

University of Strathclyde
Department of Electronic and Electrical Engineering

Learning to Trade Power

by

Richard W. Lincoln

A thesis presented in fulfilment of the
requirements for the degree of

Doctor of Philosophy

2010

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date: August 13, 2010

Acknowledgements

I wish to thank Professor Jim McDonald for giving me the opportunity for postgraduate study and for permitting me the freedom to pursue my own research interests. I am also very grateful to my supervisors, Professor Graeme Burt and Dr Stuart Galloway, for their guidance and scholarship. I wish to offer special thanks to my parents, my brother and my sister for their support throughout my PhD.

This thesis makes extensive use of open source software projects developed by researchers from other institutions. I wish to thank Dr Ray Zimmerman from Cornell University for his work on optimal power flow, researchers from the Dalle Molle Institute for Artificial Intelligence (IDSIA) and the Technical University of Munich for their work on reinforcement learning algorithms and artificial neural networks and Charles Gieseler from Iowa State University for his implementation of the Roth-Erev reinforcement learning method.

This research was funded by the United Kingdom Engineering and Physical Sciences Research Council through the Supergen Highly Distributed Power Systems consortium under grant GR/T28836/01.

Abstract

In Electrical Power Engineering, learning algorithms can be used to model the strategies of electricity market participants. The objective of this work is to establish if *policy gradient* reinforcement learning methods can provide superior models of market participants than previously applied *value function based* methods.

Supply of electricity involves technology, money, people, natural resources and the environment. All of these aspects are changing and the markets for electricity must be suitably researched to ensure that their designs are fit for purpose. In this thesis electricity markets are modelled as non-linear constrained optimisation problems that are solved using a primal-dual interior point method. Policy gradient reinforcement learning algorithms are used to adjust the parameters of multi-layer feed-forward neural networks that approximate each market participant's policy for selecting quantities of power and prices that are offered in the simulated marketplace.

Traditional reinforcement learning methods that learn a value function have been previously applied in simulated electricity trade, but are largely restricted to discrete representations of a market environment. Policy gradient methods have been proven to offer convergence guarantees in continuous environments, such as in robotic control applications, and avoid many of the problems that mar value function based methods.

The benefits of using policy gradient methods in electricity market simulation are explored and the results demonstrate their superior trading ability when operating in large constrained networks. By advancing the use of learning methods in electricity market simulation, this work provides the opportunity to revisit previous research in the field and creates the possibility for policy gradient methods to be used in decision support and automated energy trading applications.

Contents

Abstract	iv
List of Figures	viii
List of Tables	ix
1 Introduction	1
1.1 Research Motivation	1
1.2 Problem Statement	2
1.3 Research Contributions	3
1.4 Thesis Outline	4
2 Background	6
2.1 Electric Power Supply	6
2.2 Electricity Markets	8
2.2.1 The England and Wales Electricity Pool	10
2.2.2 British Electricity Transmission and Trading Arrangements	12
2.3 Electricity Market Simulation	13
2.3.1 Agent-Based Simulation	14
2.3.2 Optimal Power Flow	15
2.4 Reinforcement Learning	20
2.4.1 Value Function Methods	21
2.4.2 Policy Gradient Methods	24
2.4.3 Roth-Erev Method	26
2.5 Summary	28
3 Related Work	29
3.1 Custom Learning Methods	29
3.1.1 Market Power	29
3.1.2 Financial Transmission Rights	34
3.2 Simulations Applying Q-learning	34
3.2.1 Nash Equilibrium Convergence	34
3.2.2 Congestion Management Techniques	36
3.2.3 Gas-Electricity Market Integration	37
3.2.4 Electricity-Emissions Market Interactions	37
3.2.5 Tacit Collusion	38
3.3 Simulations Applying Roth-Erev	39

