Dynamic Correlation Approach to Early Stopping in Artificial Neural Network Training. Macroeconomic Forecasting Example

Krzysztof Michalak
Wroclaw University of Technology
Institute of Applied Informatics
michalak@zacisze.wroc.pl

Rafal Raciborski
Universite Libre de Bruxelles
European Center for Advanced Research
in Economics and Statistics
rracibor@ulb.ac.be

Abstract

Neural networks are widely used in time-series forecasting. One of the issues that arise in neural networks applications is that when a neural network is trained for too long the quality of the predictions tends to deteriorate. To overcome this problem various methods of early stopping are employed. This paper proposes a new approach to early stopping issue in neural network training. In the approach presented the validation series is chosen based on its mean dynamic correlation with forecasted series. The approach is verified by application to macroeconomic data where suitable sets of series are commonly available.

1. Introduction

Neural networks are commonly used in non-linear modeling and forecasting. When macroeconomic data is concerned neural approach is often regarded as an alternative to linear regression models (for example VAR methods) and frequency domain methods [9].

In typical time-series forecasting scenario a time series X_t is given with values $X_1, \ldots, X_{t_0}$ known to the predictor at the current time instant t_0. The aim of the prediction is to estimate future values of the series $X_{t_0+1}, \ldots$. One of the measures used for evaluating accuracy of the predictions is RMSE - a square root of the mean squared error [4] calculated over some interval over which predictions are made. RMSE measure is used in this paper for comparison of various predictors.

When using a multilayer perceptron for prediction usually a number of future values $X_{t_0+1}, \ldots, X_{t_0+l}$ is predicted based on some recent values of the time series $X_{t_0-i+1}, \ldots, X_t$. Obviously, in such case i equals tne number of input neurons and l often called the forecast horizon equals the number of output neurons.

One of the problems that arise when neural networks are used for prediction is overfitting [2]. When a neural network is trained for too large number of iterations the quality of forecasts tends to deteriorate even though the in-sample error is still decreasing. Example of overfitting scenario is presented on Figure 1.

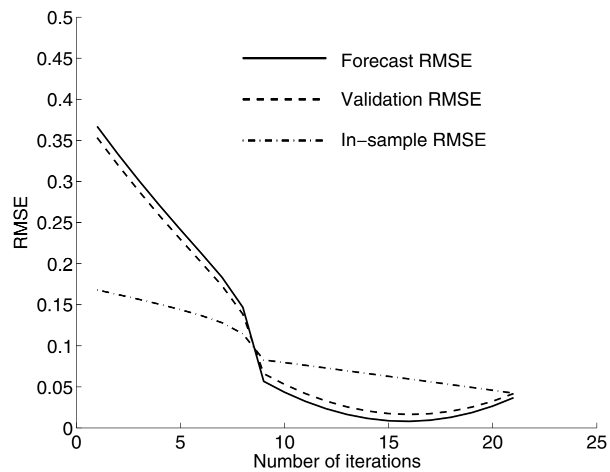

Figure 1. Example of overfitting scenario

In this case network training should be stopped at the iteration 16 where the forecast RMSE reaches the minimum. Clearly, it is important to stop network training when the expected forecast error is lowest not when the in-sample error is minimal. To achieve this various early stopping methods are developed usually based on calculation of the prediction error on some validation data. The most common approach is to divide historical data into the training set $X_1, \ldots, X_{t-v}$ and the validation set $X_{t-v+1}, \ldots, X_t$. The network is trained using only the training set and the training process is stopped at the iteration at which the forecast error in the validation set is minimal.

Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications (ISDA'05)
0-7695-2286-06/05 $20.00 © 2005 IEEE

2. Dynamic Correlation Approach

One of the disadvantages of the early stopping method presented in the previous section is that some, usually most recent, part of data is not used for network training. Another possibility for early stopping detection is to use the whole historical part of the predicted time series for training and perform validation on some other time series. In the case of macroeconomic forecasting, sets consisting of several time series are usually available. These time series represent various macroeconomic indices calculated for a given country [6, 7]. If future values of one of the series are to be predicted other series from the same set could be used for validation.

