A Computational Methodology for Modelling the Dynamics of Statistical Arbitrage

Andrew Neil Burgess

Decision Technology Centre

Department of Decision Sciences

A thesis submitted to the University of London for

the degree of Doctor of Philosophy

UNIVERSITY OF LONDON

LONDON BUSINESS SCHOOL

To my parents, Arnold and Carol.

Acknowledgements

Thanks to my supervisor, Paul Refenes, for bringing me to LBS, keeping me in bread and water, helping me with ideas and even reading the bloody thing!

Thanks to past and present colleagues at the Neuroforecasting Unit/Decision Technology Centre. Especially Yves Bentz, Peter Bolland, Jerry Connor, Stefania Pandelidaki, Neville Towers and Peter Schreiner; for discussions, hard work (not all the time), company and six good years. Also all the ex-CRLers who showed that even a real job can still be fun.

Thanks to the visitors to the LBS group, Fernando and Paul, for good times and hard work in London, Helsinki and Melbourne.

Thanks for the people who helped keep it real, especially to Pratap Sondhi for ideas and support when he was at Citibank at the beginning; the other sponsors of the Neuroforecasting Club and then the Decision Technology Centre; Botha and Derick down in S.A. for trusting me to build a trading system; and David and Andrew for all their efforts on behalf of New Sciences.

Also to whoever first decided to hold NIPS workshops at ski resorts; and to the regulars at NNCM/CF conferences: John Moody, Andreas Weigend, Hal White, Yaser Abu-Mostafa, Andrew Lo and Blake LeBaron in particular for their enthusiasm and inspiration.

Finally, my love and thanks to Deborah, who had to put up with me whilst I was writing up - and whose photographs of chromosomes had to compete for computer time with my equity curves.

Abstract

Recent years have seen the emergence of a multi-disciplinary research area known as "Computational Finance". In many cases the data generating processes of financial and other economic time-series are at best imperfectly understood. By allowing restrictive assumptions about price dynamics to be relaxed, recent advances in computational modelling techniques offer the possibility to discover new "patterns" in market activity.

This thesis describes an integrated "statistical arbitrage" framework for identifying, modelling and exploiting small but consistent regularities in asset price dynamics. The methodology developed in the thesis combines the flexibility of emerging techniques such as neural networks and genetic algorithms with the rigour and diagnostic techniques which are provided by established modelling tools from the fields of statistics, econometrics and time-series forecasting.

The modelling methodology which is described in the thesis consists of three main parts. The first part is concerned with constructing combinations of time-series which contain a significant predictable component, and is a generalisation of the econometric concept of cointegration. The second part of the methodology is concerned with building predictive models of the mispricing dynamics and consists of low-bias estimation procedures which combine elements of neural and statistical modelling. The third part of the methodology controls the risks posed by model selection and performance instability through actively encouraging diversification across a "portfolio of models". A novel population-based algorithm for joint optimisation of a set of trading strategies is presented, which is inspired both by genetic and evolutionary algorithms and by modern portfolio theory.

Throughout the thesis the performance and properties of the algorithms are validated by means of experimental evaluation on synthetic data sets with known characteristics. The effectiveness of the methodology is demonstrated by extensive empirical analysis of real data sets, in particular daily closing prices of FTSE 100 stocks and international equity indices.

Table of Contents

Introduction: A Methodology for Statistical Arbitrage

This part of the thesis introduces the concept of statistical arbitrage and examines the issues which will be developed in the rest of the thesis. Chapter 1 consists of a brief introduction which outlines the scope, motivation and organisation of the thesis as well as summarising the major contributions which it makes to the current state of the art. Chapter 2 describes the recent advances in computational modelling and computational finance which provide the background to the thesis. Chapter 3 describes the opportunities for statistical arbitrage which are presented by the advances in modelling methodology, assessing the strengths and weaknesses of existing modelling methods and highlighting the outstanding issues which the methodology presented in the thesis is designed to address. Finally Chapter 4 presents an overview of the methodology and a "route map" to the rest of the thesis.

Part I: A Cointegration Framework for Statistical Arbitrage

In this part of the thesis we describe the first of the three parts of our methodology for statistical arbitrage modelling. This consists of a framework for constructing combinations of time-series which contain a potentially predictable component. An overview of the methodology described in this part of the thesis is presented in Section 4.2.

Chapter 5 describes a framework for modelling combined asset prices, which is inspired by the econometric concept of cointegration. The cointegration framework is used to generate potential statistical mispricings, by performing stochastic detrending of asset prices with respect to other, related, asset prices. The objective of this pre-processing is to generate combinations of time-series which are largely immunised against market-wide uncertainties and which enhance the potentially predictable components of asset price dynamics. The basic methodology is extended to time-varying relationships and also to high-dimensional problems where the number of assets is numbered in the tens or hundreds.

Chapter 6 describes a range of tests which are designed to identify potential predictability in the mispricing time-series. The tests include standard autocorrelation tests, cointegration tests for stationarity, and variance ratio tests for deviations from random walk behaviour as well as novel tests based upon the shape of the variance ratio profile as a whole. The strengths and weaknesses of the various tests are examined under controlled circumstances using Monte-Carlo simulations. Modifications of the predictability tests are presented which correct for the bias induced by the construction procedure. The final section of the chapter contains the results of applying the tests to combinations of FTSE 100 equity prices.

Chapter 7 describes a set of "implicit statistical arbitrage" (ISA) trading strategies which are designed to directly exploit any mean-reverting component in the mispricing dynamics, bypassing the intermediate stage of constructing an explicit forecasting model. The underlying assumption of the ISA strategies is that future price changes will be such as will tend (on average) to reduce the mispricing between a given "target" asset and the associated "synthetic asset". The ISA rules are used to perform an empirical evaluation of the statistical mispricing methodology. The high-dimensional version of the mispricing construction methodology is evaluated with respect to models of the daily closing prices of FTSE 100 stocks; the adaptive version of the methodology is evaluated on the French CAC 40 and German DAX 30 stock market indices.

Part II: Forecasting the Mispricing Dynamics using Neural Networks

In this part of the thesis we describe the second of the three parts of our methodology for statistical arbitrage modelling. This consists of algorithms, tools and procedures for supporting the construction of predictive models of the dynamics of statistical mispricing time-series. The methodology is designed to address the particularly hard problems that arise in the context of building predictive models in investment finance. An overview of the methodology described in this part of the thesis is contained in Section 4.3.

Chapter 8 provides the motivation and general framework for our predictive modelling methodology. It firstly presents a general formulation of the model estimation process before moving on to discuss the particularly hard nature of the problems which arise in the case of building predictive models of asset price dynamics. These problems include high noise, low degree of prior knowledge, small sample sizes and potential time-variation (nonstationarity) in the underlying data-generating processes. An "equivalent kernels" perspective is used to highlight the similarities between neural modelling and recent developments in non-parametric statistics. This in turn motivates the use of neural estimation methods in conjunction with statistical testing procedures as a means of achieving both flexibility and parsimony in an attempt to overcome the "bias-variance dilemma".

Chapter 9 describes our methodology for model-free variable selection, which is based upon methods from non-parametric statistics and is intended to distinguish which variables from the information set should be included in the modelling procedure proper. The purpose of this "pre-selection" stage is to reduce the complexity, and hence **variance**, of the modelling process as a whole. The flexibility of the tests provides an important role in retaining the largest possible amount of *relevant* information upon which to condition the forecasting models. In particular, the tests are capable of identifying both nonlinear dependencies and interaction effects and thus avoid discarding variables which would be passed over by standard linear methods.

Chapter 10 describes our methodology for the actual estimation of low-bias forecasting models. This task is performed through novel algorithms which balance the flexibility of *neural networks* with the noise-tolerance and diagnostic procedures of *statistical regression*. Statistical testing and selection procedures are employed within a rigorous modelling framework which automatically optimises the specification of the neural network. Our integrated approach to the model estimation problem combines the two aspects of variable selection and architecture selection, within a common framework based upon the statistical significance tests which are developed in Chapter 9. Within this common framework we describe three alternative algorithms which aim to optimise both variable selection and model complexity using the constructive, deconstructive and regularisation-based approaches to model building.

