7-11-2008

# The Problem of Neural Networks in Business Forecasting: An Attempt to Reproduce the Hill, Oâ€™Connor and Remus Study

Lin Zhao
*Case Western Reserve University*, lin.zhao@case.edu

Fred Collopy
*Case Western Reserve University*

Miles Kennedy
*Case Western Reserve University*

芽|Sprouts Working Papers on Information Systems ISSN 1535-6078

# The Problem of Neural Networks in Business Forecasting: An Attempt to Reproduce the Hill, Oâ  Connor and Remus Study

Lin Zhao
Case Western Reserve University, USA

Fred Collopy
Case Western Reserve University, USA

Miles Kennedy
Case Western Reserve University, USA

**Abstract**

The results from applying neural networks to business forecasting have been mixed. Among the most encouraging efforts is that of Hill, Oâ  Connor and Remus (1996). In that study, neural networks produced forecasts that were significantly better than those produced by traditional methods for quarterly and monthly series, and no worse for annual series. We have attempted to reproduce that study. The pattern of our results matches that of the original study and supports many of its conclusions, but we were not able to obtain nearly the magnitude of the improvements reported there. We observe that the stopping rules proposed in the original study have exceedingly high variance associated with them, and do not seem to be reasonable stopping rules for such applications. We conclude that the original studyâ  s networks, and in particular its stopping rules, are not described in a way that permits reproducing them. This undermines the value of these systems in practice, since it is difficult both to reproduce them and to identify when and why they are failing to produce results like those obtained in research settings. At the same time, Hill et al.â  s conclusion that neural networks represent a good approach to extrapolating nonlinear and discontinuous series is supported.

**Keywords:** Keywords: Time series, replication, conditional analysis, extrapolation, and M-Competition

**Permanent URL:** http://sprouts.aisnet.org/3-18



**Reference:** Zhao, L., Collopy, F., Kennedy, M. (2003). "The Problem of Neural Networks in Business Forecasting: An Attempt to Reproduce the Hill, Oâ  Connor and Remus Study," Case Western Reserve University, USA . *Sprouts: Working Papers on Information Systems*, 3(18). http://sprouts.aisnet.org/3-18

# The Problem of Neural Networks in Business Forecasting: An Attempt to Reproduce the Hill, O'Connor and Remus Study

## Introduction

One of the ways in which science progresses most rapidly is through the replication and extension of important studies. Yet replications in the management sciences are rare (Berthon, Pitt, Ewing, & Carr, 2002; Hubbard & Vetter, 1996). There has been a considerable amount of interest in the application of artificial neural networks to management problems. In the area of business forecasting there have been over 40 studies in which neural networks have been applied and compared with other methods over the period from 1990 through 2003. At each of the past 13 meetings of the International Symposium on Forecasting there has been a track dedicated specifically to developments in neural networks for forecasting. Yet, each of the neural networks reported on in these articles and at these symposia is unique. We have been unable to find a single instance in which a neural network from one study was applied in another. (Michael Nelson joined Hill, Remus, and O'Connor on a study that examined the role of seasonality in neural network forecasts but this was more of an extension than either a new application or replication.) We attempted to remedy this omission by reproducing an important result in neural network research. A reproduction study is a replication study in which the original data are used. In this study, we used the same series and attempted to create networks that were as similar to those in the original study as we could.

## Hill, O'Connor, and Remus's study

Of all of the attempts to apply neural networks to the problem of extrapolating time series, the most successful is probably that of Tim Hill, Marcus O'Connor and William Remus, reported in Management Science in 1996. In that study, the authors used simple neural networks to forecast multiple horizons for a subset of the time series from the M-Competition (Makridakis et al., 1982). They compared forecasts from their network to those from well-established alternative methods. They used absolute percentage errors (APEs) to make the comparisons. And, they produced 1451 forecasts. They complied, in other words, with all of the procedures for making forecasting comparisons that Collopy, Adya, and Armstrong (1994) argued are essential for drawing meaningful conclusions. In addition, they did analysis of the conditions under which the neural network performed well. In short, their research design was exemplary.