One of the measures that can be used to quantitatively describe similarity between two time series is the dynamic correlation proposed in [3]. Given any two time series X_t and Y_t one can calculate the dynamic correlation $\rho_{XY}(\lambda)$ of these two series. Note that the dynamic correlation is defined in the frequency domain and is in fact a function of frequency λ. In discrete case if X and Y consist of N+1 samples measured at Δ intervals, the frequency λ falls in range $[-\frac{1}{2}\varphi, \frac{1}{2}\varphi]$ where $\varphi = \frac{1}{\Delta}$ is the sampling frequency.

To obtain easily comparable values a mean dynamic correlation ρ_{XY} over all the frequencies can be calculated

$$\rho_{XY} = \frac{\sum_{k=-\frac{N}{2}}^{\frac{N}{2}} \rho_{XY}(\frac{k}{N}\varphi)}{N+1} \quad . \tag{1}$$

Values of mean dynamic correlation ρ_{XY} fall within the range $[-1, 1]$.

In this paper the mean dynamic correlation ρ_{XY} is used to select a validation series from a given time series set. Suppose we have a set of a number k of time series $X_t^1, \ldots, X_t^k$. If the predictions are made for the time series X_t^1 a validation series can be chosen from $X_t^2, \ldots, X_t^k$. Based on previously performed tests on various macroeconomic time series sets it is proposed that the series X_t^m satisfying:

$$\forall n \in \{1, \ldots, k\} \;\; \rho_{X^1X^m} \leq \rho_{X^1X^n} \tag{2}$$

that is for which the lowest value of the mean dynamic correlation is achieved should be used as a validation set.

3. Experiments

To validate the approach proposed in this paper experiments on industrial production indices were performed. All series were obtained from the EuroStat [10]. From these series first-order 12-month differences were calculated. All series were normalized to [0, 1]. The results were compared to the results obtained by Heravi et al [5] using neural and linear models so data from the same period was used (from 01.1978 to 12.1995 for Germany and from 01.1986 to 12.1995 for the UK). Data for France from the required period was not available so tests for this country were not performed. Data used is summarized in Table 1. In this table codes of validation series chosen based on the lowest mean dynamic correlation criteria are also shown. Each series contained 204 samples in the case of Germany and 108 samples in the case of the UK. Last 24 samples in each series were used to calculate out-of-sample prediction RMSE. The mean dynamic correlation used to choose the validation series was calculated excluding these 24 samples.

Predicted series (X)	NACE code	Validation series (Y)	ρ_{XY}
Germany			
Food products	DA15	DG24	-0.577
Chemicals	DG24	DA15	-0.577
Basic Metals	DJ27	DA15	-0.321
Fabricated Metal	DJ28	DA15	-0.469
Machinery	DK29	DA15	0.078
Electrical Machinery	DL31	DA15	-0.247
Vehicles	DM34	DA15	0.041
Electricity and Gas	E	DA15	0.241
United Kingdom			
Food products	DA15	E	-0.920
Chemicals	DG24	E	-0.669
Basic Metals	DJ27	E	-0.785
Fabricated Metal	DJ28	E	-0.420
Machinery	DK29	E	-0.437
Electrical Machinery	DL31	DA15	-0.316
Vehicles	DM34	DA15	-0.583
Electricity and Gas	E	DA15	-0.920

