Abstract—This paper uses Artificial Neural Networks (ANN) for Short Term Load Forecasting (STLF) for a residential area in Yanbu Industrial City (YIC), an industrial city in the western coast of the Kingdom of Saudi Arabia (KSA). Three years data were collected for a residential substation in this city. In recent years, load forecasting raised large interest in power area in KSA. This is due to the increasing rise in number of population, expansion in residential construction, economic growth rate and the rapid developments in the Kingdom. STLF is an important study in the area of system operation and planning. The daily load behavior is affected by many factors such as social, religious, official occasions and environmental conditions. In this paper, two ANN models are proposed which are next hour and next day load forecasting. For next day load forecasting, the load is forecasted using ANN model and by iterative using of next hour model. The obtained results for ANN next hour model yield accurate results. For next day load forecasting, the two models yield satisfactory results. Comparative study is conducted to prove the effectiveness of the models proposed. The results obtained in this work are compared with other published work using different method applied to the same data.

Keywords—load forecasting, neural networks

I. INTRODUCTION

The sharp rise in energy consumption, national income, growth rate and future governmental plan make the load forecasting in KSA to be an important topic. KSA is a developing country. The rapid urbanization, economic developments, the substantial increase in oil revenue and the new infrastructure projects cause the electricity demand to grow rapidly. Generation of electricity, transmission and distribution in KSA are very important subjects for the decision makers in the Kingdom making the studies and research in the field of electricity demand to be very useful.

In YIC, an industrial city at the western coast of KSA, the load consumption is mainly affected by cooling appliances during summer season since this season is characterized by relatively high temperature and humidity levels. When the temperature increases, the load consumption increases due to the heavy usage of air conditioners and cooling devices.

In general, the problem of KSA load forecasting is very complicated and challenging. Many complex factors affect the load level due to:

- The rapid growth in the economy, larger diversity in commercial activities, and high rate of population rise of the Kingdom.
- The large gap between maximum and minimum temperature values over the year seasons. In July, August till September period, the temperature hits usually high values.
- Special holidays and school days depend on lunar Hijra calendar. Both holidays have a major factor on the load consumption. The most famous religious festivals are: Ramadan (fasting month for Muslims), Eid feter (end of Ramadan), and the first two weeks of the month of Dul-Hijjah (Hajj pilgrimage to Makah). During the month of Ramadan, load consumption changes due to changing in social and commercial activities. Moreover, religious holidays and schools are cyclic but irregular to some extent. The process of demand forecasting becomes more difficult than forecasting common behaviors people’s activities. [1]

Several methods which are varying in the complexity of estimation approaches and functional forms have been proposed in the literature to improve the accuracy of load forecasting. The methods of STLF can be categorized into two main categories: artificial intelligence including Fuzzy Logic Inference [2,3,4], Expert systems [5,6,7], Particle Swarm Optimization (PSO) [8,9,10], and wavelets [11,12,13]; and statistical methods including regression methods [14,15,16], similar day approaches [17,18,19], time series [20,21,22] and support vector regression (SVR) [23,24,25].

The objective of this paper is to build ANN models to reflect several load influential factors including people activities and environmental conditions. The proposed models are next hour and next day load forecasting. In addition to weather conditions, people activities and official occasions are considered such as Ramadan and Eid days. Hourly load, temperature and humidity data were provided for three consecutive years from 2009 to 2011 by Marafiq Company in Yanbu. The data for 2009 and 2010 will be used for training while the 2011 data will be used to test the models.

Ahmad Al-Subhi
Electrical Power Engineering Technology Department
Yanbu Industrial College
Yanbu, Saudi Arabia

Chokri Belhadj Ahmad
Electrical Engineering Department
King Fahd University of Petroleum & Minerals
Dhahran, Saudi Arabia
The paper is structured as follows. Section two will talk about different people activities and weather conditions available in YIC which are the main influential factors on the load consumption. Section three shows the ANN proposed models. Section four shows and analyzes the results obtained from the proposed models followed by a comparative study with other published work in section five and finally, conclusions and recommendations are presented in section six.

