FORECASTING WITH LEADING ECONOMIC INDICATORS – A NON-LINEAR APPROACH

Timotej JAGRIC*

Abstract:
Leading economic indicators have a long tradition in forecasting future economic activity. Recent developments, however, suggest that there is scope for adding extensions to the methodology of forecasting major economic fluctuations. In this paper, the author tries to develop a new model, which would outperform the forecast accuracy of classical leading indicators model. The use of artificial neural networks is proposed here. For demonstration a case study for Slovene economy is included. The main finding is that, at the twelve months forecasting horizon, a stable and improved forecast accuracy could be achieved for in- and out-of-sample data.

Keywords: leading economic indicators, neural network, forecasting, aggregate economic activity

JEL Classification: C45, E37

1. Introduction

There is variety of important issues associated with the problem of business cycle forecasting, especially regarding forecast methodology and forecast evaluation. Overall, we can say that macroeconomic forecasting has a fairly poor reputation (see Granger, 1996). Still, even with the recognition that forecasting business cycles is a very difficult task, we find some hopeful signs for future progress. Our research on forecasting has focused on development of new approach in forecasting with leading indicators. In our further analysis, we developed a model for monthly forecasts, which represents a new approach in current methodology of classical leading economic indicators, which was developed at National Bureau of Economic Research (NBER).

---

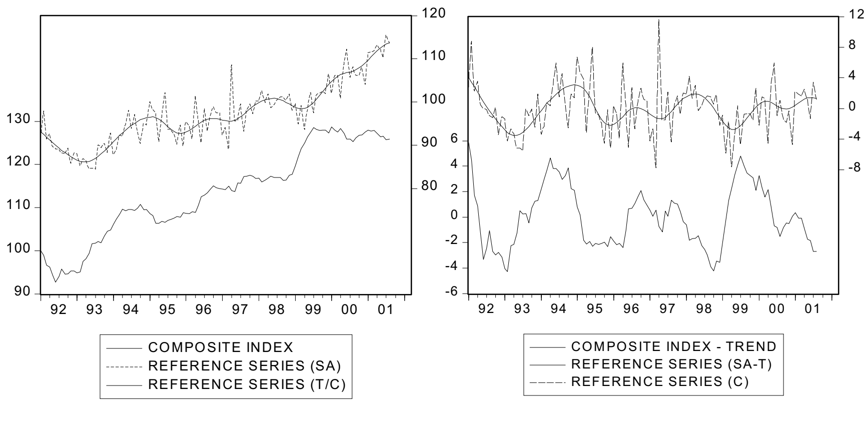
*) Faculty of Economics and Business, University of Maribor, Slovenia, Razlagova 14, 2000 Maribor, Slovenia (e-mail: timotej.jagric@uni-mb.si).

**) This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network (GDN). Additional funds for grantees in the Balkan countries have been provided by the Austrian Government through WIIW, Vienna. All opinions expressed are those of the authors and have not been endorsed by CERGE-EI, WIIW, or the GDN. The author would like to thank Department for Macroeconomic Analysis of the Central Bank of Slovenia, Statistical Office of Slovenia and Institute of Macroeconomic Analysis and Development for their help in the process of developing the database.

The original NBER model was built solely within linear framework. With recent developments in non-linear time-series analysis, several authors have begun to examine the forecasting properties of non-linear models in economics. Probably the largest share of economic applications of non-linear models can be found in the field of prediction of time-series in capital markets. Usually, linear models of financial time-series perform poorly and linear univariate models consistently give evidence for a random walk (see Meese, Rogoff, 1983; Lee et al., 1993). In a study more comparable to ours, Jaditz, Riddick and Sayers (1998) use financial variables to forecast industrial production. They estimated a non-linear non-parametric nearest-neighbor regression model. Also superior results over linear models achieved Tkacz (2001), by forecasting Canadian Gross Domestic Product (GDP) growth. Tiao and Tsay (1994) show that a simple threshold autoregressive model is superior to an AR(2) representation for GDP growth. Maasoumi, Khotanzad and Abaye (1994) show that the 14 macroeconomic series in the Nelson and Ploser (1982) study are non-linear processes rather than unit root processes.

Consistent with this view, we find that in the case of Slovenia, index of leading indicators (SLOLEI) is a weak predictor of economic activity. The forecast of economic activity is based on calculation of composite index. We observed the behavior of the index in the period from I: 1992 to IX: 2001. Turning points were determined with trend-cycle, which was estimated with program X11ARIMA (see Statistics Canada, 1999). The model forecasted eight of ten reference points correctly (see Figure 1). More important, however, is that there is no stable forecast horizon. The turning points were forecasted between two and ten months ahead, with average of six months. The behavior of the composite index after year 1998 deviates strongly from the behavior of the reference series and there are also stronger shifts of the forecast horizon. This indicates that the model does not capture important changes in the economy.

Figure 1
Classical NBER Leading Indicator Model for Slovenia (SLOLEI – composite index)



Another important issue, that has to be considered, is the prediction of future value of reference series. The classical model of leading indicators can only provide a sign for a turning point in aggregate economic activity. It is not possible to exactly define when the turning point will occur, neither how strong the following contraction or expansion will be. Therefore a reliable composite index of leading indicators should possess following properties (see Fritsche, Stephan, 2000): the movements in the index should resemble those in the business cycle reference series; the relationship between the reference series and the indicator should be statistically significant; the forecasting performance should be stable over time.