Table 1. Data summary

Network structure was established as follows. First, the ranges for the possible number of input neurons i and the number of hidden neurons h were assumed so that values commonly found in the literature [1, 5, 9] were covered. Possible range for the number of input neurons was further narrowed down by calculating the correlation dimension D_{corr} of each time series and following the theorem by Takens [8] requiring the number of inputs of the neural network to be no less than $2D_{corr} + 1$. Considering all these criteria ranges of possible numbers of neurons were determined as $i \in [6, 12]$ and $h \in [1, 6]$. Forecast horizons of 1, 3, 6 and 12 months were considered so $l = 1, 3, 6, 12$. To limit the number of possible network structures for each time series and for each forecast horizon l the optimal number of input and hidden neurons was chosen such that the minimal in-sample forecast RMSE was reached anywhere between 10 and 500 training iterations. Numbers of input and hidden neurons established in this process are summarized in Table 2. In each case activation of hidden neurons was $tanh$ and of the output neurons was sigmoid.

Series	$l = 1$	$l = 3$	$l = 6$	$l = 12$
Germany				
DA15	$i = 9$	$i = 7$	$i = 7$	$i = 7$
	$h = 4$	$h = 6$	$h = 6$	$h = 5$
DG24	$i = 9$	$i = 7$	$i = 11$	$i = 7$
	$h = 6$	$h = 6$	$h = 6$	$h = 6$
DJ27	$i = 6$	$i = 7$	$i = 10$	$i = 7$
	$h = 5$	$h = 6$	$h = 6$	$h = 5$
DJ28	$i = 9$	$i = 10$	$i = 6$	$i = 7$
	$h = 5$	$h = 4$	$h = 6$	$h = 6$
DK29	$i = 8$	$i = 8$	$i = 6$	$i = 8$
	$h = 6$	$h = 6$	$h = 6$	$h = 6$
DL31	$i = 8$	$i = 7$	$i = 6$	$i = 7$
	$h = 3$	$h = 6$	$h = 6$	$h = 4$
DM34	$i = 6$	$i = 6$	$i = 9$	$i = 6$
	$h = 6$	$h = 5$	$h = 4$	$h = 4$
E	$i = 11$	$i = 9$	$i = 10$	$i = 8$
	$h = 6$	$h = 3$	$h = 5$	$h = 2$
United Kingdom				
DA15	$i = 9$	$i = 10$	$i = 7$	$i = 6$
	$h = 6$	$h = 6$	$h = 4$	$h = 12$
DG24	$i = 7$	$i = 7$	$i = 6$	$i = 6$
	$h = 6$	$h = 6$	$h = 6$	$h = 5$
DJ27	$i = 9$	$i = 6$	$i = 6$	$i = 6$
	$h = 6$	$h = 3$	$h = 6$	$h = 3$
DJ28	$i = 6$	$i = 6$	$i = 6$	$i = 6$
	$h = 6$	$h = 5$	$h = 4$	$h = 6$
DK29	$i = 6$	$i = 6$	$i = 9$	$i = 6$
	$h = 6$	$h = 5$	$h = 6$	$h = 5$
DL31	$i = 6$	$i = 7$	$i = 11$	$i = 7$
	$h = 5$	$h = 2$	$h = 5$	$h = 2$
DM34	$i = 7$	$i = 8$	$i = 6$	$i = 6$
	$h = 5$	$h = 6$	$h = 6$	$h = 6$
E	$i = 11$	$i = 8$	$i = 9$	$i = 11$
	$h = 6$	$h = 6$	$h = 5$	$h = 5$

Table 2. Optimal network sizes

Using a neural network with the structure described above two methods of early stopping were tested on each series. First, sets of input vectors of length i and corresponding output vectors of length l were constructred from the series using sliding window technique. The last $25 - l$ vector pairs were used for out-of-sample testing thus covering the period of 24 months. From the remaining samples 16 last vector pairs constituted the validation set. The rest of the samples were used for network training.

