Causal-Retro-Causal Systems induce a Dynamics on a Manifold

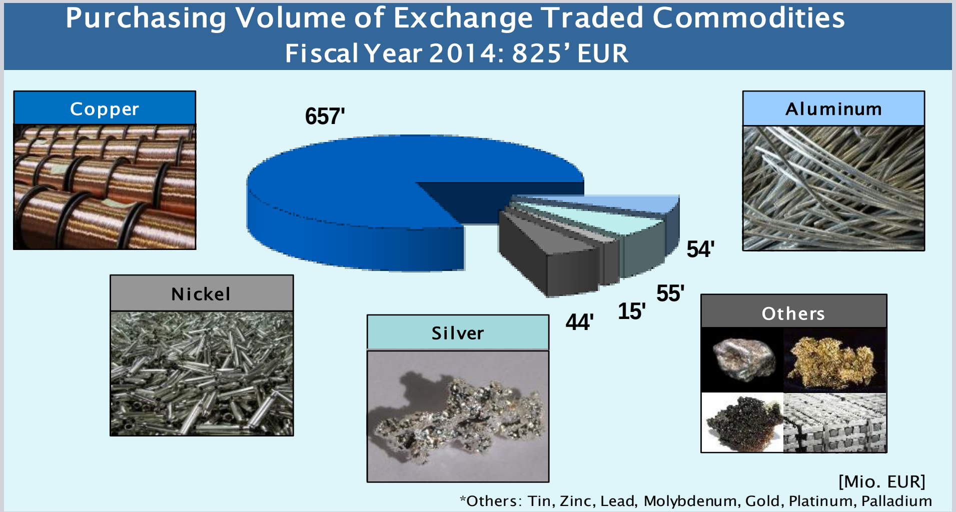
Hans Georg Zimmermann

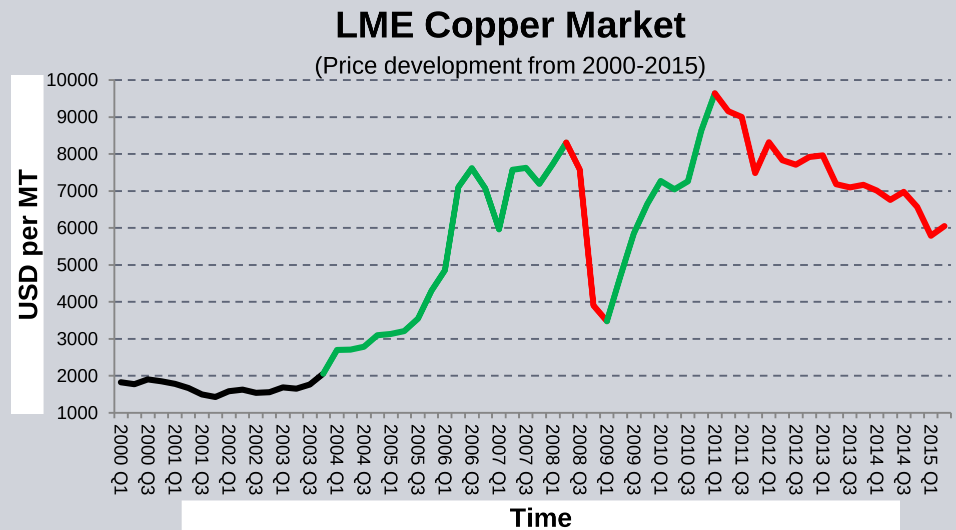
Fraunhofer Gesellschaft, IIS SCS Analytics, Nürnberg

Email: hans.georg.zimmermann@scs.fraunhofer.de

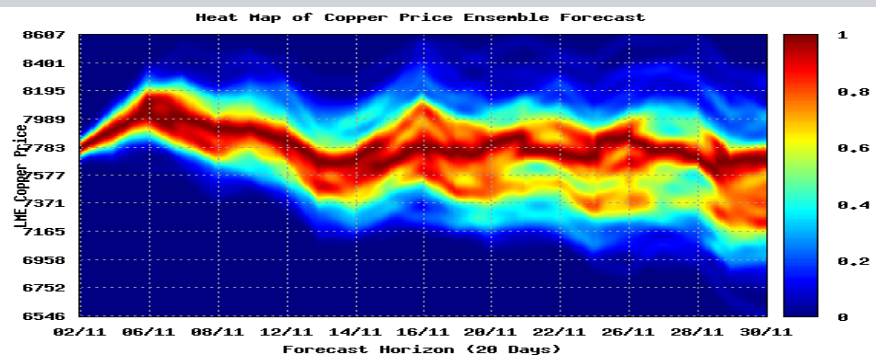
Fraunhofer IIS

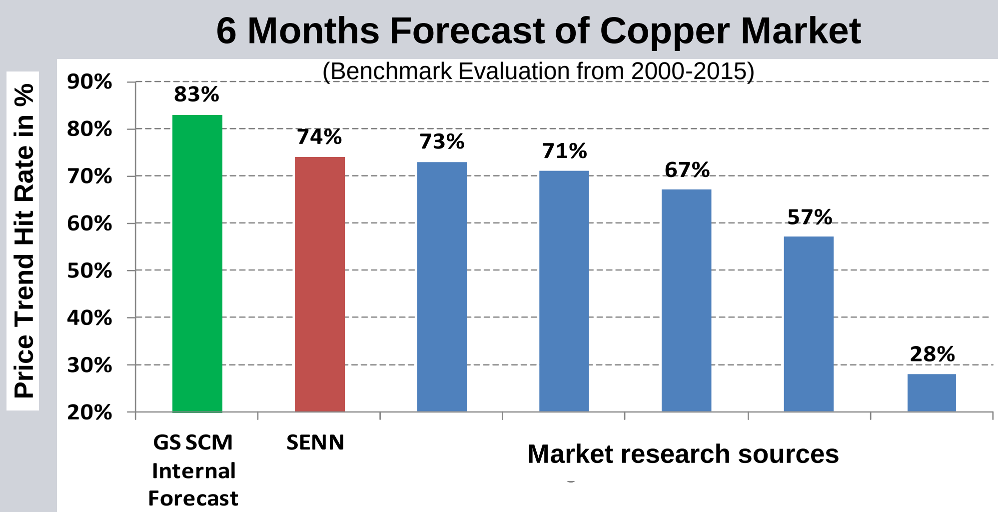
Forecasting Commodity Prices with Neural Networks (Siemens)





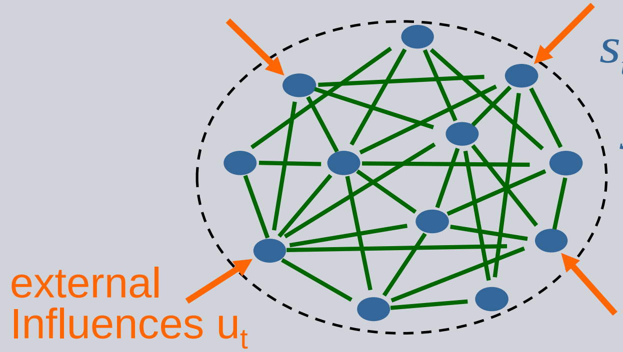
- LME Copper: 79% of total purchasing volume
- Optimal procurement based on market forecasts
- Our forecasts include an uncertainty analysis



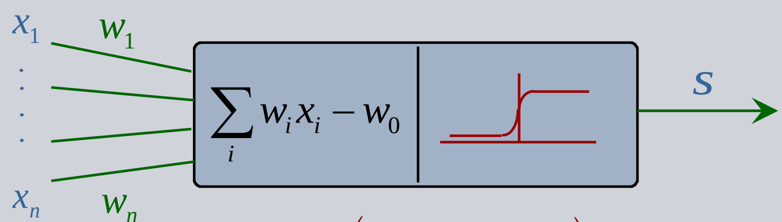


Fraunhofer IIS

Neural Networks - From Decision Models to Markets



$$s_{t+1} = f(s_t, u_t)$$
$$y_t = g(s_t)$$



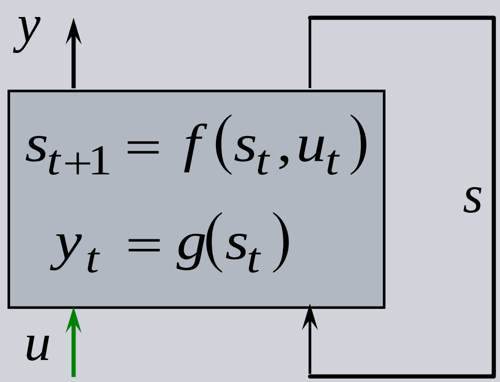
$$s = f\left(\sum_i w_i x_i - w_0\right)$$

A neuron is a decision model combining information filtering & aggregation & switching.

A neural network can be seen as interaction of decisions and hence, as a market model.