Our approach differs from previous studies in several ways. First, we try to modify original NBER model with the aim of overcoming the deficiencies of the model. Second, our focus is on constructing multivariate neural network forecasting model. Third, our model is used for monthly forecasts.

The decision to focus on the neural networks arises directly from the features of these models, as described by Bishop (1995). First, neural networks are data-driven and can "learn" from, and adapt to, underlying relationships. This is useful in contexts where one does not have any a priori beliefs about functional forms. Second, when properly specified they are universal functional approximates, implying that they can approximate functional forms to any given degree of accuracy. Finally, neural networks are non-linear, which seems to be a case for many macroeconomic time-series.

The paper is organized as follows. In section 2 the structure of the database and the selection of reference series is explained. The selected input variables of the model are presented in section 3. The suggested model is explained in section 4. In section 5 of the paper we present the results. Section 6 contains the conclusion.

2. Data

Important step in the process of constructing the model is development of a broad database, which should cover all crucial fields of economic activity. The database, which was used in the model, includes 365 time series. To ensure sufficient transparency, the time series are classified in categories, which are presented in Table 1. The database can be divided in two major groups of time series: time series representing Slovene economic activity, time series representing foreign economic activity.

Since Slovenia has got independence in October 1991, the time series started in January 1992. At the moment the database covers the period I: 1992 – VIII: 2001. In the final version of the database, all time series were transformed into growth rates. All series, which are presented in current prices, were converted using consumer price index (CPI 1999 = 100).

The aim of the model is to forecast a selected reference variable, which is the benchmark that indicates fluctuations in the economic activity. The variable must have the advantage of being a monthly reported variable, available for many countries, and measure the real sector of the economy. There are two alternative strategies for obtaining a time series that represents current business activity on a monthly level (see Dias, 1994): either adopt a single series as the variable of interest or use a function of several variables. Both approaches have long traditions in empirical macroeconomics. For example, the empirical literature on the monthly money-income relationship focuses on the predictability of monthly industrial production. Alternatively, Burns and Mitchell (1946) constructed a reference series by averaging several different major aggregate time series; this reference series was then used to date their reference cycles.

The construction of leading economic indicators demands a monthly and up-to-date series. Another important issue is the fact, that in Slovenia time series cannot be longer then nine years, since Slovenia got independence in October 1991. This makes it difficult to determine, whether the selected time series has the characteristics of coincident indicator. What is more, the aim of this research was not to con-

Table 1
Database Structure

Code	Category	Number of series included
S01	Industrial production	27
S02	Construction	6
S03	Trade	6
S04	Tourism	15
S05	Transport	8
S06	Export	23
S07	Import	23
S08	Balance of payments	10
S09	Employment	10
S10	Wages	21
S11	Unemployment	5
S12	Labor costs and productivity	16
S13	Money aggregates	5
S14	Bank claims	36
S15	Bank liabilities	45
S16	Interest rates (active)	10
S17	Interest rates (passive)	12
S18	Exchange rates	15
S19	International liquidity	3
S20	Government expenditure	10
S21	Prices	16
S22	Consumption	4
S23	Foreign activity indicators	22
S24	Indicators of German economic activity	17
	Total	365

struct a perfect composite coincident indicator, but to construct composite leading indicator. Therefore we selected monthly index of total industrial production. Extensive analysis (see Jagric, 2002) of such reference variable also gave support to our decision, since it was discovered, that industrial production has same cyclical characteristics as gross domestic product in Slovenia.

3. Scoring System for Business Cycle Indicators

To construct a forecasting model, a selection of input variables is needed. In our study we extended the use of criteria employed by NBER, by adding some elements of Stock-Watson approach in the scoring system. The scoring of each series reflects our desire not only to make as explicit as possible the criteria for selecting indicators but also to increase the amount of information available to the user in order to aid in evaluating their current behavior. The scoring plan includes five major elements: economic significance, statistical adequacy, promptness of publication, smoothness, conformity and timing. When the subheads under these elements are counted, eight different properties of series are rated in all.

A high score for economic significance is accorded to a series that succeed to measures a variable, which has an important role in the analysis of business cycle movements. A series that represents a strategic process more broadly is rated higher than one more narrowly defined. Such broadly defined series is also less likely to shift as a result of technological developments, changing consumer tastes, and other similar factors.

Statistical adequacy reflects the requirement that a series continue to measure the same economic process during future business cycle fluctuations, when the selected indicators are put to the hard test of current usage. The main element is the stationarity. This characteristics of a time series is important due to demands of further econometric testing. Other elements are: coverage of time unit, measure of revisions and comparability throughout the period.

For short-run business cycle forecasting a leading indicator can be useful only, if it is up to date. Series that are released promptly, therefore, are assigned higher scores than those that lag in publication.

The smoothness criterion is the same as in original NBER scoring plan. Since the start of a new cyclical phase can be discerned more promptly in a series, which is smoother than in one, which is irregular, smoother series are given higher ratings. Due to the fact, that we only use monthly series, only months of cyclical dominance (MCD) value was used to measure the smoothness. The MCD value is reported by X11ARIMA program.

Conformity of an indicator to past business cycles and timing of its turning points relative to those in aggregate economic activity are obviously essential qualities in an indicator. Since in our case the time-series can only cover a period of up to eight years, we could not apply the NBER approach of a probability test. Therefore we used a criterion, which is based on Granger causality (see Jagric, 2001). This enabled us to introduce econometric testing into scoring system. Econometric testing was performed on all series in the database four-times. First we tested the series presented as annual growth rates (original and seasonally adjusted) and than as monthly growth rates (original and seasonally adjusted). This was necessary, since many researchers found out, that long term relationship between two time series may depend on data transformation procedures (see Charemza, Deadman, 1992).

