From Time Series to Market Modeling

Hans-Georg Zimmermann and Anton Schäfer

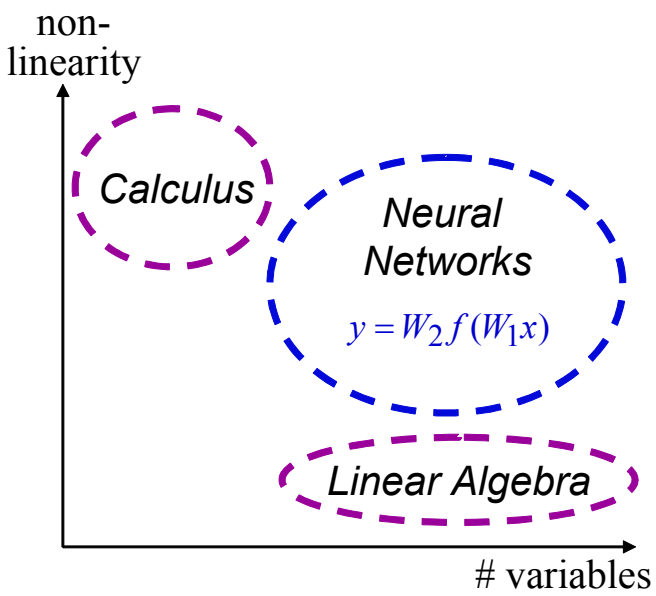
Siemens AG
Corporate Technology (CT IC 4)
University of Ulm
Department of Optimization and Operations Research

Email : Hans_Georg.Zimmermann@siemens.com

Corporate Technology: About 1,800 Researchers and Developers Worldwide …

Neural Networks in Nonlinear Regression

Basics on Neural Networks

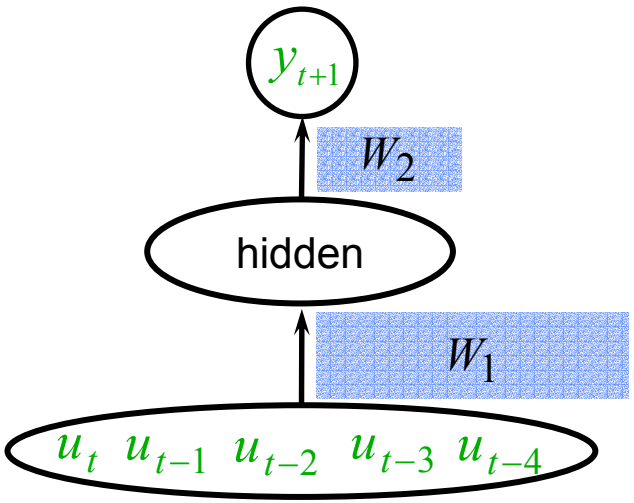


Basics on Forecasting

Dependent on past & present data we quest for a function modeling the shift to the future.

$$y_{t+1} = h(u_t, u_{t-1}, \ldots)$$

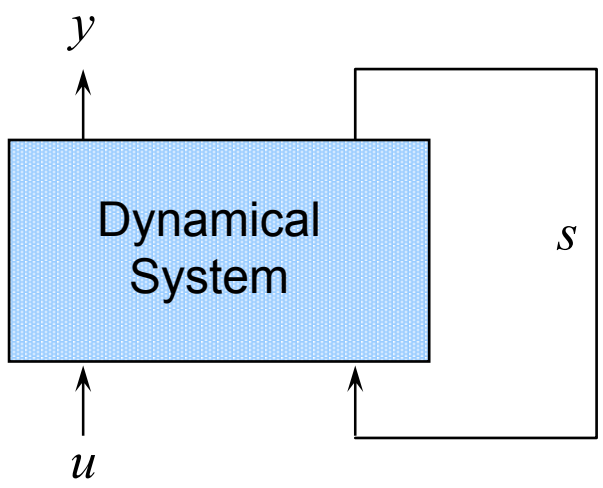
$$\sum_{t=1}^{T} \left(y_t - y_t^d\right)^2 \rightarrow \min_h$$



Existence Theorem:
(Hornik, Stinchcombe, White 1989)

A 3-layer network can approximate any continuous function on a compact domain.

