Modelling and Forecasting with Kohonen Networks and Linguistic Equations

Seija Juuso* and Esko K. Juuso**
*Department of Economics, University of Oulu, POB 4300, FIN-90410 Oulu, Finland
E-mail: Seija.Juuso@oulu.fi
**Control Engineering Laboratory, Infotech Oulu and Department of Process Engineering, University of Oulu, POB 4300, FIN-90410 Oulu, Finland
E-mail: Esko.Juuso@oulu.fi

ABSTRACT: Productivity-augmenting effects of education on economic output has been studied in order to understand better the technological change as a source of long run economic growth. Cross-sectional time series of Finnish manufacturing data from 1975 to 1994 provide a basis for modelling of the regional economics. Nonlinear multivariable models have been developed by using *Kohonen networks* as a clustering method, and generating L*inguistic Equation (LE)* models from these neurons. The resulting LE model, which can be used to any direction, represents the Kohonen network very accurately. Most cases in the time series are also well represented. The differences in the interactions, which can be seen as an increased fuzziness of the model, are related to changes in regional economic conditions, which should be handled separately. The LE model can be considered as a new type of neural network, *linguistic Kohonen network*, where each neuron weight has also a linguistic meaning. This network can be adapted to changing conditions, e.g. different regions and industry, by adjusting membership definitions.

KEYWORDS: modelling, forecasting, linguistic equations, neural networks, regional economics, classification

INTRODUCTION

Productivity-augmenting effects of education on economic output analysed within a fuzzy production function framework have been studied in [Juuso 1995a]. This research analyses the role of education in economic growth and is an attempt to understand better the technological change as a source of long run economic growth. An interesting question is in which extend regional growth rates can differ if human capital and incentives to adopt new technology differ across regions. The linear regression analysis was applied to study relationships of education and economic output in the Finnish provinces from 1975 to 1987. A measure of educational expansion, expenditures of comprehensive school, is included explicitly as an additional input in the Cobb-Douglas production function of two basic factors, labour and capital. Panel data - cross-sectional time series - provide the base to estimate a random coefficient regression model [Swamy 1970]. It is assumed that the coefficients in regression equation to be estimated are random across units but follow the same distribution with the same mean and the same variance-covariance matrix. Additionally the model with a varying intercept and fixed slopes is estimated to measure and compare productivity efficiency in the manufacturing industries of the Finnish provinces.

In economic systems, the inherent ambiguity or fuzziness of assembled data is influential. Since the validity of methods to measure variables can also be questionable, the exact modelling of these systems may be very difficult. To avoid these problems, the method of linguistic equations and the fuzzy set theory were used to determine the possible values of estimated parameters and to analyse their relations. The sets based on statistics are some of the most naturally fuzzy sets since any measurement based on statistics must make some allowance for deviations from value obtained by the measurement. Another property of statistically based sets is that they are naturally quantitative. These fuzzy sets provide a good model for the definition of confidence intervals [Juuso 1995a].

The transition from acceptability of a particular sample mean to unacceptability is gradual and is modelled by a fuzzy confidence set. Instead of determining the exact boundaries as in an ordinary set, a fuzzy set allows no sharply defined boundaries. The estimated regression coefficients are represented as fuzzy numbers and their optimal membership functions derived from the Gaussian probability density function. Finally, the numeric values of membership functions are replaced by the linguistic values and the linguistic equation framework is used to combine the set of linguistic relations of estimated regression coefficients.

On the basis of the fuzzy Cobb-Douglas -type regression model developed in [Juuso 1995a], a positive relationship of education and economic output was found. Some evidence of variations in productivity - augmenting effects of education across provinces and manufacturing industries was also pointed out in this analysis. The fuzzy Cobb-Douglas -type regression model can be used in several applications using fuzzy mathematics. The structure of the constructed linguistic model is extendable, e.g. other variables can be included. The model gives a good basis to tune membership functions and provides a flexible environment for combining linguistic rules with more efficient modelling methods to elucidate the question of the effects of education on the economic output.