Chapter 11 describes an empirical evaluation of the methodology from Part I and Part II of the thesis used in combination. The model-free variable selection procedures and neural model estimation algorithms are applied to the problem of forecasting the dynamics of the statistical mispricings generated by the first part of our methodology. The objective of this exercise is to generate *conditional statistical arbitrage* (CSA) models in which nonlinearities, interaction effects and time-series effects can all be captured and exploited without being explicitly prespecified by the modeller. An empirical evaluation is presented of a set of CSA models which are based upon statistical mispricings between the constituent stocks of the FTSE 100 index.

Part III: Diversifying Risk by Combining a Portfolio of Statistical Arbitrage Models

In this part of the thesis we describe the third part of our statistical arbitrage methodology. This addresses the implementation issues which arise in the context of applying predictive models to risk-averse decision-making in general and statistical arbitrage in particular. The methodology aims to reduce the risks which are inherent in the modelling process itself, thus increasing the extent to which the predictive information is efficiently exploited and increasing the likelihood of achieving successful statistical arbitrage strategies.

Chapter 12 describes our methodology for diversifying model risk. This is achieved through the use of model combination techniques which take into account the two equally important objectives of maximising return and minimising risk. Furthermore, the methodology emphasises the importance of using selection criteria which are as similar as possible to the ultimate performance measure (i.e. after-costs trading performance) rather than the traditional statistical criteria based upon forecasting accuracy alone. The traditional approach to model combination is used in conjunction with the risk-averse optimisation techniques of modern portfolio theory in order to achieve a “portfolio of models” approach. This approach is evaluated with respect to the conditional statistical arbitrage models described in the previous part of the thesis.

Chapter 13 represents a less developed solution to a much more ambitious task, namely that of integrating all stages of the modelling process in a single optimisation procedure. The objective of this approach is to reduce, and ultimately eliminate, the various inefficiencies which arise through the use of “multi-stage” approaches to modelling; for instance, when pre-processing, predictive modelling, and trading rule implementation as treated as separate rather than inter-dependent tasks. This chapter describes a population-based algorithm in which an entire set of models is generated in the context of a “joint” optimisation procedure. Through the use of optimisation criteria in which individual models are evaluated in terms of the *added-value* they provide to an existing portfolio, the algorithm actively encourages diversification and hence maximises the consequent opportunities for risk-reduction. The approach is evaluated with respect to both artificial and real-world problems.

Conclusions and Bibliography

In this final part of the thesis we present a summary of our main conclusions and a bibliography. Chapter 14 presents the main conclusions of the thesis and discusses directions for further work; it contains an evaluation of the contribution which our current methodology makes to the state of the art, suggestions for future extensions and refinements of the methodology, and a discussion of the scope of possible practical applications of our work both in statistical arbitrage specifically and in the broader context of investment finance in general.

14. Conclusions

In this thesis we have proposed and developed an integrated framework which enables the use of recent advances in computational modelling as a means of exploiting small but consistent regularities in asset price dynamics. We adopt a holistic perspective in which our methodology is based on an extensive analysis of the obstacles which arise in financial forecasting and the manner in which they influence the effectiveness of the modelling process as a whole. We have addressed the weaknesses of existing methodology by combining computational modelling techniques from a number of different fields. Within our methodological framework we apply the different techniques only to the parts of the modelling process for which they are inherently suitable, thus maximising the strengths of the various techniques whilst minimising the effect of their weaknesses.

Within our integrated modelling framework, we have developed specific tools and techniques which represent significant advances upon the existing state of the art. We have consistently exploited the *flexibility* which is offered by emerging modelling techniques such as neural networks and genetic algorithms whilst ensuring that this flexibility is employed in an *appropriate* manner. This has been achieved by placing the emerging modelling techniques in the context of, and in partnership with, the methodological rigour and diagnostic techniques which are provided by established modelling tools from the fields of statistics, econometrics and time-series forecasting.

In particular, we have developed extensions of the econometric methodology of cointegration which are suitable for use in cases where the parameters of the underlying relationship are time-varying and for cases where the number of time-series involved in the analysis in numbered is tens or hundreds rather than units. We have developed novel tests for identifying deterministic components in time-series behaviour which are based upon the joint distribution of a set of variance ratio statistics. We have demonstrated, through controlled simulations, that our new tests are sensitive to a *wider range* of deviations from random behaviour than are standard predictability tests. By means of a computationally intensive approach based upon Monte-Carlo simulation, we have generalised the applicability of both our new tests and existing predictability tests to the case where the time-series under analysis represents the result of a cointegration-based pre-processing procedure. The advantage of the simulation-based approach is that the actual empirical distribution of the test statistics can be determined under equivalent experimental parameters to those which are present in a given analysis, thus accounting for any artefacts which are induced by the pre-processing stage, automatically adjusting for sample size effects and avoiding any inefficiencies which would be incurred through the use of incorrect theoretical assumptions.

Through the use of an "equivalent kernels" perspective taken from nonparametric statistics, we have achieved a synthesis which includes both traditional parametric regression modelling and neural network learning. In particular, this approach allows us to compute the "degrees of freedom" which are contained in a neural network model and use this as the basis of a variety of statistical significance tests for neural network models and components within such models. We have developed three variant algorithms for neural model estimation which combine the low-bias of neural modelling techniques with the low-variance of statistical modelling procedures. Through controlled simulations, we have verified the properties of the algorithms and demonstrated that such a combined approach is vital in the case of modelling highly noisy time-series in which few *a priori* assumptions can be made about the nature of the underlying data-generating process.

We have generalised the model combination approaches of statistical forecasting in order to achieve diversification, and hence reduction, of the model risk which applies in the case of trading model-based strategies. Our "portfolio of models" approach extends the ensemble approach to forecasting with ideas from portfolio theory in order to provide a means of maximising the expected returns whilst simultaneously minimising the level of risk of a combined set of trading strategies as a whole. We have built upon this approach in order to develop a population-based algorithm, which exploits the particular strengths of genetic (and evolutionary optimisation) algorithms as a means of jointly optimising the whole set of models within a population. We have demonstrated, through the use of controlled simulations, that this approach can overcome the *criterion* risk which arises in cases where complex models are optimised in multiple stages and on an *individual* rather than *collective* basis. In particular, the population-based algorithm can be used to generate a portfolio of complementary models by actively encouraging diversification within the population and thus maximising the benefits which can be achieved through the portfolio approach.

We have applied these various methodological developments from a particular perspective which we refer to as "statistical arbitrage". We consider statistical arbitrage as a generalisation of traditional "riskless" arbitrage strategies which are based on predefined relationships between financial assets, typically between derivative instruments such as options and futures contracts and the “underlying” assets upon which the derivatives are based. From our statistical arbitrage perspective, we apply our extended cointegration methodology to identify statistical “fair price” relationships between sets of related asset prices. Just as deviations from theoretical no-arbitrage relationships are considered “mispricings” which represent potential opportunities for riskless arbitrage, we likewise consider deviations from the analogous statistical fair-price relationships as potential opportunities for statistical arbitrage, and refer to them as “statistical mispricings”.

We have demonstrated that our methodology is applicable to real-world problems by performing extensive experimental evaluations from the statistical arbitrage perspective. This approach can, perhaps, be considered the purest method of evaluating the added-value which is provided by a computational modelling approach to investment finance. This is because the profits and losses of the resulting models are almost entirely independent of the underlying movements in the market as a whole and instead reflect only the informational advantage, if any, which is provided by the models themselves. Furthermore, the significance of the resulting performance can be evaluated not only from a statistical perspective but also from a practical perspective in which the economic advantages of the models can be assessed after factors such as transaction costs have been taken into account. In principle, the risk and return of our strategies can either be multiplied through leverage (up or down) and/or overlayed as a market-timing component on top of a more traditional trading strategy. Thus the benefits of our approach are potentially of value to active fund managers in general, as well as arbitrageurs and hedge funds in particular.

From this perspective, the results of our empirical evaluations can be taken as being highly promising whilst at the same time not conclusive. The results are highly promising because they indicate that significant levels of profitability can be achieved, at acceptable levels of risk, even after transaction costs have been taken into account. At the same time we believe that the results are not conclusive because the true tests of a trading methodology cannot be evaluated in a research environment using historical data but must ultimately be performed in a true trading environment using real prices, real trading costs and real trading infrastructure. Having made this caveat, we do believe that our experimental results demonstrate significant potential. Our first set of extensive experiments were based upon an implicit assumption of mean-reverting behaviour in the time-series dynamics of statistical mispricings between daily closing prices of FTSE 100 constituents. During a 200 day out-of-sample period, a set of these “implicit statistical arbitrage” strategies produced a backtested performance of between 7 and

10%, assuming typical institutional levels of transaction costs, with the corresponding Sharpe Ratios of between 2 and 3 demonstrating a high degree of consistency within this performance. More advanced "conditional statistical arbitrage" strategies based upon low-bias neural models of the mispricing dynamics achieved a collective annualised out-of-sample performance of 21% return and 2.45 Sharpe Ratio even at a moderately high level of transaction costs (0.5%). This annualised out-of-sample performance was further improved to 26.6% return and 3.40 Sharpe Ratio by means of the "portfolio of models" approach. Additional experimental results of statistical arbitrage models between international equity market indices indicate that the methodology has real potential in these cases also.