The analysis of conditions is one of the most important contributions of Hill et al. Such condition analysis has been recommended in the literature generally (Chamberlin, 1897; Greenwald, Leippe, Pratkanis, and Baumgardner, 1986) and in the forecasting literature specifically (Armstrong, 1988). Few studies have actually done it though. In their analysis, Hill et al. determined that their neural network performed especially well 1) on discontinuous series, 2) on longer horizon forecasts, and 3) on nonlinear series where there were fewer than 100 observations in the historical data.

The neural network used by Hill et al. was a modification of a general network described by Rumelhart and McClelland (1986). It used a simple architecture with one hidden layer. The number of nodes in the input layer was determined experimentally for each of the three intervals (yearly, quarterly and monthly) of data. There were three input nodes for annual data, four for quarterly data, and nine for monthly data. All three models had one output node. All three models started with two hidden nodes. This number was doubled until there was no significant increase in forecasting power. This resulted in architectures that included two hidden nodes in the annual and quarterly models and four in the monthly model. The networks parameters were estimated using back-propagation and a sigmoid transfer function.

Hill et al.'s training procedure consisted of eight learning phases. Each of the phases ended after 200 epochs or in response to one of three other stopping rules. When one of those conditions was met, the learning rate was reduced by half and another phase begun. An interesting aspect of their procedure is the decision to begin training with a learning rate of 10 (most texts recommend initial learning rates no greater than 1). The effect of a learning rate of 10 is that an adjustment of ten times the error is made in the direction opposite the error. This seems an extraordinary thing to do, and indeed, as illustrated in Figure 1, it does produce erratic movements during early training. We don't see that it will ultimately prove too damaging, however, as by the fourth phase, the learning rate will be a reasonable 1.25, with five phases remaining.

