APPLICATION OF STATISTICAL AND NEURAL APPROACHES TO THE DAILY LOAD PROFILES MODELLING IN POWER DISTRIBUTION SYSTEMS

Joanicjusz Nazarko, Member, IEEE
Institute of Management and Marketing
Bialystok Technical University, Poland

Zbigniew A. Styczynski, Member, IEEE
Institute of Power Transmission and HV Technology
University of Stuttgart, Germany

Abstract: Load modelling is an essential task in economic analysis, operation and planning of distribution systems. Particularly, when a Demand Side Management system is taken into account on a deregulated energy market the knowledge of load profiles is of the greatest importance. Forecasting of daily demand, based upon load models, uses comparable load research data for a different customer mix. For the given season and day of the week the shape of a daily load curve depends mainly on the customer composition. Difficulties in defining objective customer classes significantly complicate the forecasting process. Usage of statistical clustering and neural network approach makes possible to improve the load modelling accuracy. This paper presents load modelling methods useful for the long term planning of power distribution systems. Theoretical statement is illustrated by examples which correspond to Polish and German distribution systems.

Keywords: Load modelling, Load forecasting, Clustering, Statistical method, Neural network, Network planning

I. INTRODUCTION

Load modelling is an important task when considering the analysis and operation planning of electrical power distribution systems. Planning engineers use load modelling to predict load shapes in different parts of a distribution system in a process of optimisation of network development. It is evident that the quality of load modelling strongly influences the system expansion scenarios.

Particularly, a precise knowledge of load profiles is required when Demand Side Management (DSM) system has to be used on a deregulated energy market. Also an introduction of a new equipment like decentralised energy storage, which helps to match energy consumption request, needs information on load profiles [8].

The total demand met by an electric utility is a sum of individual demands of a diverse set of customers; each with a multitude of different electric devices connected to system buses. Considering a large number of customers in a power distribution system and a great deficiency of stationary measuring and recording devices at system buses, the analysis of behaviour of daily load shape is quite difficult.

The decomposition of the load into characteristic classes helps to couple some of the load variation phenomena. The standard way of analysing a large set of load curves is to perform a clustering of them so that planning engineers could look at a small number of classes (i.e. clusters) of similar curves instead of whole set of curves.

II. LOAD MODELLING

A. Problem Formulation

The electric load is classified – by the power utility – according to the typical load profile into different subjective customer classes e.g. residential, administration, commerce, industrial, street lights etc. Each class has its own peculiarities. For example, supreme consumption for residential consumers occurs during morning and evening hours, when people, being at home, use most of their domestic electrical devices. For industrial sector operating on one shift maximal consumption occurs during the daytime with a slight decrease at lunch-time [3, 4].

The principle of the load modelling is shown in Fig. 1. To each node of the distribution system of medium voltage (e.g. 10 kV) a group of consumers or a low voltage grid can be connected. They build a mix of consumers supplied from a network node and determinate the daily load curve which can be observed at the corresponding node. It was tested that for the given season and day of the week the shape of a daily load profile depends mainly on the customer composition [1,7].

The essence of load forecasting, based upon load models, is to conclude from n known load curves and corresponding consumers compositions the unknown or future expected ones for the assumed consumers mix (Fig. 1).

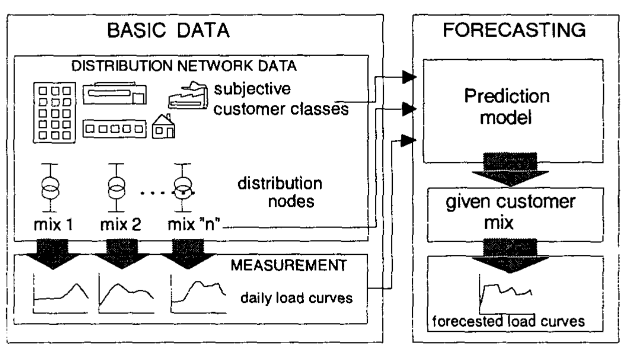

Fig. 1. Modelling process of the daily load curves

0-7803-5515-6/99/$10.00

The modelling process can be generally described as a compression relation „ $\Leftrightarrow$ “ as following:

$$\underset{i \in R}{\forall} \Gamma_i \Rightarrow (\Gamma_a \Leftrightarrow \Gamma_i) = \min(E) \qquad (1)$$

where: E – error of comparison, Γ_i - set of predicted load curves, Γ_a – set of original load research data, i – mix of customers, R – set of mixes.