Fraunhofer IIS

Modeling of Open Dynamical Systems with Recurrent Neural Networks (RNN)



$$s_{t+1} = \tanh(As_t + Bu_t)$$ *state transition*

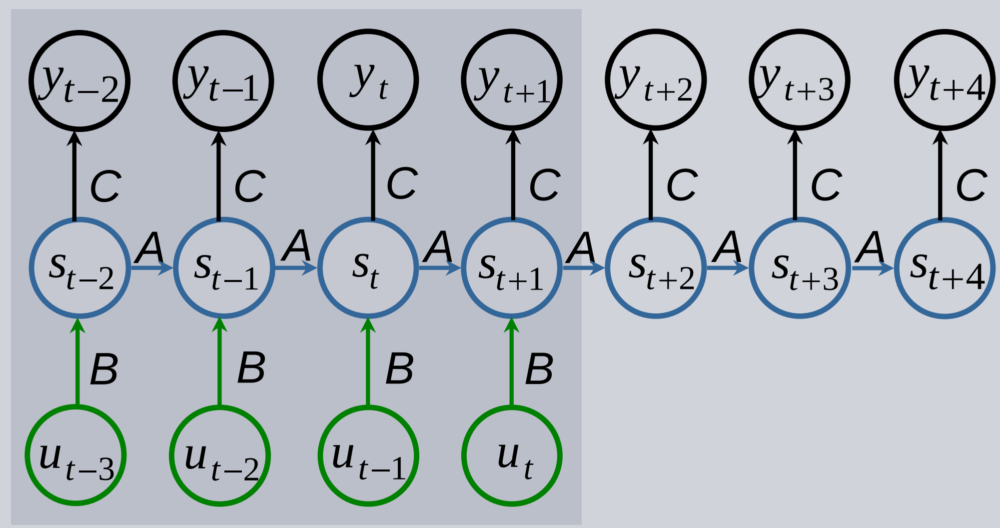
$$y_t = Cs_t$$ *output equation*

$$\sum_{t=1}^{T} (y_t - y_t^d)^2 \rightarrow \min_{A,B,C}$$ *identification*

Finite unfolding in time transforms time into a spatial architecture.

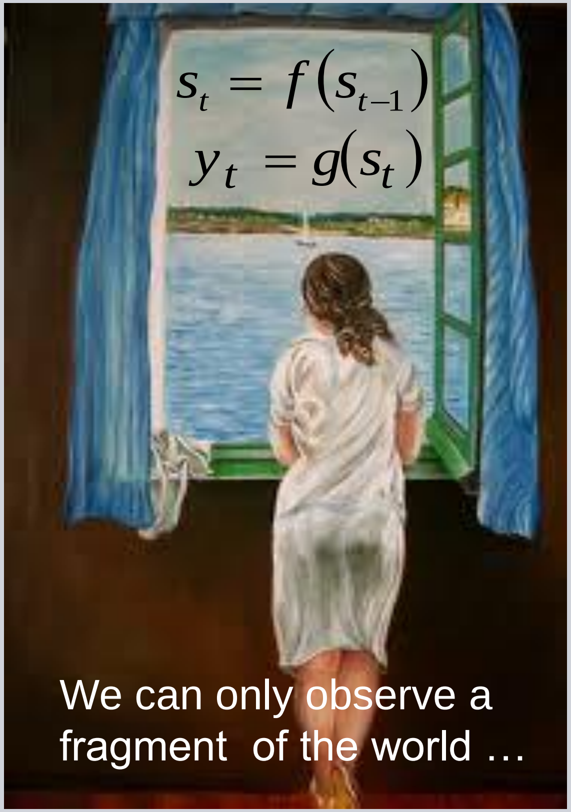
The analysis of open systems with RNNs allows a decomposition of the autonomous & external driven part.

Long-term predictability depends on a strong autonomous subsystem.



Fraunhofer IIS

Modeling Closed Dynamical Systems with Recurrent Neural Networks (HCNN)

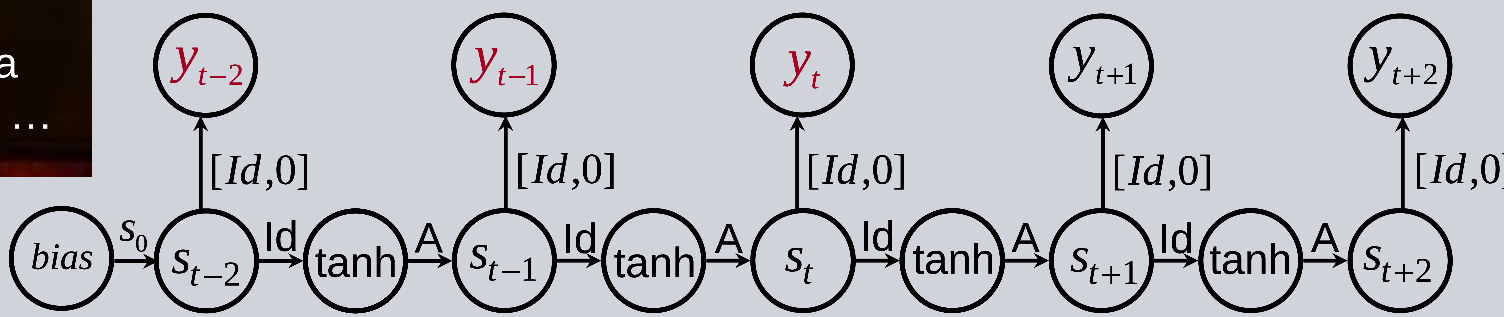


$$s_t = A\tanh(s_{t-1}) \quad , s_0$$ state transition

$$y_t = [Id, 0]s_t$$ output equation

$$\sum_{t=1}^{T}\left(y_t - y_t^d\right)^2 \to \min_{A, s_0}$$ identification

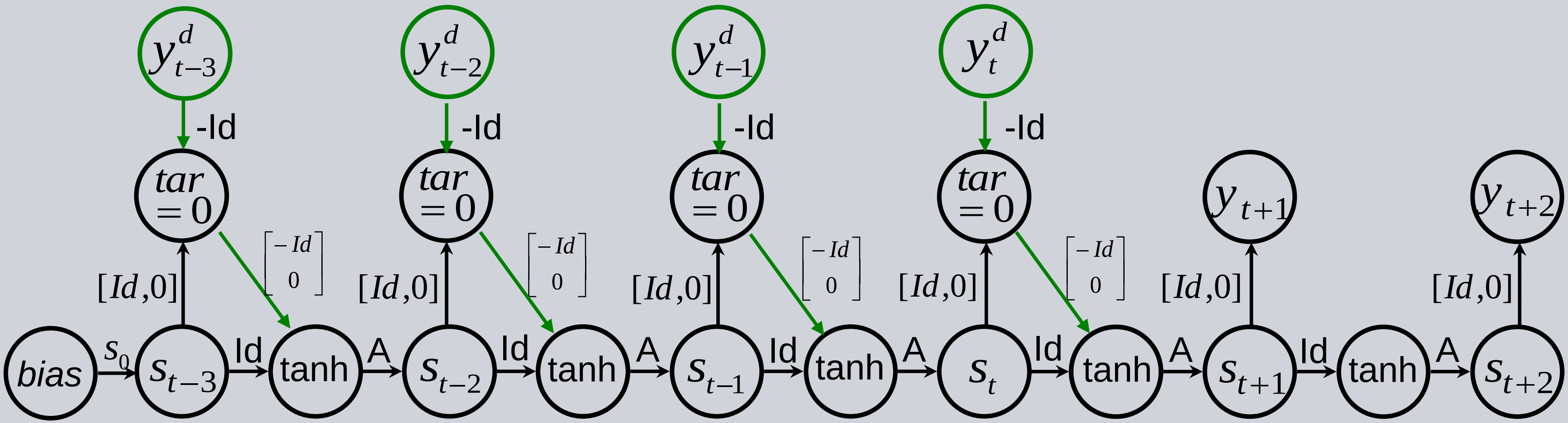
The model is unfolded along **history** → only 1 training example



… but to understand the dynamics of the observables, we have to reconstruct at least a part of the hidden states of the world. Forecasting is based on observables and hidden states.

The Role of Observations in the Identification of Dynamical Systems

Embed the original architecture into a larger architecture, which is easier to learn. After the training, the extended architecture has to converge to the original model.