II. AFFECTING FACTORS ON POWER CONSUMPTION

An extensive analysis is needed to identify the factors that affect the load consumption. This is done by referring to the academic calendar, official calendar and the weather conditions situation within the entire year. For example, a ten days period load profile of the studied residential area for years 2009 and 2010 containing different intervals like Ramadan, Eid Al-Feter day and after Eid Al-Feter days is shown in Fig. 1.

![Fig. 1: Ramadan days, Eid feter day, after Eid feter days and working day for years 2009 and 2010](image1)

So, as shown in Fig.1, an interesting period is Ramadan month which is the month of fasting. In this lunar month, people activities are totally different from any other month. This month is characterized by different dining hours, less working hours, almost no difference between week days and weekend days. Also, almost there is no difference between schools days and vacations. This is because the load is depending mainly on people activities and fasting time which are almost similar for all days of the month. So, it is clear from Fig 1 that Ramadan days have their own characteristics which are totally different from other days of the year. Because of fasting during the day, there is no presence of the period of lunch time which is the case for other days of the year. Moreover, the evening load is higher in Ramadan days with extended hours.

Eid feter day is the first day coming directly after the month of fasting and it has also a special behavior. Sharp decrease in load consumption is noted in Fig. 1. The residential load at this day is very low due to people outdoor activities especially at morning time. The load behavior at lunch time is back to normal.

After Eid feter days are the group of days coming directly after Eid Feter day. The load increases gradually because people are coming gradually back to normal behaviors. This gradual load rising continues until the starting of the next interval which is normal working days as shown in Fig.1.

In addition, another period containing Eid Hajj day, after Eid Hajj days and a working day is shown in Fig. 2. It is clear that there is a difference between these days according to people activities and weather conditions.

![Fig. 2: Before Eid hajj days, Eid hajj day, after Eid hajj days and working day for years 2009 and 2010](image2)

In Eid Hajj day, a sharp decrease in the load is observed compared to its previous and next days. For After Eid Hajj day mode, the load curve is increasing gradually in almost the same manner of after Eid Feter days mode until the beginning of working days period as shown in Fig.2.

Moreover, a sample week starting from Saturday to Friday during winter period for years 2009 and 2010 can be seen in Fig. 3.
In KSA, working days at the time of data collection were from Saturdays to Wednesdays while weekend days were Thursdays and Fridays. It is obvious that there is a difference between the shape of the load on a typical weekend day, such as Friday and a working day like Saturday or Sunday. The similarity of the days in the same interval for different years is also noted. Same analysis is done for all other intervals of the year like hot periods, mid-year vacation and summer vacation. Therefore, according to the previous figures, it can be concluded that the load consumption is mainly affected by weather conditions and people activities. All of these parameters will be considered and coded as inputs to the ANN models.

III. PROPOSED ANN MODELS

The most important point in ANN model is the selection of inputs parameters, the determination of the number of layers required and the selection of hidden neurons values. Many previous works tried to optimize the selection of these parameters values. However, there is no rule to decide about these parameters. This is highly dependent on the application, nature of data and number of samples used. Usually, the use of trial and error technique is followed in specifying the values of these parameters.

A. Next Hour Load Forecasting Model

This model is basically a two layer feed-forward ANN. It uses one hidden layer containing three neurons with tan-sigmoid transfer functions. The output layer is composed of one neuron using linear function. The ANN input includes hour, day sequence through the year, special occasion (Ramadan, Eid day, vacation ….. etc.), temperature and humidity, all these inputs are for the forecasted hour. Also, previous hours load historical data include all previous 24 hours load consumption. The details of the input vector and model structure for next hour load forecasting to forecast the load at next hour P(h), are as follows:

Inputs:
- All load values of the previous 24 hours to the forecasted hour.
- The next forecasted hour which takes one of the values from 1 to 24.
- Day sequence through the year which takes one of the values from 1 to 365.
- Code number (special day, Ramadan, week end, summer, Eid,…….etc.)
- Expected temperature of the forecasted hour.
- Expected humidity of the forecasted hour.

Output:
- Load value at next hour.