B. Subjective and Objective Customer Classes

Traditional customer classes, used in the load modelling, like residential, industrial etc. are described as subjective because the natural patterns of load curves inside the same specific class is not homogenous. They depend strongly on local conditions e.g. part of a city, intensity of production or energy demand, and can essentially vary from node to node, and from power utility to power utility.

The power utility does not usually have enough information about processes which are supplied with energy from the distribution substations. Especially for consumers classified as industrial a significant differences in the load curve behaviour are observed. For example it is difficult to define if an administration of a factory or the manufacturing plant is supplied from a network node. Additionally, the daily load curve does not only depend on the consumer composition, but also on external conditions i.e. temperature, brightness, humidity etc. The same problem arises for other classes. This is the reason why we call the mentioned classification subjective.

The use of subjective classes in load modelling increases a range of possible forecasting errors and a convergence of prediction method could be not preserved. In the most cases the accuracy of prediction is not satisfactory (Table 4).

The method of building an objective load profiles classification based on available measurements is one of possible undertakings to improve the accuracy of the prediction. The main idea is to perform a clustering of the curves. Using clustering techniques, it is possible to obtain objective classes of load profiles where each class represents a set of curves of similar shape. In this case a set of measured load profiles represents an input information which permits to classify the nodes.

The objective classes – clusters – become the base for the prediction of load curves in a considered network node when the composition of customers supplied from it is known. The clustering of load curves into objective clusters minimise the multidimensional error of curve fitting in the class. The clustering and prediction with clusters is schematically presented in Fig. 2.

III. MODELLING METHODS

A. Statistical Approach

A graphic illustration of the daily load profile is the 24-hour chronological load diagram $P_{dt} = f(t_{dt})$ (Fig. 3).

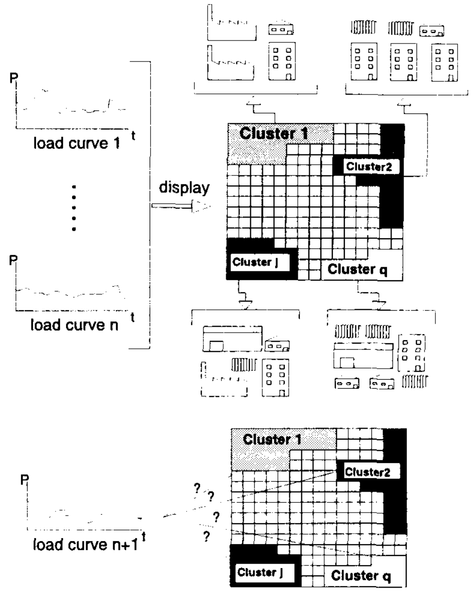

Fig. 2. Clustering of load curves

As a function of time the load pattern shows several peaks and troughs. There are four characteristic load values at the diagram: P_{dp}– the daily peak load, P_{da}– the daily average load, P_{bd}– the daily base load, P_{di}– installed capacity of the electric devices connected to the system bus.

To differ consumed electrical energy according to the level of demand and to the time of energy consumption, the division of daily load profile into the horizontal layers and the vertical columns is introduced (Fig. 3).

Three layers: base load layer, intermediate load layer, and peak load layer and four columns: morning (m), afternoon (a), evening (e), and night (n) are distinguished. Columns can correspond to the tariff periods for example. The clustering of the customers according to daily load profiles is performed on the basis of the average alignment degree for each column. To avoid the influence of the instantaneous values on the changes of power consumed by a customer, the average load of 24-hours should be taken as the reference quantity.

