System Identification, Forecasting, Control

with

Neural Networks

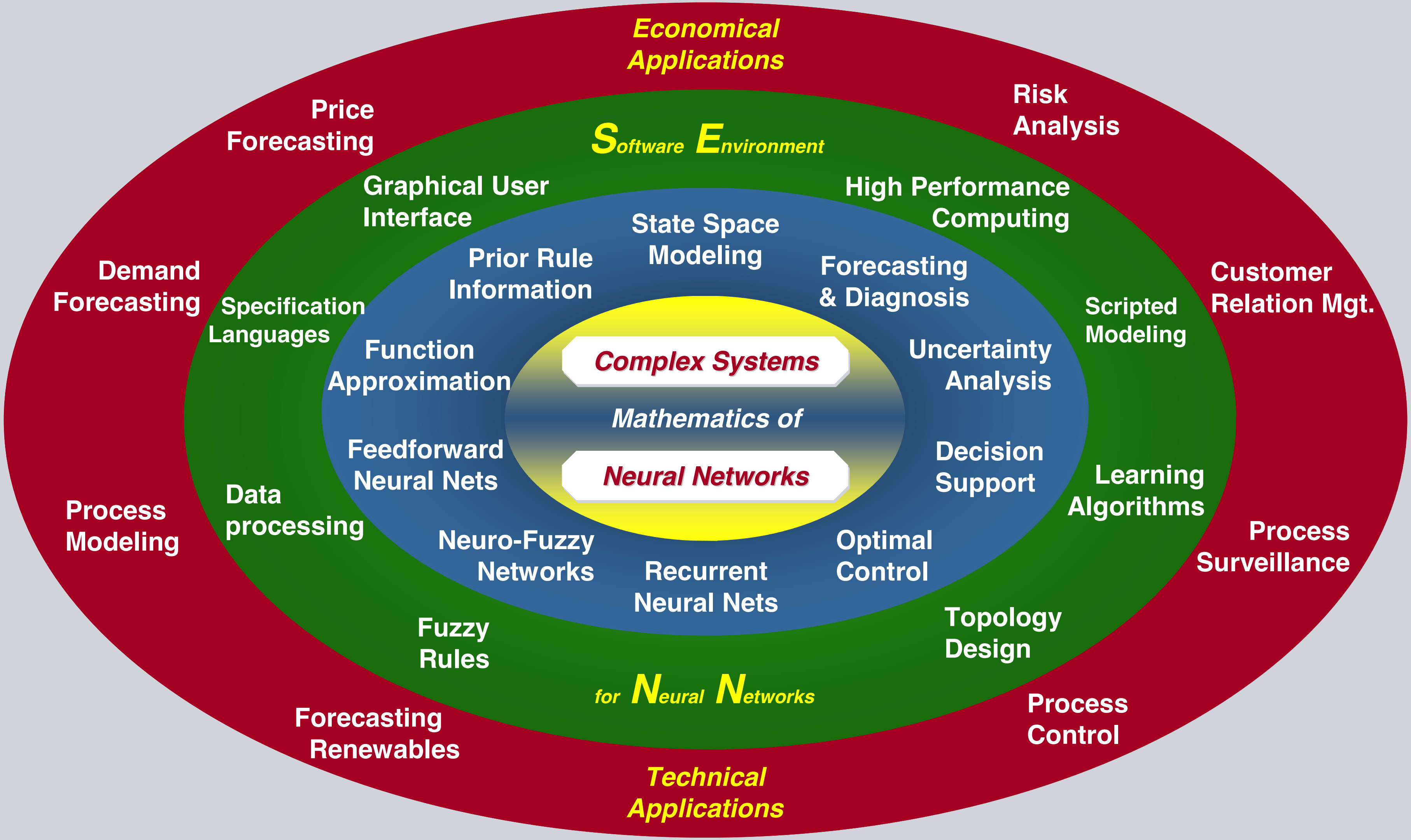
Hans Georg Zimmermann

Siemens AG, Corporate Technology

Email : Hans_Georg.Zimmermann@siemens.com

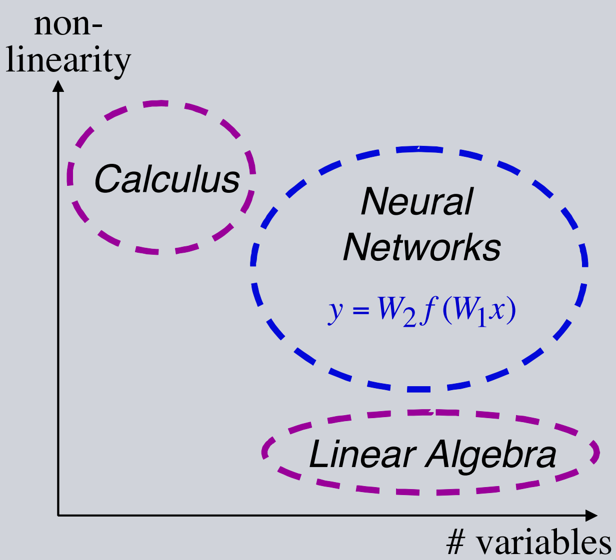
SIEMENS

Neural Networks @ Siemens: 25 Years of Research, Development, Innovation



Mathematical Neural Networks

Complex Systems



Existence Theorem:

(Hornik, Stinchcombe, White 1989)

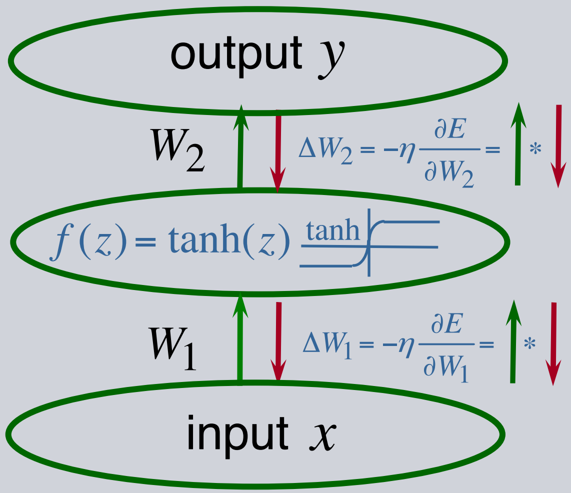
3-layer neural networks can approximate any continuous function on a compact domain.