INTELLIGENT MODELLING AND FORECASTING

Computational intelligence provides many methodologies for nonlinear multivariable modelling. Neural networks can be considered as fairly complex nonlinear black-box approximators with similar benefits and drawbacks as conventional regression techniques. Fuzzy set systems represent gradually changing nonlinear together with abrupt changes on the basis of understanding of the system behaviour. Linguistic equation approach combines various intelligent modelling techniques on a unified framework [Juuso 1999]: a close connection to fuzzy set systems was important already in the early applications, and data-driven modelling has brought the LE approach close to neural networks. The LE approach has also been successfully extended to the dynamic modelling [Juuso 1998].

In changing operating conditions, the forecasting problem consists of following subproblems:

- *Classification:* Operating conditions should be detected before forecasting since they will define the process dynamics. Neural networks have been widely used in classification, e.g. Kohonen networks are suitable for finding unknown structures in data sets. Another suitable clustering technique for these applications is fuzzy clustering.
- *Dynamic modelling:* Dynamic models are developed for each case on the basis of an appropriate data cluster. Fuzzy set systems combine these models into a smoothly operating overall model. Multilayer perceptron networks could be used for forecasting within individual cases.

This modular approach can aid in understanding the problem. Combining linguistic equations and neural networks, or more specifically Kohonen networks, for modelling and forecasting will be discussed in this paper.

NEURAL COMPUTING

In econometric terms, artificial neural network (ANN) models constitute a particular class of nonlinear parametric models. Learning corresponds to statistical estimation of model parameters [Kuan and White 1994]. Modelling architectures are similar but richer than flexible functional forms and semi-parametric approaches. Implementing these networks have shown the way to powerful new methods of computer intensive model search and optimisation in econometrics. The principal advantages of neural networks over conventional computational techniques are [Ormerod and Taylor 1991]: ability to perform complex pattern recognition tasks, self-adjusting to map functional relationships between inputs and outputs, and explicitly parallel operation.

Application areas of ANN include time series modelling and forecasting, nonparametric estimation, and learning by economic agents. Also they have been successfully used to decode deterministic chaos [Kuan and White 1994]. Neural networks seem to suit well to problems of economic modelling of time series data on the area of asset price data [Ormerod and Taylor 1991]. For spatial economic systems, conventional methods cannot estimate relationships needed in calculating profits *ex ante* [Nijkamp and Reggiani 1993]. First, because only a limited number of determinants can be analysed at the same time. Second, because the structural relationship between an asset price and its determinants can change over time. And the third, because many of the rules which drive asset prices are qualitative or fuzzy, requiring judgements, and hence by definition are not susceptible to purely quantitative analysis.

Models that are capable of unsupervised training can be valuable in explanatory work. In a case, where the existence of distinct classes in a collection of samples is hypothesised but the open question is what those classes are, it is possible to let neural networks discover salient characteristics of training data on their own. Another case is, when the classes of data are already known but there are suspects that some of the classes contain subclasses [Masters 1993]. In *Kohonen network*, the fundamental point of self-organisation of weight vectors is that knowledge is not concentrated in single neurons but their neighbourhood neurons, too. Each input pattern is passed into the competitive learning layer of the network, and the winner, i.e. the neuron with the best response to that input pattern, has its weights updated using the current learning rate. Its neighbours have their weights updated also. The learning rate for neighbours is less than that for the winner, so their weights change less.