Whilst beyond the scope of this thesis itself, certain additional evaluations and developments of the methodology have been made in the commercial world itself. The original inspiration for our methodology arose from collaborative projects between the Computational Finance Group of the Decision Technology Centre (formerly Neuroforecasting Unit) at London Business School and a number of financial institutions. An earlier version of the methodology formed the basis of the modelling work conducted in the ESPRIT research project "High frequency Arbitrage Trading" (HAT), with favourable live performance evaluations carried out by the two banks in the consortium, one bank evaluating models within the equity and equity derivatives markets and the other evaluating models within the fixed-income (and derivatives) markets. Additional developments of the methodology currently form the basis of commercial negotiations between an LBS spin-off company and a major financial information services company and data vendor.

In terms of further methodological developments we believe that our work raises many avenues for future research. In particular we have highlighted the important role played by the different sources of potential error during the modelling process, especially in the context of the high noise content and temporal instability (nonstationarity) of predictive relationships between asset prices. We believe that further analysis of the issues raised in the thesis will lead to developments concerning the identification of appropriate modelling biases for financial markets, and methods for controlling the sources of model error which are represented by model variance, data snooping, performance nonstationarity and criterion risk.

We believe that there is significant potential to develop the specific methodology described in this thesis, both in the context of statistical arbitrage modelling and also extensions to other modelling domains. Furthermore, our modelling framework as a whole should be considered as inclusive rather than exclusive. We have referred within the body of the thesis to the fact that our cointegration based approach for constructing statistical mispricings could be replaced or combined with other multivariate approaches such as principal components analysis, factor analysis and independent components analysis. Furthermore the linear fair price relationships, to which we restrict ourselves for reasons of implementational convenience, could nevertheless be extended to the nonlinear cases enabled by recent advances in these modelling techniques.

Similarly, the low-bias neural modelling methodology in the second stage of our framework could be either extended to include, or indeed be replaced by, related low-bias approaches from nonparametric statistics or machine learning. It is our belief that given appropriate underlying modelling assumptions the achievable level of performance will be limited more by the informational content of the data itself than by the differences between alternative modelling techniques. However, it would certainly be interesting to quantify the extent to which this is true, and also the circumstances under which techniques such as projection pursuit, smoothing splines and support vector machines will indeed achieve similar results to our synthesis of neural and statistical techniques.

It is at the final stage of our methodology that perhaps the largest questions remain, and the greatest potential for future development. Assuming that cases exist where financial asset price data *can* be preprocessed into a form which contains significant deterministic components, and predictive models *can* be estimated of the resulting time-series dynamics, then the most interesting and important question arises in the form of "how can this advantage best be exploited ?" We believe that we have made important first steps in this direction, in particular though our recognition of the inter-related nature of the various stages of the modelling process and our integration of these stages in the joint optimisation procedure of our population-based algorithm.

Perhaps a productive route towards future advances may be to adopt a similar philosophy to that which underlies this thesis, namely to identify the important problems which established techniques have been developed to solve, and then also to identify the manner in which overly restrictive assumptions of these techniques can be relaxed, through the appropriate application of the continual developments in computational hardware, software, data availability and modelling methodology. Another fruitful route may be to identify the important but previously unanswerable (and hence generally unasked) *questions* which these developments now make it possible to answer. If the financial markets can be thought of as an artificial ecology, then we believe that computational analogies with natural evolution, neural recognition and reinforcement learning, which have achieved such amazing results in the natural ecology, have an equivalent potential in finance which we are only just beginning to realise.

Bibliography

Abu-Mostafa, Y. S., 1990, Learning from hints in neural networks, *Journal of Complexity*, 6, 192-198

Abu-Mostafa, Y. S., 1993, A method for learning from hints, in S. Hanson *et al* (eds), *Advances in Neural Information Processing Systems 5*, Morgan Kaufmann, San Diego, 73-80

Abu-Mostafa, Y. S., 1995, Financial market applications of learning from hints, in A. N. Refenes (ed), *Neural Networks in the Capital Markets,* John Wiley and Sons Ltd., Chichester, 221-232

Akaike, H., 1973, Information theory and an extension of the maximum likelihood principle, in *Second International Symposium on Information Theory* (ed. B. N. Petrov and F. Czaki), Akademiai Kiado, Budapest, 267-81.

Akaike, H., 1974, Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes, *Annals of the Institute of Statist. Math.*, 26, 363-387

Akaike, H., 1974b, A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, AC-19, 716-723.

Amari, S., 1995, Learning and statistical inference, in *The Handbook of Brain Theory and Neural Networks*, M. A. Arbib (Ed.), MIT Press: Cambridge, MA, pp. 522-6

Amari, S., Cichocki, A., and Yang, H. H., 1996, A new learning algorithm for blind signal separation, in Touretzky, D. S., *et al* (Eds.) *Advances in Neural Information Processing Systems 8*, MIT Press, Cambridge MA, 757-763.

Amari, S., and Murata, N., 1993, Statistical theory of learning curves under entropic loss, *Neural Computation*, 5:140-153.

Anders, U. and Korn, O., 1996, Model selection in Neural Networks, *ZEW Discussion Paper 96-21*.

Anderson, J. A., and Rosenfeld, E., (eds.), 1988, *Neurocomputing: Foundations of Research*, Cambridge, Ma: MIT Press

Back, A. D., and Weigend, A. S., 1998, Discovering structure in finance using independent component analysis, in A-P. N. Refenes *et al* (Eds.) *Decision Technologies for Computational Finance*, Kluwer Academic Publishers, Dordrecht., 309-322.

Baillie, R. T., and Bollerslev, T., 1994, Cointegration, Fractional Cointegration, and Exchange Rate Dynamics, *Journal of Finance*, 49, 737-745.

Baluja, S., 1997, Genetic Algorithms and Explicit Search Statistics, in Mozer M. C., *et al* (Eds.) *Advances in Neural Information Processing Systems 9*, MIT Press, Cambridge MA., 319-325.

Barnard, G. A., 1963, New methods of quality control, *Journal of the Royal Statistical Society A*, **126**, 255-9

Bates, J. M. and Granger, C. W. J., 1969, The combination of forecasts, *Operations Research Quarterly*, **20**, 319-25

Bell, A., and Sejnowski, T., 1995, An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, **7**(6), 1129-1159.

Bellman, R. E., 1957, *Dynamic Programming*, Princeton University Press, Princeton, NJ.

Bengio, Y., 1997, Training a neural network with a financial criterion rather than a prediction criterion, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 36-48.

Bentz, Y., 1997, *Factor models in equity investment management: a review*, Working Paper, Decision Technology Centre, London Business School.

Bentz, Y., 1999, *Identifying and modelling conditional factor sensitivities: an application to equity investment management*, Unpublished PhD Thesis, London Business School.

Bentz, Y., and Connor, J. T., 1998, Unconstrained and constrained time-varying factor sensitivities in equity investment management, in A-P. N. Refenes *et al* (Eds.) *Decision Technologies for Computational Finance*, Kluwer Academic Publishers, Dordrecht., 291-308.

Bentz, Y., Refenes, A. N. and De Laulanie, J-F., 1996, Modelling the performance of investment strategies, concepts, tools and examples, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 241-258

Bergerson K., and Wunsch D. C., 1991, A commodity trading model based on a neural network-expert system hybrid, *Proc. IEEE International Conference on Neural Networks*, 1991, 1289-1293, reprinted in (Trippi and Turban, 1993)

Bilge, U., and Refenes A. N., Application of sensitivity analysis techniques to neural network bond forecasting, *Proc. First International Workshop on Neural Networks in the Capital Markets*, November 18-19, 1993, London Business School

Bishop, C. M., 1995, *Neural networks for Pattern Recognition*, Clarendon Press, Oxford

Black, F., 1976, The pricing of commodity contracts, *Journal of Financial Economics*, 3, 167-179

Black, F., and Scholes, M., 1973, The pricing of options and corporate liabilities, *Journal of Political Economy*, 81, 637-654

Bolland, P. J., 1998, *Robust neural estimation and diagnostics*, Unpublished PhD Thesis, London Business School.

