Advanced Neural Networks: Finance, Forecast, And Other Applications

Dissertation
zur Erlangung
des Titels eines Doktors der Wirtschaftswissenschaften
— rerum politicarum —
der Wirtschaftswissenschaftlichen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover

vorgelegt von
Hans-Jörg Henri von Mettenheim
geboren am 24. Juli 1981 in Hannover

Erstgutachter: Prof. Dr. Michael H. Breitner
Zweitgutachter: Prof. Dr. Karl-Peter Wiedmann

Hannover, 8. Oktober 2009

Acknowledgements

First and foremost I thank my supervisor Prof. Dr. Michael H. Breitner who encouraged me in many ways to write this dissertation. His ideas appear throughout the text and it is always interesting to discuss new topics with him. Especially, I am very grateful to him for giving me the opportunity to continue studying mathematics.

I also thank Prof. Dr. Klaus-Peter Wiedmann, my second referee, for accepting to examine my dissertation.

Many people contributed in different ways to my dissertation. It is impossible to name them all. Without making a claim of completeness I mention some of them. I thank my colleagues Karsten Sohns and Marc Klages. We often had interesting discussions which gave me new insights on how markets work. I also thank Cornelius Köpp for numerous technical advice. You three had to endure me all the time... I also thank Prof. Christian Dunis, director of the Centre for International Banking Economics and Finance at Liverpool John Moores University. His studies on neural networks helped me considerably.

I dedicate this work to my family. Your continuous support is essential for me. Thank you!

Contents

Executive Summary

The content summarized for the rushing executive in 150 words

My dissertation enhances the FAUN neurosimulator. It implements a grid computing client. With this client you can reuse spare computing capacity of user workstations. Remote wake up and shutdown saves power costs when the computers are not needed. I present and analyze a novel neural network topology, the shared layer perceptron. It is memory enabled and allows multi asset and multi step forecasts. Convergence is robust and not sensitive to meta parameters. Applications include modeling market value at risk, transaction decision support and investment. I use 25 financial time series spanning 10 years. The shared layer perceptron produces good or even very good results on equities, interest and exchange rates, and commodities. Multi step forecasts especially enable market timing with high accuracy. The distribution of returns allows to evaluate the probable path of the portfolio within confidence bands. Performance is robust over a time span of 8 years, without retraining.

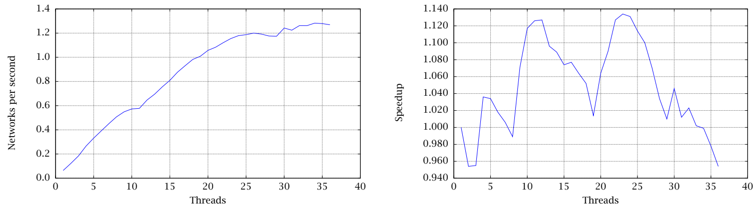
The content summarized on 8 pages

My dissertation answers the research question «Can advanced neural networks provide sustainable and economic competitive edge in today's financial markets?» I show that neural networks are indeed capable of adding value to financial applications. To achieve this requires several components working together. Figure 0.1 on the following page provides an overview.

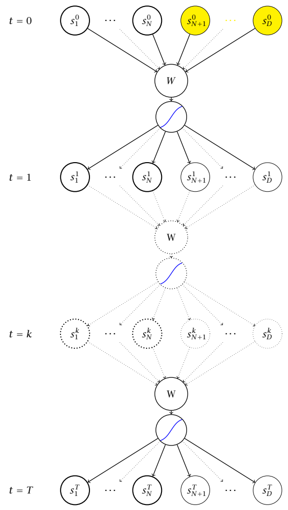
My research question considers several important aspects:

- I investigate *advanced neural networks*. This is not a standard multi layer perceptron but a quite new topology, the shared layer perceptron, that allows easy modeling of multi dimensional financial time series.
- The models should be *sustainable*, i.e., more than just a statistical fluke, more than just a lucky hit. They should be robust over time.

Chapter 2 The FAUN grid computing client offers speedups of more than 95 %.



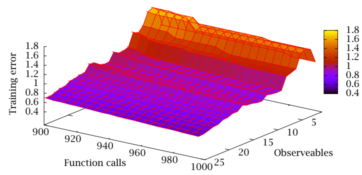
Chapter 3 We meet the shared layer perceptron. A memory enabled neural network topology for multi asset multi time step forecasts.