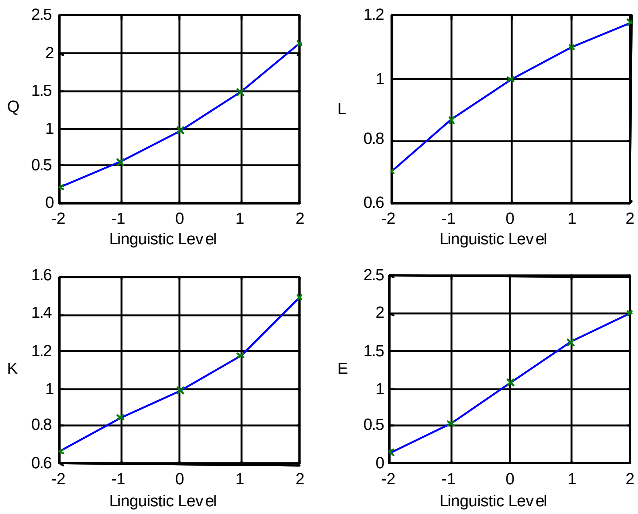

Figure 1: Membership definitions generated from the neurons of Kohonen network.

LINGUISTIC EQUATIONS

Linguistic equations provide a flexible environment for modelling and control of both data intensive and knowledge intensive applications. The model or the controller is represented by matrix equations. The system is adaptive since the meaning of the linguistic values depends on the operating conditions. Membership definitions are tuned by simulation experiments, by expert knowledge or by data from real systems. According to the original framework [Juuso 1999], a set of linguistic rules or relations can be changed into a compact equation

$$\sum_{j=1}^{m} A_{ij} X_j = 0, \tag{1}$$

where X_j is the linguistic level for the variable j, j=1…m, i.e. the linguistic values *very_low*, *low*, *normal*, *high* and *very_high* are replaced by numbers –2,-1,0,1 and 2. Multipliers A_{ij} {-1,0,1} describe the direction of interaction between variables. Linguistic models can be presented as matrix equation

$$AX = 0 \tag{2}$$

where *A* is an *n* x *m*-matrix. The LE approach has been extended to real-valued and fuzzy equations, and applied to dynamic modelling [Juuso 1998].

Knowledge is handled with interactions represented by linear equations, and nonlinearities are taken into account by membership definitions (Figure 1). The resulting models are smoothly changing nonlinear surfaces (Figure 2).

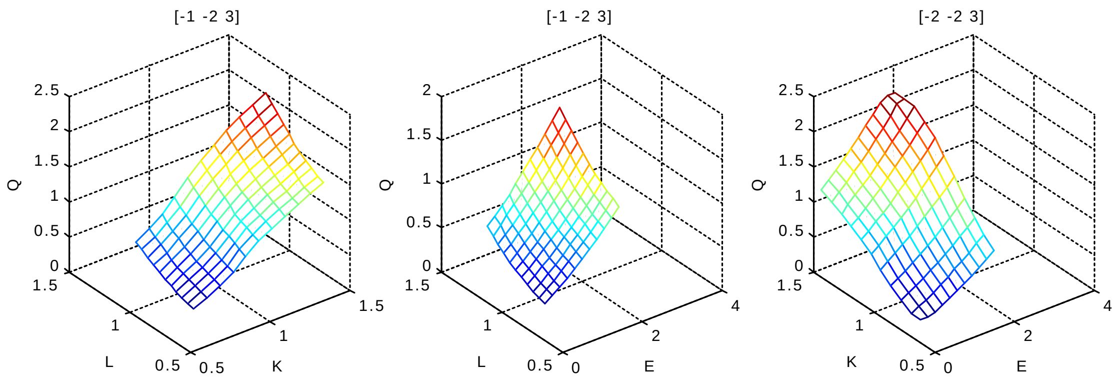

Figure 2: Model surfaces for the linguistic equation model [3 -1 -2 -2]*[Q L K E]^T.

MODELLING OF REGIONAL ECONOMICS

The three-variable Cobb-Douglas -production function was the starting point in [Juuso 1995a] where this approach was extended to a fuzzy model for the relationship of its parameters. In [Juuso 1995b] the Kohonen network was used to find relationships of variables. Output of each industry is measured as gross industrial product, which depends on labour measured as man-hours, capital measured as energy consumption and education as expenditures of comprehensive schools. Observations of each variable were divided by their respective time periodical means. The weight matrix was generalised by the linguistic equation approach. The generalisation is necessary because of limited data.

