Appears in Proceedings of
**PASE '93: Parallel Applications in Statistics and Economics**,
Mirko Novák (ed.), Special Issue of **Neural Network World, 3** (6) pp. 791-794,
VSP International Science Publishers, Zeist, The Netherlands, 1993.

# Predicting the U.S. Index of Industrial Production (Extended Abstract)

John Moody, Uzi Levin, and Steve Rehfuss*

## Introduction

Of great interest to forecasters of the economy is predicting the "business cycle", or the overall level of economic activity. The business cycle affects society as a whole by its fluctuations in economic quantities such as the unemployment rate (the misery index), corporate profits (which affect stock market prices), the demand for manufactured goods and new housing units, bankruptcy rates, investment in research and development, investment in capital equipment, savings rates, and so on. The business cycle also affects important socio-political factors such as the the general mood of the people and the outcomes of elections.

A scientific model of business cycle dynamics is not yet available due to the complexities of the economic system, the impossibility of doing controlled experiments on the economy, and the non-quantifiable factors such as mass psychology and sociology which influence economic activity. Given the absence of reliable or convincing scientific models of the business cycle, economists have resorted to analyzing and forecasting economic activity by using the empirical "black box" techniques of standard linear time series analysis.

We have developed robust predictive models of the business cycle based on neural networks which outperform the standard linear AR models used by most economists.

Economic statistics for the U.S. such as the national income and product accounts and the indices of leading, coincident, and lagging indicators have been collected and computed by the Bureau of Economic Analysis of the Department of Commerce since 1946.

The standard measures of economic activity used by economists to track the business cycle are the Gross Domestic Product (GDP)[1] and the Index of Industrial Production (IP). GDP is a broader measure of economic activity than is IP. However, GDP is computed by the Department of Commerce on only a quarterly basis, while Industrial Production is computed and published monthly.

We have focussed on the Index of Industrial Production rather than GDP for three reasons. First, being published monthly, there is more data available for Industrial Production than for GDP. Second, the IP series is more timely than GDP and is therefore watched more closely by business, financial, and

*The authors are with Nonlinear Prediction Systems, PO Box 681, Portland, OR 97207-0681 and the Department of Computer Science of the Oregon Graduate Institute, PO Box 91000, Portland, OR 97291-1000. Email correspondence may be directed to moody@cse.ogi.edu.

[1] In 1990, GDP replaced Gross National Product (GNP) as a standard measure of domestic economic activity. GNP includes so-called "factor payments" to and "factor income" from foreign sources which are not included in GDP. These factors relate to interest, dividends, and reinvested earnings by foreign subsidiaries of US companies. As such, they are not really part of the domestic economy. GDP also includes the consumption of fixed capital, and important effect which is not captured by GNP.

economic professionals for making business, trading, or policy decisions. Third, the IP series is more interesting and challenging from a time series forecasting standpoint than is GDP.

## Predicting the U.S. Index of Industrial Production

Our experiments on forecasting the monthly U.S. Index of Industrial Production include both regression and classification approaches and a variety of both linear and nonlinear techniques. In the interest of brevity, however, we present only a summary of our regression results for a standard benchmark. The results presented here are preliminary.

Our standard benchmark is to predict the monthly Index of Industrial Production for January 1980 to January 1990 for models trained on January 1950 to December 1979.

Rather than predict the actual index, which is non-stationary, we constructed models to predict the rates of return of the index on six time scales: one month, two months, three, six, nine, and twelve months. The rates of return are close enough to being stationary the enable construction of predictive models. However, we did find conclusive evidence of nonstationarity for the rates of return series.[2] A summary of our preliminary results is shown in Table 1 below.[3]

We computed and compared five predictors: a trivial predictor, a univariate linear AR model, a multivariate linear regression model, and two types of neural net models.

The trivial predictors estimate the $X$-month rate of return for any time during the test period as the average $X$-month rate of return during the training period. If the series were truly stationary, the normalized average squared error for the trivial predictors on the test set would be very close to 1.00. (Note that the normalized prediction errors are computed by dividing the average squared error by the test set variance.) However, we see in Table 1 that their performance is slightly worse (by 4% to 12%). This is due to nonstationarities (shifting means) in the rate of return series.

The univariate AR model uses past values of just the one month rate of return series for computing one step predictions. The order of the AR model, fourteen, was selected via the minimum AIC criterion. Predictions on time horizons greater than one month are obtained by iterating the one step predictions forward. With the exception of the twelve month iterated predictions, the AR(14) model outperformed the trivial predictor.