Training was performed for a fixed number of $N_{iter} = 100, 200$ or 500 iterations. At each training iteration predicted values of the series were recorded. At each iteration prediction RMSE in the validation range and in the validation series (except the last 24 samples) was also recorded. After completing N_{iter} iterations two predictions were selected for which the minimal RMSE on validation range and on validation series were reached. If both early stopping methods indicated the last iteration as the optimal one it was assumed that no overfitting occured within the maximum allowed number of iterations and the test was repeated. To compensate for the inherent fluctuations of prediction error between different network training sessions the same network was used with both methods of early stopping. Therefore the samples from validation set were not used for training even when the validation was performed on the separate series.

Average values of the out-of-sample RMSE obtained for N_{iter} = 100 using both methods in 10 consecutive runs for each series and each forecast horizon are presented in Tables 3 and 4.

The third and the fourth columns contain results for linear and neural models respectively obtained by Heravi et al. [5]. As the results presented by these authors concern non-normalized data, values of the RMSE presented in their paper were normalized by dividing by the amplitude of respective series. Two last columns contain RMSE obtained using early stopping detection on validation range of the predicted series and on the validation series chosen by minimal mean dynamic correlation respectively.

For the time series from Germany only in 4 out of 32 test early stopping using validation range performed better than early stopping using validation series. In all 28 remaining cases the new method outperformed traditional approach in some cases resulting in almost 50% decrease of the prediction RMSE. In every test neural approach using early stopping detection gave better results than linear model and neural approach without early stopping detection

For the United Kingdom the traditional method of early stopping gave better results in 9 cases out of 32. In the remaining 23 cases early stopping using validation series yielded lower RMSE values. In two cases linear models performed better than a neural approach.

In tables 5 and 6 results for $N_{iter} = 200$ are presented. To allow easier comparison columns containing results presented in [5] are included even though they are the same as for $N_{iter} = 100$.

Similarily as for $N_{iter} = 100$ in most cases the early stopping using validation series gave better results than the early stopping using validation range. The number of best results obtained for Germany using validation range is in this case 8 out of 32 tests. In the remaining 24 cases early stopping using validation series gave better results. For the United Kingdom the traditional method of early stopping yielded better results in 11 tests out of 32 compared to

Series	l	Linear [5]	Neural [5]	Validation range	Validation series
DA15	1	0.1437	0.1419	0.0422013	0.0391167
	3	0.1631	0.1777	0.0509800	0.0498109
	6	0.1971	0.2547	0.0634741	0.0652556
	12	0.2007	0.2141	0.0958120	0.0895738
DG24	1	0.2645	0.2603	0.0604186	0.0580548
	3	0.3835	0.3926	0.1078620	0.0934837
	6	0.5157	0.5157	0.1076422	0.1092451
	12	0.6124	0.6207	0.2331240	0.2129880
DJ27	1	0.1682	0.1748	0.0419796	0.0287983
	3	0.1785	0.1599	0.0434392	0.0356586
	6	0.1851	0.2054	0.0567701	0.0512500
	12	0.2298	0.2298	0.0893264	0.0810909
DJ28	1	0.1826	0.1859	0.1296378	0.0695061
	3	0.2926	0.3114	0.1304727	0.1056635
	6	0.3289	0.3718	0.2347620	0.1480580
	12	0.5148	0.5107	0.3080110	0.2374600
DK29	1	0.1597	0.2150	0.0817067	0.0512646
	3	0.1539	0.1588	0.0726779	0.0574139
	6	0.1329	0.1427	0.1191599	0.0595830
	12	0.2422	0.2368	0.1146395	0.0976333
DL31	1	0.1632	0.1588	0.1034522	0.0694671
	3	0.2158	0.2170	0.1526505	0.0879076
	6	0.2739	0.3283	0.1439650	0.1281340
	12	0.2539	0.3283	0.2089660	0.2062990
DM34	1	0.2506	0.3416	0.0676660	0.0614212
	3	0.2498	0.3634	0.0879779	0.0885787
	6	0.2263	0.2305	0.1481297	0.1171772
	12	0.2251	0.2460	0.1750130	0.1376813
E	1	0.3046	0.3588	0.0369963	0.0344279
	3	0.3056	0.4217	0.0645246	0.0577431
	6	0.3404	0.3443	0.1080795	0.1057837
	12	0.3211	0.3269	0.1258380	0.1278410

Table 3. Out-of-sample RMSE for Germany and a maximum of 100 iterations.