Since we want to find series, which have information about cyclical component of reference series, we used additional econometric procedures. Burns and Mitchell (1946) quantified comovements in terms of leads or lags at turning points of each series relative to the reference cycle and in terms of their index of conformity. More recent work has focused on the second moment of the joint distribution of the series of interest. For example, Hymans (1973) summarized cyclical timing by estimating phases in the frequency domain at business cycle frequencies. This perspective – focusing on the second moment properties of the series – is adopted here.

All tests were performed on both original and seasonally adjusted series. This allowed us to test the stability of the results. But in the final forecasting model, only original time series were used. An example of a scoring table is presented in Table 2. All mentioned categories are used two-times, since we transformed original data into two datasets using following transformation for a selected time series y_t:

$$y_t(1) = \frac{y_t - y_{t-1}}{y_{t-1}} - 1 \qquad y_t(12) = \frac{y_t - y_{t-12}}{y_{t-12}} - 1 \tag{1}$$

Table 2
Example of a Scoring Table for Selected Time-series – y_t (max. scores)

Criteria	Score for $y_t(1)$	Score for $y_t(12)$	Total score
Economic significance: – aggregation – results in other studies	1 0.5 0.5	1 0.5 0.5	2
Statistical adequacy: – stationarity – coverage of time unit – measure of revisions – comparability throughout the period	1 0.25 0.25 0.25 0.25	1 0.25 0.25 0.25 0.25	2
Promptness of publication	1	1	2
Smoothness	1	1	2
Conformity and timing: – Granger causality – cross-spectral analysis	1 0.5 0.5	1 0.5 0.5	2
Total score			10

We decided to score all time-series. The total score of time-series, theoretical lead-time, and the results of graphical analysis, were than used to form the group of leading indicators, which includes 58 time-series from database. The average lead-time is determined by cross-spectral analysis (phase and coherency) and Granger test of causality, where two criteria were used: the value of adjusted determination coefficient, and Akaike information criteria. The average lead-time is only an estimate of actual lead-time for selected time-series. The scoring system, we have used, ensured that the selected indicators posses the best characteristics among all time-series in database. They cover different fields of economic activity. Table 3 presents the list of groups, which contain the selected time-series.

Table 3
Groups of Selected Leading Indicators

Code	Group name	Number of indicators (in %)
S04	Tourism	1 (1.7)
S05	Transport	1 (1.7)
S06	Export	5 (8.6)
S07	Import	6 (10.3)
S08	Balance of payments	1 (1.7)
S10	Wages	2 (3.4)
S12	Labor costs and productivity	5 (8.6)
S13	Money aggregates	3 (5.2)
S14	Bank claims	5 (8.6)
S15	Bank liabilities	3 (5.2)
S16	Interest rates (active)	1 (1.7)
S18	Exchange rates	4 (6.9)
S19	International liquidity	1 (1.7)
S21	Prices	3 (5.2)
S23	Foreign activity indicators	9 (15.5)
S24	Indicators of German economic activity	8 (13.8)
	Total	58 (100)

4. Neural Network Model

The origin of neural networks dates back to the 1940s. McCulloch and Pitts (1943) and Hebb (1949) researched networks of simple computing devices that could model neurological activity and learning within these networks, respectively. Later, the work of Rosenblatt (1962) focused on computational ability in perceptrons, or single-layer feed-forward networks. Proofs showing that perceptrons, trained with the Perceptron Rule on linearly separable pattern class data, could correctly separate the classes, generated excitement among researchers and practitioners. This excitement, however, waned with the discouraging analysis of perceptrons presented by Minsky and Papert (1969). The analysis pointed out that perceptrons could not learn the class of linearly inseparable functions. It also stated that the limitations could be resolved if networks contained more than one layer, but that no effective training algorithm for multi-layer networks was available.

Rumelhart, Hinton, and Williams (1986) revived interest in neural networks by introducing the generalized delta rule for learning by backpropagation, which is today the most commonly used training algorithm for multi-layer networks. More complex network types, alternative training algorithms involving network growth and pruning, and an increasing number of application areas characterize the state-of-the-art in neural networks (see Bishop, 1995). But no advancement beyond feed-forward neural networks trained with backpropagation has revolutionized the field. Neural networks, more accurately called Artificial Neural Networks (ANN), are computational models that consist of a number of simple processing units that communicate by sending signals to each other over a large number of weighted connections. The processing units transport incoming information on their outgoing connections to other units. The "electrical" information is simulated with specific values stored in those weights that make these networks have the capacity to learn, memorize, and create relationships amongst data. A very important feature of these networks is their adaptive nature where "learning by example" replaces "programming" in solving problems. This feature renders these computational models very appealing in application domains where one has little or incomplete understanding of the problems to be solved, but where training data are available (see Ripley, 1996).

The scoring system, which was used, determined the selection of input variables in our model. They cover different fields of economic activity. As expected, most important category is foreign activity (including S06, S23 and S24 in Table 3). All selected variables are presented by monthly data for the period from I: 1992 to VIII: 2001. The target variable (one that is to be forecasted) is the monthly index of industrial production, which was selected as reference series in the original model. The input and target variables are not seasonally adjusted. As Stock and Watson (1998) noted, seasonal adjustment procedure applied to the data may be cleansing them of any underlying non-linearities.