The remaining three models, linear regression and two flavors or neural nets, were fit to a data set containing 30 input series. The 30 input series consisted of five filtered versions of each of six raw input series. The six raw series were the Index of Leading Indicators, housing starts, the money supply M2, the S&P 500 Index, and two variables related to interest rates.

The linear regression models were fit to make direct predictions on the various time horizons on the basis of a subset of the 30 available input variables. The variable subsets were selected via a stepwise regression algorithm and contained four to six variables. The linear regression models outperform both the AR(14) models and trivial predictors for one, two, and three month predictions, but perform significantly worse for six, nine, and twelve month predictions. The poor performance on the longer prediction horizons could be due to nonlinearities in the problem, poor extrapolation performance, or spurious regressions. The excellent performance of the neural networks suggests that nonlinearity is a likely explanation.

We fit two layer neural network regression models (with a nonlinear internal layer and a linear output layer) using two simulators, `neuz` and `proj`. The `neuz` nets were trained using stochastic gradient descent, early stopping via a validation set, and the new PCA Regularization method (see Levin, Leen,

[2] The issue of nonstationarity will be discussed in greater detail elsewhere.

[3] The table presents normalized prediction errors, which are average squared errors on the test set divided by the test set variance. Note that since these variances are different for each prediction horizon, the normalized errors for the different horizons are not directly comparable.)

and Moody 1993). The `proj` nets were trained using the Levenburg-Marquardt algorithm, and network pruning after training was accomplished via the methods described in Moody and Utans (1991). The internal layer nonlinearities for the `neuz` nets were sigmoidal, while some of the `proj` nets included quadratic nonlinearities as described in Moody and Yarvin (1991). Both simulators utilized robust error measures similar to the Huber measure.

As is apparent from Table 1, both simulators produced networks which significantly outperformed the linear regression, AR(14), and trivial predictors for all prediction horizons. The `proj` nets performed best for the one and two month predictions, while the `neuz` nets performed extremely well for three, six, nine, and twelve month predictions. Relative to the trivial predictors, the best neural nets yielded reductions in forecast errors of from 28 % to 43 %. It should be emphasized that these results are preliminary, that our work is continuing, and that we expect to obtain even better neural net results.

| Prediction Horizon (Months) | Trivial (Average of Training Set) | Univariate AR(14) Model Iterated Pred. | Multivariate Linear Reg. Direct Pred. | `neuz` Nets w/ PCA Pruning Direct Pred. | `proj` Nets w/ BDI Pruning Direct Pred. |
|---|---|---|---|---|---|
| 1 | 1.04 | 0.90 | 0.87 | 0.81 | 0.75 |
| 2 | 1.07 | 0.97 | 0.85 | 0.77 | 0.76 |
| 3 | 1.09 | 1.07 | 0.96 | 0.75 | 0.83 |
| 6 | 1.10 | 1.07 | 1.38 | 0.73 | 0.96 |
| 9 | 1.10 | 0.96 | 1.38 | 0.67 | 0.90 |
| 12 | 1.12 | 1.23 | 1.20 | 0.64 | 0.73 |

Table 1: Comparative summary of normalized prediction errors for rates of return on Industrial Production for five model types for the Jan 1950 - Dec 1979 / Jan 1980 - Jan 1990 benchmark problem. The neural network models computed by our simulators `neuz` and `proj` significantly outperform the linear models. (See text for a description of the normalized prediction errors.)

As is apparent from these results, the U.S. Index of Industrial Production exhibits significant nonlinearity, and by using state of the art neural network learning, regularization, and pruning methods, it is possible to obtain substantial forecast improvements over standard linear techniques.

## Acknowledgements

This work was supported by DARPA Contract DAAH01-92-CR361.

## References

A. U. Levin, T. Leen, and J. Moody. (1993) Fast Pruning Using Principal Components, to appear in **Advances in Neural Information Processing Systems 6**, Cowan, Tesauro, and Alspector, eds, Morgan Kaufmann Publishers, San Mateo, CA.

J. Moody and J. Utans. (1991) Principled Architecture Selection for Neural Networks: Application to Corporate Bond Rating Prediction. in **Advances in Neural Information Processing Systems 4**, Moody, Hanson, and Lippmann, eds, Morgan Kaufmann Publishers, San Mateo, CA.

J. Moody and N. Yarvin. (1991) Networks with learned unit response functions. in **Advances in Neural Information Processing Systems 4**, Moody, Hanson, and Lippmann, eds, Morgan Kaufmann Publishers, San Mateo, CA.