21 better results obtained using the new method. Values of RMSE obtained for N_{iter} are in general very similar to those obtained for N_{iter} mostly due to the fact that in most cases the mimumum of forecast RMSE is reached in fewer than 100 iterations.

Due to space limitations only results for $N_{iter} = 100$ and $N_{iter} = 200$ are presented in full detail. Results for all values of N_{iter} are summarized briefly in Table 7.

It is worth noticing that by careful design of the network structure and by using any early stopping method prediction results can be significanlty improved. Furthermore, in most cases early stopping detection using validation series produces lower RMSE values than early stopping based on validation range.

Series	l	Linear [5]	Neural [5]	Validation range	Validation series
DA15	1	0.4018	0.4084	0.1631780	0.1495970
	3	0.4570	0.4702	0.2013820	0.1932300
	6	0.5099	0.4989	0.2739090	0.2588200
	12	0.5188	0.7947	0.4202880	0.3538140
DG24	1	0.1885	0.1855	0.0622468	0.0322384
	3	0.2070	0.2018	0.0729502	0.0666841
	6	0.2029	0.2111	0.1019603	0.1028000
	12	0.2428	0.3033	0.1890800	0.1912350
DJ27	1	0.1854	0.1993	0.0773522	0.0694057
	3	0.1818	0.2572	0.0899338	0.0822206
	6	0.2100	0.2243	0.1393740	0.1183430
	12	0.2767	0.2870	0.2571560	0.2273420
DJ28	1	0.1360	0.3196	0.0704956	0.0570593
	3	0.1552	0.1826	0.0936294	0.0824598
	6	0.1893	0.1999	0.1323640	0.1203170
	12	0.2086	0.2811	0.1591350	0.1651520
DK29	1	0.0949	0.2100	0.0846247	0.0659427
	3	0.1219	0.1202	0.1346760	0.1156870
	6	0.1344	0.1318	0.1415050	0.1422280
	12	0.1000	0.0829	0.1538900	0.1477390
DL31	1	0.1717	0.2904	0.0926523	0.0937625
	3	0.1876	0.2096	0.1815780	0.1841580
	6	0.1786	0.1846	0.2310230	0.2088350
	12	0.1901	0.3199	0.2283870	0.1983140
DM34	1	0.1957	0.1971	0.0553093	0.0528085
	3	0.2424	0.2686	0.0798133	0.0593021
	6	0.3124	0.3190	0.1057203	0.0915303
	12	0.3919	0.4248	0.1791850	0.1614850
E	1	9.1406	9.3750	0.1233134	0.1205223
	3	10.2813	10.2031	0.1033929	0.1449493
	6	11.0000	11.7188	0.0832300	0.1091696
	12	8.7500	9.5469	0.0551458	0.0683686

Table 4. Out-of-sample RMSE for the United Kingdom and a maximum of 100 iterations.

4. Method Comparison

Based on the experimental results it is possible to calculate the p-value of the hypothesis that the new method statistically gives worse results than the traditional one. Let m_r and m_s denote the mean RMSE yielded by early stopping methods based on validation range and validation series respectively and n_r and n_s denote number of best results given by each method. Assume that

$$m_r \leq m_s \tag{3}$$

that is the new method statistically gives worse results than the traditional one.