Bolland, P. J., and Burgess, A. N., 1997, Forecasting volatility mispricing, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 214-224.

Bolland, P. J., and Connor, J. T., 1996, Identification of FX arbitrage opportunities with a non-linear multivariate Kalman filter, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 122-134

Bollerslev, T., 1986, Generalised autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-328.

Bosarge, W. E., 1991, Adaptive processes to exploit the nonlinear structure of financial markets, Presented at the Santa Fe Institute of Complexity Conference: *Neural Networks and Pattern Recognition in Forecasting Financial Markets,* February 15, 1991, reprinted in (Trippi and Turban, 1993)

Box, G. E. P. and Jenkins, G. M., 1970, *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, (revised edn, 1976)

Box, G. E. P. and Pierce, D. A., 1970, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*., **70**, 1509-26.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone C. J., 1984, *Classification and Regression Trees*, Wadsworth and Brooks/Cole, Monterey.

Breiman, L., 1996, Bagging predictors, *Machine Learning*, 24, 123-140

Breitung, J., 1998, Nonparametric tests for nonlinear cointegration, in A-P. N. Refenes *et al* (Eds.) *Decision Technologies for Computational Finance*, Kluwer Academic Publishers, Dordrecht., 109-123.

Broomhead, D. S., and Lowe, D., 1988, Multivariable functional interpolation and adaptive networks. *Complex Systems*, **2**, 321-355

Brown, R. G., 1963, *Smoothing, Forecasting and Prediction*, Prentice-Hall, Englewood Cliffs, NJ.

Brown, R. L., Durbin, J. and Evans, J. M., 1975, Techniques for testing the constancy of regression relationships over time (with discussion), *Journal of the Royal Statistical Society, series B*, 37, 149-192.

Bruce, A., Connor, J. T. and Martin, R. D., 1996, Prediction with robustness towards outliers, trends, and level shifts, in Refenes et al (eds), *Neural Networks in Financial Engineering*, Singapore: World Scientific

Bunn, D. W., 1975, A Bayesian approach to the linear combination of forecasts, *Operational Research Quarterly*, **26**, 325-9.

Bunn, D. W., 1989, Forecasting with more than one model, *Journal of Forecasting*, **8**, 161-166.

Burgess, A. N., 1995, Non-linear model identification and statistical significance tests and their application to financial modelling, in *IEE Proceedings of the 4th International Conference on Artificial Neural Networks*, Cambridge, 312-317

Burgess A. N, 1995b, Methodologies for neural network systems", *Proc. Neural Networks in Marketing*, London, Dec 13, 1995

Burgess, A. N., 1995c, Robust financial modelling by combining neural network estimators of mean and median, *Proc. Applied Decision Technologies*, UNICOM Seminars, Brunel University, London, UK

Burgess, A. N., 1996, Statistical yield curve arbitrage in eurodollar futures using neural networks, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 98-110

Burgess, A. N., 1997, Asset allocation across european equity indices using a portfolio of dynamic cointegration models, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 276-288.

Burgess, A. N., 1998, *A Computational Intelligence methodology for forecasting noisy, nonstationary time-series*, Working Paper, Decision Technology Centre, London Business School.

Burgess, A. N., 1998b, Controlling nonstationarity in statistical arbitrage using a portfolio of cointegration models, in A-P. N. Refenes *et al* (Eds.) *Decision Technologies for Computational Finance*, Kluwer Academic Publishers, Dordrecht., 89-107.

Burgess, A. N., 1999, *Using regularisation to improve the stability of statistical fair price relationships*, Working Paper, Decision Technology Centre, London Business School.

Burgess, A. N., 1999b, *Analysis of the expected profitability of statistical arbitrage trading strategies in the presence of mean-nonstationarity*, Working Paper, Decision Technology Centre, London Business School.

Burgess, A. N., 1999c, *Multi-stage versus joint optimisation of composite models*, Working Paper, Decision Technology Centre, London Business School.

Burgess A.N. and Bunn, D.W., 1994, The Use of Error Feedback Terms in Neural Network Modelling of Financial Time Series", Proc. NNCM 1994, Pasadena, November 16-18 1994

Burgess, A. N., and Pandelidaki, S., 1996, Etude comparative des reseaux de neurones et de la regression logistique pour identifier les opportunities de ventes croisees, *Recherche et Applications en Marketing*

Burgess, A. N., and Pandelidaki, S., 1998, A Statistical Methodology for Specifying Neural Network Models, in Aurifeille J-M and Deissenberg C., (eds) *Bio-Mimetic Approaches in Management Science*, Kluwer Academic Publishers, 139-152.

Burgess, A. N. and Refenes, A. N., 1995, Principled variable selection for neural network applications in financial time series, *Proc Quantitative Models for Asset Management*, London, 1995

Burgess, A. N. and Refenes, A. N., 1996, Modelling non-linear cointegration in international equity index futures, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 50-63

Burgess, A. N., and Refenes, A. N., 1999, Modelling non-linear moving-average processes using neural networks with error feedback: An application to implied volatility forecasting, *Signal Processing*, 74, 89-99

Candy, J. V., 1986, *Signal processing: The model-based approach*, McGraw-Hill, New York.

Choey, M., and Weigend, A. S., 1997, Nonlinear trading models through Sharpe Ratio Maximization, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 3-22.

Chow, K. V., and Denning, K. C., 1993, A simple multiple variance ratio test, *Journal of Econometrics*, vol 58, no 3, 385-401

Cochrane, J. H., 1988, How Big is the Random Walk in GNP?, *Journal of Political Economy*, vol 96, no. 5, 893-920

Connor, J. T., Martin, R. D., and Atlas, L. E., 1994, Recurrent neural networks and robust time series prediction, *IEEE Transactions on Neural Networks*, March 1994, 240-254

Connor, J. T., Bolland, P. J., and Lajbcygier, P., 1997, Intraday modelling of the term structure of interest rates, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 225-232.

Crites, R. H., and Barto, A. G., 1996, Improving elevator performance using reinforcement learning, in Touretzky, D. S., *et al* (Eds.) *Advances in Neural Information Processing Systems 8*, MIT Press, Cambridge MA, 430-436.

Davis, L., (Ed.), 1987, *Genetic algorithms and simulated annealing*, Pitman, London.

Davis, L., (Ed.), 1991, *Handbook of genetic algorithms*, Van Nostrand Reinhold, New York.

Deb, K., 1989, *Genetic algorithms in multimodal function optimization*, Masters Thesis and TCGA Report No. 89002, University of Alabama, Tuscaloosa, AL.

De Jong, K. A., 1975, An analysis of the behaviour of a class of genetic adaptive systems. *Dissertation Abstracts International*, 36 (10), 5140B.

Dempster, A. P., Laird, N. M, and Rubin, D. E., 1977, Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society B.*, 39, 1-38.

Dickey, D. A., 1976, *Estimation and Hypothesis Testing in Nonstationary Time Series*, Ph.D. dissertation, Iowa State University, Ames.

Dickey, D. A. and Fuller, W. A., 1979, Distribution of the estimators for autoregressive time-series with a unit root, *Journal of the American Statistical Association*, 74, 427-431.

Diebold, F. X., 1998, *Elements of Forecasting*, South-Western College Publishing, Cincinnati, Ohio

Dueker, M., and Startz, R., 1995, Maximum Likelihood Estimation of Fractional Cointegration wiith an Application to the Short End of the Yield Curve, Working paper, October 1995, University of Washington.

Duncan, D. B. and Horn, S. D., 1972, Linear dynamic recursive estimation from the viewpoint of regression analysis, *Journal of the American Statistical Association*, 67, 815-821

Dutta, S., and Shashi, S., 1988, Bond rating: A non-conservative application of neural networks, *in Proc. ICNN-88 San Diego*, 24-27 July 1988, Vol. II, 443-450

Eckbo, B. E., and Liu, J., 1993, Temporary Components of Stock Prices: New Univariate Results, *Journal of Financial and Quantitative Analysis*, Vol. 28, NO. 2, 161-176

Efron, B., and Tibshirani, R. J., 1993, *An Introduction to the Bootstrap,* Chapman and Hall, New York.

