.S. 2008. Long-term energy consumption forecasting using genetic programming. Mathematical and Computational Applications, 13, (2) 71–80.

Kermanshahi, B. & Iwamiya, H. 2002. Up to year 2020 load forecasting using neural nets. International Journal of Electrical Power & Energy Systems, 24, (9) 789 - 797 available from: http://www.sciencedirect.com/science/article/pii/S0142061501000862

Khuntia, S. R., Rueda, J. L. &and van der Meijden, M. A. M. M. 20016. Forecasting the load of electrical power systems in mid- and long-term horizons: a review. IET Transactions on Generation, Transmission & Distribution. 10, (16), 3971–3977.

Li, H.z., Guo, S., Li, C.j., & Sun, J.q. 2013. A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. Knowledge-Based Systems, 37, (0) 378 - 387 available from: http://www.sciencedirect.com/science/article/pii/S0950705112002353

Mantegna, R.N. 1994. Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes. Physical Review E, 49, 4677–4683.

Ministry Of Electricity & Energy - Egyptian Electricity Authority 2006, Annual Report of Electric Statistics Cairo, Egypt.

Ministry of Electricity and Water 2011, Statistical Year Book on Electric Energy State of Kuwait.

Montgomery, D.C., Jennings, C.C., & Kulachi, M. 2008. Introduction to Time Series Analysis and Forecasting John Wiley & Sons, Inc.

Nie, H., Liu, G., Liu, X., & Wang, Y. 2012. Hybrid of ARIMA and SVMs for Short-Term Load Forecasting. Energy Procedia, 16, Part C, (0) 1455–1460 available from : http://www.sciencedirect.com/science/article/pii/S1876610212002391

Rahman, S. & Hazim, O. 1996. Load forecasting for multiple sites: development of an expert system-based technique. Electric Power Systems Research, 39, (3) 161–169 available from: http://www.sciencedirect.com/science/article/pii/S0378779696011145

Takiyar, S. 20015. Grid reliability enhancement by peak load forecasting with a PSO hybridized ANN model.

4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Noida, India.

Temraz, H.K., El-Naggar, K.M., & Salama, M.M. 1998. Application of noniterative least absolute value estimation for forecasting annual peak electric power demand in the presence of bad data. Canadian Journal of Electric and Computer Engineering, 23, (4) 148–153.

Valian, E., Mohanna, S., & Tavakoli, S. 2011. Improved cuckoo search algorithm for feed forward neural network. International Journal of Artificial Intelligence Applications, 2, (3) 36–43.

Walton, S., Hassan, O., Morgan, K., & Brown, M.R. 2011. Modified cuckoo search: A new gradient free optimization algorithm. Chaos, Solitons and Fractals, 44, 710–718.

Wang, J., Li, L., Niu, D., & Tan, Z. 2012. An annual load forecasting model based on support vector regression with differential evolution algorithm. Applied Energy, 94, (0) 65 - 70 available from: http://www.sciencedirect.com/science/article/pii/S0306261912000165

Willis, H.L. & Northcote-Green, J.E.D. 1984. Comparison Tests of Fourteen Distribution load Forecasting Methods. IEEE Transactions on Power Apparatus and Systems, PAS-103, (6) 1190–1197.

Wu, Q. 2010. Power load forecasts based on hybrid PSO with Gaussian and adaptive mutation and Wv-SVM. Expert Systems with Applications, 37, (1) 194–201 available from: http://www.sciencedirect.com/science/article/pii/S0957417409004308

Yang, X., Cui, Z., Xiao, R., Gandomi, A.H., & Karamanoglu, M. 2013. Swarm Intelligence and Bio-Inspired Computation: Theory and Applications Elsevier.

Yang, X.-S. & Deb, S. 2009. Cuckoo Search via Levy Flights. World Congress on Nature & Biologically Inspired Computing, NaBIC 2009 210–214.

Yang, X.-S. & Deb, S. 2010. Engineering optimisation by cuckoo coucko search. International journal of mathematical modeling and numerical optimisation. 1, (4), 330–343.

Yang, X.-S. 2013. Optimization and metaheuristic metahueristic algorithms in engineering. In: metaheuristic metahuersistic in water, geotechnical and transport engineerings. ElsevierElsavier. 1–23.

Submitted: 15/11/2016
Revised : 06/07/2017
Accepted : 30/07/2017

Journal of Engg. Research **Vol. 6 No. (3) September 2018 pp.** *135*

توقعات الحمل الكهربائي طويل الأجل
في أنظمة الطاقة الكهربائية الكويتية والمصرية

محمد الهاجري، محمد الرشيدي وخالد النجار

قسم تكنولوجيا الهندسة الكهربائية، كلية الدراسات التكنولوجية، الهيئة العامة للتعليم التطبيقي والتدريب، الكويت

الخلاصة

يقدم هذا البحث منهجية فعالة للتنبؤ بمتطلبات وقت الذروة سنوياً في أنظمة الطاقة الكهربائية. تم تطوير المنهجية المقترحة كطريقة تنبؤ بديلة للطرق الأخرى القائمة. وتستند الطريقة على خوارزمية بحث الوقواق cuckoo. تم استخدامها لتقليل الخطأ المرتبط بمعلمات النموذج المُقدرة للمتطلبات المتوقعة لتتبع البيانات الحقيقية للأحمال. تم استخدام البيانات الحقيقية من الشبكات الكويتية والمصرية لإجراء هذه الدراسة. تم استخدام ثلاثة نماذج تنبؤ طويلة المدى في هذا العمل البحثي لقياس متانة أداة التقدير المتطورة. تم إجراء اختبار إحصائي Durbin-Watson للتحقق من مدى ملائمة النماذج المختارة وتم تطبيق تحويل النموذج كإجراء علاجي لبيانات سلسلة زمنية سيئة الشروط عند الحاجة. تمت كتابة تقرير عن نتائج التنبؤات ومقارنتها بتلك التي تم الحصول عليها باستخدام تقنيات التنبؤ الأخرى. تمت دراسة وتقييم أداء الطريقة المقترحة. وكشفت النتائج أن خوارزمية بحث الوقواق لها إمكانات واعدة كأداة قابلة للتطبيق لتقدير المعلمات.