As the averages of 10 measurements have approximately normal distributions, the probability of getting k best results

Series	l	Linear [5]	Neural [5]	Validation range	Validation series
DA15	1	0.1437	0.1419	0.0432731	0.0377351
	3	0.1631	0.1777	0.0500890	0.0487510
	6	0.1971	0.2547	0.0646844	0.0693567
	12	0.2007	0.2141	0.0966148	0.0874995
DG24	1	0.2645	0.2603	0.0816817	0.0583767
	3	0.3835	0.3926	0.1001246	0.0751738
	6	0.5157	0.5157	0.0975287	0.0995934
	12	0.6124	0.6207	0.1922340	0.1750620
DJ27	1	0.1682	0.1748	0.0359080	0.0302467
	3	0.1785	0.1599	0.0428335	0.0361958
	6	0.1851	0.2054	0.0531818	0.0578614
	12	0.2298	0.2298	0.0823209	0.0689500
DJ28	1	0.1826	0.1859	0.1468700	0.0612068
	3	0.2926	0.3114	0.1694960	0.0928555
	6	0.3289	0.3718	0.2165630	0.1230730
	12	0.5148	0.5107	0.3044000	0.2056850
DK29	1	0.1597	0.2150	0.0816042	0.0503593
	3	0.1539	0.1588	0.0779698	0.0611495
	6	0.1329	0.1427	0.1238267	0.0689476
	12	0.2422	0.2368	0.0956626	0.0911314
DL31	1	0.1632	0.1588	0.1000497	0.0891106
	3	0.2158	0.2170	0.1353818	0.0790428
	6	0.2739	0.3283	0.1790190	0.1181280
	12	0.2539	0.3283	0.2035980	0.2063130
DM34	1	0.2506	0.3416	0.0610506	0.0500045
	3	0.2498	0.3634	0.0908866	0.0899202
	6	0.2263	0.2305	0.1460836	0.1334224
	12	0.2251	0.2460	0.1641210	0.1315357
E	1	0.3046	0.3588	0.0247423	0.0279950
	3	0.3056	0.4217	0.0471875	0.0490819
	6	0.3404	0.3443	0.1009698	0.1009703
	12	0.3211	0.3269	0.1212980	0.1272470

Table 5. Out-of-sample RMSE for Germany and a maximum of 200 iterations.

using the new method $P(k)$ would then have a Bernoulli distribution with single success probability $q < \frac{1}{2}$. Therefore upper bound on p-value of the null hypothesis that the new method statistically gives worse results than the traditional one can be calculated with respect to the results obtained as:

$$p = 1 - P(k < n_s) \tag{4}$$

$$P(k < n_s) = \sum_{i=0}^{n_s-1} \binom{n_s + n_r}{i} q^i (1-q)^i \tag{5}$$

$$P(k < n_s) \geq \frac{1}{2^{n_s+n_r}} \sum_{i=0}^{n_s-1} \binom{n_s + n_r}{i} \tag{6}$$

Series	l	Linear [5]	Neural [5]	Validation range	Validation series
DA15	1	0.4018	0.4084	0.1401940	0.1398060
	3	0.4570	0.4702	0.1965110	0.1909670
	6	0.5099	0.4989	0.2643380	0.2502950
	12	0.5188	0.7947	0.4149760	0.3551340
DG24	1	0.1885	0.1855	0.0455157	0.0233281
	3	0.2070	0.2018	0.0932723	0.0598739
	6	0.2029	0.2111	0.1046151	0.1033852
	12	0.2428	0.3033	0.1869000	0.1898810
DJ27	1	0.1854	0.1993	0.0863172	0.0627446
	3	0.1818	0.2572	0.0846010	0.0813554
	6	0.2100	0.2243	0.1363245	0.1255960
	12	0.2767	0.2870	0.2643580	0.2323000
DJ28	1	0.1360	0.3196	0.0739463	0.0575345
	3	0.1552	0.1826	0.1003860	0.0801035
	6	0.1893	0.1999	0.1357250	0.1200510
	12	0.2086	0.2811	0.1598120	0.1697570
DK29	1	0.0949	0.2100	0.0760170	0.0580950
	3	0.1219	0.1202	0.1251520	0.1130270
	6	0.1344	0.1318	0.1413090	0.1423680
	12	0.1000	0.0829	0.1458230	0.1467880
DL31	1	0.1717	0.2904	0.0778732	0.0787921
	3	0.1876	0.2096	0.1747170	0.1816120
	6	0.1786	0.1846	0.2192230	0.2050940
	12	0.1901	0.3199	0.2350270	0.2029390
DM34	1	0.1957	0.1971	0.0497617	0.0514852
	3	0.2424	0.2686	0.0651507	0.0625240
	6	0.3124	0.3190	0.1039865	0.0909482
	12	0.3919	0.4248	0.1653210	0.1593670
E	1	9.1406	9.3750	0.0912030	0.1111564
	3	10.2813	10.2031	0.0721959	0.0902810
	6	11.0000	11.7188	0.0453832	0.0601453
	12	8.7500	9.5469	0.0654558	0.0821658