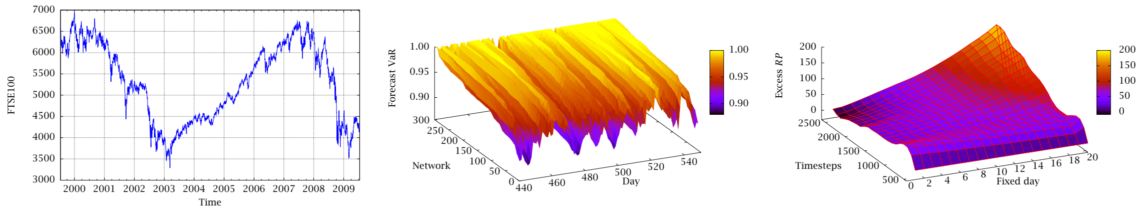
$$\vec{s}^{t+1} = \tanh(W\vec{s}^{t}).$$

$$\frac{\partial E}{\partial w_{i,j}} = \sum_{t=1}^{T} l_i^t \vec{s}_j^{t-1}.$$

$$\vec{l}^t = (1 \ominus (\vec{s}^t)^2) \otimes (W'\vec{l}^{t+1} \oplus \vec{\varepsilon}^t).$$



Chapter 4 We analyze different financial applications: market value at risk, transaction decision support, and investment.



Chapter 5 Conclusions. The shared layer perceptron topology

- is very robust. It performs regardless of asset or time span.
- adds economic value. It beats the benchmarks hands down.
- is versatile. It works well on a wide variety of financial applications.
- is easily parallelizable. It can be trained on off-the-shelf hardware.

Advanced neural networks provide sustainable and economic competitive edge in today's financial markets.

Figure 0.1: Steps towards advanced neural networks for financial applications.

- Computational requirements should be low, i.e., economic. Especially, computation should not require special high performance computers.
- The modeled applications should not be simple forecasts. They should offer real *competitive edge*.
- I focus on *financial markets*.

Everything in my dissertation is linked to the FAUN neurosimulator, FAUN = Fast Approximation with Universal Neural Networks. Since my supervisor Michael H. Breitner started the FAUN project in 1996 there has been continuous development and improvement. You will find the following highlights:

- The FAUN neurosimulator now also uses fine-grained parallelization. This allows for easily achieved speedups on dual and quad core CPUs. End users are therefore enabled to utilize their workstation to full capacity without having to deal with the increased complexity of message passing software.
- FAUN now also features coarse-grained parallelization using an easy to install grid computing client. Via the web interface it is possible to use clusters of heterogeneous workstations. Spare computing capacity gets reused. Automatic wake up and shutdown saves power costs.
- FAUN is now well-equipped to handle time series problems. It uses a very innovative shared-layer perceptron architecture. I provide a detailed analysis of the computational requirements for the gradient calculation. The gradient calculation itself is presented extensively. Using reverse accumulation and matrix algorithms allows for very efficient computation.
- The examples are designed to provide a maximum of practicality. This includes not only the standard trading application but also market value at risk modeling and transaction decision support.
- The same dataset is used for *different* application. This offers the possibility to benchmark the performance of neural networks or more standard modeling procedures in different domains. The dataset spans 10 years. It includes bear and bull cycles and is not limited to a single up or down trend where most models perform well anyway. The models are very robust and work well without retraining over a period of 8 years.