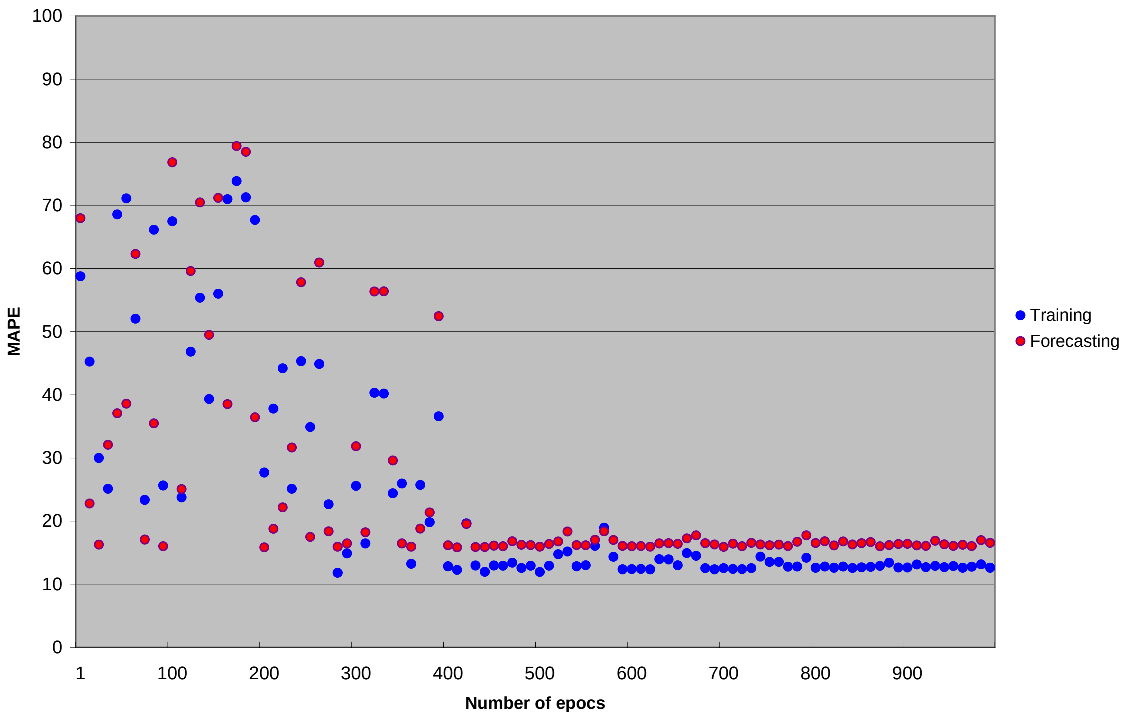


**Figure 1: Fits and forecasts are unrelated for many epochs**

In addition to stopping each phase when it reached 200 epochs, Hill et al. had three other stopping rules. We will describe our understanding of these rules in the section where we discuss our attempt to reproduce their study. We think that understanding them is at the core of producing a successful replication, and as you will see, they are difficult to interpret.

## The Performance of Neural Networks Since Hill, O'Connor, and Remus

The most important comparative study of forecasting methods in recent years was the M3-Competition, undertaken by Spyros Makridakis and Michele Hibon (2000). In that study 22 extrapolative forecasting methods, both commercial and academic, were compared on the basis of their accuracy in forecasting 3003 time series from a wide variety of settings. The methods included an automatic neural network, implemented by Balkin and Ord (2000). Balkin's PhD thesis (2000) makes reference to Hill et al., so we presume that Balkin and Ord used lessons learned from Hill et al. and others to implement the best possible general extrapolation network. That neural network's performance was generally significantly worse than that of the other methods, regardless of forecast interval (yearly, quarterly, or monthly) or horizon (Hibon and Steckler, 2003).

Neural networks have been compared with other methods in smaller studies as well. In a comparison of neural networks with Box-Jenkins and Holt-Winters exponential smoothing, Faraway and Chatfield (1998) found that neural networks often gave poorer out-of-sample forecasts. They also concluded that they were difficult to interpret. They used the well-know airline data used by Box et al. (1994) and Brown (1963) to fit several NNs. Although the fit measures (the Akaike information criterion and the Bayesian information criterion) were comparable for various NN structures, *ex ante* forecasting performance varied greatly. They concluded therefore that NN models were hard to interpret and that little guidance could be given on how many parameters to use.

Most recent studies of neural networks have identified particular conditions under which they do well. Zhang (2001), using both 240 simulated linear series and three actual time series, concluded that NNs were able to outperform Box-Jenkins' ARMA (p, q) in all but one of their cases. They found that simple NNs were often adequate in forecasting linear times series. Hu et al. (1999) investigated the out-of-sample performance of neural networks in predicting the weekly British pound/US dollar exchange rate. Using monthly data from 1976 through 1993, they concluded that neural networks were superior to random walk models when the forecast horizon is short. Using relative errors, Goh (1998) found that neural networks outperformed the univariate Box-Jenkins approach and the multiple loglinear regression on quarterly data. In a study by Leung et al. (2000), neural networks were significantly more accurate than multivariate transfer functions and random walk models for currency exchange rate forecasting of monthly data. Hwarng (2001) compared NNs with ARMA (p,q) structures on 320 generated time series using RMSE and MAPEs. He concluded that back-prop neural networks trained with a normal level of noise tend to perform better than ARMA (p, q) structures.

## What We Did

This project started with us developing a flexible implementation of a neural network for forecasting. The program, written in Visual Basic 6.0, provides dozens of switches which allowed us to change such things as the number of iterations before the net quits searching, the rate at which it learns, and the output it produces for diagnostic purposes.

After using our network on some data like that in the M-Competition, it became obvious that our network's performance was not nearly as good as that reported in Hill et al. On a close examination of their paper, we observed several potentially important differences between what we were doing and what they had done. Our implementation kept track of the neural net's parameters whenever it performed at its best on out-of-sample forecasts. We then used those parameters to produce the forecasts. Theirs, on the other hand, had a fairly sophisticated set of learning and stopping rules to determine which network parameters to use.