3.3.1	Market Power	39
3.3.2	Italian Wholesale Electricity Market	40
3.3.3	Vertically Related Firms and Crossholding	42
3.3.4	Two-Settlement Markets	43
3.4	Policy Gradient Reinforcement Learning	45
3.4.1	Financial Decision Making	45
3.4.2	Grid Computing	46
3.5	Summary	47
4	Modelling Power Trade	49
4.1	Electricity Market Model	49
4.1.1	Unit De-commitment	51
4.1.2	Power Exchange	52
4.1.3	Auction Example	53
4.2	Multi-Agent System	54
4.2.1	Environment	54
4.2.2	Task	57
4.2.3	Agent	58
4.2.4	Simulation Event Sequence	58
4.3	Summary	63
5	Nash Equilibrium Analysis	64
5.1	Introduction	64
5.2	Aims and Objectives	65
5.3	Method of Simulation	65
5.4	Simulation Results	67
5.5	Discussion and Critical Analysis	75
5.6	Summary	76
6	System Constraint Exploitation	77
6.1	Introduction	77
6.2	Aims and Objectives	77
6.3	Method of Simulation	78
6.4	Simulation Results	82
6.5	Discussion and Critical Analysis	82
6.6	Summary	82
7	Conclusions and Further Work	83
7.1	Further Work	83
7.1.1	Alternative Learning Algorithms	83
7.1.2	UK Transmission System	84
7.1.3	AC Optimal Power Flow	86
7.1.4	Multi-Market Simulation	86
7.2	Summary Conclusions	87
	Bibliography	88

A	Open Source Power Engineering Software	95
A.1	MATPOWER	95
A.2	MATDYN	98
A.3	Power System Analysis Toolbox	98
A.4	UWPFLOW	100
A.5	TEFTS	100
A.6	Distribution System Simulator	101
A.7	Agent-based Modelling of Electricity Systems	102
A.8	DCOPFJ	103
A.9	PYLON	103
B	Case Data	105
B.1	6-Bus Case	105
B.2	IEEE Reliability Test System	105

List of Figures

2.1	Basic structure of a three phase AC power system.	7
2.2	UK power station locations.	9
2.3	Pool bid structure.	11
2.4	Piecewise linear active power cost function with constrained cost variable minimisation illustrated.	11
2.5	Nominal- π transmission line model in series with a phase shifting transformer model.	17
2.6	Sequence diagram for the basic reinforcement learning model. . .	21
2.7	Multi-layer feed-forward perceptron with bias nodes.	26
3.1	One-line diagram for a stylised Italian grid model.	41
4.1	Class diagram for the power system model.	50
4.2	Class diagram for the Pyreto.	55
4.3	Class diagram for Pyreto market experiment.	59
4.4	Sequence diagram for action selection process.	60
4.5	Sequence diagram for the reward process.	61
4.6	Sequence diagram for the SARSA learning process.	62
5.1	One line diagram for six bus power system model.	66
5.2	Average markup for agent 1 and standard deviation over 10 runs.	69
5.3	Average markup for agent 2 and standard deviation over 10 runs.	70
5.4	Average markup for agent 1 and standard deviation.	71
5.5	Average markup for agent 2 and standard deviation.	72
5.6	Average reward for agent 1 and standard deviation.	73
5.7	Average reward for agent 2 and standard deviation.	74
6.1	Generator cost functions for the IEEE Reliability Test System . .	79
6.2	Hourly, daily and weekly load profile plots from the IEEE Reliability Test System	80
6.3	IEEE Reliability Test System	81
7.1	UK transmission system.	85
A.1	UKGDS EHV3 model in PSAT Simulink network editor.	99
B.1	One line diagram for six bus power system model.	106

List of Tables

4.1	Example discrete action domain.	56
5.1	Generator cost configuration 1 for 6-bus case.	66
5.2	Generator cost configuration 2 for 6-bus case.	66
5.3	Agent rewards under cost configuration 1	68
5.4	Agent rewards under cost configuration 2	68
6.1	Cost parameters IEEE RTS generator types.	78
6.2	Agent portfolios by type and location.	82
A.1	Open source electric power engineering software feature matrix. . .	96
B.1	6-bus case bus data.	105
B.2	6-bus case generator data.	107
B.3	6-bus case branch data.	107
B.4	IEEE RTS bus data.	108
B.5	IEEE RTS generator data.	109
B.6	IEEE RTS branch data.	110
B.7	IEEE RTS generator cost data.	111

Chapter 1

Introduction

This thesis examines learning algorithms in the domain of electric power trade. In this chapter the motivation for research into electricity trade is explained, the problem under consideration is defined, the principle research contributions of this thesis are stated and an outline the remaining chapters is given.