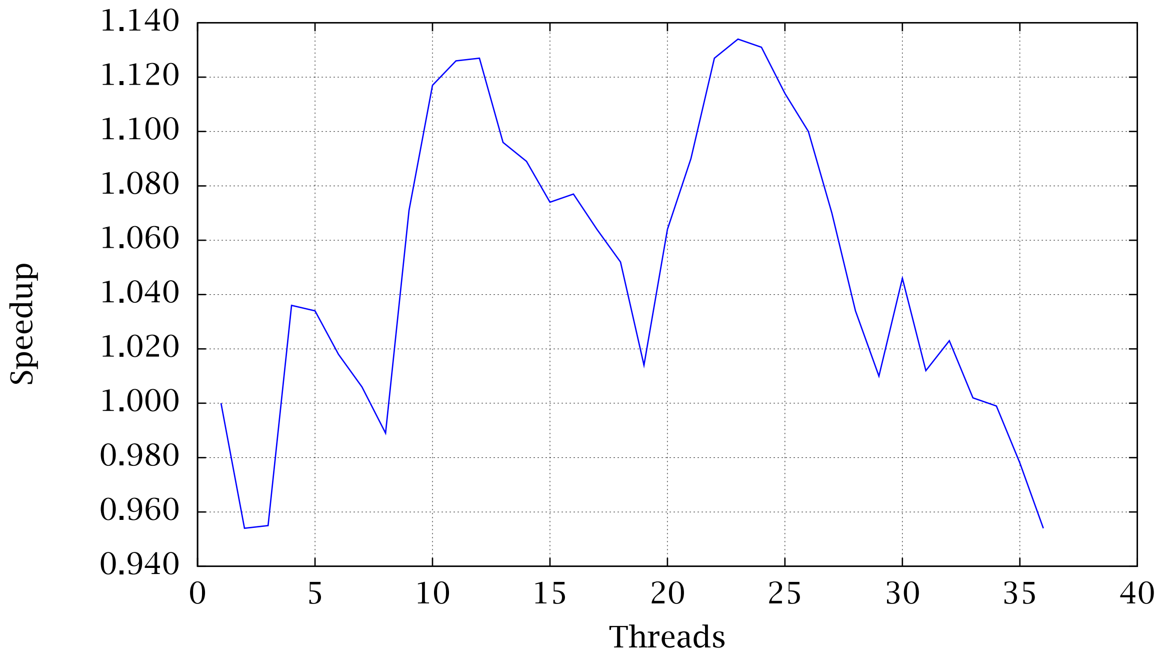


Figure 0.2: The FAUN grid computing client offers consistent speedup above 95 percent on networks of heterogeneous computers.

Grid Computing

Successfully training neural networks is also a matter of having enough computing capacity available. Neural networks are ideally suited for coarse grained parallelization. Communication requirements are low. You can distribute every single neural network to a separate thread. With the FAUN grid computing client spare computing capacity on user workstations is reused. There is no need to install specialized message passing software. The client is self contained. The achievable speedup is above 95 percent, see figure 0.2. This means that 95 percent of *theoretically* available computing power compared to a single thread is used.

It is a waste of energy to leave computers running continually. The FAUN grid computing client allows to wake up and shutdown computers remotely.

The update procedure is simple because all functionality is hosted on the server. An important feature of my client and server is that they are totally platform independent. Working combinations include the last releases of Debian and Ubuntu Linux, Windows 7, Windows Vista and Windows XP. This functionality is normally only implemented in commercial message passing software which necessitates a much more complicated setup.

The Shared Layer Perceptron Topology

The shared layer perceptron provides an elegant method to build *multi asset* and *multi step* models, see figure 0.3 on the following page. It augments the observable states $s_1, \ldots, s_N$ by hidden states $s_{N+1}, \ldots, s_D$. Hidden states allow the model to build up memory. Philosophically the shared layer perceptron acknowledges an incomplete view of the world. You do *not* assume that your «variables» are a perfect description of what happens. Rather you explicitly allow other «hidden» variables to influence your model. Training a shared layer perceptron implicitly also involves finding the right trajectory through the state space: for observable *and* hidden variables.

At each time step the state space is squeezed through the common weight matrix W and the following non linearity. This is an essential difference to standard multi layer perceptrons. Only a *single* weight matrix is used. This reduces the number of free parameters and also training times.

This topology produces at each time step *all* necessary *input* for the *next* time step. This simple mechanisms has two additional advantages. First, you *automatically* get forecasts for *all* your observables. Second, you can *reuse* the forecast at the next time step and produce multi step forecasts.

Financial Applications

My dataset includes 25 financial time series from July 1999 to July 2009, i.e., 10 years of data. The dataset is divided into four asset classes: equity indices, interest rates, currency exchange rates and commodities. Interest rates are generally proxied by using yield curves. This dataset is challenging because it includes the boom and bust of the new economy, the bull market up to the credit crisis of 2007, the subsequent sharp bear market and even a small part of the ongoing recovery. Contrary to other studies this dataset truly represents all market cycles.