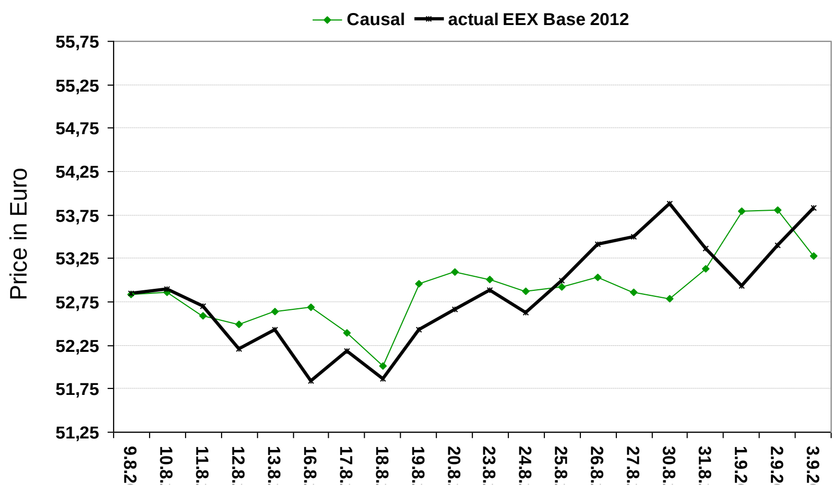


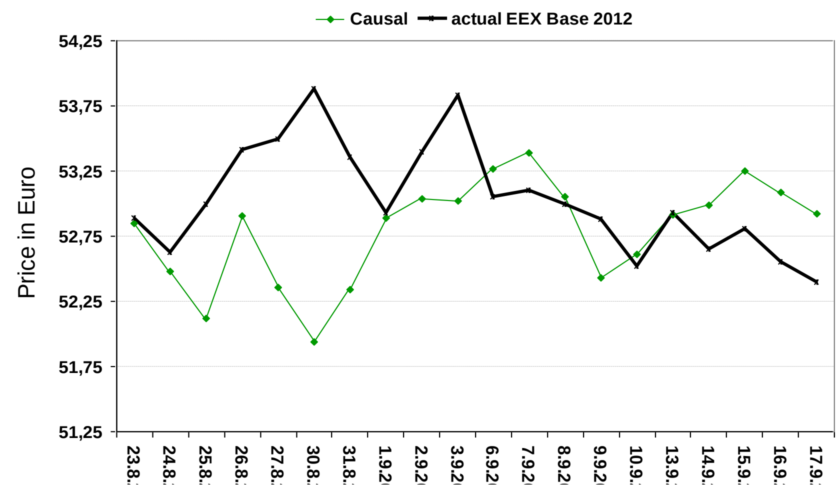
- The essential task is NOT to reproduce the past observations, but to identify related hidden variables, which make the dynamics of the observables reasonable.
- Use an architectural teacher-forcing (**ATF**) to support the learning of the HCNN. Replace expectations y_τ by observations y^d_τ: $r^{upper}_\tau = y_\tau - (y_\tau - y^d_\tau) = y^d_\tau$
- The state flow decomposes in 1) known observables, 2) reconstructed hidden variables and 3) unidentifiable random variables, which act as a net-internal regularization.

Fraunhofer IIS

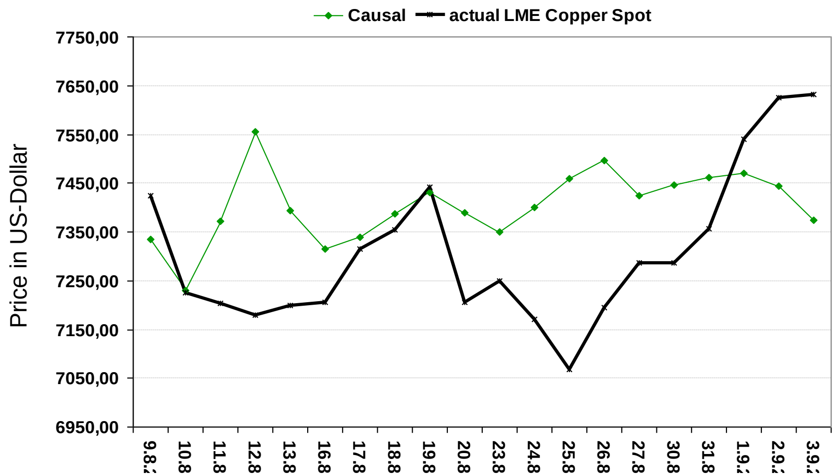
20 Days Price Forecasting with Causal Neural Networks

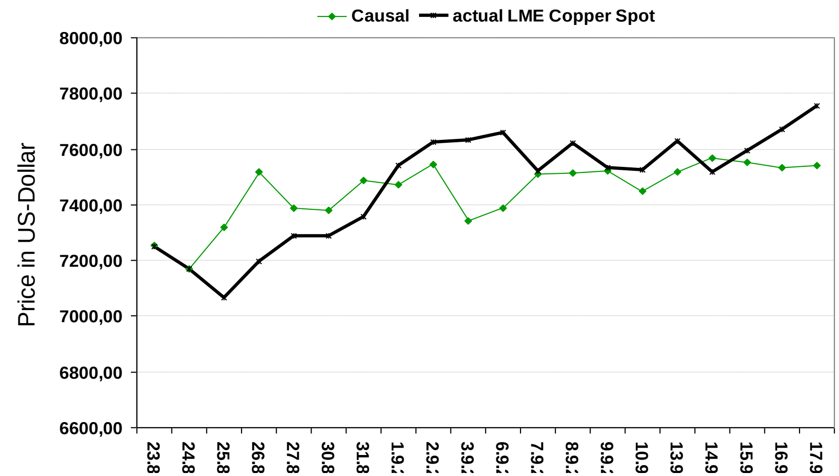
EEX Base Market





LME Copper Market





2010-08-09 – 2010-09-03

2010-08-23 – 2010-09-17

Fraunhofer IIS

Identification of Goal Oriented Dynamical Systems

When we are able to identify the goals of the agents we can try to explain the dynamics backward from their goals.

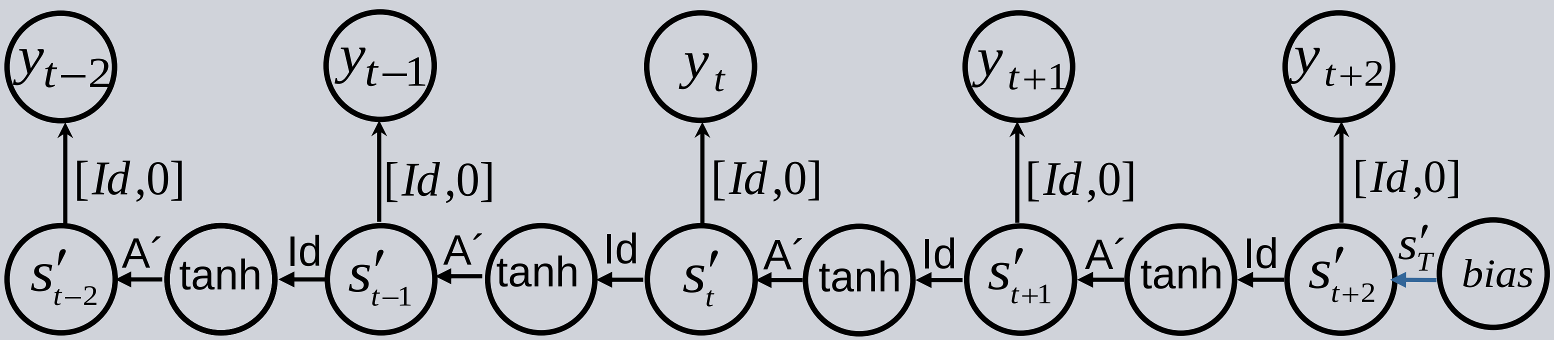
$$s'_t = A' \tanh\left(s'_{t+1}\right) \;,\; s'_T \quad \text{state transition}$$

$$y_t = \left[Id, 0\right] s'_t \quad \text{output equation}$$

$$\sum_{t=1}^{T} \left(y_t - y_t^d\right)^2 \rightarrow \min_{A', s'_T} \quad \text{identification}$$

Causal nets are given by an initial state s_0 & matrix A

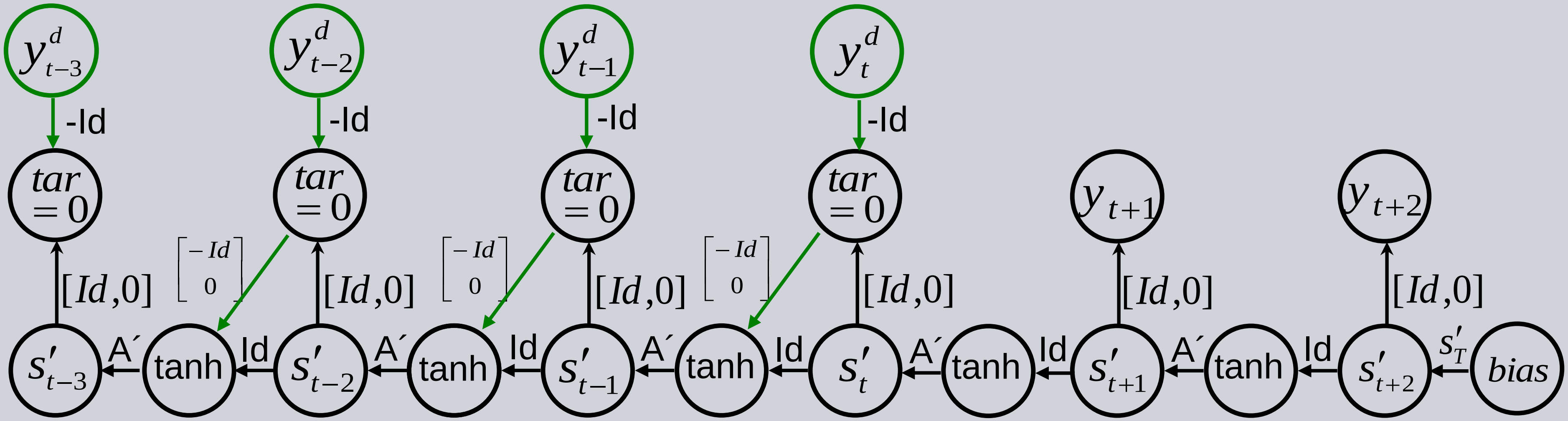
Retro-causal nets are given by a final state s_T & matrix A´



Fraunhofer IIS

The Identification of Retro-Causal Dynamical Systems

Embed the original architecture into a larger architecture, which is easier to learn. After the training, the extended architecture has to converge to the original model.