Once we decided to reproduce the study by Hill et al. we contacted the authors in hopes of obtaining their program, the data they used, the individual forecasts they produced, and their coding of the conditions (which had been done judgmentally by an expert). We have been unable to obtain a copy of their source code, so have had to do our best to reproduce it from the information contained in the *Management Science* article. We were also unable to obtain the forecasts they produced, which has limited our ability to make statistical comparisons between their results and ours.

The data are available on the internet and from the authors of the original M-Competition, so we obtained it readily. One of the authors, Marcus O'Connor was able to provide us with information about how they classified each of the series. And Michael Nelson, a graduate student of his who confirmed the basic results at the time they were working on a revision of the original paper, provided us with the executable C code used in their study and several excel spreadsheets used for data analysis in his research (Nelson, Hill, Remus, & O'Connor, 1999). We were not able to execute their C code however because the parameter files that were required to initialize the code were not available.

There is a substantial literature, much of it summarized in Pant and Starbuck (1990) that argues against using fit as a guide for forecast accuracy. For this reason, we were particularly intrigued that Hill, et al. had taken some pains to design and implement stopping rules which seemed to be aimed at avoiding the well-known problem of over-fitting. Over-fitting occurs when the degrees of freedom in a model are so vast that the method can very accurately model the historical data, including the noise in the series. Such fits tend to do a terrible job of producing out-of-sample forecasts, because in addition to extrapolating the effects of the level and the trend, they extrapolate the noise. Neural networks, with their non-linear characteristic and their essentially unlimited degrees of freedom would seem to be particularly prone to this problem. Hill et al. used only one hidden layer and a single output thereby producing networks with reasonable degrees of freedom.

Like Hill et al. we removed seasonality and scaled each series. Hill et al. trained their networks in eight phases. Though each phase had a maximum of 200 training passes, they also developed three stopping rules that could terminate a phase earlier than this. In each of the eight phases, the learning rate was reduced by half. So in the first phase the learning rate was10, and in the last it was approximately 0.08.

The first stopping rule, based on the current total error, summed the squares of the error terms over all output nodes over all training periods. (They use the phrase "all output nodes" though all of their networks have a single output node.) When this value fell below a minimum, the phase was terminated. This minimum was unique for each series, but the authors report that on average it was 10% for each output. The formula given in a footnote is confusing. (The criterion was (0.005) 2 times the number of output units times the number of patterns presented.) The decimal and the 2 are not separated by an operator, so we tried several, checking each to see if it provided an average close to the approximately 10% that the authors said should result.

Finally, we have interpreted this footnote as (0.005) / 2 * output units * patterns. This formula produces minima with an average of about 12%, not far from the average they mention. We also tried using the 10% figure itself. There was little difference in the results.

The second stopping rule, based on progress in reducing the total error, was estimated as the exponentially smoothed average of the differences in the total error between successive training passes. The smoothing coefficient was 0.9 and a stage was terminated when this value fell below a minimum. The minimum started out at 0.005 and was reduced by 25% through each phase.

The third stopping rule, based on progress made in the error term in the last training period, was estimated as the exponentially smoothed average of the differences in the error terms between successive training passes. Again, the smoothing coefficient was 0.9. A phase was terminated if the average difference exceeded a maximum value, intended to indicate the beginning of an upward trend in the error term. This maximum value was held constant at 0.00001 for all phases.

To make our results comparable to Hill et al.'s we removed the first two series each from the annual, quarterly and monthly data. (They had used these series to interactively develop their learning and stopping heuristics.) We also removed the monthly series that they said they eliminated because it had extremely large errors with all models. In short, we used the same 104 series to test our networks as they used to test theirs.

## Results

Despite months of effort and numerous attempts, we were unable to produce neural networks that were as accurate overall as those reported by Hill et al. as can be seen in Table 1. The three entries labeled NN give respectively—their results, the results of a neural network that shares their architecture, but runs through 200 epochs in each of eight phases without any other stopping rules, and a neural network that represents our best attempt to replicate their stopping rules.