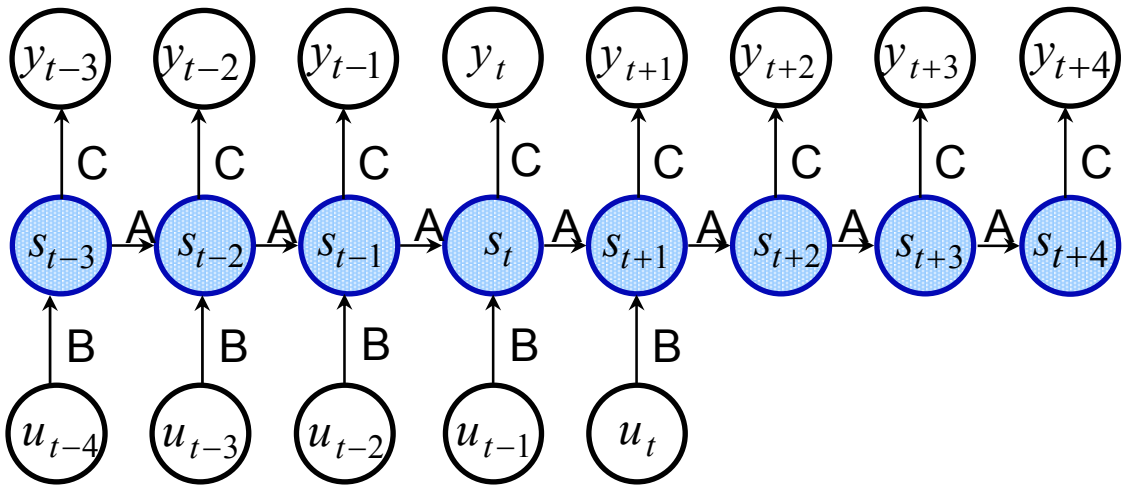
A Short Course in Recurrent Neural Network Modeling



$$s_{t+1} = f(s_t, u_t)$$ state transition

$$y_t = g(s_t)$$ output equation

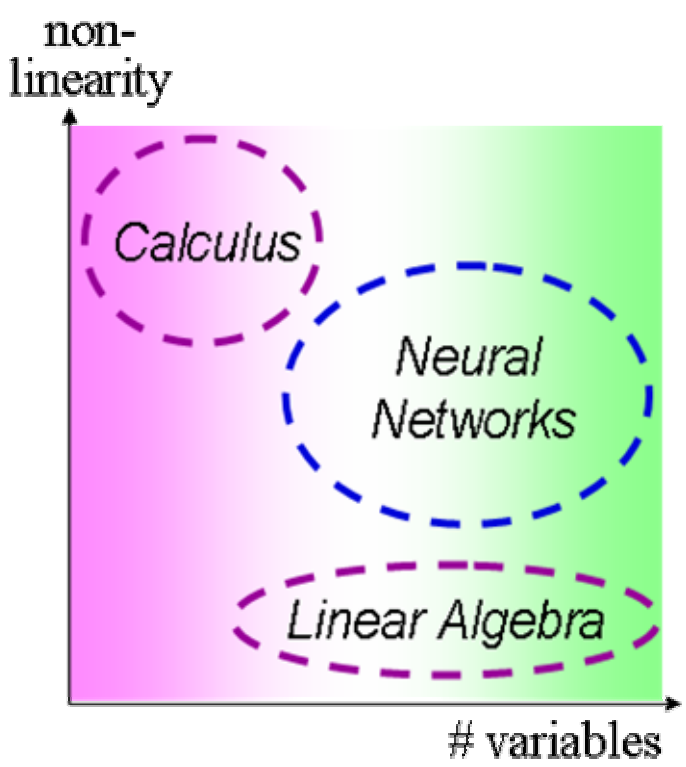
$$\frac{1}{T}\sum_{t=1}^{T}\left(y_t - y_t^d\right)^2 \rightarrow \min_{f,g}$$ identification



Unfolding in time allows the learning of memory dependent systems.

Overshooting enforces the autonomous part of the dynamics.

From Small to Large Neural Networks



What we have done so far:

- Feedforward Neural Networks
- Recurrent Networks (dim(s) < 15)

What we have missed so far:

- Time series analysis is an approach to model a single variable. More natural is a modeling of <u>all (interrelated) system variables</u>.
- Modeling of interacting dynamical systems is only feasible with large networks.
- Simple enlargement of small neural networks fails because of e.g. overfitting.