B. Next Day Load Forecasting Model

For next day forecasting, the model used is a three-layer feed-forward ANN, with two hidden layers containing ten and five neurons, respectively with tan-sigmoid transfer functions. Similar to next hour model, the output layer is made of one neuron with linear function behavior. The details of the input vector and model structure related to the forecasted hour, h of day, d are as follows:

Inputs:
- Load, temperature and humidity values at the same hour of the previous day, previous two days, previous three days and previous week.
- Hour from 1 to 24.
- Expected temperature of each forecasted hour.
- Expected humidity of each forecasted hour.
- Code number (special day, Ramadan, week end, summer, Eid,…….etc.)
- Day sequence through the year.

Output:
- Load value of hour, h of next day, d.

C. Iterative Next Day Load Forecasting Model

Next day load forecasting model can be also modeled by repetitive use of the ANN next hour model. This is achieved by forecasting the load of next hour at a time. After that, this load is aggregated to the series, so that the forecasts for the later hours will be based on the forecasts of the earlier ones. Next hour models are heavily dependent on recent hourly loads. Since these load values are forecasted and not measured, forecasting errors are accumulating and the error values will increase accordingly. However, the error is not always increasing with the increase number of the forecasted hours. This depends on the nature of data to be analyzed, inputs parameters and the forecasted hour. The model gives good results when forecasting next 4 to 6 hours.

IV. RESULTS AND ANALYSIS

The load forecast is compared to the actual load data and the error is calculated. The Mean Absolute Percentage Error (MAPE) is used to evaluate the performance of the model. It is defined as:
\[ \text{MAPE} = \frac{1}{k} \sum \left[ \frac{|P_{\text{actual}}(k) - P_{\text{forecasted}}(k)|}{P_{\text{actual}}(k)} \right] \times 100 \]  

(1)

Where \( P_{\text{actual}}(k) \) is the actual load, \( P_{\text{forecasted}}(k) \) is the forecasted load and \( k \) is the number of data points. In this data source, the used data for training are 2009 and 2010, while the forecasted year is 2011 in the basis of next hour and next day load forecasting. Typical four intervals selected from different seasons exposed to different characteristics will be shown and forecasted. Each interval is a one week period starting from Saturday to Friday.

A. Results of Next Hour Load Forecasting Model

For next hour model, comparison between the ANN model output and the actual load for the four intervals are shown in Figs. 4 and 5.

Therefore, as shown from Figs. 4 and 5, next hour model performance is tested for one week in four different seasons; cold week working period, hot week working period, Ramadan and summer vacation, which will evaluate the model performance across different load profiles. The performance of the model gives accurate results for implementation. MAPE for all four intervals, respectively, are 0.4973 %, 0.4826 %, 0.3542 % and 0.3509 % in different load and weather conditions, different calendar times, day types and through special events. This indicates that the model is valid for application through the whole time horizon.

B. Results of Next Day Load Forecasting (ANN Model)

Comparison between the ANN model output and the actual load for the four intervals are shown in Figs. 6 and 7.

As shown, the obtained results from this model are very satisfactory and ranging within good limits. MAPE values for the four typical intervals, respectively, are 2.5818 %, 2.5173 %, 1.4494 % and 2.5247 % which show a good indication for model validity.

C. Results of Next Day Load Forecasting (Iterative Next Hour ANN Model)

The same forecasted periods using ANN model will be forecasted using the iterative ANN next hour model. The results are shown in Figs. 8 and 9.
So, forecasting next day load using iterative model achieved satisfactory results. The MAPE values for all intervals, respectively, are 2.9563 %, 2.6228 %, 2.8272 % and 2.3754 %. With the exception of few hours where MAPE values are relatively high, the results show the effectiveness of this model especially when trying to forecast next 10 to 12 hours. Table I shows MAPE results obtained from the two next day forecasting models.