NEURAL NETWORK MODELLING

The Kohonen network can be considered as a clustering method, which is easy to interpret. The training is very fast since fairly small networks can be used. The modelling was started in Matlab® environment for the panel data of the education and the economic output in the Finnish provinces from 1975 to 1987 [Juuso 1995b]. The Kohonen network represents well most of the cases. However, it can ignore cases with very few data points in the data set.

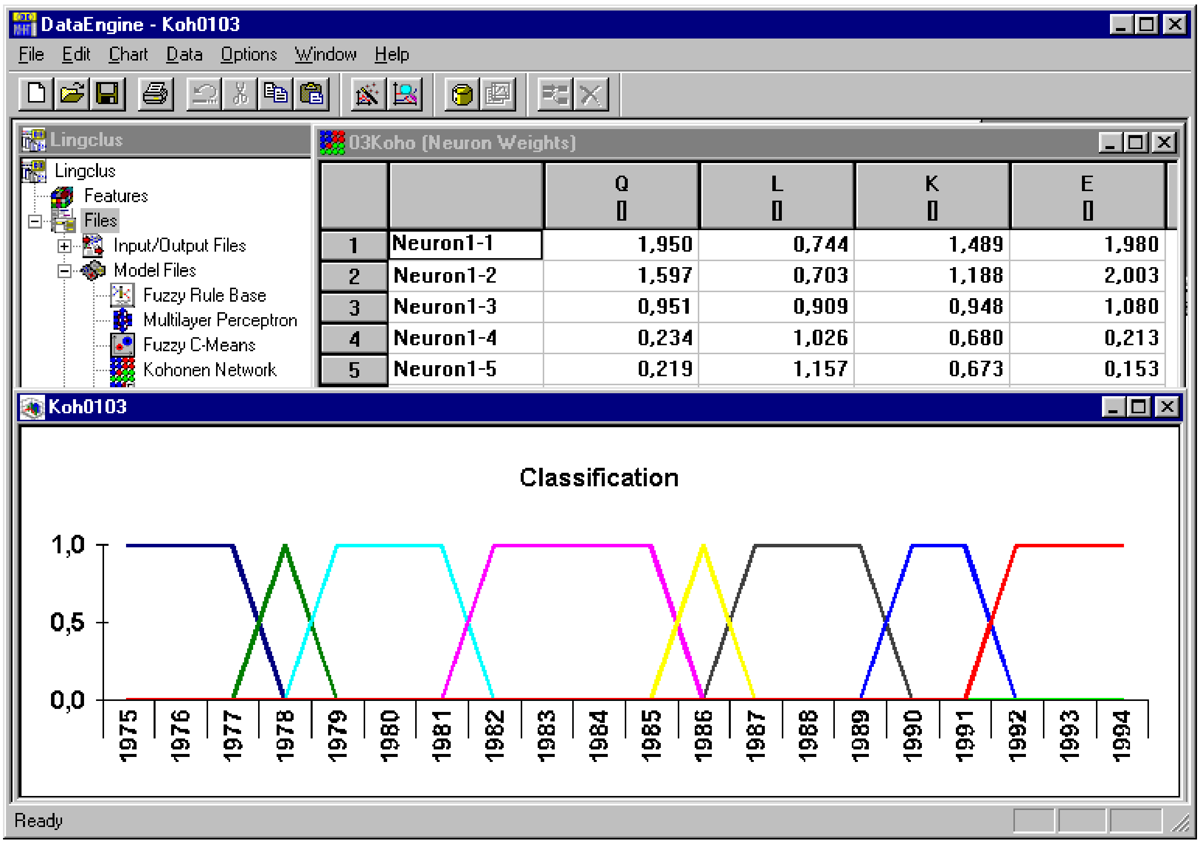

Figure 3: Classification with a Kohonen network.