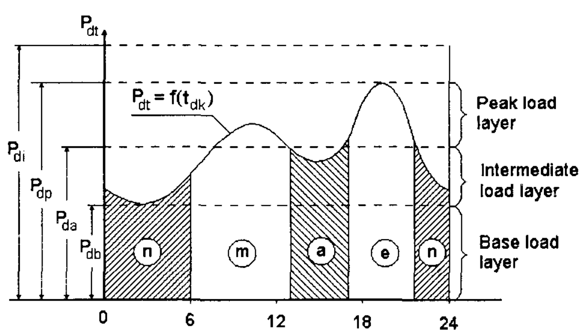

Fig. 3. Daily load profile: division into layers and columns; n – night, m – morning, a – afternoon, e – evening

The average alignment degree for the column is defined as a ratio of the average load in the column to the daily average load:

$$l_c = \frac{P_{ja}}{P_{da}} \tag{2}$$

where: c – the column index (n, m, a, e), P_{ja} – the average load in the column j, P_{da} – the daily average load.

The load diagrams are considered as similar ones when their average alignment degrees for each column have similar values.

To distinguish the particular classes of daily load profiles the four-dimensional space with l_m, l_a, l_e, and l_n co-ordinates is constructed. Then it is divided into clusters with Δl_m, Δl_a, Δl_e, and Δl_n sides. Each 24-hour load profile is represented by a point in the space with co-ordinates corresponding to alignment degrees calculated for the profile.

The customers for which the points corresponding to their load profiles are grouped at the same cluster are regarded as the same customer class. The exemplary figure for the two-dimensional space is presented in Fig. 4.

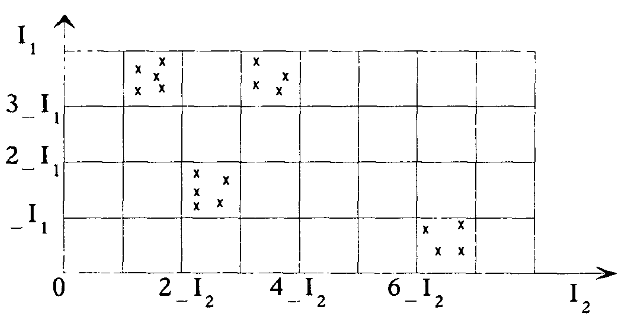

Fig. 4. The example of clustering of daily profiles on the plane I_1, I_2

B. Neural Network Approach

A number of neural network methods are widely used for load forecasting [2,7, 9]. The main advantages of using neural networks in prediction algorithms are:

- automatic learning of behaviours included in the training data,
- automatic generalisation of the information.

These features make the use of neural networks forceful and effective for the load modelling in distribution systems.

The procedure of the application of neural network to the load modelling in the medium voltage network nodes can be subdivided into three steps:

- classification of consumers into characteristic classes,
- training of an appropriated neural network with selected, known load curves,
- use of the trained neural network to forecast load demand sequences for given input data.

IV. TEST CALCULATIONS

A. Test Data

For a Polish municipal distribution system of 15 kV the measurements of daily load profiles were made at 26 randomly selected 15/0.4 kV distribution transformers. Measurements took place in December. Transformer loads were recorded every 15 minutes in the course of 2 weeks. Transformer ratings are from 50 to 630 kV A and their peak loads are in a range of 29,6 to 312,0 kW. Some subjective classes were chosen for the classification of consumers in order to compare the results of calculation.

In Germany load curves were measured at nine selected buses in a 10 kV network of a distribution utility. Consumer mix and load curves were well known at each bus. On this base any consumer can be classified into one of the following five groups: household (HH), commerce (CM), industry (IN), business and service (BS), storage heating (SH).

B. Statistical clustering

The data recorded for work days were used to test the worked out clustering method. For each daily load profile four alignment coefficients were calculated according to Eq. 2. Then a computer program that is a practical implementation of the described algorithm was used to make clustering of daily load profiles.

The results of clustering of the load research data are presented in Table 1.