<table>
<thead>
<tr>
<th>Season</th>
<th>Week</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>ANN</td>
<td>2.5818 %</td>
</tr>
<tr>
<td></td>
<td>iterative</td>
<td>2.9563 %</td>
</tr>
<tr>
<td>Hot</td>
<td>ANN</td>
<td>2.5173 %</td>
</tr>
<tr>
<td></td>
<td>iterative</td>
<td>2.6228 %</td>
</tr>
<tr>
<td>Ramadan</td>
<td>ANN</td>
<td>1.4494 %</td>
</tr>
<tr>
<td></td>
<td>iterative</td>
<td>2.8272 %</td>
</tr>
<tr>
<td>Summer</td>
<td>ANN</td>
<td>2.5247 %</td>
</tr>
<tr>
<td></td>
<td>iterative</td>
<td>2.3754 %</td>
</tr>
</tbody>
</table>

Some sources of errors may be caused by a change in the load profile which is reflected to several possible reasons. The most affecting factor in this situation is the sudden and unpredicted change of weather conditions like sand storm, rainfall or significant change in wind direction. Also, another source of error may include sudden social occasion that cannot be identified from the calendar and as a result of that, it is not considered in the model.

V. MODELS COMPARISON WITH OTHER PUBLISHED WORK

In this section, analysis and results of the models used in this work are compared with a published work that studied STLF based on abductive networks. The paper to compare with is reference [26]. Both approaches are applied to the same set of data. The data source used in this analysis consists of measured hourly load consumption in MW and temperature in Fahrenheit for the Puget power utility, Seattle, USA, over the period from 1 January 1985 to 31 December 1990. Years from 1985 to 1989 are used for training to forecast next hour and day loads for the year 1990. Some load characteristics and data nature will be demonstrated. The concept and history of abductive networks in addition to the analysis and data processing approach are shown in details in [26]. The obtained results from abductive networks model in [26] are compared with the obtained results from the models used in this work; both applied to the same set of data.

A. Nature of the Data: Overview

The data consists of measured hourly load and temperature data for the Puget power utility, Seattle, USA, over the period 1 January 1985 to 31 December 1990. The nature of people activities and weather variables in USA are completely different from KSA. Unlike KSA, the working days for USA are from Mondays to Fridays while weekend days are Saturdays and Sundays. Also, some holidays and special days are present in USA and not present in KSA like New Year’s Day, Labor Day, Thanksgiving Day and Christmas day.

Regarding weather conditions, the seasonally change from winter to spring to fall to summer is noted clearly in USA, whereas in KSA, only summer and winter load curve behavior are present. As in KSA, the effect of week days and weekend days is very significant. Fig. 10 shows a typical week starting from Monday to Sunday for each season.

Therefore, as expected, week days in every season are different in behavior and higher in consumption than weekend days. In special days analysis, there are special activities for people to do at these days. For example, in Thanksgiving Day, human behavior is completely different from any other normal day. So, every special day will be coded and processed as inputs to the ANN models in the same manner done previously for KSA data.

All special days data in USA were found from reference [27]. However, some of them are not official and cannot be identified. So, there is a problem of how to specify whether these days are special or not. The solution to this is simply by plotting these days, and if any day seems...
to have a special behavior, it will be clear and considered as a special day with specific code. If not, it will be considered as a normal day. Fig. 11 shows load profiles of some special days in Seattle, USA, 1989. For example, Columbus Day will be considered as a normal day since it does not have special characteristics that are different from its previous and next days. However, Thanksgiving Day has a unique and special behavior which is different from its previous and next days that supports the decision to consider it as a special day. Same approach is applied to all other days.

B. Next Hour Load Forecasting Models Comparison

For next hour load forecasting modeled in this work, the inputs parameters and values arrangement follow the same approach done for KSA data. For abductive networks model, the input parameters are shown in details in [26]. In comparing the results of both next hour models (abductive networks and ANN models), table II lists the MAPE values for all hours, giving the overall value for the evaluation year as 1.14 % reached by the abductive networks model and 0.4539 % reached by ANN model which indicates the effectiveness of such models for STLF.

From table II, it is clear that next hour load forecasting model using ANN has accurately forecasted next hour load for the year 1990. The performance of the model gives accurate results for implementation and indicates the validity of the proposed model. Therefore, ANN next hour load forecasting model used in this work gives better results than the abductive networks model. The inputs used in the ANN model include all the previous hours (from h-1 up to h-24) loads. The model includes the most up to date load variation and the impact of the surrounding environment contribution.