What we have found so far:

There exist large networks without counterpart in the set of small networks **<u>and</u>** which have new features.

From Standard to Normalized Recurrent Neural Networks

In the standard RNN the dynamics is described by 3 matrices (A, B, C):

$$\tau \le t: \quad s_{\tau+1} = \tanh(\boldsymbol{A} s_\tau + c + B u_\tau)$$

$$t < \tau: \quad s_{\tau+1} = \tanh(\boldsymbol{A} s_\tau + c)$$

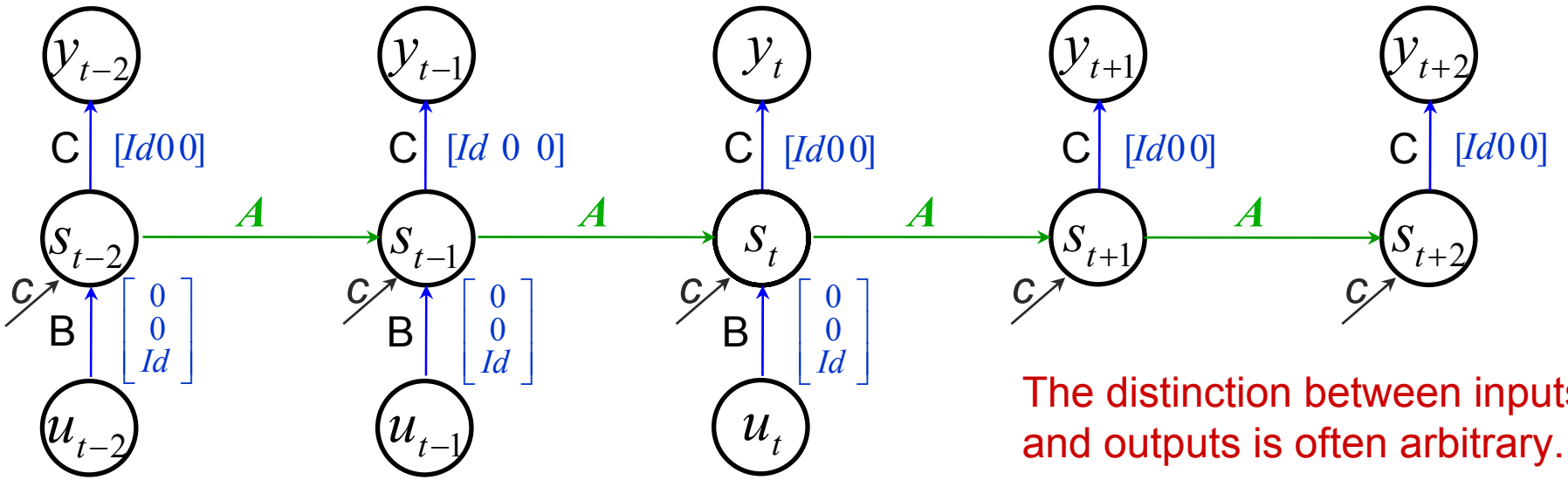
$$y_\tau = C s_\tau$$

W.l.o.g. a dynamics can & should be modeled by only one matrix A:

$$\tau \le t: \quad s_\tau = \tanh\left(\boldsymbol{A}\, s_{\tau-1} + c + \begin{bmatrix} 0 \\ 0 \\ \mathrm{Id} \end{bmatrix} u_\tau \right)$$

$$\tau > t: \quad s_\tau = \tanh\left(\boldsymbol{A}\, s_{\tau-1} + c \right)$$

$$y_\tau = \begin{bmatrix} \mathrm{Id} & 0 & 0 \end{bmatrix} s_\tau$$



The distinction between inputs and outputs is often arbitrary.

Modeling the Dynamics of Observables

Inputs and targets are merged to observables (y_t).