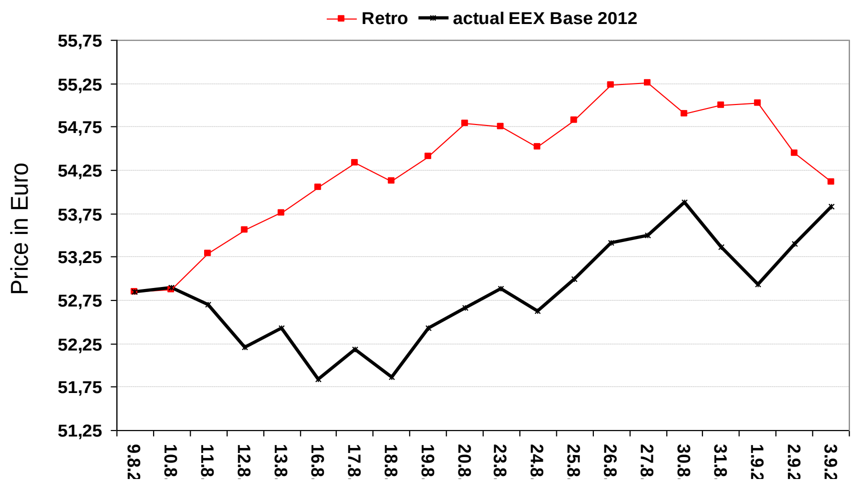


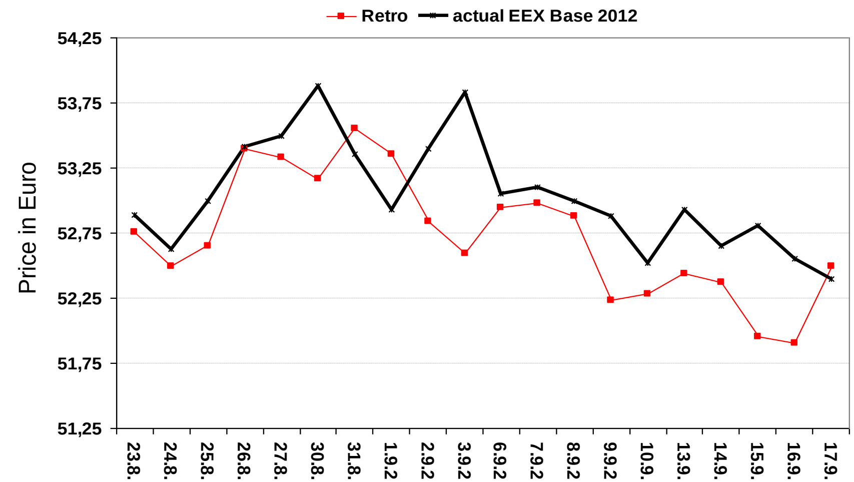
- The essential task is NOT to reproduce the past observations, but to identify related hidden variables, which make the dynamics of the observables reasonable.
- Use a form of an architectural teacher-forcing (**ATF**) to support the learning of the HCNN. Replace expectations y_τ by observations $y^d_\tau \rightarrow \quad r^{upper}_\tau = y_\tau - (y_\tau - y^d_\tau) = y^d_\tau$
- Observations of states are a cut in the flow of the forward dynamics as well as in the back-ward error flow $\quad \varepsilon^{upper}_\tau = \varepsilon_{\tau+1} - \varepsilon_{\tau+1} + error_\tau \approx y_\tau - y^d_\tau$

Fraunhofer IIS

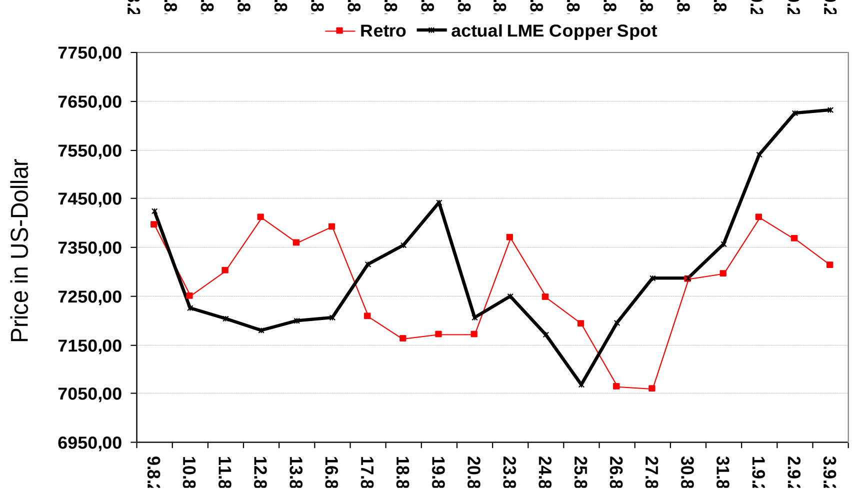
20 Days Price Forecasting with Retro-Causal Neural Networks

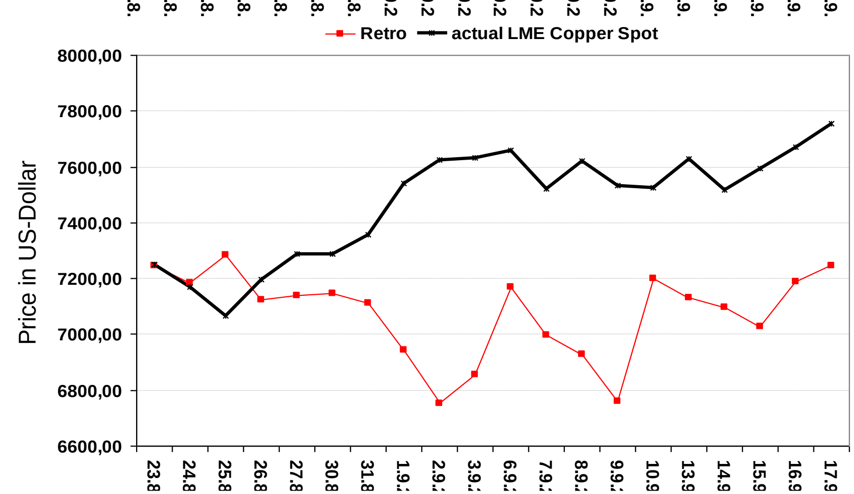
EEX Base Market





LME Copper Market





2010-08-09 – 2010-09-03

2010-08-23 – 2010-09-17

Fraunhofer IIS

Comparison of Causal Forecasts and Retro-Causal Forecasts

Causal Forecasts

Retro-Causal Forecasts

Insights: Causal and retro-causal forecasting is complementary.

Ex ante there is no chance to know which approach is the appropriate.

We should find a way to combine the best of both approaches.