The modelling has been continued in *DataEngineâ* with extended data including also the years 1988-1994. The data pre-processing was repeatedly performed in constructing the data sets for analysing both different regions and different industry. Merging, selecting columns, scaling and splitting are included to the data pre-processing card. Training, testing and recalling was done interactively by changing network parameters and training parameters. Two-dimensional Kohonen network with 2x3 neurons was used.

The example corresponding to the regional data of the whole industry was very suitable for generalisation with linguistic equations. Also larger networks were studied for direct application of the Kohonen network. The inputs and results can be incorporated into the card. The panel data was analysed with the Kohonen network, which classifies the data points to different classes described by neuron weights. Only one class is selected at a time. In the example shown in Figure 3, the selected classes are active for one to four years. Although the meanings of the neurons can be easily found for small systems, the transformation to fuzzy and linguistic systems will improve understanding of the system.

LINGUISTIC EQUATION MODELLING

The meaning of each neuron is important to understand, and therefore transformations to fuzzy set systems and linguistic equation models have been developed. A fuzzy model with 25 linguistic rules was generated from the weight matrix of the Kohonen network by classifying the network weights into fuzzy numbers corresponding to linguistic labels *very small*, *small*, *normal*, *large* and *very large* [Juuso 1995b]. The class *normal* for each variable was selected to represent the group of neurons, which is most common in data. For linguistic equations, integer numbers -2, -1, 0, 1 and 2 replaced these labels. The resulting linguistic equation was fuzzy, i.e. three different interaction vectors were developed. Introducing a finer partition can reduce fuzziness, but it is also possible that the interactions depend on the region or on the industry.

Fuzziness of the equations can be removed from interaction matrices by tuning the membership functions [Juuso 1999]. Clusters generated by fuzzy clustering or by Kohonen networks can be used as input data for automatic linguistic equation modelling. The present *FuzzEqu* system implemented in Matlab® is based on the real-valued equations: nonlinearities are handled by membership definitions, and linguistic equations are generated from input data transformed by these definitions. As in the previous study [Juuso 1995b], the interaction vector was [1 -1 -1 -1] for output, labour, capital and education, respectively. The directions of interactions are consistent with econometric theories. Since the strengths of interactions are different, a vector [3 -1 -2 -2] improved the fitting. Figure 1 shows the membership definitions after tuning with this interaction matrix. The resulting model surfaces are smoothly changing (Figure 2).

The resulting LE model can be used to any direction, i.e. any of the variables can be calculated if the others are known. For each variable, the value calculated from other three variables was compared to the value in the data. The model represents the Kohonen network very accurately (Figure 4). The accuracy is lowest for the labour, which has the smallest coefficient in the interaction matrix.