1.1 Research Motivation

The average total demand for electricity in the United Kingdom (UK) is approximately 45GW and the cost of buying 1MW for one hour is around £40 (Department of Energy and Climate Change, 2009). This equates to yearly transaction values of £16 billion. The value of electricity to societies becomes particularly apparent when supply fails. The New York black-out in August 2003 involved a loss of 61.8GW of power supply to approximately 50 million consumers. The majority of supplies were restored within two days, but the event is estimated to have cost more than \$6 billion (Minkel, 2008; ICF Consulting, 2003).

Quality of life for a person has been shown to be directly proportional to that person's electricity usage (Alam, Bala, Huo, & Matin, 1991). The world population is currently 6.7 billion and forecast to exceed 9 billion by the year 2050 (United Nations, 2003). As people endeavour to improve their quality of life, finite primary energy fuel resources will become increasingly scarce and market mechanisms (e.g. auctions), where the final allocation is based upon the claimants' willingness to pay for the goods, provide a device for efficient allocation of resources in short supply.

Commercialisation of large electricity supply industries began two decades ago in the UK. The inability to store electricity, once generated, in a commercially viable quantity prevents trade as a conventional commodity. Trading mecha-

nisms must allow shortfalls in electric energy to be purchased at short notice from quickly dispatchable generators. Different market structures that facilitate this have been implemented in countries and states around the world. Designed correctly, a competitive electricity market promotes efficiency and drives down costs to the consumer, while design errors can allow market power abuse and elevated market prices.

The value of electricity to societies makes it infeasible to experiment with radical changes to trading arrangements on real systems. A practical alternative is to study abstract mathematical models (with sets of simplifying approximations and assumptions) and, where possible, to find analytical solutions using digital computer programs. Competition is a fundamental component of all markets, but accurate participant strategy models are difficult to create. Reinforcement learning methods can be used to represent the adaptive behaviour of competing players and have been shown to be capable of learning complex strategies (Tesauro, 1994).

1.2 Problem Statement

Individuals participating in an electricity market (be they representing generating companies, load serving entities, firms or traders etc.) must utilise multi-dimensional data, mostly continuous in nature. Certain data, such as demand forecasts, exhibits a degree of uncertainty and other market information, such as the bids of competitors, is hidden.

Traditional reinforcement learning methods associate a value with each available action in a given state. When these values are stored in look-up tables, these methods become restricted by Bellman's Curse of Dimensionality (Bellman, 1961) and can not be applied to complex problems with high-dimensional state and action spaces. When used with function approximation techniques (e.g. artificial neural networks) to allow operation in continuous environments, these methods have been shown to have poor convergence properties, even in simple problems (Gordon, 1995; Baird, 1995; Tsitsiklis & Roy, 1994).

Policy gradient reinforcement learning methods do not attempt to approximate a value function, but use function approximation techniques to represent a policy for selecting actions and search directly in the space of its parameters. They do not suffer from many of the problems that mar value-function based methods in high-dimensional domains. They have strong convergence properties, do not require that all states be continuously visited and work with state and

action spaces that are continuous, discrete or mixed. Policy performance may be degraded by uncertainty in state data, but the learning methods need not be altered. Policy gradient methods have been successfully applied in many operational settings (Sutton, McAllester, Singh, & Mansour, 2000; Peters & Schaal, 2006; Moody & Saffell, 2001; Peshkin & Savova, 2002).

It is proposed in this thesis that agents which learn using policy gradient methods may outperform those using value function based methods in simulated competitive electricity trade. It is further proposed that policy gradient methods may operate better under dynamic electric power system conditions, achieving greater profit by exploiting constraints to their financial benefit. This thesis will use electricity market simulation techniques to compare value function based and policy gradient learning methods and explore these proposals.

1.3 Research Contributions

The research presented in this thesis pertains to the academic fields of Electric Power Engineering, Artificial Intelligence and Economics. The principle contributions made by this thesis in these fields are:

- The first application of policy gradient reinforcement learning methods in simulated energy trade.
- The first application of a non-linear optimal power flow formulation in agent based electricity market simulation.
- A new stateful formulation of the Roth-Erev reinforcement learning method.
- Simulation results which show how policy gradient reinforcement learning methods converge more slowly than value function based methods when learning simple power trade policies.
- Simulation results which show how agents using policy gradient reinforcement learning methods achieve greater profitability than those using value function methods when competing to supply electric power on equal terms.
- An implementation of a multi-agent system for electricity market simulation with discrete and continuous sensor and action space representations.
- The idea of applying Neuro Fitted Q Iteration and $GQ(\lambda)$ in simulations of competitive energy trade.