### Data Period Used

Despite the fact that our results do not conform to those reported in Hill et al., in important respects they do match the pattern of their results. Like Hill et al.'s, our networks did better than most of the reference methods on quarterly and monthly data, even though the magnitude of the improvements was nowhere near as great for ours as for theirs. Unlike theirs, our networks did not do as well as the other methods on annual data.

### Functional Form of the Series

Hill et al. argued from prior literature that neural networks are particularly well suited to modeling nonlinear patterns and time series containing discontinuities. To examine this they conducted further conditional analysis on the monthly time series. They compared the performance of the neural network forecasts to those from the other methods. For these comparisons they used the average of the APEs from the six traditional methods they examined. Since that included a graphically-based judgmental method for which we do not have the forecasts, we could not construct their reference average series by series. So we made

**Our neural networks were less accurate than those of Hill, O'Connor and Remus using the MAPE criterion**

| | Annual | Quarterly | Monthly |
|---|---|---|---|
| D. Exponential Smoothing | 15.9 | 18.7 | 15.2 |
| Box-Jenkins | 15.7 | 20.6 | 16.4 |
| D. Holt's | 12.1 | 26.9 | 19.2 |
| Graphical Judgment | 12.5 | 20.5 | 16.3 |
| Six Methods (CombiningA) | 12.6 | 21.2 | 16.7 |
| Naïve | 16.4 | 20.0 | 27.0 |
| Reference Average | 15.0 | 22.6 | 17.1 |
| NN Hill, O'Connor & Remus | 14.2 | 15.3 | 13.6 |
| NN (1600 epochs) | 19.0 | 17.4 | 15.7 |
| NN (with Stopping Rules) | 24.0 | 21.0 | 16.6 |

**Table 1. Our neural networks were less accurate than those of Hill, O'Connor and Remususing the MAPE criterion**

comparisons between our 1600 epoch neural networks and Combining A, from the original M-Competition. It has performed well and been used as a benchmark in many studies.

Hill et al. separated the monthly series into four groups—linear without discontinuities (n=37), nonlinear without discontinuities (n=14), linear with discontinuities (n=11), and nonlinear with discontinuities (n=3). Because of the small number of series in the last group it was not analyzed. They refer to the three groups that they did analyze as "linear", "nonlinear", and "discontinuous" respectively. Like Hill et al. we compared the differences in the performance for each of the three functional forms.

For the linear group the APE difference was –2.08; overall the neural network was significantly better than Combining A ($t=-3.23$, $df=665$, $p<0.001$). This is opposite the result in Hill et al. but consistent with their thesis that neural networks would perform well for nonlinear series, even if not for linear series. For the nonlinear group the APE difference score was 1.04. For this group there was no significant difference in this group between the models ($t=1.08$, $df=251$, $p<0.29$). Hill et al. also found no difference for this group. For the discontinuous group the APE difference score was 11.63; neural networks were significantly more accurate than Combining A ($t= 2.13$, $df=197$, $p<0.04$). This replicates what Hill et al. found.

An ANOVA comparing between group differences revealed a significant difference due to the functional form of the series ($F=11.40$, $df=2$, $p<0.001$). As in Hill et al., post-hoc Scheffe tests confirmed at the 0.05 level that the major difference resulted from the superior performance of the neural network model in discontinuous series. Overall then, the pattern of results is remarkably like that hypothesized and found in Hill et al.

**Accuracy at Each Horizon**

Again arguing from prior research, Hill et al. hypothesized that as the forecasting horizon got longer, the relative advantage of neural networks would improve. They divided the results into three equal horizon blocks (1-6, 7-12, and 13-18 months ahead). Using the same categories as before they compared the relative performance over these three blocks.

For discontinuous series, the neural network model was significantly more accurate than Combining A during the 13-18 period (t=2.23, df=65, p<0.03), but less accurate in the 1-6 period (t=-1.69, df=65, p<0.10) and not very different in the 7-12 period (t=0.87, df=65, p<0.40). The difference was significant across the three horizons blocks (F=4.54, df=2, p<0.02). This pattern of results is consistent with their hypothesis.