Since neural networks can perform essentially arbitrary non-linear functional mappings between sets of variables, a single neural network could, in principle, be used to map the raw input data directly onto the required final output values. In practice, such an approach will generally give poor results. For many practical applications is therefore the choice of pre-processing one of the most significant factors in determining the performance of the final system.

There is a key assumption, which is implicit in our approach to time-series prediction, which is, that the statistical properties of the generator of the data are time-independent. Provided this is the case, then de-trending will map the time-series problem onto a static function approximation problem, to which a feed-forward network can be applied. In our case many selected input variables (including the target variable) show an underlying trend. As there is no universal trend function, which could be applied to all variables, we decided to use one-year and monthly growth rates (in decimal). This linear transformation also ensured that the values of variables do not differ significantly from each other and are between –1 and 1.

Another important reason for pre-processing is phenomenon known as the "course of dimensionality". If we are forced to work with a limited quantity of data, as we are in practice, then the dimensionality of the input space can rapidly lead to the point where the data are very sparse, in which case they provide a very poor representation of the mapping. Therefore our goal is, to map vectors $\boldsymbol{x}^n$ in a d-dimensional space ($x_1, \dots, x_d$) onto vectors $\boldsymbol{z}^n$ in an M-dimensional space ($z_1, \dots, z_M$), where $M < d$. To achieve this goal, we use unsupervised linear transformation technique known as principal component analysis (PCA), where a set of data are summarized as a linear combination of an orthonormal set of vectors. The data $\boldsymbol{x}_i$, $i = 1, \dots, n$, are summarized using the approximating function

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{V}) = \boldsymbol{u} + (\boldsymbol{x}\boldsymbol{V})\boldsymbol{V}^T \qquad (2)$$

where $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{V})$ is a vector-valued function, $\boldsymbol{u}$ is the mean of the data $\{x_i\}$, and $\boldsymbol{V}$ is a d x m matrix with orthonormal columns. The mapping $\boldsymbol{z}_i = \boldsymbol{x}_i\boldsymbol{V}$ provides a low-dimensional projection of the vectors $\boldsymbol{x}_i$ if $m < d$. The principal component decomposition estimates the projection matrix $\boldsymbol{V}$ which minimizes the empirical risk

$$E_M(\boldsymbol{x},\boldsymbol{V}) = \frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{x}_i - \boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{V})\right\|^2 \tag{3}$$

subject to the condition that the columns of $\boldsymbol{V}$ are orthonormal. The parameter matrix $\boldsymbol{V}$ and the projection vectors $\boldsymbol{z}$ are found using the singular value decomposition of the n x d data matrix $\boldsymbol{X}$, given by

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \tag{4}$$

where the columns of $\boldsymbol{U}$ ($\boldsymbol{V}$) are the eigenvectors of $\boldsymbol{X}\boldsymbol{X}^T$ ($\boldsymbol{X}^T\boldsymbol{X}$). The matrix $\boldsymbol{\Sigma}$ is diagonal and its entries are the square roots of the nonzero eigenvalues of $\boldsymbol{X}\boldsymbol{X}^T$ or $\boldsymbol{X}^T\boldsymbol{X}$. Let us assume that the diagonal entries of the matrix $\boldsymbol{\Sigma}$ are placed in decreasing order along the diagonal. These eigenvalues describe the variance of each of the components. To produce a projection with dimension $m < d$, which has maximum variance, all but the first m eigenvalues are set to zero. Then the decomposition becomes

$$\boldsymbol{X} \cong \boldsymbol{U}\boldsymbol{\Sigma}_m\boldsymbol{V}^T \tag{5}$$

The m-dimensional projection vectors are given by

$$\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{V}_m \tag{6}$$

where $\boldsymbol{Z}$ is an n x m matrix whose rows correspond to the projection $\boldsymbol{z}_i$ for a given data sample $\boldsymbol{x}_i$ and $\boldsymbol{V}_m$ is a d x m matrix constructed from the first m columns of $\boldsymbol{V}$.

The estimated principal components provide a linear approximation that represents the maximum variance of the original data in a low-dimensional projection. They also provide the best low-dimensional linear representation in the sense that the total sum of squared distances from data points to their projections in the space are minimized.

More challenging was to determine the design of neural network. Changes to the architecture can fundamentally alter the forecasts produced by the network, even when no changes are made to the inputs, outputs or sample size. Various techniques have been developed for optimizing the architecture. It is important to distinguish between two distinct aspects of the architecture selection problem. First, we need a systematic procedure for exploring some space of possible architectures, and this forms the subject of this selection. Second, we need some way of deciding which of architectures considered should be selected. This is determined by the requirement of achieving the best possible generalization. In our case we first formed some basic requirements for the network:

– we will use feed-forward back-propagation network, since this type of network is mostly used in forecasting applications;

– due to the type of input and output data, we will only use two types of transfer function: pure linear and tan-sigmoid activation function;

– all neurons will have a bias;

– the output-layer will only have one neuron, since we try to predict the future value of one reference series;

– the network will not have more than three layers of neurons, since such network is a universal approximator.

The above requirements have greatly reduced the space of possible architectures. Therefore we could apply a simplified tiling algorithm (see Mezard, Nadal, 1989). We build the network in successive layers with each layer having fewer units than the previous layer. When a new layer was constructed, a single unit, called the master unit is added. Then step-by-step additional units are added. In every step the network is trained and the forecasting performance is estimated. The whole process is repeated until a larger network does not sufficiently contribute to the forecasting performance.