- The idea of using data from the National Grid seven year statement to simulate the UK electricity market.

The publications that have resulted from this thesis are: Lincoln, Galloway, and Burt (2009).

1.4 Thesis Outline

The presentation of this thesis is organised into nine chapters. Chapter 2 provides background information on electric power supply, wholesale electricity markets and reinforcement learning. It describes how optimal power flow formulations can be used to model electricity markets and defines the reinforcement learning methods that are later examined.

In Chapter 3 the research in this thesis is described in the context of previous work that is related in terms of application field and methodology. Publications on agent based electricity market simulation are reviewed with emphasis on the reinforcement learning methods utilised. Previous applications of policy gradient learning methods in other market settings are also discussed.

Chapter 4 describes the power exchange auction market model and the multi-agent system used to simulated electricity trade. It defines the association of learning agents with portfolios of generators, the process by which offers are submitted and the reward calculation process. Finally, it explains how look-up tables, used with value function based methods, and artificial neural networks, used for policy function approximation, are structured.

An experiment that examines the convergence to a Nash equilibrium of systems of multiple electric power trading agents is reported in Chapter 5. A six bus test case is described and results for four learning algorithms are presented and compared.

Chapter 6 examines the ability of agents to learn policies for exploiting constraints in simulated power systems. The 24 bus model from the IEEE Reliability Test System provides a complex environment with dynamic loading conditions.

The primary conclusions drawn from the results in this thesis are summarised in Chapter 7.2. Shortcomings of the approach are noted and the broad implications are addressed. Some ideas for further work are also outlined. Alternative reinforcement learning methods that could be used in a similar study are listed. A model of the UK transmission system, constructed from data in the National Grid Seven Year Statement, is described and ideas for how the model could be combined with the advances made in this thesis are explained. Opportunities

provided by the use of AC optimal power flow in agent based electricity market simulation are also explored. Finally, the possibilities for using policy gradient methods in multi-market studies that include gas and emissions trade are recognised.

Bibliography

- Alam, M. S., Bala, B. K., Huo, A. M. Z., & Matin, M. A. (1991). A model for the quality of life as a function of electrical energy consumption. Energy, 16(4), 739-745.
- Amerongen, R. van. (1989, May). A general-purpose version of the fast decoupled load flow. Power Systems, IEEE Transactions on, 4(2), 760-770.
- Application of Probability Methods Subcommittee. (1979, November). Ieee reliability test system. Power Apparatus and Systems, IEEE Transactions on, PAS-98(6), 2047-2054.
- Auer, P., Cesa-Bianchi, N., Freund, Y., & Schapire, R. E. (2003). The non-stochastic multiarmed bandit problem. SIAM Journal of Computing, 32(1), 48-77.
- Baird, L. (1995). Residual algorithms: Reinforcement learning with function approximation. In Proceedings of the twelfth international conference on machine learning (p. 30-37). Morgan Kaufmann.
- Bellman, R. E. (1961). Adaptive control processes – A guided tour. Princeton, New Jersey, U.S.A.: Princeton University Press.
- Bhatnagar, S., Sutton, R. S., Ghavamzadeh, M., & Lee, M. (2009). Natural actor-critic algorithms. Automatica, 45(11), 2471–2482.
- Bishop, C. M. (1996). Neural networks for pattern recognition (1st ed.). Oxford University Press, USA. Paperback.
- Bower, J., & Bunn, D. (2001, March). Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the england and wales electricity market. Journal of Economic Dynamics and Control, 25(3-4), 561-592.
- Bower, J., Bunn, D. W., & Wattendrup, C. (2001). A model-based analysis of strategic consolidation in the german electricity industry. Energy Policy, 29(12), 987-1005.
- Bunn, D., & Martoccia, M. (2005). Unilateral and collusive market power in the electricity pool of England and Wales. Energy Economics.