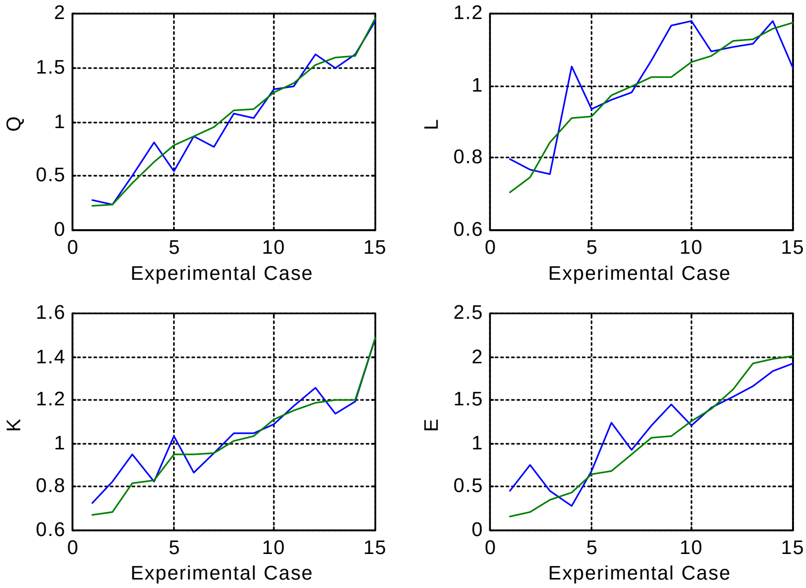

Figure 4: Testing of the LE model with the neurons of the corresponding Kohonen network.

Most cases in the complete panel data are also well represented (Figure 5). The differences in the interactions, which can be seen as an increased fuzziness of the model, are related to changes in regional economic conditions, which should be handled separately. Detecting these special conditions could be done with a case-based reasoning -type fuzzy linguistic equation system similar to the web break sensitivity indicator for paper machines [Juuso *et al.* 1998].

The resulting LE model can be considered as a new type of neural network, *linguistic Kohonen network*, where each neuron weight has also a linguistic meaning. The linguistic neuron model shown in Figure 6 has an interesting interpretation: high economic output is related to a high value for both the capital and the education cost and a low value for the labour cost. This network can be adapted to changing conditions, e.g. different regions and industry, by adjusting membership definitions.

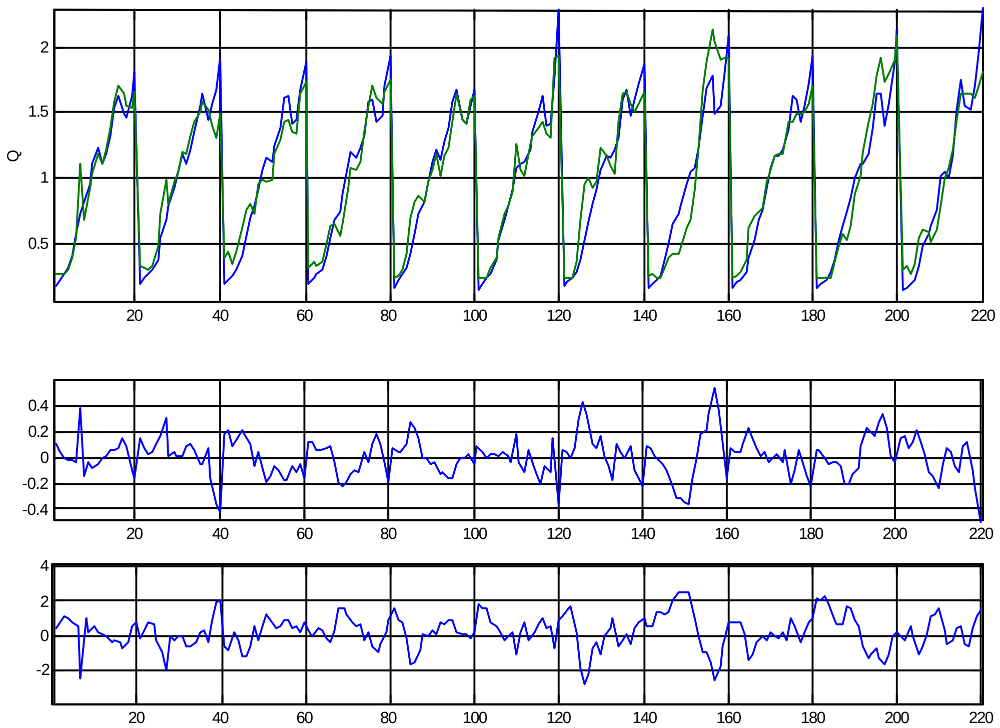

Figure 5: Testing the linguistic equation model with the panel data.