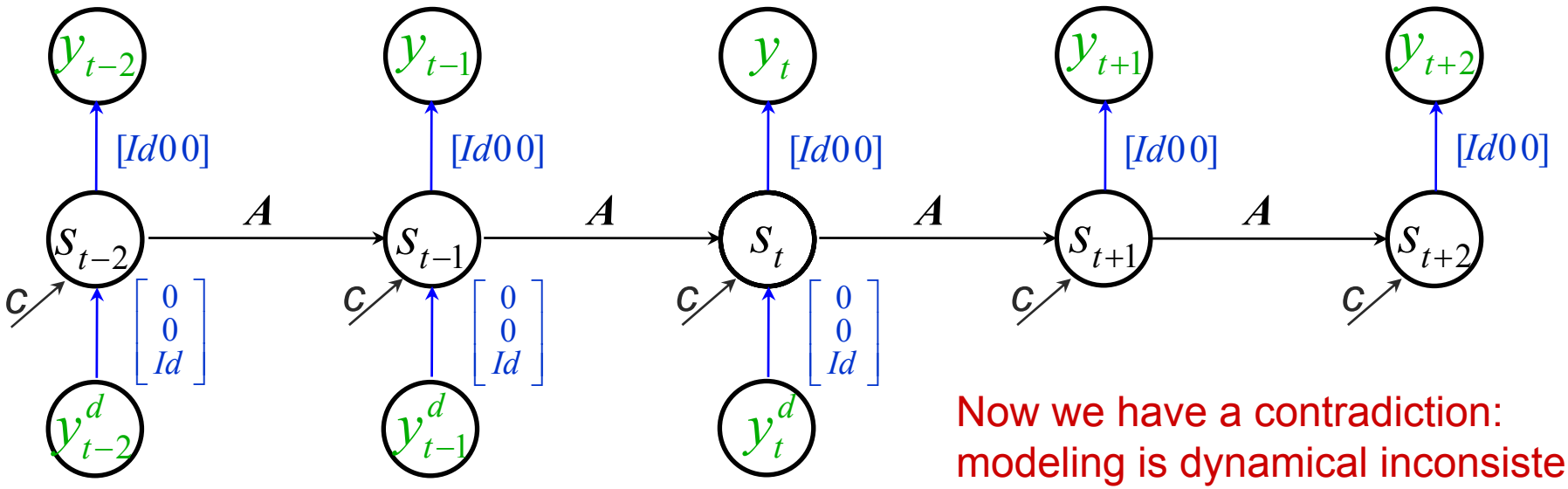
Normalized networks allow the required high-dim. modeling.

Due to the network design, the input to output relation is delayed.

$$\tau \le t: \quad s_\tau = \tanh\left(As_{\tau-1} + c + \begin{bmatrix} 0 \\ 0 \\ \mathrm{Id} \end{bmatrix} y_\tau^d\right)$$

$$\tau > t: \quad s_\tau = \tanh(As_{\tau-1} + c)$$

$$y_\tau = \begin{bmatrix} \mathrm{Id} & 0 & 0 \end{bmatrix} s_\tau$$



Now we have a contradiction: modeling is dynamical inconsistent.

Dynamical Consistent Neural Networks

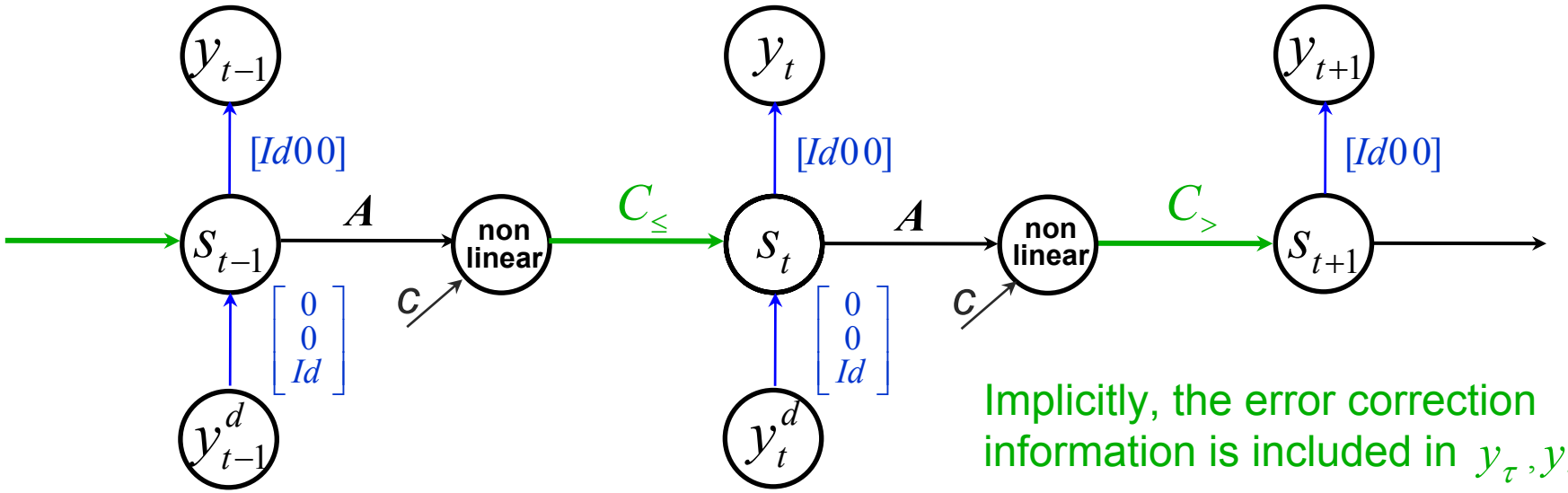
Missing future observations can be substituted by the models own expectations.