Nonlinear Regression

Based on data identify an input-output relation

$$y = W_2 f(W_1 x)$$

$$E = \sum_{t=1}^{T} \left(y_t - y_t^d \right)^2 \rightarrow \min_{W_1, W_2}$$



Neural networks imply a **Correspondence** of *Equations*, *Architectures*, *Local Algorithms*.

Neural Networks are No Black Boxes

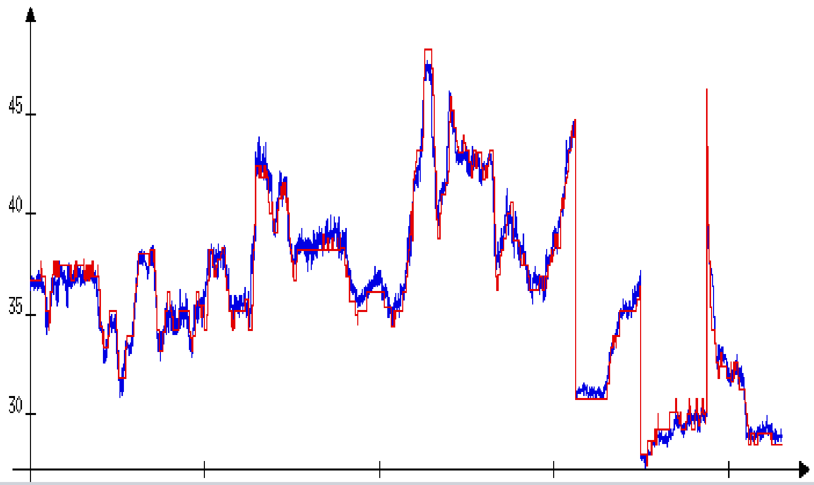
Application: Modeling of a Gas Turbine

- Inputs: 35 sensor measures and control variables of the turbine
- Output: NO_x emission of the gas turbine