To train our network, we use quasi-Newton approach, since it has a significant advantage over the conjugate gradient method, which is normally used for training. The only potential disadvantage of this approach is, that in large networks we could encounter prohibitive memory requirements. The quasi-Newton approach involves generating a sequence of matrices $\mathbf{G}^{(\tau)}$, which represents increasingly accurate approximation to the inverse Hessian $\boldsymbol{H}^{-1}$, using only information on the first derivates of the error function. From the Newton formula

$$\boldsymbol{w}^* = \boldsymbol{w} - \boldsymbol{H}^{-1}\boldsymbol{g} \tag{7}$$

where $\boldsymbol{w}$ is the weight vector, we see that the weight vectors at steps τ and $\tau + 1$ are related to the corresponding gradients by

$$\boldsymbol{w}^{(\tau+1)} - \boldsymbol{w}^{(\tau)} = -\boldsymbol{H}^{-1}(\boldsymbol{g}^{(\tau+1)} - \boldsymbol{g}^{(\tau)}) \tag{8}$$

which is known as the quasi-Newton condition. The approximation $\boldsymbol{G}$ of the inverse Hessian is constructed so as to satisfy this condition also. The two most commonly used update formulae are the Davidson-Fletcher-Powell (DFP) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) procedures. We used the BFGS procedure, since this is generally regarded as being superior (see Bishop, 1995):

$$\boldsymbol{G}^{(\tau+1)} = \boldsymbol{G}^{(\tau)} + \frac{\boldsymbol{p}\boldsymbol{p}^T}{\boldsymbol{p}^T\boldsymbol{v}} - \frac{\left(\boldsymbol{G}^{(\tau)}\boldsymbol{v}\right)\boldsymbol{v}^T\boldsymbol{G}^{(\tau)}}{\boldsymbol{v}^T\boldsymbol{G}^{(\tau)}\boldsymbol{v}} + \left(\boldsymbol{v}^T\boldsymbol{G}^{(\tau)}\boldsymbol{v}\right)\boldsymbol{u}\boldsymbol{u}^T \tag{9}$$

where we have defined the following vectors:

$$\boldsymbol{p} = \boldsymbol{w}^{(\tau+1)} - \boldsymbol{w}^{(\tau)} \quad \boldsymbol{v} = \boldsymbol{g}^{(\tau+1)} - \boldsymbol{g}^{(\tau)} \quad \boldsymbol{u} = \frac{\boldsymbol{p}}{\boldsymbol{p}^T\boldsymbol{v}} - \frac{\boldsymbol{G}^{(\tau)}\boldsymbol{v}}{\boldsymbol{v}^T\boldsymbol{G}^{(\tau)}\boldsymbol{v}} \tag{10}$$

Derivation of this expression can be found in many standard texts on optimization methods such as Polak (1971), or Luenberger (1984). It is straightforward to verify by direct substitution that BFGS expression satisfy the quasi-Newton condition.

One of the problems that occur during neural network training is called overfitting. The error on the training set is driven to a very small value, but when new data is presented to the network the error is large. The network has memorized the training examples, but it has not learned to generalize to new situations. One method for improving network generalization is to use a network that is just large enough to provide an adequate fit. If we use a small enough network, it will not have enough power to overfit the data (see Hagan et al., 1996). Since we do not know how large a network should be in our application, we selected regularization. The technique of regularization encourages smoother network mappings by adding a penalty Ω to the error function to give

$$\tilde{E} = E + v\Omega \tag{11}$$

Here E is a standard mean sum of squares error function, and the parameter v controls the extent to which the penalty term (regularizer) $\boldsymbol{\Omega}$ influences the form of solution. In the training process we decided to use a simple form of regularizer called weight decay. It consists of the sum of the square of the adaptive parameters in the network

$$\Omega = \frac{1}{2}\sum_i w_i^2 \tag{12}$$

where the sum runs over all weights and biases. In conventional curve fitting, the use of this form of regularizer is called ridge regression. It has been found empirically that a regularizer of this form can lead to significant improvements in network generalization (see Hinton, 1987).

5. Results

In the process of network architecture selection, we tested every form for 6, 9, 12, and 15 months forecast (all calculations were performed with Matlab v.6.0 R12). We developed special program, which automatically supervised the testing and recorded the performance. The performance was measured with different criteria: total mean square error for in- and out-of-sample data, determination coefficient for regression between target variable and estimated variable, and comparison of a spectrum for target and estimated variable.

Testing was performed on two computers. Major testing was performed on IBM Server X220 with Pentium X-III 866 MHz. The stability of results was tested on the second computer with Pentium III 500 MHz. Testing required about 600 computing hours on the server since every estimation of a possible architecture was repeated max. 5000 times (the number of repeated estimations depended on the complexity of a network). To avoid idle processing time, the developed program controlled this process. After each cycle the program reported final results, which were then tested manually on the second computer.

It has frequently been argued that statistical error measures do not measure the right thing. Therefore many authors developed additional evaluation methods (see Stekler, 1991; Leitch, Tanner, 1991). In the experimental phase of testing, we discovered that even if a network scores high, the results may not be stable or are not well suited for forecasting. Therefore we developed a new testing routine for networks with two or more layers:

– first we developed a smoothed series based on inverse Fourier transform of reference series. Such series includes only cyclical components, which we try to forecast;

– we add an additional neuron to the output layer with the same characteristics as already present neurons. This allows us to forecast two time-series at the same time. Estimation procedure is therefore changed;

– we test such network and compare the results with the results from major testing. If there are no major differences, the results are treated as stable and reliable.