$$s_\tau = \begin{pmatrix} y_\tau \\ h_\tau \\ \left\{ \begin{matrix} y_\tau^d \\ y_\tau \end{matrix} \right\} \end{pmatrix} = \begin{pmatrix} \text{expectations} \\ \text{hidden states} \\ \left\{ \begin{matrix} \tau \le t: \text{ observations} \\ \tau > t: \text{ expectations} \end{matrix} \right\} \end{pmatrix}$$

$$\tau \le t: \quad s_\tau = \begin{bmatrix} \mathrm{Id} & 0 & 0 \\ 0 & \mathrm{Id} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tanh\left(\boldsymbol{A}\, s_{\tau-1} + c \right) + \begin{bmatrix} 0 \\ 0 \\ \mathrm{Id} \end{bmatrix} y_\tau^d$$

$$\tau > t: \quad s_\tau = \begin{bmatrix} \mathrm{Id} & 0 & 0 \\ 0 & \mathrm{Id} & 0 \\ \mathrm{Id} & 0 & 0 \end{bmatrix} \tanh\left(\boldsymbol{A}\, s_{\tau-1} + c \right)$$

$$y_\tau = \begin{bmatrix} \mathrm{Id} & 0 & 0 \end{bmatrix} s_\tau$$



Implicitly, the error correction information is included in y_τ, y_τ^d

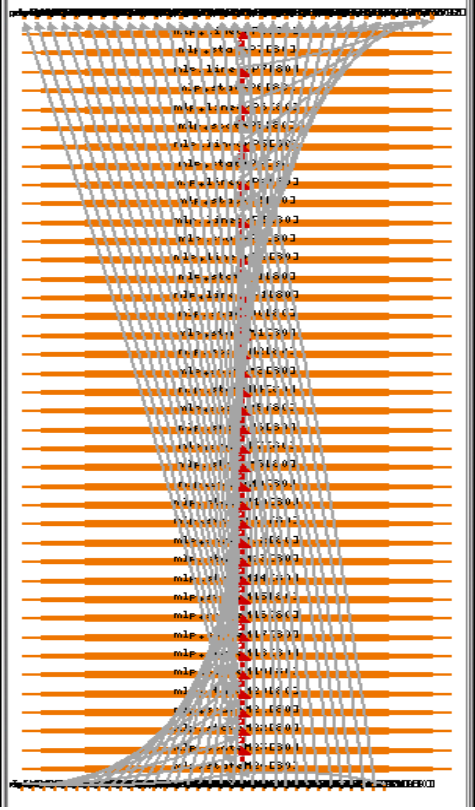
Load Forecasting by DCNN

NeuroSimulator <5>

Project View Actions Help

Energy Load Forecast dependent on load, weather and energy prices

(15 min. time grid, state dimension = 80)



NormSum 1

relative forecast error (15 min.)

Epochs

419 0.00534776

NormSum 2

relative forecast error (30 min.)

Epochs

418 0.00926424
419 0.00926424

NormSum

relative forecast error (45 min.)