The first application models market value at risk. We are interested in the worst expected portfolio value over the next 10 days. Figure 0.4 on page 12 shows a sample forecast for the FTSE 100 index. We are interested in modeling the *worst* returns as closely as possible. It turns out that the shared layer perceptron beats the benchmark historical simulation for *every* asset on a time span of 110 days. It still beats the benchmark *without* retraining on 8 years except for 5 cases. This allows institutions to reduce the margin of safety to an appropriate level.

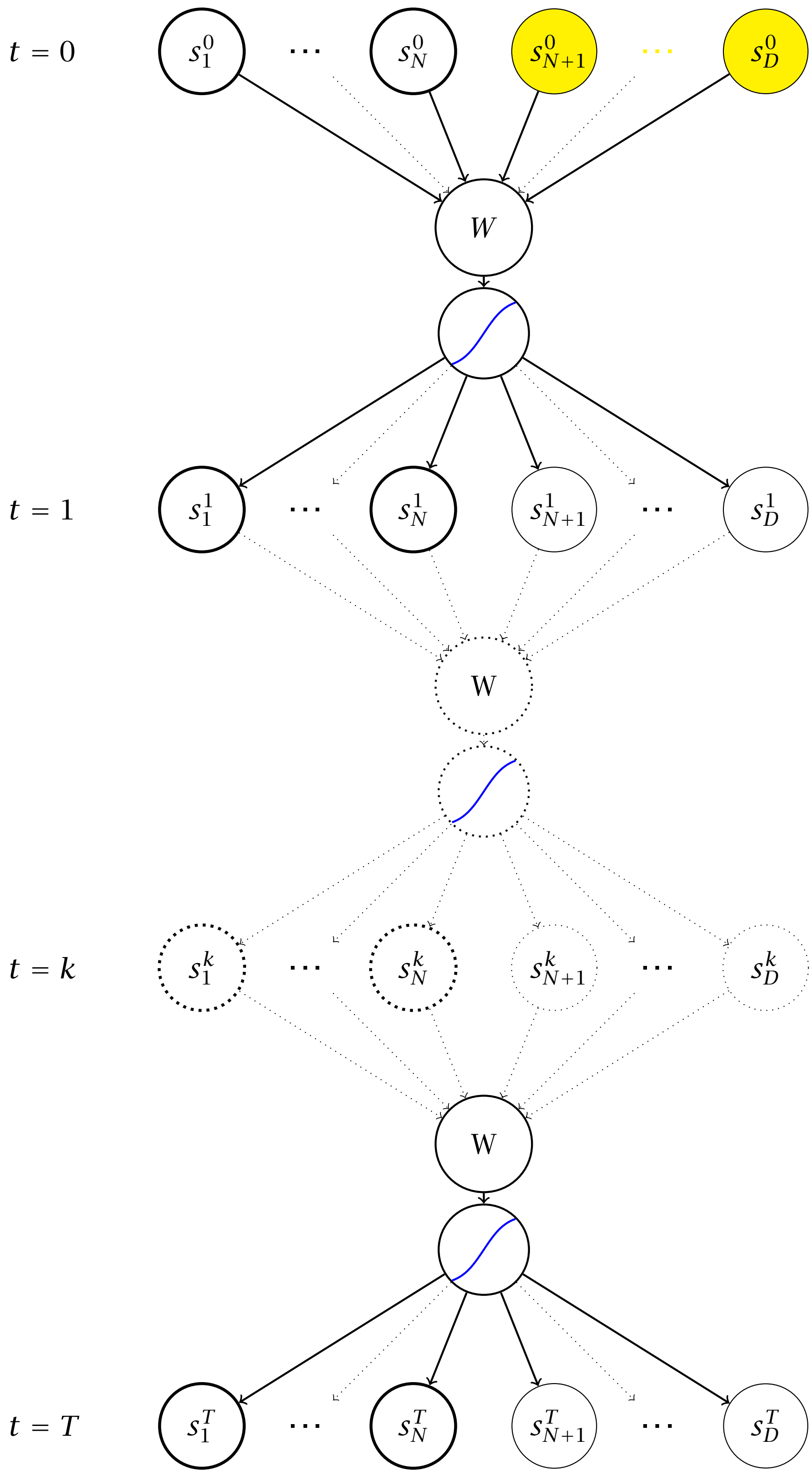


Figure 0.3: The shared layer perceptron for multi asset multi step models.