Class	Customers	Factors	
1	Industrial, three shifts	l_n	0.88 ÷ 1.02
		l_m	0.96 ÷ 1.19
		l_a	0.79 ÷ 1.08
		l_e	0.84 ÷ 1.07
2	municipal-services, bank, post-office, hospitals, ambulatory, office buildings	l_n	0.66 ÷ 0.84
		l_m	0.90 ÷ 1.27
		l_a	1.00 ÷ 1.29
		l_e	0.93 ÷ 1.27
3	municipal-living, blocks of flats eleven and four-storey boarding-school, shops, street lighting	l_n	0.48 ÷ 0.67
		l_m	0.65 ÷ 0.86
		l_a	0.75 ÷ 1.05
		l_e	1.76 ÷ 1.99
4	municipal-living, small one-shift industry, blocks of flats eleven and four-storey, hotel, shops, street lighting	l_n	0.54 ÷0.80
		l_m	0.74 ÷ 1.17
		l_a	0.91 ÷ 1.25
		l_e	1.24 ÷1.87
5	living, blocks of flats four-storey	l_n	0.34 ÷ 0.41
		l_m	0.69 ÷ 0.83
		l_a	1.19 ÷ 1.27
		l_e	1.82 ÷ 2.03
6	broadcast transmitter	l_n	0.63 ÷ 0.71
		l_m	1.14 ÷ 1.19
		l_a	1.11 ÷ 1.18
		l_e	1.11 ÷ 1.13

Table 1. Range of average values of alignment coefficients for characteristic classes of customers

The tendency of grouping points in several areas was observed. For small length of the cluster sides ($\Delta l_c \leq 0.2$) a great number of classes which consist only of a single profile was obtained. The best results of grouping (several classes) were obtained for the following values:

$\Delta l_m = 0.3$; $\Delta l_a = 0.3$; $\Delta l_e = 0.4$; $\Delta l_n = 0.3$;

Several objective clusters were obtained (Table 1). They contain similar load curves and can be also related to the subjective customers classes. For each class of customer composition a typical (average) daily load demand can be modelled.

The developed construction method of the typical load profile for the particular class bases on the discrete Fourier transformation [4, 5]. Figures 5 and 6 present the typical load profiles for two selected customer classes.

As a accuracy measure of real shapes approximation by typical profiles the relative mean-square deviation was taken

$$\eta = \sqrt{\frac{1}{96}\sum_{k=0}^{95}\left[\overline{P}_r(k) - \overline{P}_j(k)\right]^2} \qquad (3)$$

where: $\overline{P}_r(k)$ – the normalised real load curve,

$\overline{P}_j(k)$ – the typical load profile of *j* class of customers.

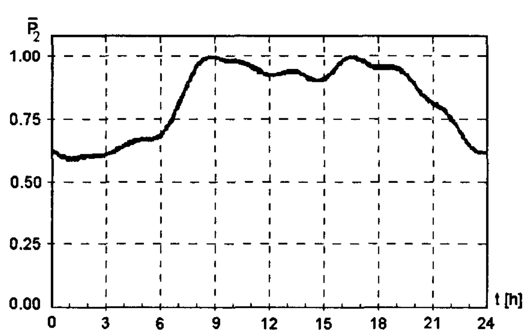

Fig. 5. Typical daily load profile for class 2: municipal-commercial customers

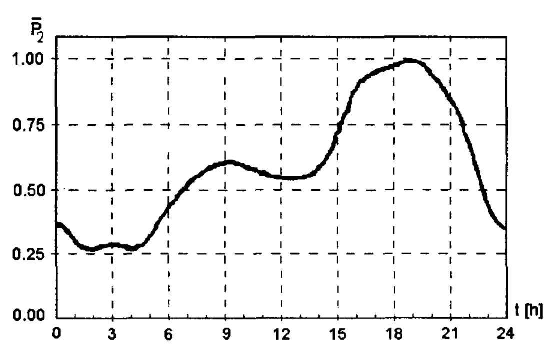

Fig. 6. Typical daily load profile for class 4: municipal-living customers

The results of the calculations of relative mean-square deviation for selected daily load profiles of the particular customers classes show that η value is of several - dozen percent and it does not generally exceed the value of 20%. Bigger error values were stated in cases when the receivers of large, short time power demand have considerable participation in the substation load. The typical daily load profiles are utilised for the forecasting of real runs.