Table 6. Out-of-sample RMSE for the United Kingdom and a maximum of 200 iterations.

$$p \leq 1 - \frac{1}{2^{n_s+n_r}} \sum_{i=0}^{n_s-1} \binom{n_s + n_r}{i} \tag{7}$$

Table 8 presents upper p-value bounds for the null hyposesis calculated with respect to the results summarized in Table 7.

The overall upper p-value bound calculated by summing all the results together ($n_r = 52$, $n_s = 140$) is 0.00000000008176. This indicates with very high statistical significance that the new method gives on average better results than the traditional one.

Country	N_{iter}	Validation range	Validation series
Germany	100	4	28
	200	8	24
	500	9	23
United Kingdom	100	9	23
	200	11	21
	500	11	21

Table 7. Number of best results obtained using each method.

Country	N_{iter}	Upper p-value bound
Germany	100	0.00000965059735
	200	0.00350018334575
	500	0.01003080350347
United Kingdom	100	0.01003080350347
	200	0.05509208259173
	500	0.05509208259173

Table 8. Upper p-value bounds obtained in the tests.

5. Conclusion

The early stopping method proposed in this paper produces lower prediction errors than the traditional method of early stopping based on validation range. Results presented in the paper imply that this improvement is of a high statistical significance. The new method can be useful in forecasting macroeconomic time series as in this case data sets consisting of several time series are commonly available.

Some further study is neccessary to provide more theoretical insight into the relation between the dynamic correlation and usefulness of correlated time series for the early stopping detection.

In practical applications of the predictor using the early stopping method based on the dynamic correlation criteria it would be desirable to use full historical sample set for training. This issue was ommited in this paper to allow for easier comparison of the early stopping methods.

References

[1] J. M. Binner, A. Gazely, and T. Elger. Dynamic neural network based inflation forecasts for the uk. *Global Business & Economics Review - Anthology*, pages 546–564, 2004.

[2] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[3] C. Croux, M. Formi, and L. Reichlin. A measure of comovement for economic variables: Theory and empirics. *CEPR Discussion Paper*, (2339), December 1999.

[4] M. M. Gupta. *Static and Dynamic Neural Networks: From Fundamentals to Advanced Theory*. Wiley-IEEE Press, 2003.

[5] S. Heravi, D. R. Osborn, and C. R. Birchenhall. Linear versus neural network forecasts for european industrial production series. *International Journal of Forecasting*, 20:435–446, 2004.

[6] P. D. Munoz. *Methodology of Short-term Business Statistics Interpretation and Guidelines*. Office for Official Publications of the European Communities, 2002.

[7] R. J. Shapiro. *1997 Economic Census. Manufacturing Industry Series*, November 1999.

[8] F. Takens. Distinguishing deterministic and random systems. *Dynamical Systems and Bifurcations*, 1981.

[9] G. Tkacz. Neural network forecasting of canadian gdp growth. *International Journal of Forecasting*, (17):57–69, 2001.

[10] Eurostat web page. *http://europa.eu.int/comm/eurostat*.