For nonlinear series we found no significant difference across the forecast horizons (F=2.84, df=2, p<0.07) as did Hill et al. We also found none at any of the three horizon blocks. They found a significant difference at the 7-12 horizon block, but that did not really support any of their hypotheses.

For linear series, we found significant differences at the 7-12 horizon block. They found significant differences in the 1-6 and the 12-18 horizon blocks. Neither of us found any difference in the comparative advantage among the three forecast horizons (F=2.47, df=2, p<0.09).

Overall then, we found much the same thing that Hill et al. did. For the series on which the differences are most substantial, those with discontinuities, there was a significant improvement in the comparative advantage of neural networks for longer horizon forecasts.

**Number of Observations Required**

Hill et al. speculated that the number of observations available to calibrate the model might affect the relative accuracy of neural networks. For linear series, their ANOVA revealed no effect on the difference scores due to the amount of historical data. For the nonlinear series there was a significant difference. We found significant differences both for the linear (F=5.44, df=2, p<0.01) and for the nonlinear (F=10.56, df=2, p<0.001). Clearly the differences on the nonlinear series were more significant than on the linear series. So in that way the pattern of our results matches that of Hill et al.

## Differences Between the Two Studies

Our results were remarkably like those reported by Hill et al. with the disturbing exception that our overall accuracy was nowhere near as good at their's. This is of course extremely frustrating, given the efforts we have made to replicate and understand their results.

There are several possible explanations for our failure to achieve the results reported by Hill, et al. One is that we have done something wrong in our attempts to reproduce their results. The second is that they did something wrong, either in implementing their networks, or in reporting on the output from them. A final possibility is that they were luckier than us. Neural networks make use of randomly assigned initial weights, so it is theoretically possible that on a number of series their runs benefited from "better" initial choices than ours made. This is an unlikely explanation, so in the remainder of this section we concentrate on the other two.

The most likely way in which we may have failed is in our interpretation of their learning and stopping heuristics. As we have noted, we did find the material in their Appendix 2 difficult to understand, and even found some of it counter-intuitive. In particular, starting with 10 as the learning rate seems problematic. And many training phases proceeded through relatively few epochs because of the first stopping rule.

Another possibility is that Hill et al. made errors. It is not uncommon in efforts like this to produce thousands of runs, each with attendant documentation and output. This is, indeed, one of the reasons for replications such as the present one. We can only be certain that the results work has been adequately described, when it is possible for others to reproduce the results based solely on the final document. In this case, we were unable to do that.

## Implications and Recommendations

The most important implication of our research is that it is considerably harder to obtain substantial improvements in extrapolative forecasting with neural networks than might be assumed reading the earlier studies. This results in part from the “black box” character of neural networks. Since we cannot understand precisely which of the features in Hill et al.’s networks account for their superior performance it is difficult to know where to put one’s energy when re-implementing them in other environments.

Given the pattern of results obtained in the two studies, though, it does appear that neural networks are an appropriate extrapolation technique for nonlinear and discontinuous series. This is important, because those have been difficult to forecast successfully with other methods. For series that can be identified as nonlinear and discontinuous, it may be possible to gain significant improvements over other methods even by using very simple neural network architectures and stopping rules.

We also believe our research has implications for replication research. First, it is clear that important studies such as Hill et al. must be replicated. If a number of researchers all made efforts to reproduce and understand these results, we would come eventually to appreciate which factors were most important in attaining them. Replications should be done soon after the initial studies, so that the needed materials can be obtained and authors consulted on details of the work. We suggest that all of the relevant materials should be maintained on web sites so that it is available to researchers who wish to reproduce the result.

As for practice-oriented recommendations, we can only recommend that those building forecasting systems based around neural networks proceed cautiously. Clearly, over-fitting is a hazard to be avoided. Though the heuristics provided in this study represent one approach to this problem, it is not clear that they obtain the result desired. Other approaches should be explored, and until some robust ones establish themselves in the literature, practitioners should remain somewhat skeptical about extrapolative forecasts produced by neural networks.