C. Neural Load Modelling

The neural network used for load modelling is based on the model with six inputs: time of day and five inputs which represent customer composition (normalised share for each characteristic class). The output is represented by the actual load demand sequence. The load curve consists of 49 values, mean half-hourly load measurements.

The influence of four seasons: spring, summer, fall, winter and three characteristic days of the week: week day, Saturday and Sunday are not taken into account by supplementary inputs in the model, but it is considered by the separate training of the neural network with different sets of data for special combinations of seasons and days of the week.

The results obtained by using such specialised networks were much more satisfactory than those given by generalised network with 8 entries and 12 times higher number of training patterns (see [9] and Table 2).

In practice, what was mentioned in Section 2, consumers classified into the same class do not have exactly the same load demand pattern [9]. Due to this the neural network is always trained with slightly misleading information at the beginning. In order to avoid a hyper-sensitive reactions caused by two contradictory pieces of information, the network should never be trained with equal or similar input patterns when the corresponding load values are different.

To avoid this problem, supplementary artificial data sets were created by the weighted superposition of already existing data. Tests have shown that the superposition can be a means to obtain more general data and significant better training results. New simulations were effected with 20 measured and mixed consumer combinations. The influence of the supplementing mixed data on the calculation results is shown in Fig. 7.

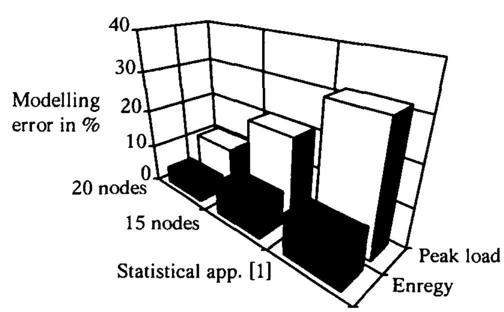

Fig. 7. Errors in load modelling with neural network

For the comparison of the method accuracy two parameters are observed. There are:

- the error of the peak load prediction E_L

$$E_L = \frac{100}{n}\sum_{k=1}^{n}\frac{\left|P'_{k\,max} - P_{k\,max}\right|}{P_{k\,max}} \qquad (4)$$

and

- the error of the daily energy calculation E_E

$$E_E = \frac{100}{n}\sum_{k=1}^{n}\sum_{i=1}^{48}\frac{\left|P'_{k,i} - P_{k,i}\right|}{P_{k,i}} \qquad (5)$$

where: P_k' and P_k – predicted and measured value of load respectively, n – number of sets used for comparison.

It should be noticed that the increase of used mixed customers combinations up to 20 nodes (9 real and 11 mixed) decreases drastically (twice) the value of prediction error (Fig. 7). However, in both cases with neural network the predication errors are smaller in comparison with statistical approach [1].

The simulation software SNNS was used to perform calculations [6]. It provides almost twenty training algorithms. The Backpropagation-Momentum-algorithm (BM) and the Resilient-Propagation-algorithm (RP) for neural network learning were implemented. Better results were obtained however by the use of the BM-algorithm (Table 3).

Two criteria are used to judge the quality of a trained neural network: its capacity of "learning" and its capacity of "generalising". "Learning" is a measure of the precision with which the neural network reproduces already trained data, whereas "generalising" is a measure of the quality of the network response to untrained input data. As a selection parameter "usefulness" has been introduced. It is calculated as the sum of values of both mentioned criteria. The results obtained for the different network structures are listed in Table 3.

Neural Network Structure (neuron per layer)	Network Number (network type 1)/training's algorithm 2))
6-12-6-3-1	1 (SC/BM) / 8 (F/BM) / 18 (F/RP)
6-15-9-6-1	2 (SC/BM) / 5 (F/BM)
8-24-16-8-4-1	4 (SC/BM) / 6 (F/BM) / 27 (F/BM)
6-12-6-1	12 (F/RP) / 13 (F/BM)
6-5-4-3-2-1	14 (F/RP)
8-16-8-4-1	15 (SC/BM)
8-16-8-1	19 (SC/BM) / 21 (F/RP) / 26 (F/BM)
6-18-12-6-3-1	20 (F/RP)
1) F- full, SC- shortcut	2) BM-Backprop. Momentum, RP-Resilient Prop.