The LE model developed in *FuzzEqu* Toolbox was transferred to *DataEngine*® for further use in analysis of regional data and forecasting. The analysis can lead to spotting trends in movements of properties of regions and industry. The cases with high fuzziness of the LE model are located to certain linguistic neurons by the linguistic Kohonen network. The dynamics of the system is different in these cases, i.e. the production is changing faster or slower.

For forecasting, dynamic LE models are used separately for each data set in the panel data. The normal model is modified for special cases depending on the region or on the industry: several classes detected by linguistic Kohonen network can be handled in the same way. The model structures are adopted from process control applications where the process dynamics has been simulated fairly accurately with dynamic LE models [Juuso 1998].

The LE model can be transformed into a knowledge-based fuzzy system by locating membership functions into suitable linguistic values. Clusters developed by fuzzy clustering techniques can be processed in the same way as the Kohonen neurons. The combination of pre-processing, statistics and intelligent data analysis leads to an efficient tool in nonlinear multivariable modelling. To integrate different approaches a *Linguistic Equation PlugIn* is in progress for *DataEngine*®.

CONCLUSIONS

Nonlinear multivariable modelling of regional economics was extended with *Kohonen networks* and *Linguistic equations (LE)*: Kohonen networks were used for clustering, and the resulting neurons are used as an input data for generation and tuning of linguistic equations. In this way the system can be kept in a compact form. The resulting LE model, which can be used to any direction, represents the Kohonen network very accurately. Most cases in the time series are also well represented. The differences in the interactions, which can be seen as an increased fuzziness of the model, are related to changes in regional economic conditions, which should be handled separately. The LE model can be considered as a new type of neural network, *linguistic Kohonen network*, where each neuron weight has also a linguistic meaning. This network can be adapted to changing conditions, e.g. different regions and industry, by adjusting membership definitions.