After extensive testing of possible forms of network architectures, we selected a neural network, which can be represented with following equations:

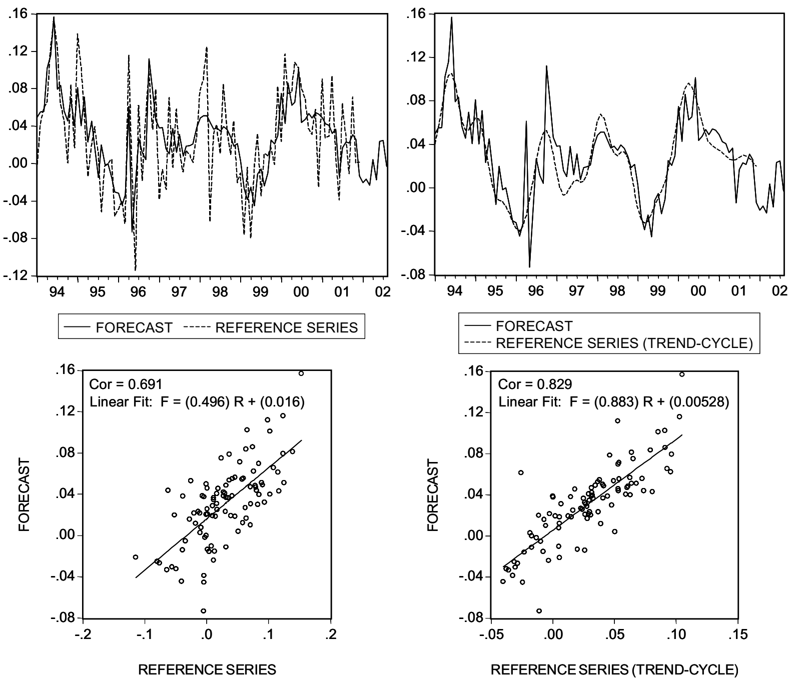
$$\begin{gathered} \boldsymbol{y} = \boldsymbol{f}^{3,1}\left(\boldsymbol{w}_{3,2}\boldsymbol{f}^{2,2}\left(\boldsymbol{w}_{2,6}\boldsymbol{f}^{1,3}\left(\boldsymbol{w}_{1,12}\boldsymbol{x}\right)\right) + \boldsymbol{b}_{3,1}\right) \\ f^i = \text{purelin}(n) = n \qquad i = 3,1 \\ f^j = \text{tansig}(n) = \frac{2}{1+e^{-2n}} - 1 \quad j = 2 \end{gathered} \tag{13}$$

where $\boldsymbol{y}$ is output vector, $\boldsymbol{f}^{i,z}$ is matrix of z activation functions of neurons in layer i, and $\boldsymbol{w}_{i,q}$ is matrix of q weights in layer i. Only output layer of neurons has a bias vector, which are represented by $\boldsymbol{b}_{i,z}$. As it can be seen, we selected a three-layer network. The input layer has three neurons (each with pure linear activation function), hidden layer has two neurons (each with non-linear tansig activation function), and the output layer has one neuron (with pure linear activation function). The selected model performed best for 9, 12 and 15 months forecast. Forecasts for longer periods did not produce good results – as we expected, since we did not have long time-series.

By using principal component analysis we were able to reduce the input space significantly. Instead of 58 input variables, we used only seven principal components, since we eliminated those principal components that contribute less than 10 % to the total variation in the data set of input variables.

The forecast performance of the new model of leading indicators for Slovenia is presented in Figure 2. The parameters of the model were estimated on the data from I: 1993 to VIII: 1997. The data after VIII: 1997 were not presented to the model in the estimation phase and can be therefore treated as out-of-sample data. As the upper two graphs in Figure 2 suggest, the model capture the dynamics of the reference series well. All turning points were detected and the forecasted value follows the dynamics of the reference series. The period from 1996 till 1998 is also well forecasted, since the model detected two peaks.

Figure 2
Performance of Neural Network Model



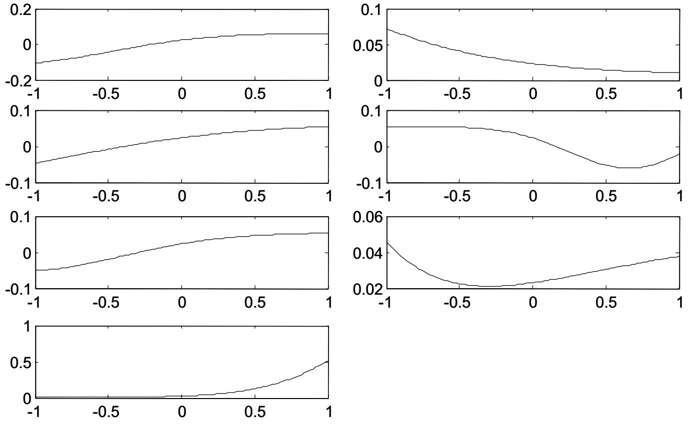
Additional measures of forecast performance:
– Mean Squared Error = 0.001505
– Mean Absolute Error = 0.030141

The statistical properties of the estimated model are presented in lower two graphs. We performed post-regression analysis, where we compared original (target) data with forecasted (estimated) data for in- and out-of-sample data. As it can be seen, the model captures best the business cyclical frequencies of the reference series. The high frequencies are not detected. This was expected, since we use principal components – non-measurable indicators, which do not have all the information of original leading indicators. In addition, we only searched for linkage between reference and leading indicators in business cycle frequencies. Additional measures of forecast performance have not an important informative value, since we did not develop competitive models. Presented mean squared and absolute errors were used in the process of estimating the parameters (weights) of the model.