Fraunhofer IIS

Combining Causal & Retro-Causal Neural Networks (CRCNN)

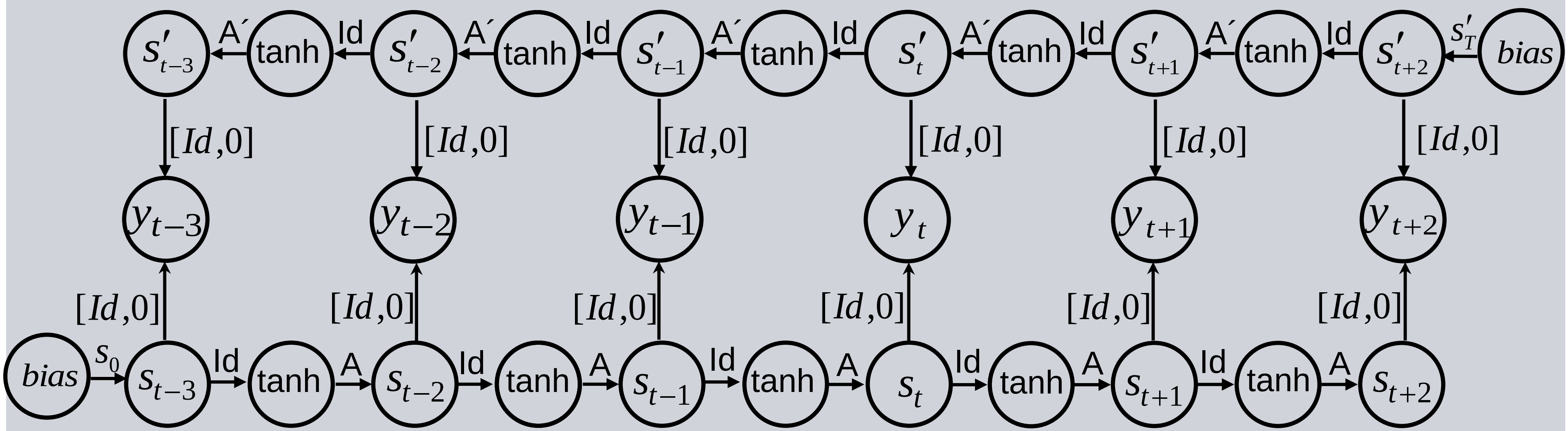
- We explain observations by a symmetric superposition of causal & retro-causal subnets.
- Both subnets are universal learners, but the more appropriate branch learns faster and reduces the error flow of the opposite branch.
- In non-unique optimization, we have to have attention on the path to the optimum!!!

$$s_t = A \tanh(s_{t-1}) \quad , s_0 \qquad \text{causal transition}$$

$$s'_t = A' \tanh(s'_{t+1}) \quad , s'_T \qquad \text{retro transition}$$

$$y_t = [Id, 0] s_t + [Id, 0] s'_t \qquad \text{output equation}$$

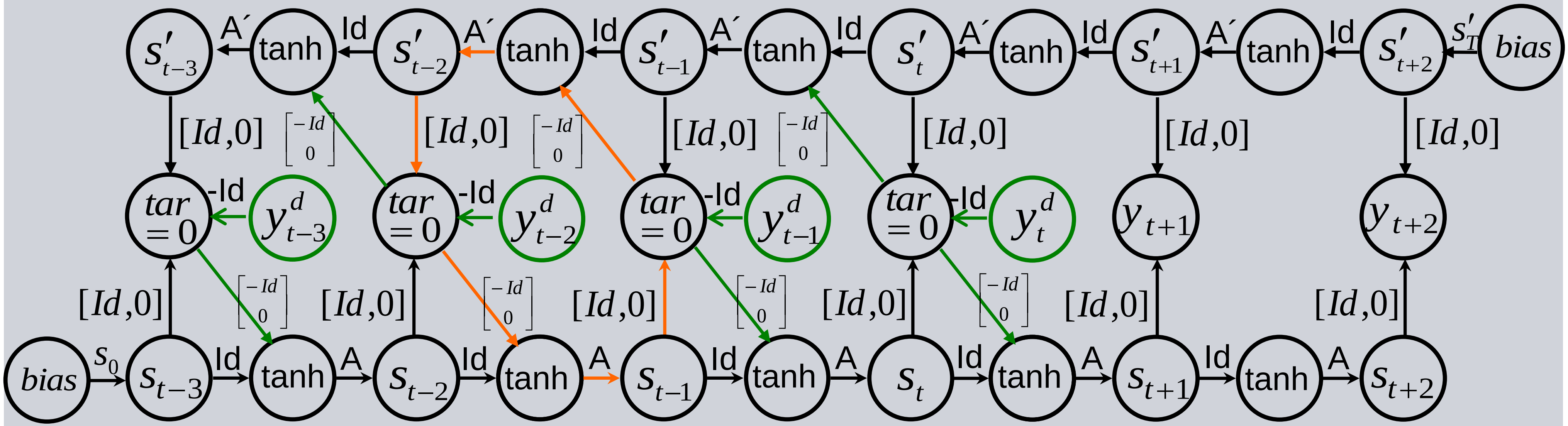
$$\sum_{t=1}^{T} \left(y_t - y_t^d\right)^2 \rightarrow \min_{A, s_0, A', s'_T} \qquad \text{identification}$$



Fraunhofer IIS

System Identification of Causal-Retro-Causal Networks

Problem 1: Obviously we do not know the future data for the observables. Thus skip the teacher forcing beyond present time.

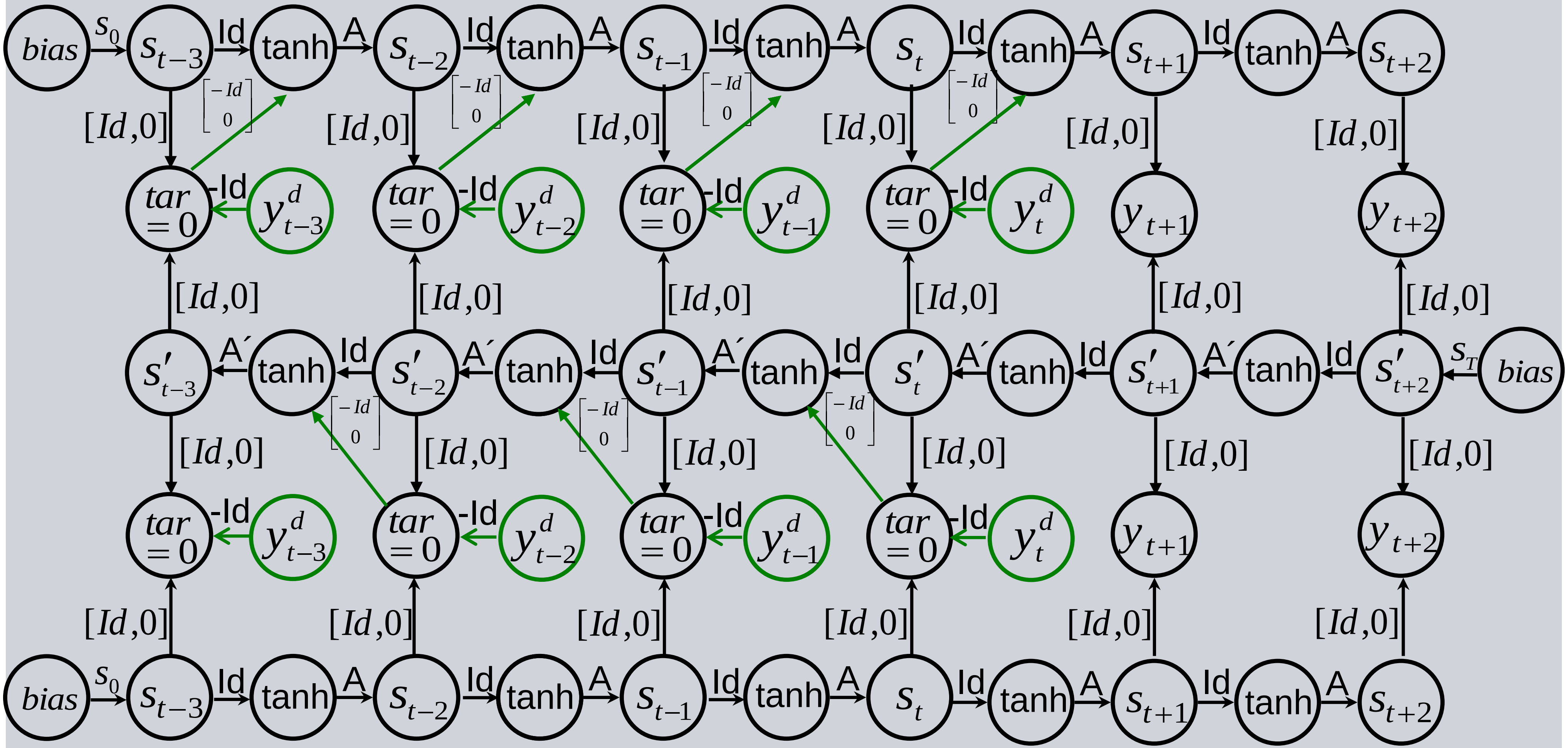


Problem 2: At every time step the network has a fix point loop (equality constraint), defining a manifold.

In a joint effort we have to identify the dynamics and the manifold, which is implicitly identified by causal-retro-causal architecture.