C. Next Day Load Forecasting Models Comparison

For next day load forecasting modeled in this work, the inputs parameters and values arrangement follow the same approach done for KSA data. For abductive networks model, the input parameters are shown in details in [26].

Table II: Performance of next-hour load forecasting models over the evaluation year

<table>
<thead>
<tr>
<th>Forecasted Hour, h</th>
<th>MAPE, %</th>
<th>Abductive Networks model</th>
<th>ANN model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.14</td>
<td>0.3376</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.01</td>
<td>0.3762</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.93</td>
<td>0.4142</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.88</td>
<td>0.4609</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.08</td>
<td>0.8437</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.27</td>
<td>1.2241</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2.08</td>
<td>0.5297</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.55</td>
<td>0.6398</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.28</td>
<td>0.3797</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.82</td>
<td>0.297</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.94</td>
<td>0.2903</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.7</td>
<td>0.3982</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.8</td>
<td>0.323</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.69</td>
<td>0.3461</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.7</td>
<td>0.3835</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.77</td>
<td>0.4024</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.27</td>
<td>0.4207</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.31</td>
<td>0.4066</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1.48</td>
<td>0.3769</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.25</td>
<td>0.4022</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1.59</td>
<td>0.4228</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>1.23</td>
<td>0.4653</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1.2</td>
<td>0.3675</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1.29</td>
<td>0.3855</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.14</td>
<td>0.4539</td>
<td></td>
</tr>
</tbody>
</table>
Full-day load curves were forecasted using abductive networks models for some selected days of the evaluation year. The results of these forecasted days using abductive networks are shown in Figs. 12a, 13a and the results using ANN model for the same days are shown in Figs. 12b, 13b.

In addition, as done previously for KSA data, next day load forecasting can be modeled using repetitive use of next hour load model. Fig. 14 shows the above mentioned sample days forecasted using this method.

Table III shows the MAPE comparison between the three models for some selected days for next day load forecasting.

Table III: MAPE comparison between the three models for each day for next day load forecasting

<table>
<thead>
<tr>
<th></th>
<th>8-Aug</th>
<th>11-Aug</th>
<th>12-Aug</th>
<th>3-Sep</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>1.1376</td>
<td>1.2683</td>
<td>1.5542</td>
<td>3.3031</td>
</tr>
<tr>
<td>Iterative</td>
<td>2.5549</td>
<td>1.3484</td>
<td>2.0204</td>
<td>3.6159</td>
</tr>
<tr>
<td>Abductive</td>
<td>1.73</td>
<td>2.3</td>
<td>1.97</td>
<td>3.48</td>
</tr>
</tbody>
</table>

For abductive networks model and ANN model, it is clear that the forecasting accuracy is the best for the working day and poorest for the holiday due to the fewer examples of holiday load patterns encountered during training. Moreover, the results obtained using iterative forecasting model are very satisfactory.
VI. CONCLUSION
This paper presents ANN models for residential STLF applied to different types of loads subject to distinct people activities and weather conditions. Two ANN models are proposed, the first is the next hour load forecasting and the second is the next day load forecasting with two approaches to forecast the load at next day. The next hour load forecasting model uses the historical load data up to the previous hour to the forecasted one. The obtained results are accurate. The MAPE for the evaluation year gives very low values showing the validity and good accuracy of the realized next hour load model.

Next day model approach uses different inputs parameters utilizing the latest available data and load values up to the same previous day’s hour. Also, next day model by iterative using of next hour model is shown. The obtained results from both models are satisfactory.

As an extension of this work, further development models may include several different weather parameters data like wind speed, wind direction, rainfall, clear and cloudy sky conditions. Also, this work is only to forecast next hour and next day loads. So, next week load forecasting could be a continuation on this work. Moreover, load type could be extended to include aggregated load, commercial and industrial loads.

ACKNOWLEDGMENT
The authors acknowledge the support of King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia as well as the support of Royal Commission of Yanbu, Saudi Arabia.

REFERENCES