- Bunn, D. W., & Oliveira, F. S. (2003). Evaluating individual market power in electricity markets via agent-based simulation. Annals of Operations Research, 57-77.
- Carpentier, J. (1962, August). Contribution à l'étude du Dispatching Economique. Bulletin de la Society Francaise Electriciens, 3(8), 431-447.
- Department of Energy and Climate Change. (2009). Digest of United Kingdom Energy Statistics 2009. In (chap. 5). National Statistics – Crown.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. The American Economic Review, 88(4), 848-881.
- Ernst, D., Minoia, A., & Ilic, M. (2004, June). Market dynamics driven by the decision-making of both power producers and transmission owners. In Power Engineering Society General Meeting, 2004. IEEE (p. 255-260).
- Fausett, L. (Ed.). (1994). Fundamentals of neural networks: architectures, algorithms, and applications. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
- Gieseler, C. (2005). A Java reinforcement learning module for the Repast toolkit: Facilitating study and implementation with reinforcement learning in social science multi-agent simulations. Unpublished master's thesis, Department of Computer Science, Iowa State University.
- Glimn, A. F., & Stagg, G. W. (1957, april). Automatic calculation of load flows. Power Apparatus and Systems, Part III. Transactions of the American Institute of Electrical Engineers, 76(3), 817-825.
- Goldfarb, D., & Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. Mathematical Programming, 27, 1-33.
- Gordon, G. (1995). Stable function approximation in dynamic programming. In Proceedings of Twelfth International Conference on Machine Learning (p. 261-268). Morgan Kaufmann.
- Grainger, J., & Stevenson, W. (1994). Power system analysis. New York: McGraw-Hill.
- Guo, M., Liu, Y., & Malec, J. (2004, October). A new Q-learning algorithm based on the metropolis criterion. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 34(5), 2140-2143.
- ICF Consulting. (2003, August). The economic cost of the blackout: An issue paper on the northeastern blackout. (Unpublished)
- IEEE Working Group. (1973, November). Common format for exchange of solved load flow data. Power Apparatus and Systems, IEEE Transactions on,

92(6), 1916-1925.

- Kallrath, J., Pardalos, P., Rebennack, S., & Scheidt, M. (2009). Optimization in the energy industry. Springer.
- Kienzle, F., Krause, T., Egli, K., Geidl, M., & Andersson, G. (2007, September). Analysis of strategic behaviour in combined electricity and gas markets using agent-based computational economics. In 1st European workshop on energy market modelling using agent-based computational economics (p. 121-141). Karlsruhe, Germany.
- Kietzmann, T. C., & Riedmiller, M. (2009). The neuro slot car racer: Reinforcement learning in a real world setting. Machine Learning and Applications, Fourth International Conference on, 0, 311-316.
- Kirschen, D. S., & Strbac, G. (2004). Fundamentals of power system economics. Chichester: John Wiley & Sons.
- Krause, T., & Andersson, G. (2006). Evaluating congestion management schemes in liberalized electricity markets using an agent-based simulator. In Power Engineering Society General Meeting, 2006. IEEE.
- Krause, T., Andersson, G., Ernst, D., Beck, E., Cherkaoui, R., & Germond, A. (2004). Nash Equilibria and Reinforcement Learning for Active Decision Maker Modelling in Power Markets. In Proceedings of 6th IAEE European Conference 2004, modelling in energy economics and policy.
- Krause, T., Beck, E. V., Cherkaoui, R., Germond, A., Andersson, G., & Ernst, D. (2006). A comparison of Nash equilibria analysis and agent-based modelling for power markets. International Journal of Electrical Power & Energy Systems, 28(9), 599-607.
- Li, H., & Tesfatsion, L. (2009a, July). The ames wholesale power market test bed: A computational laboratory for research, teaching, and training. In IEEE Proceedings, Power and Energy Society General Meeting. Alberta, Canada.
- Li, H., & Tesfatsion, L. (2009b, March). Capacity withholding in restructured wholesale power markets: An agent-based test bed study. In Power systems conference and exposition, 2009 (p. 1-11).
- Lincoln, R., Galloway, S., & Burt, G. (2009, May). Open source, agent-based energy market simulation with Python. In Proceedings of the 6th International Conference on the European Energy Market, 2009. EEM 2009. (p. 1-5).
- Maei, H. R., & Sutton, R. S. (2010). $G_q(\lambda)$: A general gradient algorithm for temporal-difference prediction learning with eligibility traces. In In proceedings of the third conference on artificial general intelligence. Lugano,

Switzerland.