Epochs

418 0.0131443
419 0.0131443

NormSum 3

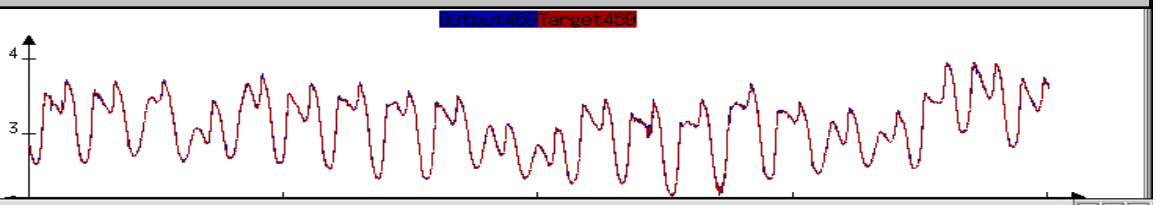
relative forecast error (60 min.)

Epochs

418 0.0168802
419 0.0168802

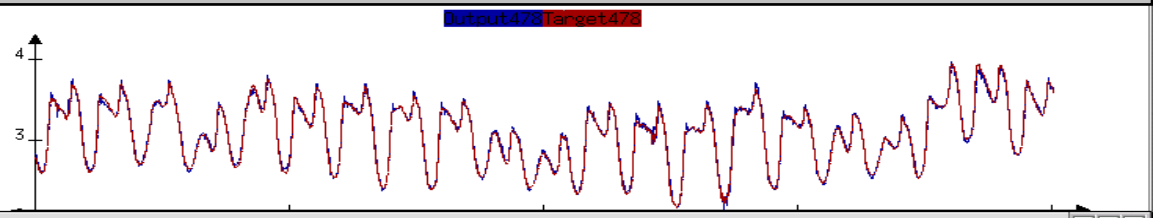
Comparison <17>

Testperiod: 15 min. Load Forecast: Model Output vs. Target



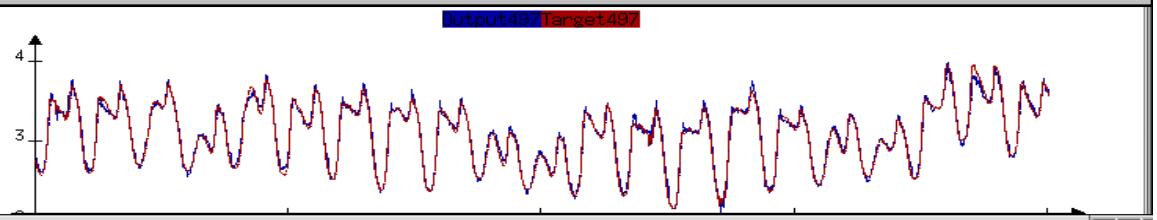
Comparison <18>

Testperiod: 30 min. Load Forecast: Model Output vs. Target



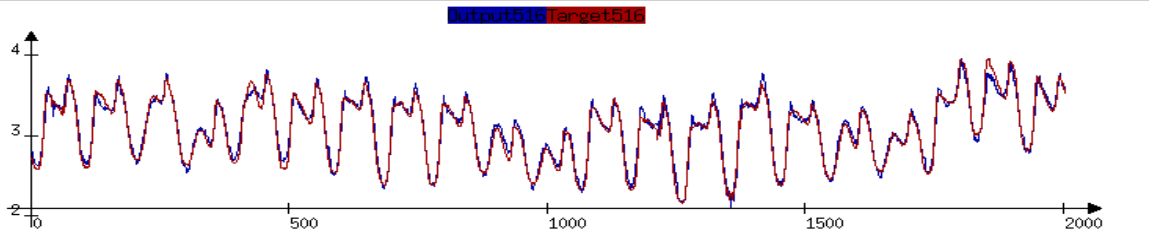
Comparison <20>

Testperiod: 45 min. Load Forecast: Model Output vs. Target

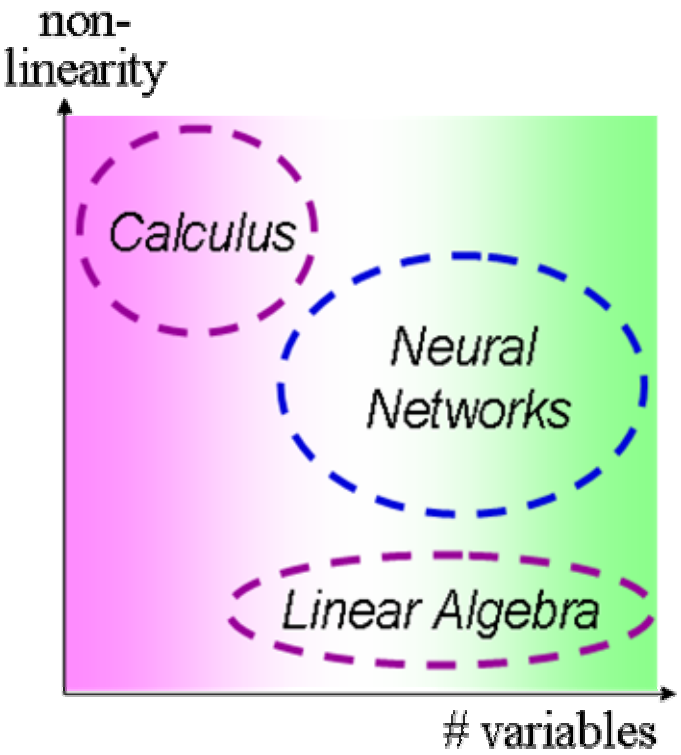


Comparison <19>

Testperiod: 60 min. Load Forecast: Model Output vs. Target



A Duality between Small and Large Networks



Small Recurrent Networks:

- Data over constrains the degrees of freedom. This results in a unique regression model.
- The lack of randomness causes difficulties with local minima.

Medium Recurrent Networks:

- Data is not sufficient to estimate a unique model.
- We have a set of solutions. – Limited randomness does not average out in generalization.

Large Recurrent Networks:

- Data only constrains the eigendynamics of the large recurrent network.
- The eigenactivity of the network can be seen as a self-created noise, which is a regularization.

Selected References I

1. *Zimmermann, H.G.:* ***Neuronale Netze als Entscheidungskalkül,*** In: Neuronale Netze in der Ökonomie; Eds.: Rehkugler, H., Zimmermann, H.G., p. 1-87; Franz Vahlen; 1994; ISBN 3-8006-1871-0
2. *Neuneier, R.; Zimmermann, H.G.*: ***How to Train Neural Networks,*** In: Neural Networks: Tricks of the Trade; Eds.: Orr G. B., Müller K. R.; p. 373-423; Springer, 1998; ISBN 3-540-65311-2
3. *Zimmermann, H.G.; Neuneier, R.*: ***Neural Network Architectures for the Modeling of Dynamical Systems,*** In: A Field Guide to Dynamical Rec. Networks; Eds.: Kolen J. F., Kremer S. C.; IEEE Press; 2000; ISBN 0-780-35369-2