Elman, J. L., 1990, "Finding Structure in Time" *Cognitive Science*, 14

Engle, R. F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of the UK inflation, *Econometrica*, 50, 987-1007.

Engle, R. F., and Brown, S. J., 1986, Model Selection for Forecasting, *Applied Mathematics and Computation*, 20, 313-327

Engle, R. F. and Granger, C. W. J., 1987, Cointegration and error-correction: representation, estimation and testing, *Econometrica*, 55, 251-276

Engle, R. F., Lilien, D. M. and Robins, R. P., 1987, Estimating time-varying risk premia in the term structure: the ARCH-M model, *Econometrica*, 55, 391-407.

Engle, R. F., and Yoo, S., 1989, *A Survey of Cointegration*, mimeo, San Diego: University of California

Engle, R. F and Yoo, S., 1991, "Cointegrated economic time series: an overview with new results", pp 237-66 in R.F Engle and C W J Granger (eds) *Long-Run Economic Relationships;* Oxford University Press

Fahlman, S. E. and Lebiere, C., 1990, The cascade-correlation learning algorithm, in D. S. Touretsky (ed.), *Advances in Neural Information Processing Systems 2*, Morgan Kaufmann, San Mateo, CA, 525-532

Faust, J., 1992, When are variance ratio tests for serial dependence optimal?, *Econometrica*, Vol. 60, No. 5, 1215-1226.

Frean, M, 1990, The upstart algorithm: a method for constructing and training feedforward neural networks. *Neural Computation* **2** (2), 198-209

Freund, Y., 1990, Boosting a Weak Learning Algorithm by Majority, *Proceedings of the Third Workshop on Computational Learning Theory*, Morgan-Kaufmann, 202-216

Friedman, J.H. and Stuetzle, W., 1981. Projection pursuit regression. *Journal of the American Statistical Association*. Vol. 76, pp. 817-823.

Friedman, J.H., 1991. Multivariate Adaptive Regression Splines (with discussion). *Annals of Statistics*. Vol 19, num. 1, pp. 1-141.

Frisch, R. and Waugh, F. V., Partial time regressions as compared with individual trends. *Econometrica*, 1, 387-401

Fuller, W. A., 1976, *Introduction to Statistical Time Series*, New York: Wiley.

Gardner, E. S., 1985, Exponential smoothing: The state of the art (with discussion), *Journal of Forecasting*, 4, 1-38.

Geisel, M. S., 1974, Bayesian comparison of simple macroeconomic models, *Money, Credit and Banking*, (1974), 751-72

Geman, S., Bienenstock, E., and Dorsat, R., 1992, Neural networks and the bias/variance dilemma, *Neural Computation* **4**(1), 1-58

Goldberg, D. E., 1989, *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, Reading MA.

Goldberg, D. E., 1998, *From Genetic and Evolutionary Optimization to the Design of Conceptual Machines*, IlliGAL Report No. 98008, University of Illinois.

Goldberg, D. E., 1999, *Genetic and Evolutionary Algorithms in the Real World*, IlliGAL Report No. 99013, University of Illinois.

Gonzalez-Miranda, F., 1993, Modelling implied volatilities using neural networks, *presented at First International Workshop on Neural Networks in the Capital Markets*, November 18-19, 1993, London

Gonzalez-Miranda, F., and Burgess, A. N., 1997, Modelling market volatilities: the Neural Network Perspective, *European Journal of Finance*, Vol. 3, No., 2, 137-1257.

Granger, C. W. J., 1983, Cointegrated variables and error-correcting models , *UCSD Discussion Paper*.

Granger, C. W. J., 1989, Combining Forecasts – Twenty Years Later, *Journal of Forecasting*, **8**, 167-173.

Granger, C. W. J., and Joyeux, R., 1980, An Introduction to Long-Memory Models and Fractional Differencing, *Journal of Time Series Analysis*, 1, 15-39.

Granger, C. W. J., and Hallman, 1991, Nonlinear Transformations of Integrated Time-Series, *Journal of Time Series Analysis*, 12:207-224.

Greene, W. H., 1993, *Econometric Analysis*, Prentice-Hall, New Jersey.

Gregory A. W., 1991, *Testing for cointegration in linear quadratic models*, mimeo, Dept of Economics, Queen's University, Ontario

Grudnitski, G. and Do, A. Q., 1995, Important factors in neural networks' forecasts of gold futures prices, in Refenes (1995), 163-173

Haerdle, W., 1990. *Applied nonparametric regression*. Cambridge University Press.

Hall, A., 1989, Testing for a unit root in the presence of moving average errors, *Biometrika*, **76**, 49-56.

Hansen L., K., and Salamon, P., 1990, Neural network ensambles, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10), 993-1001

Hargreaves, C. P., 1994, *Nonstationary Time Series Analysis and Cointegration*, Oxford University Press

Hargreaves, C. P., 1994b, “A Review of Methods of Estimating Cointegrating Relationships”, in *Nonstationary Time Series Analysis and Cointegration,* Oxford University Press (Ed. C. P. Hargreaves), pp. 87-131

Harik, G., Cantu-Paz, E., Goldberg, D. E., and Miller, B. L., 1997, The Gambler’s Ruin Problem, Genetic Algorithms and the Sizing of Populations, *Proceedings of the 1997 IEEE Conference on Evolutionary Computation*, IEEE Press, New York, 7-12.

Harris D. and Inder B, 1994, “A test of the null hypothesis of cointegration”, pp 133-152 in *Nonstationary Time Series Analysis and Cointegration,* Oxford University Press (Ed. C. P. Hargreaves)

Harrison, P. J. and Stevens C. F., 1971, A Bayesian approach to short-term forecasting, *Operational Research Quarterly*, 22, 341-362

Harrison P. J. and Stevens, C. F., 1976, Bayesian forecasting, *Journal of the Royal Statistical Society, series B*, 38, 205-247.

Harvey, A. C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

Harvey, A. C., 1993, *Time Series Models*, second edition, Harvester Wheatsheaf, London.

Harvey, A. C. and Phillips, G. D. A., 1979, Maximum likelihood estimation of regression models with autoregressive moving average disturbances, *Biometrika*, 66, 49-58.

Hassibi, B. and Stork, D. G., 1993, Second order derivatives for network pruning: Optimal Brain Surgeon, *Advances in Neural Information Processing Systems 5*, Morgan Kaufmann, 164-171

Hastie, T.J. and Tibshirani, R.J., 1990. *Generalised Additive Models*. Chapman and Hall, London.

Hatanaka, M., 1975, On the global identification of the dynamic simultaneous equations model with stationary disturbances, *international Economic Review*, 16, 545-554

Hebb, D. O., 1949, *The organization of behaviour*, New York: Wiley.

Hinton, G., E., 1987. Learning translation invariant recognition in massively parallel networks. In J. W. de Bakker, A. J. Nijman and P. C. Treleaven (eds.), *Proceedings PARLE Conference on Parallel Architectures and Languages Europe*, pp. 1-13. Berlin: Springer-Verlag.

Hinton, G. E., 1989. Connectionist learning procedures. *Artificial Intelligence* **40**, 185-234.

Hoerl, A. E., and Kennard, R. W., 1970a, Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics* 12, 55-67

Hoerl, A. E., and Kennard, R. W., 1970b, Ridge regression: Application to nonorthogonal problems, *Technometrics* 12, 69-82

Holland, J. H., 1975, *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, MI.

Holland, J. H., and Reitman, J. S., 1978, Cognitive systems based on adaptive algorithms, In Waterman D. A., and Hayes-Roth, F. (Eds.), *Pattern directed inference systems*, Academic Press, New York, 313-329.

Holt, C. C., 1957, Forecasting seasonals and trends by exponentially weighted moving averages, *ONR Research Memorandum No 52, Carnegie Institute of Technology*.

Horn, J., Goldberg, D. E., and Deb, K., 1994, Implicit Niching in a Learning Classifier System: Nature's Way. *Evolutionary Computation*, 2(1), 37-66.

Horn, J., and Goldberg, D. E., 1996, Natural Niching for Evolving Cooperative Classifiers, in Koza, J. R. *et al*(Eds.), *Genetic Programming: Proceedings of the First Annual Conference 1996*. MIT Press, Cambridge MA.

Horn, J., Nafpliotis, N., and Goldberg, D. E., 1994, A Niched Pareto Genetic Algorithm for Multiobjective Optimization, *Proceedings of the First IEEE Conference on Evolutionary Computation*, IEEE Press, Piscataway, NJ.