The complex non-linear functional form of the network makes it very difficult to interpret the estimated network weights. In linear regression models, the values of the estimated coefficients provide a direct measure of the contribution of each variable to the model's output. Because of these difficulties, neural networks are sometimes called "a black box" (see Gonzalez, 2000). This problem, however, can be greatly alleviated by applying the sensitivity analysis proposed by Refenes, Zapranis and Francis (1994). Their method consists in charting the value of the output for a range of values of given input, while all other inputs are fixed at their sample mean. If the value of the output remains relatively stable for different values of the input in question, we can assume that this input does not contribute significantly to the predictive power of the network.

The results of the sensitivity analysis are presented in Figure 3. For every graph we allowed only one variable to change its value. All other variables were set to zero. As expected, the sensitivity analysis suggests that there is strong non-linear connection between input variables and reference series. Since we used seven principal components as input variables, we have no direct information about relationship between original input variables and reference series. The results of Figure 3 also suggest that all seven principal components have a strong influence on the forecasted value of reference series. Therefore a reduction of number of input variables would probably reduce the forecast performance of the model.

Figure 3
Sensitivity Analysis



Note: The original leading indicators were reduced to seven principal components, which represent 90 % of total information in 58 indicators.

6. Conclusion

Policy and investment decisions are made with an eye toward future economic conditions, and a forecasting model that can correctly forecast directional changes in the business cycle would be a boon to policymakers, the business community, and the general public. Our research targets this issue by developing a new forecasting model.

More than three hundred series have been evaluated in the process of selecting leading indicators, which are limited to the role of indicators of short-run movements in aggregate economic activity only. The list of 58 leading indicators we have chosen suggests, that foreign economic activity has an important impact on Slovene economy. We find important that almost half of selected indicators are present also in leading indicators for OECD countries. Serious problem represent time-series, which are shorter than nine years. In that case the characteristics of cyclical component of such time-series are not determined with high reliability and were therefore not included in the database.

The shortcoming of selected indicators is non-presence of three indicators, which are usually used in OECD countries: retail inventories, index of stock exchange, and prices of primary commodities. Retail inventories were not included, since Statistical Office of the Republic of Slovenia does not collect these data on monthly basis. The prices of primary commodities were not selected for two reasons: firstly, we found significant relationship between reference series and these variables for the period 1997 – 2001 only; secondly, significance was found in this period only when high price changes have occurred. We also tested the index of stock exchange, but we did not discover any significant relationship with reference variable.

The forecasting power was tested with in- and out-of-sample data for the period from I: 1993 to XII: 2001. We find that the suggested model has overcome some major deficiencies of classical leading indicators developed by NBER:

– first, the model was able to correctly forecast all reference points in sample and out-of-sample data;

– second, the model can forecast a future value of reference series. In contrast to the SLOLEI model, the final model makes a forecast of future value of reference series. To produce good forecast results, we estimated the model on data, which were not seasonally adjusted;

– third, the model has fixed forecast horizon. We estimated the model for 9, 12 and 15 months forecasts.

The sensitivity analysis suggests that there are some non-linear relationships between reference variable and selected leading indicators. This explains why we were able to improve the forecast performance of original model. There are, however, numerous structural changes going on in Slovenia and the model should be closely followed and re-estimated as more data become available in order to capture ongoing changes in transition process.

In our opinion the lack of economic structure is the main weakness of neural networks in forecasting applications. Due to the black box nature of neural networks, users of forecasts may feel some uneasiness if they are unable to give proper economic interpretation to the estimated relationships – especially as in our case, where we introduced unmeasurable indicators. On the other hand the economic theory does not always yield a specific functional form that is to be used for empirical verification of the theory. In such cases a neural networks have great advantage over traditional methods. A researcher can start with large network and prune it to the most efficient form.

Large scale econometric testing gives a researcher also an opportunity to discover or to provide additional support to some regularities, which can not be seen when testing only on few time series. There are two finding, which we considered during the testing:

– differenciation of time-series can eliminate the presence of nonstationarity, but also has its drawbacks (see Charemza, Deadman, 1992). It may also eliminate long-term relationships between economic variables;

– Stock and Watson (1998) noted that seasonal adjustment procedure applied to the data may be cleansing them of any underlying non-linearities.

Since these findings can seriously affect the final results, we performed all test on following datasets: original time-series using monthly and annually growth rates; seasonally adjusted time series using monthly and annually growth rates. Multivariate spectral analysis confirmed that differenciation does have an important impact on long-term relationship between variables. Additionally to this finding we also discovered that we could not improve the results, if we used seasonally adjusted time series in the final model. Seasonal adjustment procedure did also show a strong impact on the results of Granger causality test.

Additionally to above mentioned findings, we also came to the same conclusion as Stekler (1991) and Leitch and Tanner (1991). They argued that statistical error measures do not measure the right thing. Therefore they suggest a development of additional economic evaluation criteria, which depends on the application of investigated model. We followed the same principle and constructed evaluation criteria, which should provide additional support when deciding on the best model.