4. *Zimmermann, H.G.; Neuneier, R.; Grothmann, R.*: ***Modeling of the German Yield Curve by Error Correction Neural Networks***. In: Leung, K., Chan, L.-W. and Meng, H., Eds. Proceedings IDEAL 2000, p. 262-7, Hong Kong.
5. *Zimmermann, H.G.; Neuneier, R.; Grothmann, R.*: **Active Portfolio Management based on Error Correction Neural Networks,** In: Proceedings of Neural Information Processing (NIPS), Vancouver, Canada, Dec. 2001**.**
6. *Zimmermann, H.G.; Neuneier, R.; Grothmann, R.*: **Modeling of Dynamical Systems by Error Correction Neural Networks,** In: Modeling and Forecasting Financial Data, Techniques of Nonlinear Dynamics; Eds. Soofi, A. and Cao, L., Kluwer Academic Pub., 2002; ISBN 0-792-37680-3.
7. *Zimmermann, H.G.; Neuneier, R.; Grothmann, R.:* ***Undershooting: Modeling Dynamical Systems by Time Grid Refinements***, In: Proceedings of European Symposium on Artificial Neural Networks (ESANN) 2002, Belgium.
8. *Zimmermann, H.G.; Tietz, Ch.; Grothmann, R.*: **Yield Curve Forecasting by Error Correction Neural Networks & Partial Learning,** In: Proceedings of European Symposium on Artificial Neural Networks (ESANN) 2002, Belgium.
9. *Zimmermann, H.G.; Mueller A.; Erdem C. and Hoffmann R.*: **Prosody Generation by Causal-Retro-Causal Error Correction Neural Networks**. In: Proceedings Workshop on Multi-Lingual Speech Communication, ATR, 2000.
10. *Siekmann, S.; Neuneier, R.; Zimmermann, H.G.; Kruse R.*: ***Neuro Fuzzy Systems for Data Analysis,*** In: Computing with Words in Information / Intelligent Syst. 2; Eds.: Zadeh L. A., Kacprzyk J.;p. 35-74; Physica Verlag; 1999
11. *Zimmermann, H.G.; Neuneier, R.; Grothmann, R.:* ***Multi Agent Market Modeling of FX-Rates*** In: Advances in Complex Systems (ACS); Special Issue Vol. 4, No. 1; Hermes Science Publ. (Oxford) 2001.
12. *Zimmermann, H.G.; Neuneier, R.; Grothmann, R.*: **An Approach of Multi-Agent FX-Market Modelling based on Cognitive Systems,** In: Proc. of the Int. Conference on Artificial Neural Networks (ICANN); Springer, Vienna 2001.

