On the Method Development for Electricity Load Forecasting

Jelena Milojković and Vančo Litovski

Abstract—The way of development a method for systematic prediction of electricity load at suburban level is described. Steps that were passed are described and the evolution of the method is illustrated by proper examples. Prediction of electricity load at annual, monthly, daily and hourly level as well as peak consumption is considered. Comparisons with some other’s results are given, too.

Key words—prediction, artificial neural networks, electronics, electricity load.

I. INTRODUCTION

Prediction in everyday life and especially prediction in industry and business is of great importance. Economists are among the most interested professionals for prediction. Frequently, however, the rules underlying and governing the phenomenon to be predicted are not defined by human behavior only but by other conditions of technical and environmental nature, too. Here is why we are trying to do prediction based on our knowledge of both domains including the mathematical models involved.

Natural and social phenomena are mostly of stochastic nature and that is why most of the prediction methods are looking for values that characterize such set of events: trend, periodicity and dispersion. In many cases, however, we are after a value given deterministically since variance may be not informative enough for taking industrial or environmental action. For that reason, alike economists, we are considering deterministic prediction based on series of samples representing the behavior of the phenomena in the past.

As for the length of the series, when statistical methods are applied one has to have a relative large number of consecutive samples in order to catch all the stochastic properties of the phenomenon. In industrial and many other social conditions, however, frequently, no such data are available. Furthermore, in many cases old data are of no importance due to emergence of new governing rules that fundamentally change the flow of the phenomenon. For example, a set of new buildings connected to an existing suburban power source, may change both the average amount of power drawn and to introduce new shape of the consumption profile due to the habits of the new settlers. For that reason very old, and even medium old, data may be of no use or even misleading and making the prediction more difficult. One is to use as fresh data as possible. Of course, the set must be long enough to capture the periodic nature of the phenomenon being it daily, weekly, monthly or related to a season.

There are frequently doubts as to whether to include environmental data, such as the temperature, into the set for prediction. These data may be known from measurements (already performed) or may be predicted by some other method (by some environmental agency, for example) and used as known data for the future. For the first case, we think that the environmental data already influenced our phenomenon and accordingly that their influence is already incorporated into the main data set. In the second case, we think that the environmental data, predicted elsewhere or by a separate procedure, have a built in error which we do not need to implant into our solution.

There are several methods of deterministic prediction in practical use. We are not reviewing them here since our goal is to report on the development of our method which is based on implementation of artificial neural networks (ANN).

In the next we will go through our publications and illustrate the way how we started with a simple one-input-one-output feed-forward ANN and finished by two complex ANN structures dedicated to prediction. In addition we will explain why one needs at least two solutions in order to have reliable deterministic prediction.

II. THE DEVELOPMENT OF THE PREDICTION METHOD BASED ON ANN

We first got involved in prediction when considering the subject of electronic waste in Serbia [1]. It came out that there is no systematic way of forecasting of the amount of electronic waste to be found in the literature. In addition, the methods implemented were far of being based known scientific methods. The main reason for that was the lack of data for a longer period in the past.

That inspired us to start with the implementation of ANN that are known as universal approximators. Namely, by using ANN for approximation of a function represented by a set of equidistantly taken samples one automatically solves one of the biggest problems in approximation: the choice of the approximating function. Furthermore, ANNs are known as very successful interpolators which are frequently cited as a generalization property of ANNs. The question was, however,
does ANN can extrapolate. That we checked in [2]. To predict the amounts of electronic waste we implemented two structures: a feed-forward ANN as depicted in Fig. 1, and a new structure named feed-forward accommodated for prediction -FFAP.

![Fig.1 A fully connected feed-forward neural network with one hidden layer of neurons and multiple input and output terminals](image1)

![Fig 2. Fully connected feed-forward artificial neural network with one input signal, one hidden layer of neurons and multiple outputs - the FFAP structure](image2)

The feed-forward network was implemented with one input being the time and one output being the quantity to be extrapolated. The FFAP structure has the same structure at the input while at the output it has three categories of output terminals. One group (or a single) terminal(s) is used to approximate the past values of the function i.e. these terminals are approximating the given function displaced backwards. The second group (or single) of terminal(s) is approximating the future responses i.e. it learns the function displaced in forwards. Finally, one terminal is learning the actual function. To illustrate Table 1 is presenting a data set which was available the last item used as unknown during the training of the ANN. That value was to be predicted.

![TABLE I
AVAILABLE DATA](image3)

```
<table>
<thead>
<tr>
<th>Training set No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>t (year)</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>f(t) past</td>
<td>7.03</td>
<td>8.67</td>
<td>10.0</td>
</tr>
<tr>
<td>f(t) present</td>
<td>8.67</td>
<td>10.0</td>
<td>9.33</td>
</tr>
<tr>
<td>f(t) future</td>
<td>10.0</td>
<td>9.33</td>
<td>9.85</td>
</tr>
</tbody>
</table>
```

From that table the training data for the FFAP network were extracted and organized in the form depicted in Table 2. Note, for training the ordinary feed-forward network we had five lessons while for the training of the FFAP we have only three. The FFAP methods needs two samples to create the first past value (here: 2-7.03) and the last future one (here: 4-9.85). To get the prediction, after training, the ordinary feed-forward ANN was exited with \( t=6 \), while the FFAP ANN was exited by \( t=5 \). The result obtained by the FFAP ANN, being 9.81 (meaning that the prediction error was 2.5%) was so encouraging that we kept the FFAP idea all the time doing, however, proper improvements to meet new requirements.