Table 3. Tested neural network structures of the neural load predictor [9]

Its shortcut connection structure (6-12-6-3-1) – with which the best results were obtained – represents a particularity among the feed forward networks. Here a layer is not only connected to its neighbour layers, but to all other layers.

Fig. 8 shows modelled load curves for week-days in spring for different consumer compositions. The results of the neural load predictor have been compared with real measured data as well as with the results of the statistical method [1]. The chosen consumer compositions were not learned by the neural network.

A tendency of the neural network to predict a little too smooth curves (Fig. 9) can be observed whereas the other method sometimes predicts peaks that don't occur in reality. However, it has to be known that also the measured data can be an exceptional case which may differ from the representative average data.

For a quantitative judgement of the modelling quality an error calculation has been performed. The average error of the estimation, corresponding to the load curve prediction by the neural network for untrained load curves, of the peak load (7,1%) and total energy (3,1%) [6] is about twice lower than when using the statistical method.

Also the error dispersion for 60 unlearned load curve, corresponding to peak load and daily energy, is more than twice lower than with using of the statistic method.

If the measured data set of load profile is analysed by the neural clustering method some objective customers group can be determined, analogously to the presented statistical approach. Here the Kohonen-Map has been used. Specific for this neural network is that during the learning process the clustering accrues by the self organisation of the map. Different dimensions of the map have been tested.

Using a classification share factor (SF), the membership of a load curve to the objective classes can be determined. The value of the SF factor is equal to the sum of output values of neurones in a class divided by the number of them. The new objective classification could be not interpreted as a normalised one.

In Table 4 the results of a training and recognising processes for measuring load curves are presented. Due to the training the range of the SF values is quite wide.

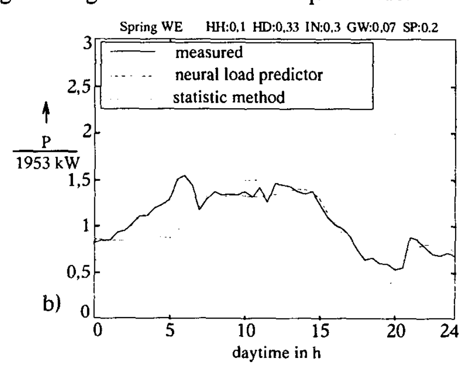

Fig. 8. Predicted load curves during work days in spring compared to the measured load. Spring, workday, HH:0,1, CM:0,33 IN: 0,3 BS: 0,07 SH:0,2

	Class				
	1	2	3	4	5
Train.	0.115	0.112	0.105	0.100	0.168
for 20 nodes	0.764	0.925	0.850	0.782	0.929
Test	0.366	0.188	0.166	0.164	0.184
for 10 nodes	0.657	0.838	0.658	0.592	0.777

Table. 4. SF values for clustering with five classes (13x13-Kohonen-Map)

On the other hands the values of SF factor by the test are more narrow because of the learning effect.

To compare the performance of neural network and statistical algorithms calculations were made for five synthetic classes of customers using a 13x13 Kohonen-Map.

Fig. 9 presents results of the load modelling with statistical and neural approach taking into account normalised and objective classification. Map of 13x13 gives better results than the classification with statistical and neural methods for subjective classes. In the case of a Kohonen-Map of smaller size (5x5) results are quite similar to the statistical method.

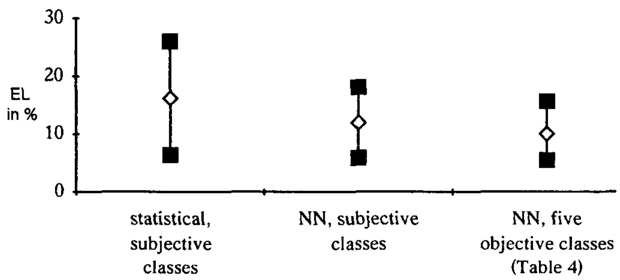

Fig. 9. Peak load prediction errors for 10 nodes by different load predictors

Use of the objective classification improves the recognition accuracy and improves results of the prediction also in comparison to the ordinary neural network.