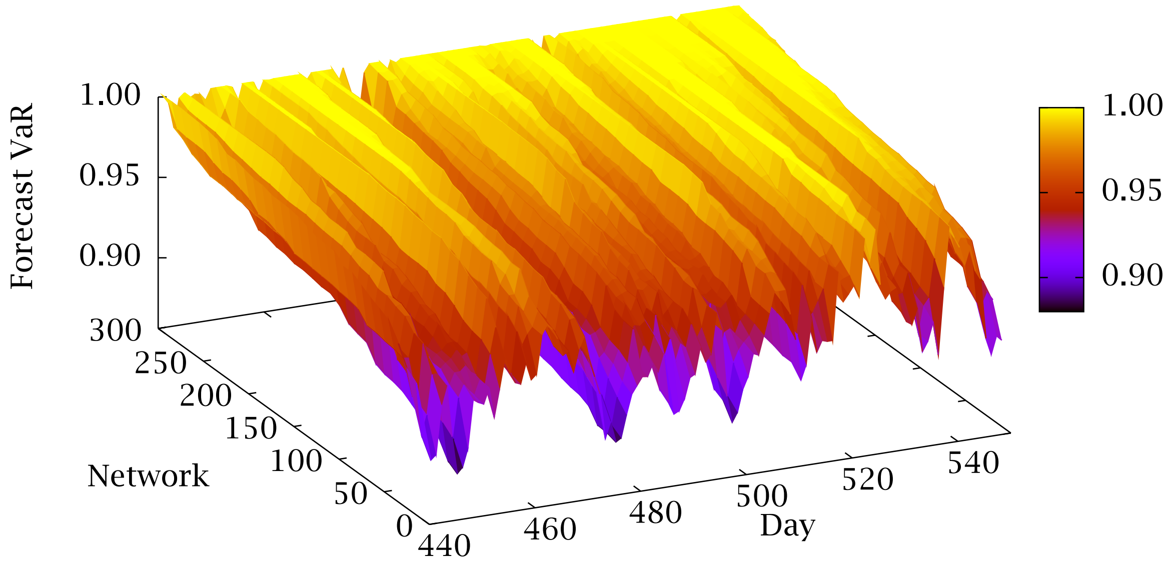


Figure 0.4: The shared layer perceptron topology models the probably worst portfolio value over the next 10 days for the FTSE 100 equity index.

Figure 0.5 on the following page shows a 20 days ahead ensemble forecast for the Baltic Exchange Dry Index. The target is to find an appropriate low entry point within the next month to secure low freight rates. You note that the shared layer perceptron appropriately models the target: first down, then flat, then slightly up again. It does not *exactly* find the lowest price. However, the suggested lowest forecast is a sensible entry point. It is located *before* the index rises again. This models the typical challenge of a corporate treasurer: regular investments on a monthly basis. Again, the shared layer perceptron beats *every* fixed day strategy for *every* asset on 110 days. It is still very successful without retraining on 8 years.

The last application focuses on correctly forecasting the sign of next day returns. I benchmark the shared layer perceptron against a naive strategy and a moving average strategy. The shared layer perceptron performs well or very well across a broad range of assets. Results are especially satisfactory on equities and currencies. It does not always beat the benchmark strategies. But it is at worst second best and shows very consistent returns. The benchmarks, however, show fabulous gains followed by catastrophic losses. The shared layer perceptron works robustly on the shorter and longer time span.