- McCulloch, W., & Pitts, W. (1943, December 21). A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biology, 5(4), 115-133.
- Micola, A. R., Banal-Estañol, A., & Bunn, D. W. (2008, August). Incentives and coordination in vertically related energy markets. Journal of Economic Behavior & Organization, 67(2), 381-393.
- Micola, A. R., & Bunn, D. W. (2008). Crossholdings, concentration and information in capacity-constrained sealed bid-offer auctions. Journal of Economic Behavior & Organization, 66(3-4), 748-766.
- Minkel, J. R. (2008, August 13). The 2003 northeast blackout—five years later. Scientific American.
- Momoh, J., Adapa, R., & El-Hawary, M. (1999, Feb). A review of selected optimal power flow literature to 1993. I. Nonlinear and quadratic programming approaches. Power Systems, IEEE Transactions on, 14(1), 96-104.
- Momoh, J., El-Hawary, M., & Adapa, R. (1999, Feb). A review of selected optimal power flow literature to 1993. II. Newton, linear programming and interior point methods. Power Systems, IEEE Transactions on, 14(1), 105-111.
- Moody, J., & Saffell, M. (2001, July). Learning to trade via direct reinforcement. IEEE Transactions on Neural Networks, 12(4), 875-889.
- Moody, J., Wu, L., Liao, Y., & Saffell, M. (1998). Performance functions and reinforcement learning for trading systems and portfolios. Journal of Forecasting, 17, 441-470.
- Naghibi-Sistani, M., Akbarzadeh-Tootoonchi, M., Javidi-D.B., M., & Rajabi-Mashhadi, H. (2006, November). Q-adjusted annealing for Q-learning of bid selection in market-based multisource power systems. Generation, Transmission and Distribution, IEE Proceedings, 153(6), 653-660.
- National Electricity Transmission System Operator. (2010, May). 2010 National Electricity Transmission System Seven Year Statement (Tech. Rep.). National Grid Electricity Transmission plc.
- Nicolaisen, J., Petrov, V., & Tesfatsion, L. (2002, August). Market power and efficiency in a computational electricity market with discriminatory double-auction pricing. Evolutionary Computation, IEEE Transactions on, 5(5), 504-523.
- Nicolaisen, J., Smith, M., Petrov, V., & Tesfatsion, L. (2000). Concentration and capacity effects on electricity market power. In Evolutionary Computation. Proceedings of the 2000 Congress on (Vol. 2, p. 1041-1047).

- Overbye, T., Cheng, X., & Sun, Y. (2004, Jan.). A comparison of the AC and DC power flow models for LMP calculations. In System sciences, 2004. Proceedings of the 37th annual hawaii international conference on (p. 9-).
- Peshkin, L., & Savova, V. (2002). Reinforcement learning for adaptive routing. In Neural Networks, 2002. IJCNN 2002. Proceedings of the 2002 International Joint Conference on (Vol. 2, p. 1825-1830).
- Peters, J., & Schaal, S. (2006, October). Policy gradient methods for robotics. In Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on (p. 2219-2225).
- Peters, J., & Schaal, S. (2008). Natural actor-critic. Neurocomputing, 71(7-9), 1180-1190.
- Rastegar, M. A., Guerri, E., & Cincotti, S. (2009, May). Agent-based model of the Italian wholesale electricity market. In Energy Market, 2009. 6th International Conference on the European (p. 1-7).
- Riedmiller, M. (2005). Neural fitted Q iteration - first experiences with a data efficient neural reinforcement learning method. In In 16th european conference on machine learning (pp. 317-328). Springer.
- Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: the rprop algorithm.
- Robbins, H. (1952). Some aspects of the sequential design of experiments. Bulletin American Mathematical Society, 58(5), 527-535.
- Roth, A. E., Erev, I., Fudenberg, D., Kagel, J., Emilie, J., & Xing, R. X. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior, 8(1), 164-212.
- Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., et al. (2010). PyBrain. Journal of Machine Learning Research, 11, 743-746.
- Schweppe, F., Caramanis, M., Tabors, R., & Bohn, R. (1988). Spot pricing of electricity. Dordrecht: Kluwer Academic Publishers Group.
- Sharpe, W. F. (1966, January). Mutual fund performance. Journal of Business, 119-138.
- Sharpe, W. F. (1994). The Sharpe ratio. The Journal of Portfolio Management, 49-58.
- Stott, B., & Alsac, O. (1974, May). Fast decoupled load flow. Power Apparatus and Systems, IEEE Transactions on, 93(3), 859-869.
- Sun, J., & Tesfatsion, L. (2007a). Dynamic testing of wholesale power market designs: An open-source agent-based framework. Computational Economics,