V. Conclusions and Further Work

Load modelling in power distribution systems is a complex problem to solve because of stochastic character of the process, deficiency of measurement data and subjective classification of customers. It is of interest to many researches and a lot of methodologies were proposed to forecast load demand. The knowledge of load profiles helps to understand the power consumption behaviour in order to properly operate, analyse and plan distribution systems. It gains particular importance on deregulated local energy markets.

Presented in this paper method based on statistic clustering and neural network modelling can be a useful tool for load curve prediction in distribution areas. To adjust the model the representative measurement data from examined area is required at the first step. The use of synthetic (objective) customer classes as a basis for load modelling can improve accuracy of load forecasting. More precise results of modelling are obtained when using specialised models for different seasons and days of the weak than with a general model.

Further work should be concentrated on:

- investigations of a new structure of clustering (e.g. fuzzy clustering)
- investigations on correlation between subjective and objective customer groups
- development of hybrid (e.g. fuzzy and neural) methods of load modelling and prediction.

VI. References

[1] Handschin E., Dörneman Ch.: *Bus load modelling for forecasting*. IEEE Transactions on Power Systems, Vol. 3, No. 2, May 1988.

[2] Kanda M., Hayashi Y., Iwamoto S., Furuya S.: Long Term Maximum Load Forecasting Using Fuzzy Neural Networks. Conf. Intelligent System Application to Power Systems, Montpellier, pp. 783-790, 1994.

[3] Markusheva N.S.: *Avtomatizirovannoe upravlenie rezimami elektrosetei 6-20 kV* (in Russian). Energia, Moskva 1980.

[4] Nazarko J.: *Modeling of electrical power distribution systems*. Bialystok Technical University Publisher, Bialystok, 1993.

[5] Shumway R.H.: Applied Statistical Time Series Analysis. Prentice-Hall Int., Inc., 1988.

[6] SNNS- Stuttgart Neural Network Simulator. User Manual, Uni Stuttgart, 1993.

[7] Srinivasan, D., Liew, A.C.: Forecasting Daily Load Curves Using a Hybrid Fuzzy-Neural Approach. IEE Proc.-Gener. Transm. Distrib., No.6, pp. 561-567, 1994.

[8] Styczynski, Z., Feser, K.: Planning of Medium Voltage Power System with Energy Storage for Load Levelling. Proceedings 11th PSCC, Avignon, pp.741-748,1993.

[9] Styczynski Z., Welfonder T., Freund H.: Die Nutzung eines Neuronale-Netze-Verfahrens zur Lastmodellierung für die Netzplanung. Elektrizitätswirtschaft 4/96. pp. 182-188.

VII. Acknowledgement

This work is partly supported by the State Committee for Scientific Research (KBN) under contract W/IZM/7/98 and by the European Community under contract ICOP-2140.

VIII. Biographies

Joanicjusz Nazarko received his Ph.D. and D.Sc. degrees in Electrical Engineering from the Warsaw University of Technology in 1983 and 1992 respectively. He is currently an Associate Professor of Electrical Engineering at the Bialystok Technical University, Poland and he holds a Chair of Informatics and Logistics. He is also a visiting lecturer at the Warsaw University of Technology. His research activity is centred on automation power distribution with emphasis on modelling and analysis of distribution systems in uncertain conditions. He is the author over 80 papers. He is a member of IEEE, IEE and CIGRE.

Zbigniew A. Styczynski after half a year at the computer centre of the TU Wroclaw joined the Institute of Power Systems. In 1977 he received his Dr.-Ing and in 1985 his Dr.-habil from the Technical University of Wroclaw. From 1986 he was Assistant Professor and from 1987 was also a deputy head for research in the Institute of Power Systems the TU Wroclaw. In 1991 he joined the University of Stuttgart (Institute Power Transmission and High Voltage Technique) where he gives lectures on using the expert system in power supply. His main research activities concern power network planning and optimisation, and intelligent computing in power system. He is a member of IEEE, CIGRE, VDE, SEP, and author of about 70 papers.