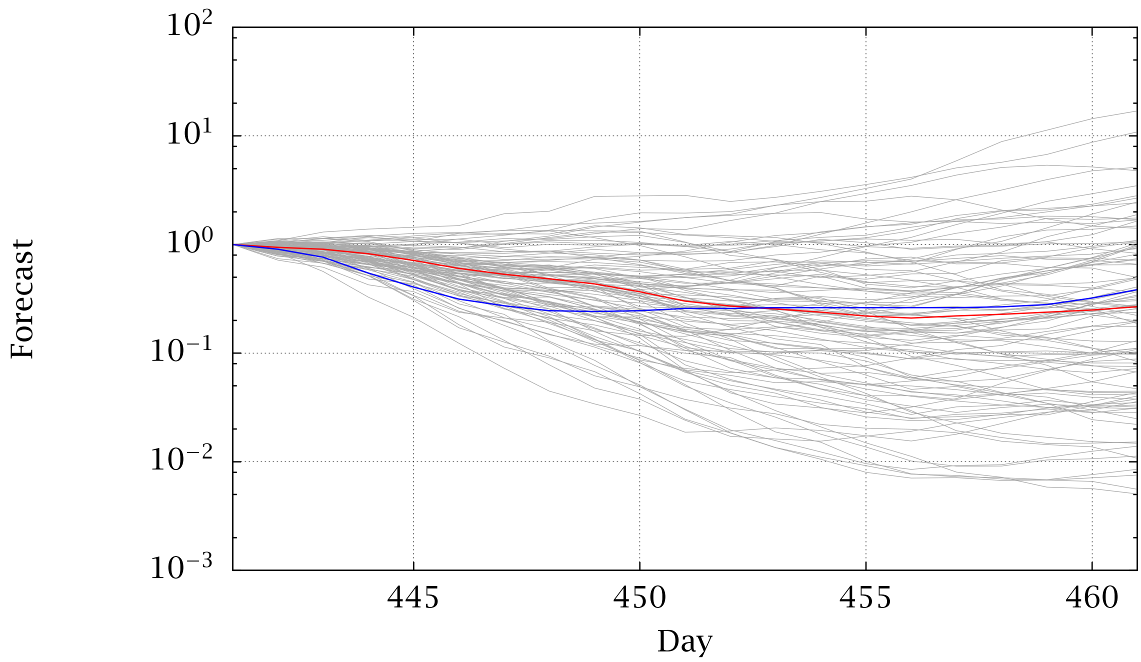


Figure 0.5: The shared layer perceptron forecasts the path of the Baltic Exchange Dry Index over the next 20 days.

Conclusions

I show that the shared layer perceptron is a very robust model. It performs well over different asset classes. It also adapts to different market circumstances and shows consistent performance for long and short time spans without retraining. The shared layer perceptron offers a *unique* way to model a market ensemble:

- The multi step forecasts offer a complete view on the portfolio value path.
- A *single* model is used. With an expert topology you get every percentile of the underlying distribution *for free*.
- You will be more confident to use a model that works well over a broad range of assets. The shared layer perceptron works for *all* inputs *by design*.

Training the networks using coarse grained parallelization and the FAUN grid computing client provides a cost efficient and failsafe path to neural network modeling. Using the client does not require additional setup. The shared layer perceptron topology adds value to financial applications. I recommend it as an important addition in the modeler's and forecaster's toolbox.

Curriculum Vitae and Publications

Contact

Hans-Jörg von Mettenheim
Institut für Wirtschaftsinformatik
Königsworther Platz 1
30167 Hannover, Germany

tel. +49.511.762-4982
email `mettenheim@iwi.uni-hannover.de`

Personal Details

Born July 24th, 1981 in Hannover, Germany
Married in 2007 to Dorothee von Mettenheim
1 son, Johannes-Michel, born 2008
French and German nationality

Career

until July 1999 Leibnizschule Hannover
until July 2003 Studying economics at Leibniz Universität Hannover
until August 2008 Studying mathematics at Leibniz Universität Hannover
August 2003 till present PhD student and stipendiary, Institut für Wirtschaftsinformatik, Leibniz Universität Hannover

Degrees

July 1999 General qualification for university entrance
July 2003 Diploma in economics
August 2008 Diploma in mathematics

Languages

French mother tongue
German father tongue
English fluent
Czech entry level
Latin großes Latinum
Ancient Greek Graecum
Ancient Hebrew Hebraicum

Awards

2006 Best student presentation award for *Dynamic Games with Neurosimulators and Grid Computing: The Game of Two Cars Revisited*, coauthored with Michael H. Breitner, at the 12th International Symposium on Dynamic Games and Applications, July, 3rd–6th, 2006, Sophia Antipolis, France
2009 Award of the Hanover Center of Finance, sponsored by Norddeutsche Landesbank, for the diploma thesis *Derivative Design: Computation and Optimization of the Black-Scholes Equation Solution with a Crank-Nicolson Scheme and SQP-Methods*