![Fig 3. One month long record of an electricity load at suburban level](image4)

The success of the FFAP network inspired us to enter the subject of prediction of electrical loads. At the beginning we tried to predict the load in near future e.g. two hours [3] [4] [5] [6]. It came out, however, that one may benefit from the quasi-periodic shape of the load function as depicted in Fig. 3.
Two types of periodicity may be observed in the data depicted in Fig. 3. First, we have daily periodicity that follows the daily habits of the consumer. To incorporate the first one we simply span the training set over a full day. Twenty-four samples of the load taken every two hours are presented to the network to learn as output values denoted $p_n$, $i-j$ in Fig. 4. Here $n$ stands for the week; $i$ for the lesson, there 20 lessons as the window is moving exposing five values while starting at the 24th sample backward; $k$ for the number samples in future and for one step ahead prediction is $k=1$.) To incorporate the second property we took (four times) one sample of the same time and the same day of the previous weeks as the instant when prediction is to be made and put them as input signals. In that way the new structure named Extended FFAP was created.

The results obtaining were really encouraging.

Alternatively, to improve the performance of the ordinary feed-forward ANN, in [7], we examined the capacities of time delayed ANN (depicted in Fig. 5) and the time controlled recurrent (TCR) (Fig. 6) which is, in fact, our construct and, to our knowledge, cannot be found elsewhere in the literature. The prediction results obtained by the TCR ANN were equally good as the ones obtained by the FFAP ANN. That was confirmed in the implementations for prediction in microelectronics [8] [9].

The subject of multistep prediction we firstly considered in [10] [11] where prediction in microelectronics was sought. Then in [12] [13] [14] we did multistep electricity load prediction. Two concepts were checked. Firstly, the two step ahead prediction was done in two steps. In the first one we made a one step prediction and then we used the result as a known data to proceed to the next step. That approach gave not very successful result. In the second approach we simply predicted two steps ahead skipping the first one. That came out to be much better approach.

Long term (one year) prediction of electricity loads for large areas such as provinces or states has the property of slow change and is not a specific challenge for the prediction algorithm. That was probably the reason why the so called Grey-theory was implemented to that problem. It is, in fact, a method of finding the coefficients of an exponential polynomial which, as one may expect, is working best when the polynomial is of first order. In [15] [16] [17] we implemented TCR and EFFAP ANNs to show that our approach is reaching much better results both in one step and in multistep forecasting.

Prediction is a walk in the dark. To rely on it one has to have a least one reference. That is why, in our opinion, for a good results, it is necessary to have at least two prediction that support each other. Considering the EFFAP solution as a good one we tried to implement the TCR in a similar way so that a new structure emerged named Extended TCR (ETCR) [6]. It is depicted in Fig.7. The meaning of the quantities is the same as in Fig.4. Its first implementation was for short term prediction of electricity loads at suburban level. The results were as good as the ones obtained with EFFAP structure. Having two equally valuable but different results one is to decide which one is to be accepted as a final prediction. We went for the average [18]. It was implemented to short term (two hours) prediction of electricity loads at suburban levels. The first results were as shown in Fig. 8. As can be seen one always get smaller average prediction error than the largest of the two entering the calculation. The benefit is especially
enhanced when the entering prediction errors are of opposite signs as for the 4th hour in Fig. 8.

This concept was later implemented in prediction of the daily peak value of electricity loads at suburban level [19] [20] [21] [22].

Fig. 8. Prediction results of the EFFAP and the ETCR network and their average

Fig. 9 depicts the relative prediction error obtained for a thirty consecutive predictions (one month). As can be seen the daily peak load was predicted with an error not exceeding 15% of its value.

It happens in everyday life and in nature that two related phenomena are governed by the same cause. In such a case one may attempt to predict both phenomena based on a single input data set. An attempt to do so for the peak load value and the time when it arises was done in [19] [22]. To illustrate the structure of the appropriate ETCR ANN used for that purpose is depicted in Fig. 10. In fact we doubled the input and the output neurons while keeping the number of the neurons in the hidden layer.

Fig. 10. ETCR. Extended time controlled recurrent ANN approximating two functions simultaneously

III. CONCLUSION

An historical overview of the development of a method for one- and multi-step prediction in electrical engineering is described. As a result two new structures of ANN were developed and always used concurrently for the same prediction if supporting each other. The results so obtained are averaged so smoothing the final prediction error curve.

ACKNOWLEDGMENTS

This research was partly supported by the Serbian Ministry of Education, Science and Technology development under the contract TR32004.

REFERENCES