Fraunhofer IIS

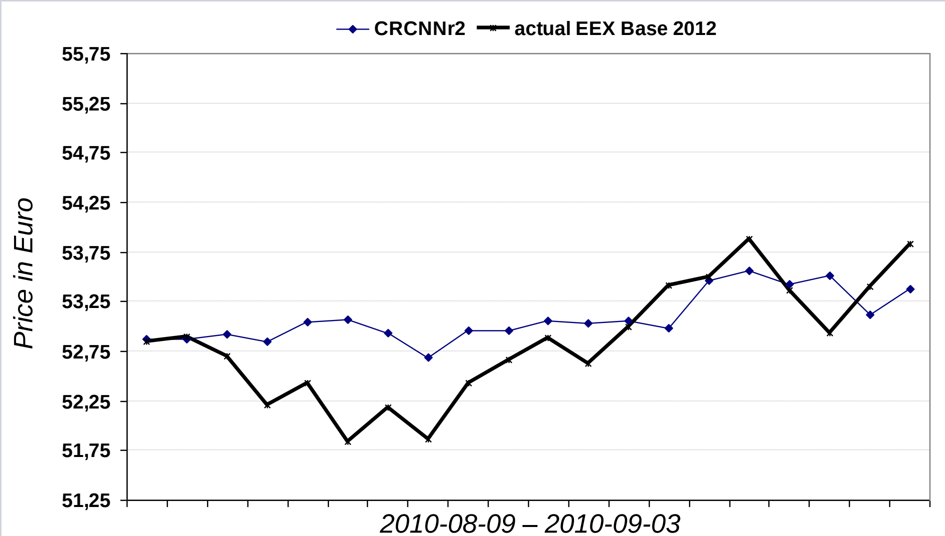
CRCNN3 Learning of Manifold & Dynamics

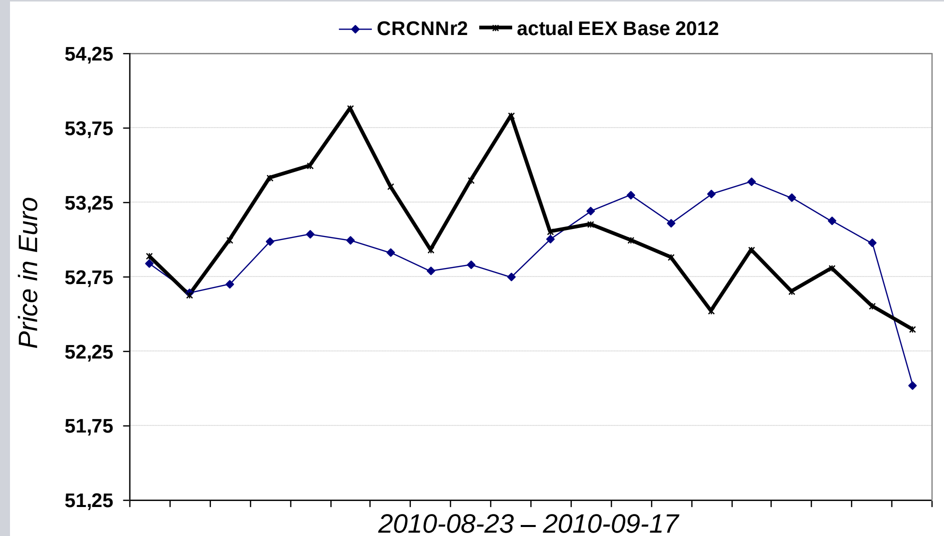


Only one branch has no teacher forcing, 2 Causal & 1 Retro.

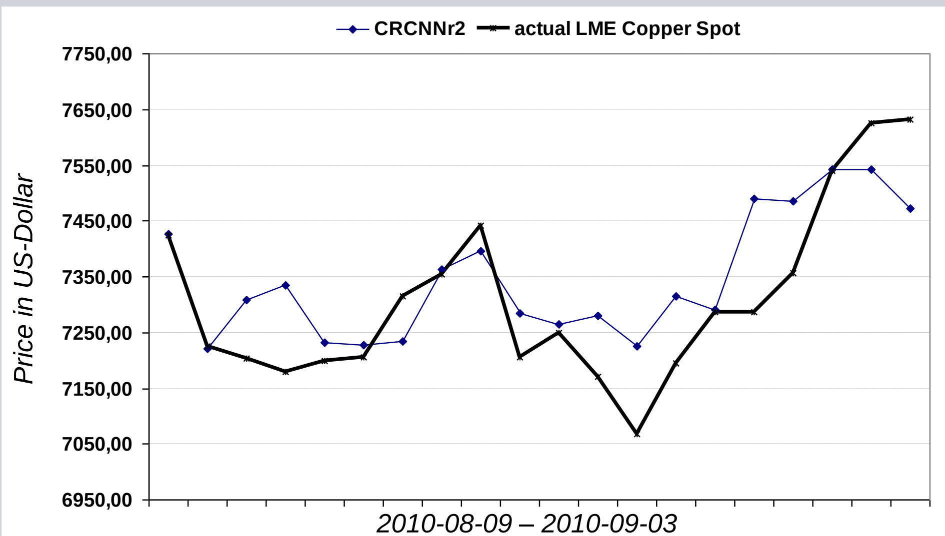
Causal-Retro-Causal Neural Networks (CRCNN)

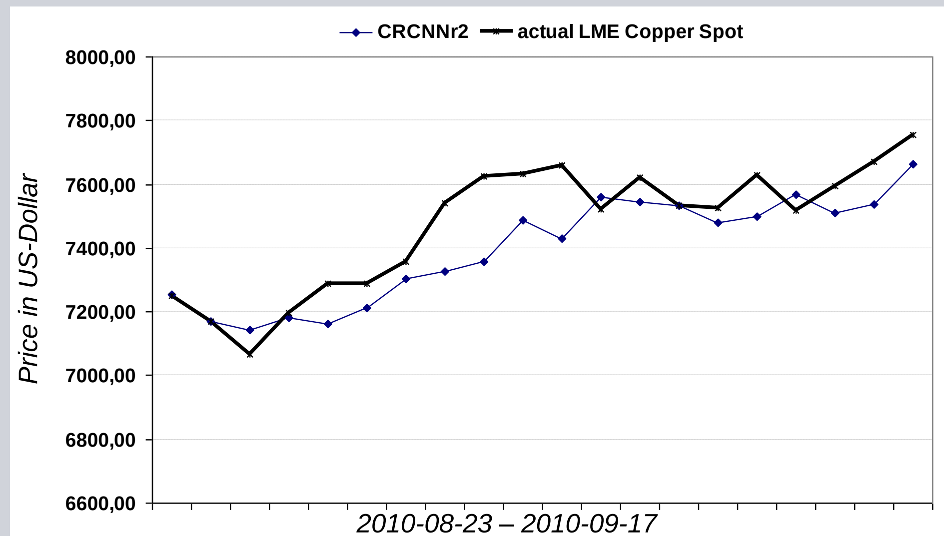
EEX Base Market





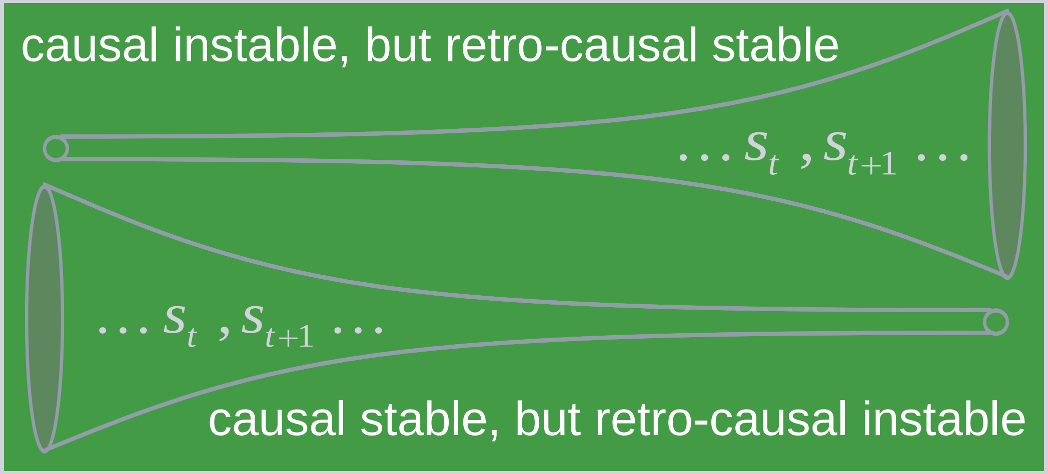
LME Copper Market





Fraunhofer IIS

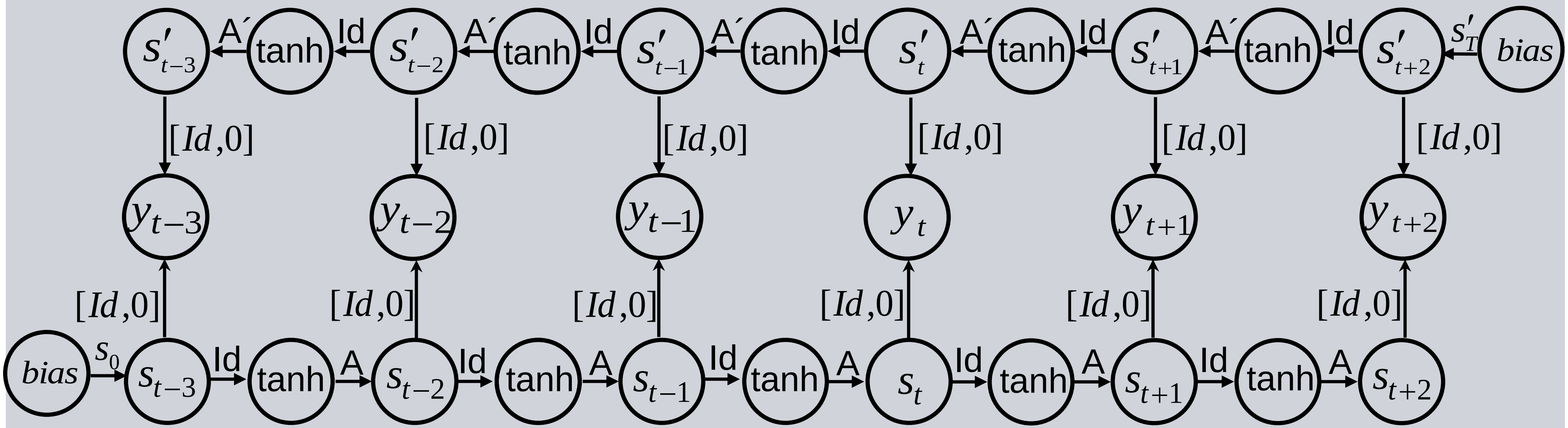
The Modeling of Partial Chaotic Systems with CRCNNs