Hornik, K., Stinchcombe, M. and White. H., 1989, Multilayer feedforward networks are universal approximators, *Neural Networks* **2** (5), 359-366

Hull, J. C., 1993, *Options, Futures and other derivative securities*, Prentice-Hall

Hutchinson, J., Lo, A. and Poggio, T., A non-parametric approach to pricing and hedging derivative securities via learning networks, *Journal of Finance*, XLIX:3 (July 1994).

Jacobs, B. I., and Levy, K. N., 1988, Disentangling Equity Return Regularities: New Insights and Investment Opportunities, *Financial Analysts Journal*, May-June 1988, 18-43.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J. and Hinton, G. E., 1991, Adaptive mixtures of local experts, *Neural Computation*, 3, 79-87

Johansen, S., 1988, Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12, 131-154.

Johansen, S., 1991, "Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models", *Econometrica*, Vol 59, pp 1551-81

Johansen, S. and Juselius, K., 1990, Maximum likelihood estimation and inference of cointegration - with applications to the demand for money, *Oxford Bulletin of Economics and Statistics*, 52, 169-210

Jolliffe, I. T., 1986, *Principal Component Analysis*, Springer-Verlag, New York.

Jordan, M. I. and Xu, L., 1995, Convergence results for the EM approach to mixtures of experts architectures, *Neural networks*, **8**(9), 1409-1431.

Juels, A., and Wattenberg, M., 1996, Stochastic Hillclimbing as a Baseline Method for Evaluating Genetic Algorithms, in Touretzky, D. S., *et al* (Eds.) *Advances in Neural Information Processing Systems 8*, MIT Press, Cambridge MA, 430-436.

Kalman, R. E., 1960, A new approach to linear filtering and prediction problems, *Journal of Basic Engineering, Transactions ASME, Series D,* 82, 35-45.

Kalman, R. E. and Bucy R. S., 1961, New results in linear filtering and prediction theory, *Journal of Basic Engineering, Transactions ASME, Series D,* 83, 95-108

Kohonen, T., 1982, Self-organized formation of topologically correct feature maps. *Biological Cybernetics* **43**, 59-69. Reprinted in Anderson and Rosenfeld (1988)

Krogh, A., and Vedelsby, J., 1995, Neural network ensembles, cross-validation and active learning, in G. Tesauro et al (eds), *Advances in Neural Information Processing Systems 7,* MIT press, Cambridge, Ma., 231-238

Kwok T-Y., and Yeung, D-Y., 1995, *Constructive neural networks for regression problems: A survey*, Technical Report HKUST-CS95-43, Hong Kong University of Science and Technology.

Lajbcygier, P., Boek, C., Palaniswami, M. and Flitman, A., 1996, Neural network pricing of all-ordinaries SPI options on futures, *in* Refenes et al. (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 64-77

Lapedes A. and Farber, R., 1987, Nonlinear signal processing using neural networks, *Proc. IEEE Conference on Neural Information Processing Systems - Natural and Synthetic*

Le Cun, Y., Denker, J. S., and Solla, S. A., 1990, Optimal brain damage. In D. S. Touretzky (Ed.) *Advances in Neural Information Processing Systems 2*, 598-605. San Mateo, CA: Morgan Kaufmann

Lee, T-H., White, H. and Granger, C. W. J., 1993, Testing for neglected nonlinearity in time series models, *Journal of Econometrics*, 56, 269-290

Lo A. W. and MacKinlay A. C., 1988, Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test, *The Review of Financial Studies*, 1988, Vol 1, No. 1, pp. 41-66

Lo A. W. and MacKinley A. C., 1989, The size and power of the variance ratio test in finite samples: A Monte Carlo Investigation, *Journal of Econometrics*, **40**, 203-238.

Lo, A. W. and MacKinlay, A. C., 1990, Data-Snooping Biases in Tests of Financial Asset Pricing Models, *Review of Financial Studies*,Vol.3, No.3

Lo, A. W., and MacKinley, A. C., 1995, *Maximising predictability in the stock and bond markets*, Working Paper No. 5027, National Bureau of Economic Research

Lo, A. W., and MacKinley, A. C., 1999, *A Non-Random Walk down Wall Street*, Princeton University Press.

Lobo, F. G., and Goldberg, D. E., 1997, Decision making in a hybrid genetic algorithm, in Back, T. (ed.), *Proc. Of the IEEE Conference on Evolutionary Computation*, pp. 121-125. New York: IEEE Press.

Lowe, D., 1995, On the use of nonlocal and non positive definite basis functions in radial basis function networks, *Proceedings of the Fourth IEE Conference on Artificial Neural Networks*, pp. 206-211

Lyung, G. M. and Box, G. E. P., 1978, On a measure of lack of fit in time series models. *Biometrika*, **65**, 297-304.

MacKay, D. J. C., 1992. A practical Bayesian framework for backpropagation networks. *Neural Computation* **4** (3), 448-472

Mahfoud, S. W., 1995, *Niching Methods for Genetic Algorithms*, PhD Thesis, University of Illinois.

Mann, H. B., and Wald, A., 1943, On the statistical treatment of linear stochastic difference equations, *Econometrica*, **11**, 173-220.

Markellos, R. N., 1997, *Nonlinear Equilibrium Dynamics*, Working Paper 97/6, Department of Economics, Loughborough University, UK.

Markowitz, H. M., 1952, Portfolio Selection, *Journal of Finance*, 7, 77-91.

Markowitz, H. M., 1959, *Portfolio Selection: Efficient Diversification of Investments*. John Wiley and Sons, New York.

McCulloch, W. S., and Pitts, W., 1943, A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, 5, 115-133

Meier, D. C., Pfeifer, R., Demostene, R. and Scheier C., 1993, Is mean-reversion on stock indices a linear artifact ?, *Proc. First International Workshop on Neural Networks in the Capital Markets,* November 18-19, 1993, London Business School

Merton, R. C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica*, 41, 867-887.

Mezard, M., and J. P. Nadal, 1989, Learning in feedforward layered networks: The tiling algorithm. *Journal of Physics, A* **22**, 2191-2203

Michalewicz, Z., 1996, *Genetic Algorithms + Data Structures = Evolution Programs* (third edition), Springer-Verlag.

Miller, B. L., and Goldberg, D. E., 1996, Optimal Sampling for Genetic Algorithms, in *Artificial Neural Networks in Engineering (ANNIE) '96, Vol. 6,* pp. 291-298, New York, ASME Press.

Miller, B. L., and Shaw, M. J., 1996, Genetic Algorithms with Dynamic Niche Sharing for Multimodel Function Optimization, *Proceedings of the IEEE Conference on Evolutionary Computation*, IEEE Press, Piscataway NJ, 786-791.

Minsky, M., 1954, *Neural nets and the brain-model problem*, Unpublished doctoral thesis. Princeton University.

Minsky, M., 1959, Some methods of artificial intelligence and heuristic programming. In *Mechanisation of thought processes: Proceedings of a symposium held at the National Physical Laboratory, November 1958. Vol 1. pp. 3-28*

Minsky, M., and Papert, S., 1969, *Perceptrons*, Cambridge, MA: MIT Press

Mitchell, M., 1996, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge MA.

Moody J. E., and Darken C. J., 1989, Fast learning in networks of locally-tuned processing units, *Neural Computation*, **1** (2), 281-294

Moody J. E, 1992, The effective number of parameters: an analysis of generalisation and regularization in nonlinear learning systems, in J. E. Moody, S. J. Hanson and R. P. Lipmann (eds), *Advances in Neural Information Processing Systems 4*, 847-54, Morgan Kaufmann, San Mateo, US

Moody, J. E., 1995, *Economic forecasting: challenges and neural network solutions*, keynote talk presented at the International Symposium on Artificial Neural Networks, Hsinchu, Taiwan, December 1995. (ftp: neural.cse.ogi.edu/pub/neural/papers)

Moody, J. E., and Rognvaldsson, T. S., 1997, Smoothing regularizers for projective basis function networks, in Mozer et al (eds) *Advances in Neural Information Processing Systems 9*, MIT press, Cambridge, Mass., 585-591

Moody, J. E., and Saffell, M., 1999, Minimizing Downside Risk via Stochastic Dynamic Programming, to appear in Y. S. Abu-Mostafa *et al* (Eds.) *Computational Finance – Proceedings of the Sixth International Conference,* Leonard N. Stern School of Business, January 1999.