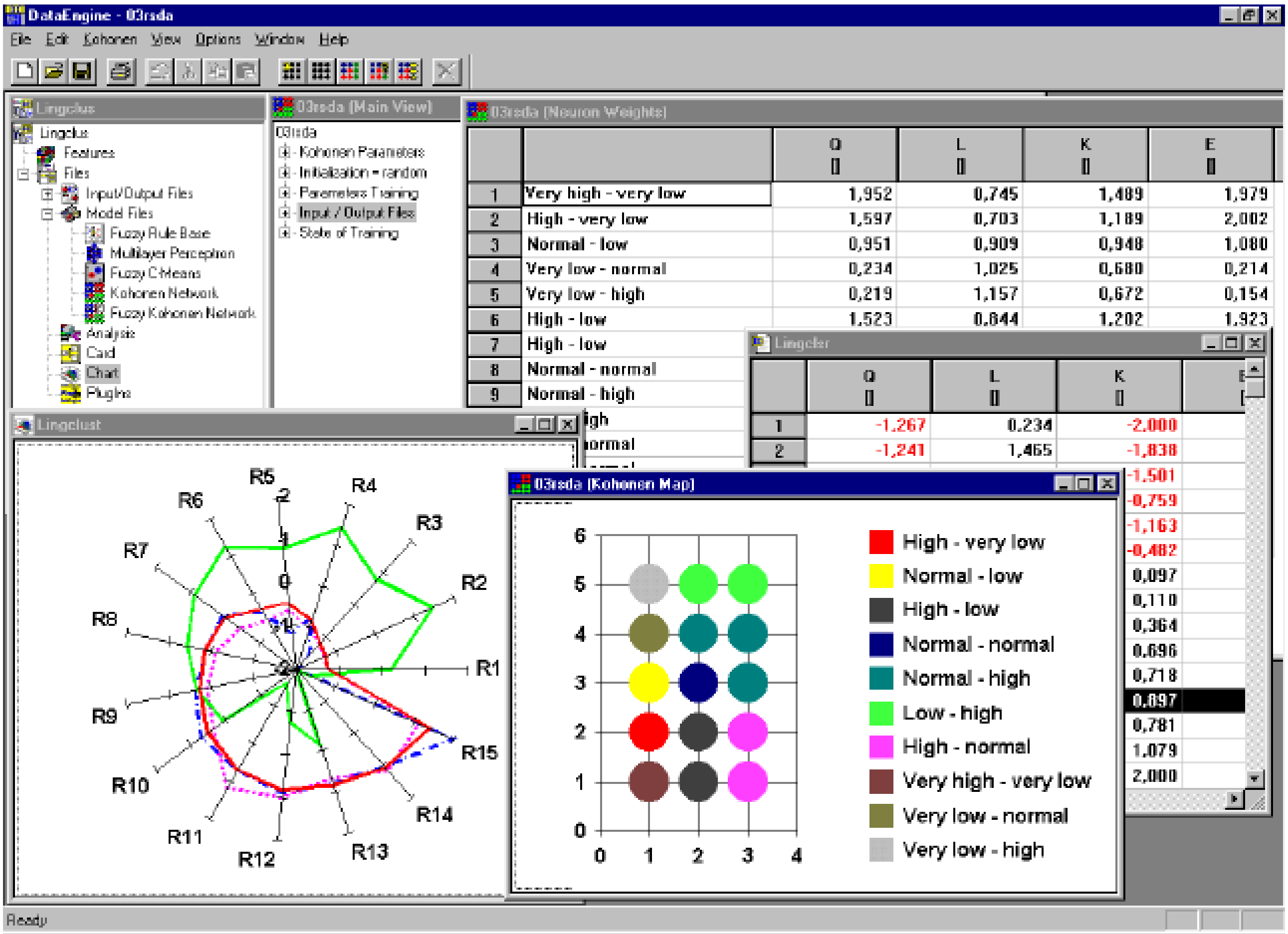

Figure 6: Linguistic Kohonen network in *DataEngine*®.

REFERENCIES

Juuso, S. 1995a. Productivity-augmenting Effects of Education on Economic Output Analyzed within a Fuzzy Production Function Framework, M.Sc. thesis (in Finnish), Department of Economics, University of Oulu, 124 p.

Juuso, S. 1995b. Neural Networks and Neural Network Modelling with Matlab, University of Manchester, School of Economic Studies, 28 p.

Juuso, E. K. (1999). Fuzzy Control in Process Industry: The Linguistic Equation Approach. In: Verbruggen, H. B., H.-J. Zimmermann and R. Babuska, editors, Fuzzy Algorithms for Control, International Series in Intelligent Technologies, pp. 243-300. Kluwer, Boston.

Juuso, E. K. 1998. Intelligent Dynamic Simulation of a Lime Kiln with Linguistic Equations. Proceedings of ESM'98, Manchester, June 16-19, 1998, Volume 2, pp. 23-28.

Juuso, E. K., T. Ahola, and K Oinonen 1998. Case-Based Reasoning with Linguistic Equations for Web Break Sensitivity Indicator. Preprints of TOOLMET'98 Symposium, April 16-17, Oulu, pp. 97-106.

Kuan, C.-M. and H. White 1994. Artificial Neural Networks: An Econometric Perspective. Econometric Reviews, Vol. 13(1), pp. 1-94.

Masters, T. 1993. Practical Neural Network Recipes in C++. Academic Press, Inc., pp. 493.

Nijkamp, P. and A. Reggiani 1995. Non-linear Evolution of Dynamic Spatial Systems: The Relevance of Chaos and Ecologically-based Models. Regional Science and Urban Economics, Vol. 25, pp. 131-151.

Ormerod, P., J.-C. Taylor and T. Walker 1991. Neural Networks in Economics in: Money and Financial Markets, Blackwell, Oxford, pp. 341-353.

Swamy, P.A.V.B. 1970. Efficient Inference in a Random Coefficient Regression Model. Econometrica V0l.38, pp. 311-323.