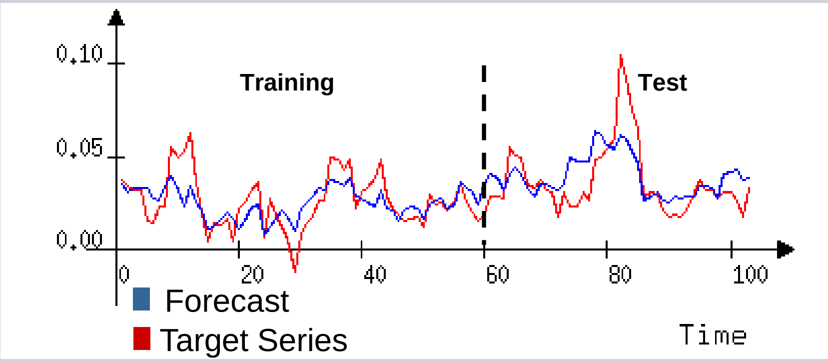
Tubes describe ensembles of trajectories. If we have a causal instable (chaotic) sub-dynamics, it is retro-causal stable and vice versa.

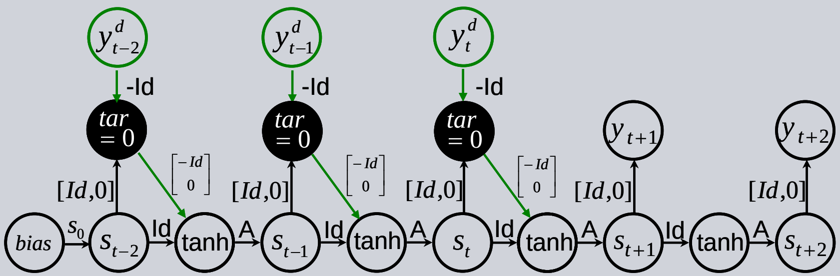
Successful learning of **CRCNNs** enforces a decomposition such that causal and retro-causal branches are both stable.

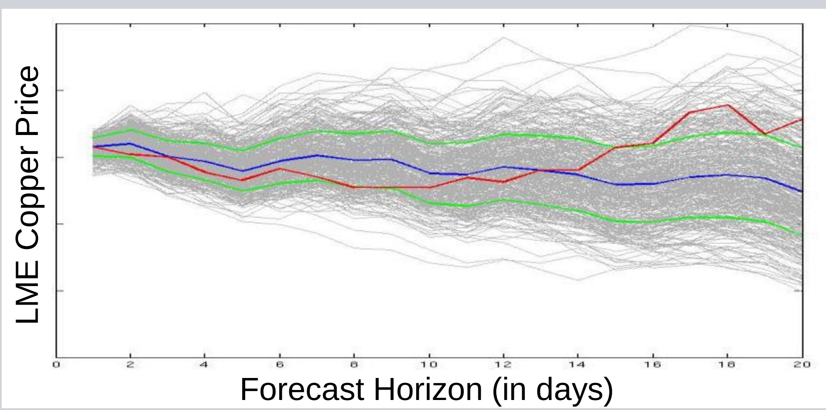
CRTNN learning doesn´t allow such a split.



Approaches to Model Uncertainty in Forecasting







1 Measure uncertainty as volatility (variance) of the target series. The underlying forecast model is a constant. Thus sin(ωt) can be highly uncertain!??

2 Build a forecast model. The error is interpreted as uncertainty in form of additive noise. The width of the uncertainty channel is constant over time.

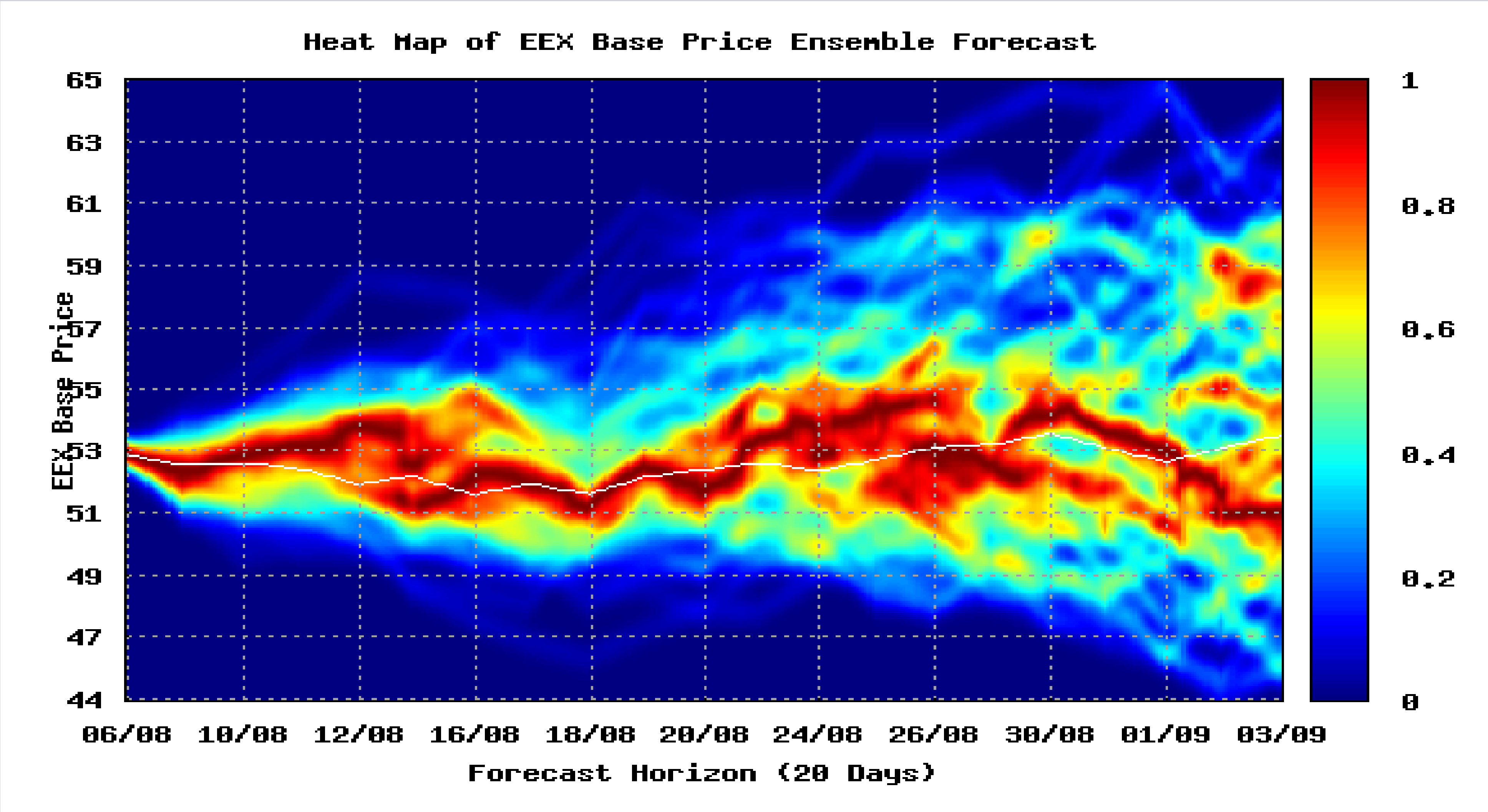
3 Describe uncertainty as a diffusion process (random walk). The diffusion channel widens over time, e.g. scaled by the one-step model error.

For large systems 2 & 3 fail: We have to learn to zero error → the uncertainty channel disappears.

4 One large model doesn't allow to analyze forecast uncertainty, but an ensemble forecast shows the characteristics of an uncertainty channel: Given a finite set of data, there exist many perfect models of the past data, showing different future scenarios caused by different estimations of the hidden states.