- 30(3), 291-327.
- Sun, J., & Tesfatsion, L. (2007b, June). Open-source software for power industry research, teaching, and training: A DC-OPF illustration. In Power Engineering Society General Meeting, 2007. IEEE (p. 1-6).
- Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. MIT Press. Gebundene Ausgabe.
- Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y. (2000). Policy gradient methods for reinforcement learning with function approximation. In Advances in neural information processing systems (Vol. 12, p. 1057-1063).
- Tellidou, A., & Bakirtzis, A. (2007, November). Agent-based analysis of capacity withholding and tacit collusion in electricity markets. Power Systems, IEEE Transactions on, 22(4), 1735-1742.
- Tesauro, G. (1994). Td-gammon, a self-teaching backgammon program, achieves master-level play. Neural Computation, 6(2), 215-219.
- Tesfatsion, L., & Judd, K. L. (2006). Handbook of computational economics, volume 2: Agent-based computational economics (handbook of computational economics). Amsterdam, The Netherlands: North-Holland Publishing Co.
- Tinney, W., & Hart, C. (1967, November). Power flow solution by Newton's method. Power Apparatus and Systems, IEEE Transactions on, 86(11), 1449-1460.
- Tsitsiklis, J. N., & Roy, B. V. (1994). Feature-based methods for large scale dynamic programming. In Machine learning (p. 59-94).
- United Nations. (2003, December 9). World population in 2300. In Proceedings of the United Nations, Expert Meeting on World Population in 2300.
- U.S.-Canada Power System Outage Task Force. (2004, April). Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations (Tech. Rep.). North American Electric Reliability Corporation.
- Veit, D., Weidlich, A., Yao, J., & Oren, S. (2006). Simulating the dynamics in two-settlement electricity markets via an agent-based approach. International Journal of Management Science and Engineering Management, 1(2), 83-97.
- Vengero, D. (2008). A gradient-based reinforcement learning approach to dynamic pricing in partially-observable environments. Future Generation Computer Systems, 24(7), 687-693.
- Visudhiphan, P. (2003). An agent-based approach to modeling electricity spot markets. Unpublished doctoral dissertation, Department of Electrical En-

- gineering and Computer Science, Massachusetts Institute of Technology.
- Visudhiphan, P., & Ilic, M. (1999, February). Dynamic games-based modeling of electricity markets. In Power Engineering Society 1999 Winter Meeting, IEEE (Vol. 1, p. 274-281).
- Wang, H., Murillo-Sanchez, C., Zimmerman, R., & Thomas, R. (2007, Aug.). On computational issues of market-based optimal power flow. Power Systems, IEEE Transactions on, 22(3), 1185-1193.
- Wang, J., Koritarov, V., & Kim, J.-H. (2009, July). An agent-based approach to modeling interactions between emission market and electricity market. In Power Energy Society General Meeting, 2009. PES 2009. IEEE (p. 1-8).
- Weidlich, A., & Veit, D. (2006, July 7-10). Bidding in interrelated day-ahead electricity markets - insights from an agent-based simulation model. In Proceedings of the 29th IAAE International Conference.
- Weidlich, A., & Veit, D. (2008, July). A critical survey of agent-based wholesale electricity market models. Energy Economics, 30(4), 1728-1759.
- Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. In Machine learning (p. 229-256).
- Wood, A. J., & Wollenberg, B. F. (1996). Power Generation Operation and Control (second ed.). New York: Wiley, New York.
- Yao, J., Adler, I., & Oren, S. S. (2008). Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network. Operations Research, 56(1), 34-47.
- Yao, J., Oren, S. S., & Adler, I. (2007). Two-settlement electricity markets with price caps and cournot generation firms. European Journal of Operational Research, 181(3), 1279-1296.
- Zimmerman, R. (2010, March 19). MATPOWER 4.0b2 User's Manual (Version 4.0b2 ed.) [Computer software manual]. School of Electrical Engineering, Cornell University, Ithaca, NY 14853.
- Zimmerman, R., Murillo-Sánchez, C., & Thomas, R. J. (2009, July). MATPOWER's extensible optimal power flow architecture. In IEEE PES General Meeting. Calgary, Alberta, Canada.