Moody, J. E., and Utans, J., 1992, Principled architecture selection for neural networks: Application to corporate bond rating prediction, in J. E. Moody, S. J. Hanson and R. P. Lipmann (eds), *Advances in Neural Information Processing Systems 4*, Morgan Kaufmann Publishers, San Mateo, CA, 683-690

Moody, J. E., and Wu, L., 1994, Statistical analysis and forecasting of high frequency foreign exchange rates, *Proc. Neural Networks in the Capital Markets*, Pasadena, November 16-18, 1994

Moody, J. E., and Wu, L., 1997, Optimization of trading systems and portfolios, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 23-35.

Moody, J. E., and Wu, L., 1997b, What is the "true price"? - state space models for high-frequency FX data, in Weigend et al (eds) *Decision Technologies for Financial Engineering*, World-Scientific, Singapore, 346-358.

Moody, J. E., Wu, L., Liao, Y., and Saffell, M., 1998, Performance Functions and Reinforcement Learning for Trading Systems and Portfolios, *Journal of Forecasting*, Vol. 17, 441-470.

Morgan, N., and Bourlard, H., 1990, Generalisation and Parameter Estimation in Feedforward Nets: Some Experiments" in *Advances in Neural Information Processing Systems 2*, D. Touretzky, ed., pp. 630-637. San Mateo, CA: Morgan Kaufmann.

Morris, P. A., 1974, Decision analysis expert use, *Management Science*, **20**, 1233-41.

Murata, N., Yoshizawa, S., and Amari, S., 1993, Learning curves, model selection and complexity of neural networks, in Hanson et al (eds) *Advances in Neural Information Processing Systems 5*, Morgan Kaufmann, San Mateo, CA.

Muth, J. F., 1960, Optimal properties of exponentially weighted forecasts, *Journal of the American Statistical Association*, 55, 299-305

Nelson, C. R., 1972, The prediction performance of the FRB-MIT-PENN model of the US economy, *American Economic Review,* **62**, 902-917

Nelson, C. R. and Kang, H., 1981, Spurious periodicity in inappropriately detrended series. *Econometrica*, 49, 741-751

Nelson, C. R. and Kang, H., 1984, Pitfalls in the use of time as an explanatory variable. *Journal of Business and Economic Statistics*, 2, 73-82

Nelson, D., 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica,* 59(2), 347-370.

Nerlove, M. and Wage, S., 1964, On the optimality of adaptive forecasting, *Management Science*, 10, 207-224.

Newbold, P., 1974, The exact likelihood function for a mixed autoregressive moving average process, *Biometrika*, **61**, 423-426.

Newbold, P., and Granger, C. W. J., 1974, Experience with forecasting univariate time-series and the combination of forecasts, *Journal of the Royal Statistical Society A,* **137**, 131-47.

Niranjan, M., 1997, Sequential tracking in pricing financial options using model-based and neural network approaches, in Mozer et al (eds) *Advances in Neural Information Processing Systems 9*, MIT press, Cambridge, Mass., 960-966.

Opitz D., and Shavlik, J., 1999, A Genetic Algorithm Approach for Creating Neural Network Ensembles, in Sharkey, A. J. C. (Ed.), *Combining Artificial Neural Networks*, Springer-Verlag, London., 79-99.

Pantula, S, 1991, Asymptotic distributions of unit-root tests when the process is nearly stationary, *Journal of Business and Economic Statistics*, **9**, 63-71

Park J. Y., 1989, “Canonical cointegrating regressions”, *Econometrica,* Vol. 60, pp 119-143

Park J. Y., Ouliaris S., and Choi B., 1988, “Spurious Regressions and Tests for Cointegration”, mimeo, Cornell University

Parker, D. B., 1985, Learning logic. Technical Report TR-47, Cambridge, MA.: MIT Center for Research in Computational Economics and Management Science.

Phillips, P. C. B., 1986, Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33, 311-340.

Phillips, P. C. B., 1987, Time series regression with a unit root, *Econometrica*, 55, 277-302.

Phillips P. C. B. and Ouliaris S., 1988, “Testing for cointegration using Principal Components Methods”, *Journal of Economic Dynamics and Control*, Vol 12, pp 105-30

Phillips, P. C. B. and Ouliaris S., 1990, Asymptotic properties of residual based tests for cointegration, *Econometrica*, 58, 165-193.

Phillips, P. C. B. and Perron, 1988, Testing for a unit root in time series regression, *Biometrika*, **75**, 335-346

Poggio T., and Girosi, F, 1990, Regularisation algorithms for learning that are equivalent to multilayer networks, *Science* **247**, 978-982

Refenes, A. N., 1992, Constructive learning and its application to currency exchange rate prediction, in (Trippi and Turban, 1993)

Refenes, A. N., 1995, (ed) *Neural Networks in the Capital Markets*, John Wiley and Sons Ltd., Chichester.

Refenes A. N., and Azema-Barac, M., 1994, Neural network applications in financial asset management, *Neural Computing and Applications*, 2, 13-39.

Refenes, A. N., Bentz, Y., Bunn, D. W., Burgess, A. N., and Zapranis, A. D., 1994, Backpropagation with discounted least squares and its application to financial time-series modelling, *Proc. Neural Networks in the Capital Markets*, Pasadena, November 16-18, 1994.

Refenes, A. N., Bentz, Y., Bunn, D. W., Burgess, A. N., and Zapranis, A. D., 1997, Backpropagation with discounted least squares and its application to financial time-series modelling, *Neurocomputing*, Vol. 14, No. 2, 123-138.

Refenes, A. N., Bentz, Y. and Burgess, N., 1994, Neural networks in investment management, *Journal of Communications and Finance*, 8, April 95-101

Refenes, A. N., Burgess A. N., Bentz, Y., 1997, Neural Networks in Financial Engineering: a study in Methodology, *IEEE Transaction on Neural Networks*, Vol. 8, No. 6, 1222-1267.

Refenes A.N., and Vithlani S., 1991, Constructive learning by specialisation. In *Proc. ICANN-1991, Helsinki, Finland*.

Refenes, A. N., Zapranis, A. D. and Francis, G., 1995, Modelling stock returns in the framework of APT: A comparative study with regression models, in (ed.) Refenes (1995), 101-125

Refenes, A. N., and Zaidi, A., 1995, Managing exchange-rate prediction strategies with neural networks, in (Refenes, 1995), 213-219

Rehfuss, S., Wu, L., and Moody, J. E., 1996, Trading using committees: A comparative study, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 612-621.

Renner, R. S., 1999, *Improving Generalisation of Constructive Neural Networks using Ensembles*, PhD Thesis, Florida State University.

Ripley, B.D., 1994, Neural networks and related methods for classification, *Journal of the Royal Statistical Society*, B, 56, No 3, 409-456.

Ripley, B. D., 1996, *Pattern Recognition and Neural Networks*, Cambridge University Press, England.

Rosenblatt, F., 1959, Two theorems of statistical separability in the perceptron, In *Mechanisation of thought processes: Proceedings of a symposium held at the National Physical Laboratory, November 1958*. Vol 1. 421-456

Rosenblatt, F., 1962, *Principles of neurodynamics*,. New York: Spartan.

Ross, S. A., 1976, The Arbitrage Pricing Theory of Capital Asset Pricing, *Journal of Economic Theory*, **13**, 341-360

Rumelhart, D., E., Hinton, G. E., and Williams R., J., 1986. Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland and the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Volume 1:Foundations, pp 318-362. Cambridge MA: MIT Press.

Said, S.E., and Dickey D. A., 1984, Testing for unit roots in autoregressive moving-average models of unknown orders, *Biometrika*, Vol. 71, pp 599-607

Sargan, J. D. 1964, Wages and prices in the United Kingdom: A study in econometric methodology, in *Econometric Analysis for National Economic Planning*, P. E. Hart *et al* (Eds.), London: Butterworth.

Sargan, J. D. and Bhargava, A., 1983, Testing residuals from least squares regression for being generated by the Gaussian random walk, *Econometrica*, 51, 153-174

Schapire, R. E, 1990, The strength of weak learnability, *Machine Learning*, 5(2), 197-227

Schoenenberg, E., 1990, Stock price prediction using neural networks: A project report, *Neurocomputing*, 2, 17-27

Schreiner, P., 1998, *Statistical Arbitrage in Euromark Futures using Intraday Data*, unpublished M.Sc. Thesis, Department of Mathematics, King's College London.

Schwarz, G., 1978, Estimating the Dimension of a Model, *Annals of Statistics*, 6, 461-464.