Selected References II

13. *Zimmermann, H.G.; Neuneier, R.*, Tietz, Ch. and Grothmann, R.: ***Market Modeling based on Cognitive Agents,*** In: Proceedings of the Int. Conference on Artificial Neural Networks (ICANN), Madrid, 2002.
14. *Zimmermann, H.G. and Neuneier, R.:* ***The Observer – Observation Dilemma in Neuro-Forecasting***, In: Advances in Neural Information Processing Systems, Vol. 10, MIT Press, 1998.
15. *Zimmermann, H.G. and Neuneier, R.:* ***Combining State Space Reconstruction and Forecasting by Neural Networks.*** In: G. Bol, G. Nakhaeizadeh and K. H. Vollmer (Eds.): Datamining und Computational Finance, 7. Karlsruher Ökonometrie-Workshop, Wissenschaftliche Beiträge 174, Physica-Verlag, 1999, pp. 259-267.
16. *Zimmermann, H.G. and Neuneier, R.:* ***Modeling Dynamical Systems by Recurrent Neural Networks***, in: Ebecken, N. and Brebbia, C.A. (Eds.) Data Mining II, WIT Press, 2000, pp. 557-566.
17. *Zimmermann, H.G., Neuneier, R. and Grothmann, R.:* ***Multi-Agent Modeling of Multiple FX-Markets by Neural Networks****. IEEE Transactions on Neural Networks Special Issue 12(4):1-9, 2001*.
18. *Zimmermann, H.G.:* ***Optimal Asset Allocation for a Large Number of Investment Opportunities.*** In: Proceedings Forecasting Financial Markets (FFM), London 2002.
19. *Zimmermann, H.G., Grothmann, R. and Tietz Ch.:* ***ATM Cash Management based on Error Correction Neural Networks***, in: Fichtner, W.; Geldermann, J.: Einsatz von OR-Verfahren zur techno-ökonomischen Analyse von Produktionssystemen. Verlag Peter Lang, Frankfurt, 2003
20. *Grothmann, R.:* ***Multi-Agent Market Modeling based on Neural Networks***, Ph.D. Thesis, University of Bremen, Germany, E-Lib, 2002, http://elib.suub.uni-bremen.de/publications/dissertations/E-Diss437_grothmann.pdf
21. *Zimmermann, H.G, Grothmann, R., Schäfer, A.M., Tietz, Ch.:* ***Identification and Forecasting of Large Dynamical Systems by Dynamical Consistent Neural Networks****,* in: Haykin, S.; Principe, J.; Sejnowski, T. and Mc Whirter, J. [Eds.]: New Directions in Statistical Signal Processing: From Systems to Brain, MIT Press, forthcoming 2005.
22. *Zimmermann, H.G, Grothmann, R., Schäfer, A.M., Tietz, Ch.:* ***Dynamical Consistent Recurrent Neural Networks***, in: Prokhorov, D.: Proc. of the Int. Joint Conference on Neural Networks (IJCNN), Montreal 2005.
23. *Tietz, Ch.:* ***SENN User Manual****,* Siemens AG, 2004*.*