In future work, we hope to refine the best neural network models in this research (by considering additional types of networks, different training methods, etc.) for use as forecasting tools to exploit readily available data in order to gauge future economic activity. Forecast comparisons with other models, such as a vector error-correction models may also be undertaken. As well, future projects may involve the construction of neural net models to forecast other important macroeconomic variables.

References

Bhishop, C. M. (1995), *Neural Networks for Pattern Recognition*. New York: Oxford University Press.

Burns, A. F., Mitchell, W. C. (1946), "Measuring Business Cycles". New York, NBER.

Charemza, W. W., Deadman, D. F. (1992), *New Directions in Econometric Practice: General to Specific Modelling, Cointegration and Vector Autoregression*. Aldeshot: Edward Elgar.

Dias, F. C. (1994), "A Composite Indicator for the Portuguese Economy". Estudios e Documentos de Trabalho, Working Papers No. 18-19.

Fritsche, U., Stephan, S. (2000), "Leading Indicators of German Business Cycles: An Assessment of Properties". DIW, Discussion Paper No. 207.

Gonzalez, S. (2000), *Neural Networks for Macroeconomic Forecasting: A Complementary Approach to Linear Regression Models*. Ottava: Department of Finance.

Granger, C. W. J. (1996), "Can We Improve the Predictive Quality of Economic Forecasts?" *Journal of Applied Econometrics*, 11, pp. 455-473.

Hagan, M. T., Demuth, H. B., Beale, M. H. (1996), *Neural Network Design*. Boston, MA: PWS Publishing.

Hebb, D. O. (1949), *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons.

Hinton, G. E. (1987), "Learning Translation Invariant Recognition in Massively Parallel Networks", in Bakker, J. W., Nijman, A. J., Treleaven, P. C., eds., *Proceedings of PARLE Conference on Parallel Architectures and Languages in Europe*. Berlin: Springer Verlag.

Hymans, S. (1973), "On the Use of Leading Indicators to Predict Cyclical Turning Points". *Brookings Papers on Economic Activity*, 2, pp. 339-384.

Jaditz, T., Riddick, L. A., Sayers, C. L. (1998), "Multivariate Nonlinear Forecasting: Using Financial Information to Forecast the Real Sector." *Macroeconomic Dynamics*, 2, pp. 369-382.
Jagric, T. (2001), "Money, Aggregate Economic Activity and Granger Causality Test." *Our Economy*, (5-6), pp. 67-89.
Jagric, T. (2002), "Measuring Business Cycles." *Eastern European Economics*, 40, pp. 63-87.
Lee, T. H., White, H., Granger, C. W. J. (1993). "Testing for Neglected Nonlinearity in Time Series Models." *Journal of Econometrics*, 56, pp. 269-290.
Leitch, G., Tanner, J. E. (1991), "Economic Forecast Evaluation: Profits versus the Conventional Error Measures." *American Economic Review*, 81, pp. 581-590.
Luenberger, D. G. (1984), *Linear and Nonlinear Programming* (Second ed.). Reading, MA: Addison-Wesley.
Maasoumi, E., Khotanzad, A., Abaye, A. (1994), "Artificial Neural Networks for Some Macroeconomic Series: A First Report." *Econometric Reviews*, 13, pp. 105-122.
McCulloch, W. S., Pitts, W. H. (1943), "A Logical Calculus of the Ideas Imminent in Nervous Activity." *Bulletin of Mathematical Biophysics*, 5, pp. 115-133.
Meese, R. A., Rogoff, A. K. (1983), "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, 13, pp. 3-24.
Mezard, M., Nadal, J. P. (1989), "Learning in Feedforward Layered Networks: The Tiling Algorithm." *Journal of Physics* A, 22, pp. 2191-2203.
Minsky, M., Papert, S. (1969), *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.
Nelson, C., Plosser, C. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics*, 10, pp. 139-162.
Polak, E. (1971), *Computational Methods in Optimization: A Unified Approach*. New York: Academic Press.
Refenes, A. N., Zapranis, A., Francis, G. (1994), "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models." *Neural Networks*, 7(2), pp. 375-388.
Ripley, B. D. (1996), *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.
Rosenblatt, F. (1962), *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, D. C.: Spartan.
Rumelhart, D. E., Hinton, G. E., Williams, R. J. (1986), "Learning Internal Representations by Error Propagation," in Rumelhart, D. E., ed., *Parallel Distributed Processing: Explorations in the Microstructures of Cognition, 1: Foundations*. Cambridge, MA: MIT Press.
Statistics Canada (1999), X11ARIMA version 2000. Ottawa.
Stekler, H. O. (1991), "Macroeconomic Forecast Evaluation Techniques." *International Journal of Forecasting*, 7, pp. 375-384.
Stock, J. H., Watson, M. W. (1989), "New Indexes of Coincident and Leading Economic Indicators," in Blanchard, O., Fischer, S., eds., *NBER Macroeconomics Annual*. Cambridge, MA: MIT Press.
Tiao, G. C., Tsay, R. S. (1994), "Some Advances in Non-linear and Adaptive Modeling in Time-series." *Journal of Forecasting*, 10, pp. 109-131.
Tkacz, G. (2001), "Neural Network Forecasting of Canadian GDP Growth." *International Journal of Forecasting*, 17, pp. 57-69.