Other Qualifications

C Examination Organ playing and choir direction
C additional qualification Jazz, Rock and Pop in the church
Typing entry and advanced level
Stenography entry level

Other Activities

Organist Apostelkirche Hannover
Musical Responsibility Taizé Prayer, Apostelkirche Hannover

Publications

- *Derivative Optimization and Design* in Valerie Belton, Erwin Pesch, Gerhard J. Woeginger (eds.): Proceedings of the 23rd European Conference on Operational Research, July 5–8, 2009, Universität Siegen, Bonn (with Michael H. Breitner)
- *Multi-Objective Optimization for Planning of Central IT Resources with Focus on Green IT* in Valerie Belton, Erwin Pesch, Gerhard J. Woeginger (eds.): Proceedings of the 23rd European Conference on Operational Research, July 5–8, 2009, Universität Siegen, Bonn (with Marc Klages, and Michael H. Breitner)
- *Numerical Solution of the Game of Two Cars with a Neurosimulator and Grid Computing* in Pierre Bernhard, Vladimir Gaitsgory, Odile Pourtallier (eds.): Advances in Dynamic Games and Their Applications: Analytical and Numerical Developments, Birkhäuser, Boston, May 2009 (with Michael H. Breitner)
- *Derivative Design* in Christian Dunis, Michael Dempster, Virginie Terraza (eds.): Proceedings of the 16th International Conference on Forecasting Financial Markets: Advances for Exchange Rates, Interest Rates and Asset Management, May 27–29 2009, Luxembourg (with Michael H. Breitner)
- *Industrialization of Derivative Design: Integrated Risk Management with the Financial Information System WARRANT-PRO-2* in Hans Robert Hansen, Dimitris Karagiannis, Hans-Georg Fill (eds.): Business Services: Konzepte, Technologien, Anwendungen, 9. Internationale Tagung Wirtschaftsinformatik, February 25–27 2009, Vienna, Volume 2, pp. 255–264 (with Michael H. Breitner)
- *Ganzheitliche Disposition von Strecken- und Flächenverkehren durch kombinierten Einsatz modifizierter Operations Research Verfahren* in Proceedings der Fachtagung der GOR AG Logistik und Verkehr, 13./14. 11. 2008, Frankfurt am Main (in German, with Marcus Gerasch, Michael H. Breitner, and Lothar Schulze)
- *Akzeptanz von Sicherheitsmaßnahmen: Modellierung, Numerische Simulation und Optimierung*, IWI Discussion Paper number 28, 30 pages, Institut für Wirtschaftsinformatik, Leibniz Universität Hannover, October 2008 (in German, with Matthias Paul, and Michael H. Breitner)

- *Intelligent Decision Support Systems and Neurosimulators: A Promising Alliance for Financial Services Providers* in H. Österle, J. Schelp, R. Winter: Proceedings of ECIS 2007, St. Gallen, 7th–9th June 2007 (with Michael H. Breitner, Frank Köller, and Simon König)

- *Distributed Neurosimulation* in H.-D. Haasis, H. Kopfer, J. Schönberger (eds.): Operations Research Proceedings 2005, Selected Papers of the Annual International Conference of the German Operations Research Society, Bremen, Springer Verlag, Heidelberg, published September 2006 (with Michael H. Breitner)

- *Dynamic Games with Neurosimulators and Grid Computing: The Game of Two Cars Revisited* in Proceedings of the 12th International ISDG Symposium on Dynamic Games and Applications, July 3–6, 2006, Sophia Antipolis/Riviera (with Michael H. Breitner)

- *Coarse-grained Parallelization of the Advanced Neurosimulator FAUN 1.0 with PVM and the Enhanced Cornered Rat Game Revisited*, International Game Theory Review, Volume 7, September 2005 (with Michael H. Breitner)

- *Neural Network Forecasting with High Performance Computers* in E. P. Hofer, E. Reithmeier (eds.): Proceedings of the 13th International Workshop on Dynamics and Control — Modeling and Control of Autonomous Decision Support Based Systems, Wiesensteig, May 2005 (with Michael H. Breitner)

- *Coarse-grained Parallelization of the Advanced Neurosimulator FAUN 1.0 with PVM* in Michael H. Breitner (ed.): Proceedings of the Fourth International ISDG Workshop, Goslar, May 2003 (with Michael H. Breitner)