## References

Armstrong, J. S., “Research Needs in Forecasting,” *International J. Forecasting*, 4, 3 (1988), 449-465.

Balkin, S. D. and J. K. Ord, “Automatic neural network modeling for univariate time series,” *International J. Forecasting*, 16, 4 (2000), 509-516.

Balkin, S. D., *A Statistical Implementation of Feedforward Neural Networks for Univariate Series Forecasting*, Ph.D. Dissertation, The Pennsylvania State University, State College, PA, 2000.

Berthon, P., L. Pitt, M. Ewing, and C. L. Carr, "Potential research space in MIS: A framework for envisioning and evaluating research replication, extension, and generation," *Information Systems Research*, 13, 4 (2002), 416-427.

Box, G. E. P., G. M. Jenkins and G. C. Reinsel, *Time Series Analysis 3rd Edition*, Prentice Hall, Englewood Cliffs, NJ, 1994.

Brown, R. G., *Smoothing, Forecasting and Predication of Discrete Time Series*, Prentice Hall, Englewood Cliffs, NJ, 1963.

Chamberlin, T. C., "The method of multiple working hypotheses," *J. Geology*, 5 (1897), 837-848.

Collopy, F., M. Adya, and J. S. Armstrong, "Principles for Examining Predictive Validity: The Case of Information-Systems Spending Forecasts," *Information Systems Research*, 5, 2 (1994), 170-179.

Faraway, J., and C. Chatfield, "Time series forecasting with neural networks: A comparative study using the airline data," *J. the Royal Statistical Society Series C-Applied Statistics*, 47 (1998), 231-250.

Goh, B.-H., "Forecasting residential construction demand in Singapore: a comparative study of the accuracy of time series, regression and artificial neural network techniques," *Engineering Construction & Architectural Management*, 5, 3 (1998), 261-275.

Greenwald, A. G., M. R. Leippe, A. R. Pratkanis, and M. H. Baumgardner, "Under What Conditions Does Theory Obstruct Research Progress," *Psychological Review*, 93, 2 (1986), 216-229.

Hibon M. and H. Stekler, "The M-3 Competition: Statistical Tests of the Results," Insead Working Paper, 2003/33/TM.

Hill, Tim, Marcus O'Connor and William Remus, "Neural Network Models for Times Series Forecasting," *Management Science*, 42 (1996), 1082-1092.

Hu, M. Y., G. Q. Zhang, C. Z. Jiang, & B. E. Patuwo, "A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting," *Decision Sciences*, 30, 1 (1999), 197-216.

Hubbard, R. and D. E. Vetter, "An empirical comparison of published replication research in accounting, economics, finance, management, and marketing," *J. Business Research*, 35, 2 (1996), 153-164.

Hwarng, H. B., "Insights into neural-network forecasting of time series corresponding to ARMA(p, q) structures," *Omega-International J. Management Science*, 29, 3 (2001), 273-289.

Leung, M. T., A.-S. Chen, and H. Daouk, "Forecasting exchange rates using general regression neural networks," *Computers & Operations Research*, 27, 11 (2000), 1093-1106.

Makridakis, Spyros, M. Hibon, "The M3-*Competition*: results, conclusions and implications," *International J. Forecasting* 19, 1 (2000), 5-22.

Makridakis, S., A. Anderson, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler, "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," *J. Forecasting*, 1 (1982), 111-153.

Nelson, M., T. Hill, W. Remus, and M. O'Connor, “Time series forecasting using neural networks: Should the data be deseasonalized first?” *J. Forecasting, 18, 5* (1999), 359-367.

Pant, P. N. and W. H. Starbuck, “Innocents in the Forest: Forecasting and Research Methods,” *J. Management,* 16 (1990), 433-460.

Rumelhart, D. and J. McClelland, *Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1986.

Zhang, G. P., “An investigation of neural networks for linear time-series forecasting,” *Computers & Operations Research*, 28, 12 (2001), 1183-1202.

芽|Sprouts

Working Papers on Information Systems | ISSN 1535-6078



芽|Sprouts