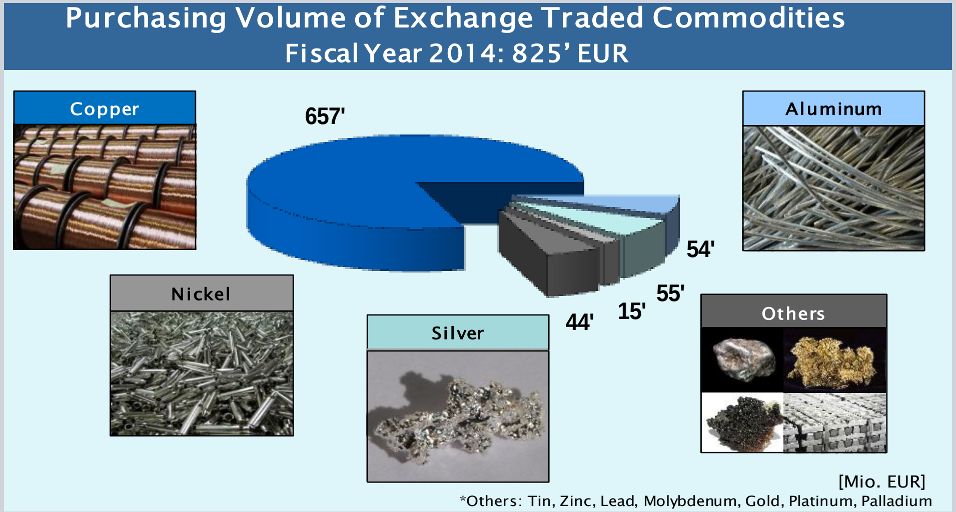
Fraunhofer IIS

EEX Electricity Price Heatmap



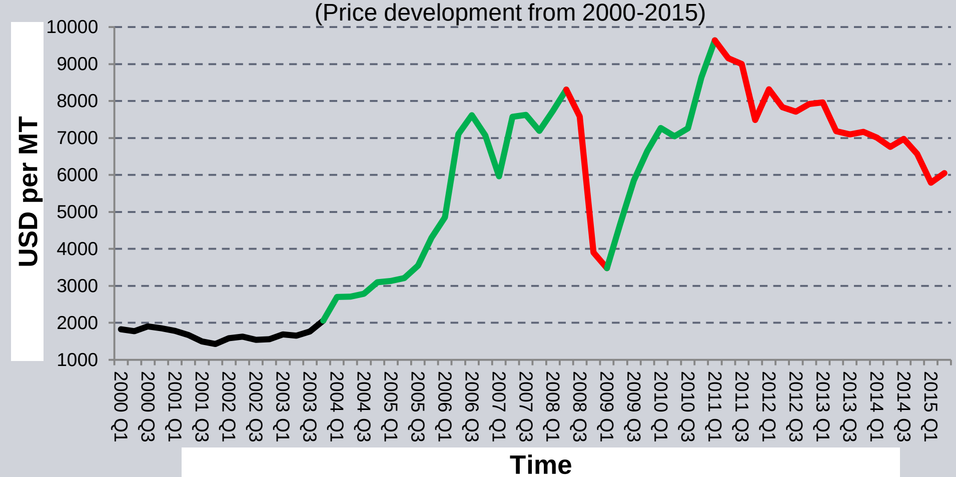
Fraunhofer IIS

Forecasting Commodity Prices with Neural Networks (Siemens)

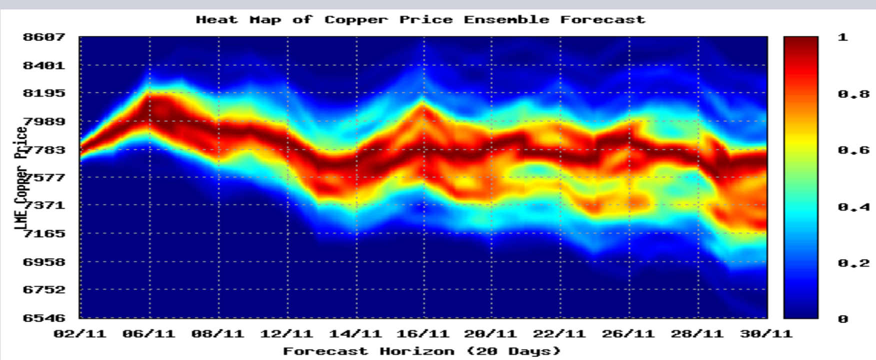


LME Copper Market

(Price development from 2000-2015)

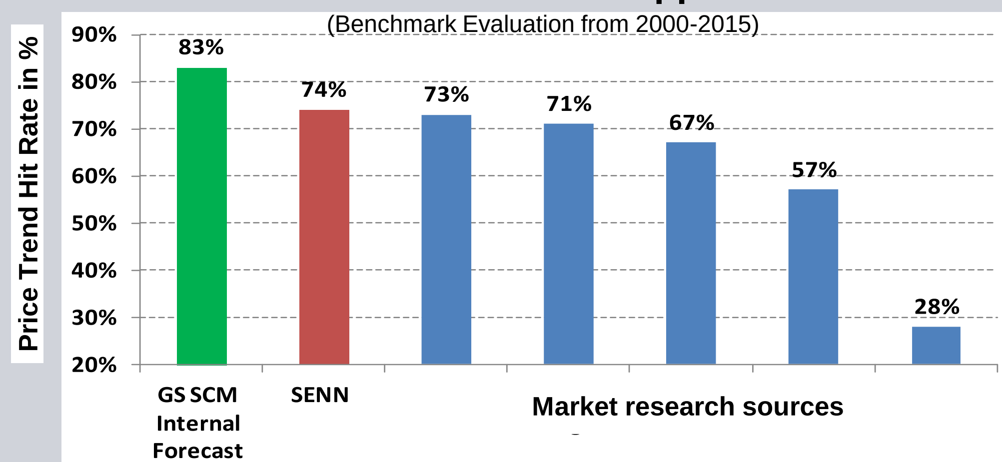


- LME Copper: 79% of total purchasing volume
- Optimal procurement based on market forecasts
- Our forecasts include an uncertainty analysis



6 Months Forecast of Copper Market

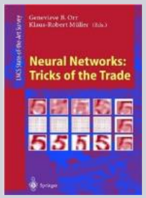
(Benchmark Evaluation from 2000-2015)



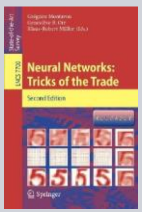
Fraunhofer IIS

Selected References

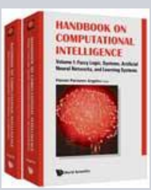
Zimmermann, H.G.: Neuronale Netze als Entscheidungskalkül, In: Neuronale Netze in der Ökonomie; Eds.: Rehkugler, H., Zimmermann, H.G., p. 1-87; Franz Vahlen; 1994; ISBN 3-8006-1871-0



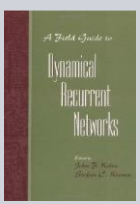
Neuneier, R.; Zimmermann, H.G.: How to Train Neural Networks, In: Neural Networks: Tricks of the Trade; Eds.: Orr G. B., Müller K. R.; p. 373-423; Springer, 1998; ISBN 3-540-65311-2



Zimmermann, H.G.; Tietz, C.; Grothmann, R.: Forecasting with Recurrent Neural Networks, In: Neural Networks: Tricks of the Trade, **Second Edition**; Eds.: Montavon, G; Orr G. B., Müller K.R.; p. 687-707; Springer, 2012; ISBN 978-3-642-35288-1

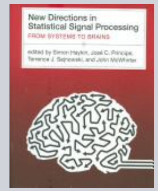


Zimmermann, H.G.; Grothmann, R.; Tietz, C.; A New View on Economics with Recurrent Neural Networks, In: Handbook on Computational Intelligence, Vol 1; Ed: Angelov, P; World Scientific 2016; ISBN 978-981-4675-02-4



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

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, 2002



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.; et. al. [Eds.]: New Directions in Statistical Signal Processing: From Systems to Brain, MIT, 2006.

Zimmermann, H.-G., Grothmann, R., Tietz Ch and von Jouanne-Diedrich, H: Market Modeling, Forecasting and Risk Analysis with Historical Consistent Neural Networks, in: Proceedings of the International Conference on Operations Research, Munich 2010, Springer, 2011

Zimmermann, H.-G., Grothmann, R., Tietz, Ch.: Forecasting market prices with causal-retro-causal neural networks. In: Lüthi, H.-J. et. al (eds.): OR Proceedings 2011: Selected Papers of the Int. Conference on Operations Research (OR 2011), Zurich, Springer, 2012

Youtube (Zimmermann, Siemens): Keynote Talk on SAIConference: System Identification and Forecasting with Recurrent Neural Networks

The CERN Lectures: Identification of Dynamical Systems with Neural Networks
http://indico.cern.ch/event/561247 http://indico.cern.ch/event/561248

University Erlangen Lecture 2019 (24 hours in german) www.fau.tv or https://www.video.uni-erlangen.de/course/id/771