Schwert, G. W., 1989, Tests for unit roots: A Monte Carlo investigation, *Journal of Business and Economic Statistics*, **7**, 147-160.

Sharkey, A. J. C., 1999, *Combining Artificial Neural Networks*, Springer-Verlag, London.

Sharpe, W. F., 1964, Capital Asset Prices: A Theory of Market Equilibrium, *Journal of Finance*, **19**, 425-442.

Sharpe, W. F., 1966, Mutual Fund Performance, *Journal of Business*, January 1966, 119-138.

Sietsma , J., and Dow, R. F. J., 1991, Creating artificial neural networks that generalize, *Neural Networks*, 4, 67-79.

Sims, C., 1980, Macroeconomics and reality, *Econometrica*, 48, 1-48.

Slutsky, E., 1927 , The summation of random causes as the source of cyclic processes. *Econometrica*, 5, 105-146

Spanos, A., 1986, *Statistical Foundations of Econometric Modelling*, Cambridge: Cambridge University Press

Srinivas, N., and Deb, K., 1995, Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms, *Evolutionary Computation*, 2(3), 221-248.

Steiner, M., and Wittkemper, H-G., 1995, Neural networks as an alternative stock market model, in (Refenes, 1995), 137-147.

Steurer, E., and Hann, T. H., 1996, Exchange rate forecasting comparison: neural networks, machine learning and linear models, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 113-121

Stock, J. H., 1987, "Asymptotic properties of least squares estimators of cointegrating vectors", *Econometrica*, Vol. 55, pp 1035-56

Stock, J. H. and Watson, M., 1988, Testing for common trends, *Journal of the American Statistical Association*, 83, 1097-1107.

Stock J. H. and Watson M., 1989, A simple MLE of cointegrating vectors in higher order integrated systems, Technical Working Paper No. 83, National Bureau of Economic Research

Sullivan, R., Timmerman, A., and White, H., 1998, *Data-Snooping, Technical Trading Rule Performance, and the Bootstrap*, UCSD Working Paper, 97-31.

Sutton, R. S., 1988, Learning to predict by the method of temporal differences, *Machine Learning*, 3, 9-44.

Tesauro, G. J., 1992, Practical issues in temporal difference learning, *Machine Learning*, 8 (3/4), 257-277.

Theil, H. and Wage, S., 1964, Some observations on adaptive forecasting, *Management Science*, 10, 198-206.

Thierens, D., and Goldberg, D. E., 1993, Mixing in genetic algorithms, *Proceedings of the Fifth International Conference on Genetic Algorithms*,38-45.

Thierens, D., 1995, *Analysis and design of genetic algorithms*, Unpublished doctoral dissertation, Catholic University of Leuven, Leuven.

Tiam, L.C., 1993, Hybrid technologies for far east markets, *Proc. First International Workshop on Neural Networks in the Capital Markets*, London, November 18-19, 1993

Titterington, D. M., 1985, Common Structure of Smoothing Techniques in Statistics, *International Statistical Review*, **53**, 2, pp. 141-170

Tjangdjaja, J., Lajbcygier, P., and Burgess, N., Statistical Arbitrage Using Principal Component Analysis For Term Structure of Interest Rates, in Xu, L. *et al* (Eds.), *Intelligent Data Engineering and Learning*, Springer-Verlag, Singapore, 43-53.

Towers, N., 1998, *Statistical Fixed Income Arbitrage*, Deliverable Report D4.6, ESPRIT project "High performance Arbitrage detection and Trading" (HAT), Decision Technology Centre, London Business School.

Towers, N., 1999, *Joint optimisation of decision policies and forecasting models*, Working Paper, Decision Technology Centre, London Business School.

Towers, N., and Burgess, A. N., 1998, Optimisation of Trading Strategies using Parametrised Decision Rules, in Xu, L. *et al* (Eds.), *Intelligent Data Engineering and Learning*, Springer-Verlag, Singapore, 163-170.

Towers, N., and Burgess, A. N., 1999a, Implementing trading strategies for forecasting models, *Proceedings Computational Finance 1999,* (in press).

Towers, N., and Burgess, A. N., 1999b, A framework for applying Reinforcement Learning to Investment Finance, *Machine Learning*, accepted for review March 1999.

Trippi, R. R., and Turban, E., 1993, *Neural networks in finance and investing*, Probus Publishing, Chicago

Utans, J, and Moody, J. E., 1991, Selecting neural network architectures via the prediction risk: Application to corporate bond rating prediction, in *Proc. of the First International Conference on Artificial Intelligence Applications on Wall Street*, IEEE Computer Society Press, Los Alamitos, CA.

Wahba, G., 1990, *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia.

Wahba, G., and Wold, S., 1975, A completely automatic French curve: fitting spline functions by cross-validation, *Communications in Statistics, Series A, 4*, 4-17.

Watkins, C., 1989, Learning from delayed rewards, PhD Thesis, Cambridge University.

Wallis, K. F., 1977, Multiple time series analysis and the final form of econometric models, *Econometrica*, 45, 1481-1497

Weigend A. S., Huberman B. A. and Rumelhart, D. E., 1990, Predicting the Future: A Connectionist Approach, *International Journal of Neural Systems* **1**, 193-209

Weigend A. S., Huberman B. A. and Rumelhart, D. E., 1992, Predicting sunspots and exchange rates with connectionist networks, in *Nonlinear modelling and forecasting*, Eds. Casdagli M. and Eubank S., Addison-Wesley

Weigend, A. S., and Mangeas, M., 1995, Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting. Technical Report CU-CS-725-94, Computer Science Department, University of Colorado at Boulder, ftp://ftp.cs.colorado.edu/pub/Time-Series/MyPapers/experts.ps

Weigend A. S., and Mangeas, M., 1996, Analysis and prediction of multi-stationary time series, in Refenes et al (eds), *Neural Networks in Financial Engineering*, World Scientific, Singapore, 597-611

Weigend, A. S., and Shi, S., 1998, *Predicting Daily Probability Distributions of S&P500 Returns*, Working Paper IS-98-23/S-98-36, Leonard N. Stern School of Business.

Weisberg S, 1985, *Applied Linear Regression*, Wiley, New York, USA

Weisweiller, R., 1986, *Arbitrage*, Wiley, New York, USA.

Werbos, P. J., 1974, Beyond regression: new tools for prediction and analysis in the behavioural sciences. Ph.D. Thesis, Harvard University, Boston, MA.

Werbos P.J., 1990, Backpropagation through time: What it does and how to do it, *Proc IEEE, Vol 78, No. 10*, Oct 1990.

White, H., 1988, Economic prediction using neural networks: The case of IBM daily stock returns, *Proc. IEEE International Conference on Neural Networks, July 1988,* reprinted in (Trippi and Turban, 1993)

White, H., 1989, An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks, in *Proc. of the International Joint Conference on Neural Networks,* Washington DC (IEEE press, New York, NY) II 451-455

White, H., 1991, Nonparametric estimation of conditional quantiles using neural networks, *Proc. Twenty-Second Symposium on the Interface*, New-York, Springer-Verlag, 190-199

White, H., 1997, *A Reality Check for Data Snooping,* San diego, NRDA Technical Report 97-01.

Williams, C. K. I., Qazaz, C., Bishop, C.M., and Zhu, H., 1995, On the relationship between bayesian error bars and the input data density, *Proceedings of the Fourth IEE Conference on Artificial Neural Networks,* pp. 160-165.

Winkler, R. C., and Makridakis, S., 1983, The combination of forecasts, *Journal of the Royal Statistical Society A*, **146**, 150-57

Winter, G., Periaux, J., Galan, M., and Cuesta, P., (eds.), 1995, *Genetic algorithms in engineering and computer science*, John Wiley and Sons, Chichester.

Winters, P. R., 1960, Forecasting sales by exponentially weighted moving averages, *Management Science*, 6, 324-342

Wolpert, D. H., 1992, Stacked generalisation, *Neural Networks*, 5(2), 241-259

Wong, M. A., 1993, *Fixed-Income Arbitrage*, John Wiley and Sons, Inc: New York.

Yule, G.U., 1921, On the time-correlation problem with special reference to the variate-difference correlation method. *Journal of the Royal Statistical Society*, 84, July, 497-526

Zapranis, A., and Refenes, A. N., 1999, *Principles of Neural Model Identification, Selection and Adequacy*, Springer, London.

Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics*, John Wiley, New York