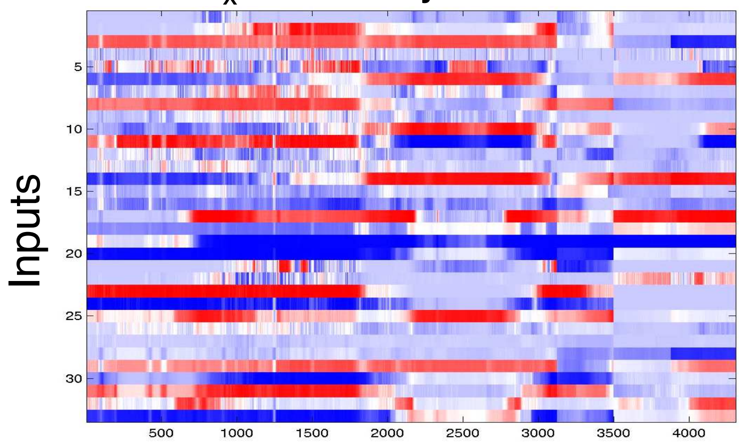
Sensitivity Analysis: Compute the first derivatives along the time series:

$$\frac{\partial\, output}{\partial\, input_i} > 0 \qquad \frac{\partial\, output}{\partial\, input_i} < 0$$

Model Output vs. actual NO_x



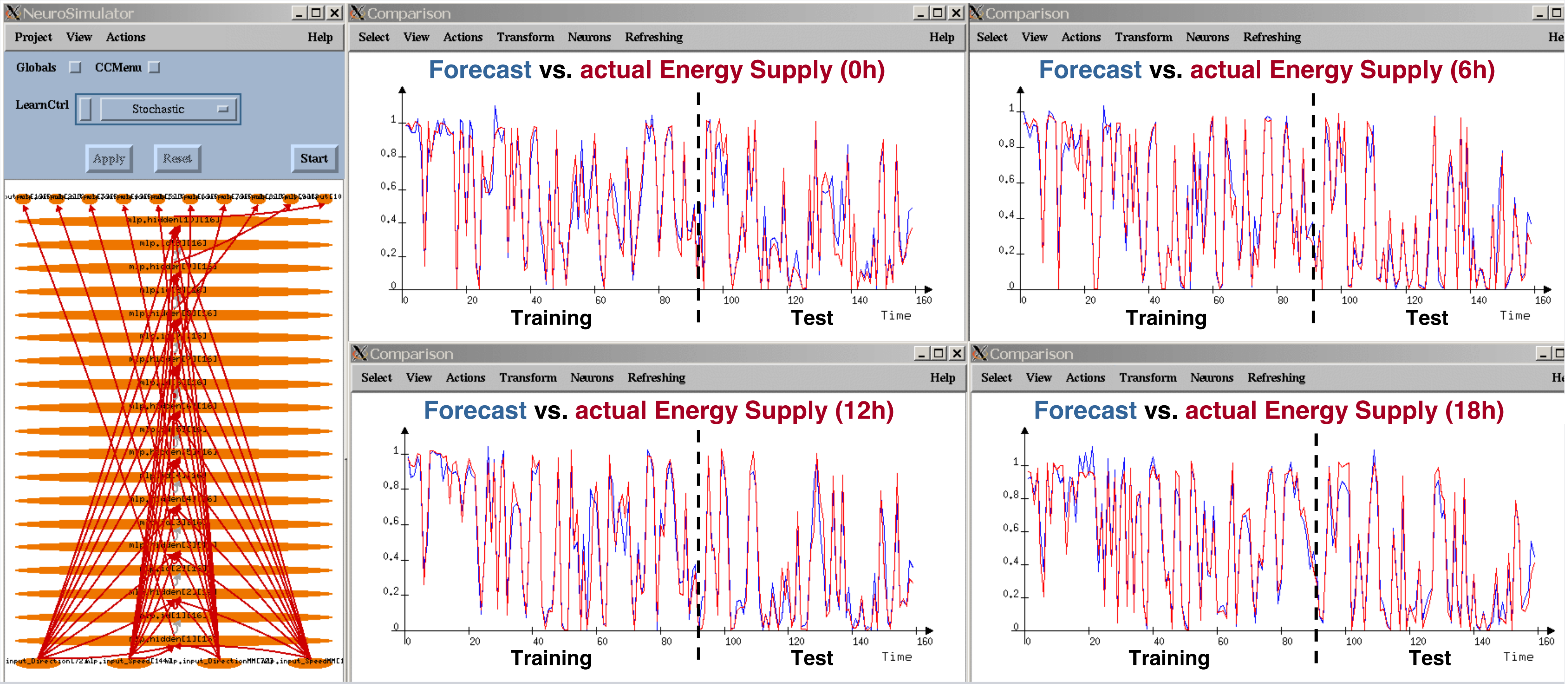
NO_x Sensitivity over Time



A classification of input-output sensitivities:

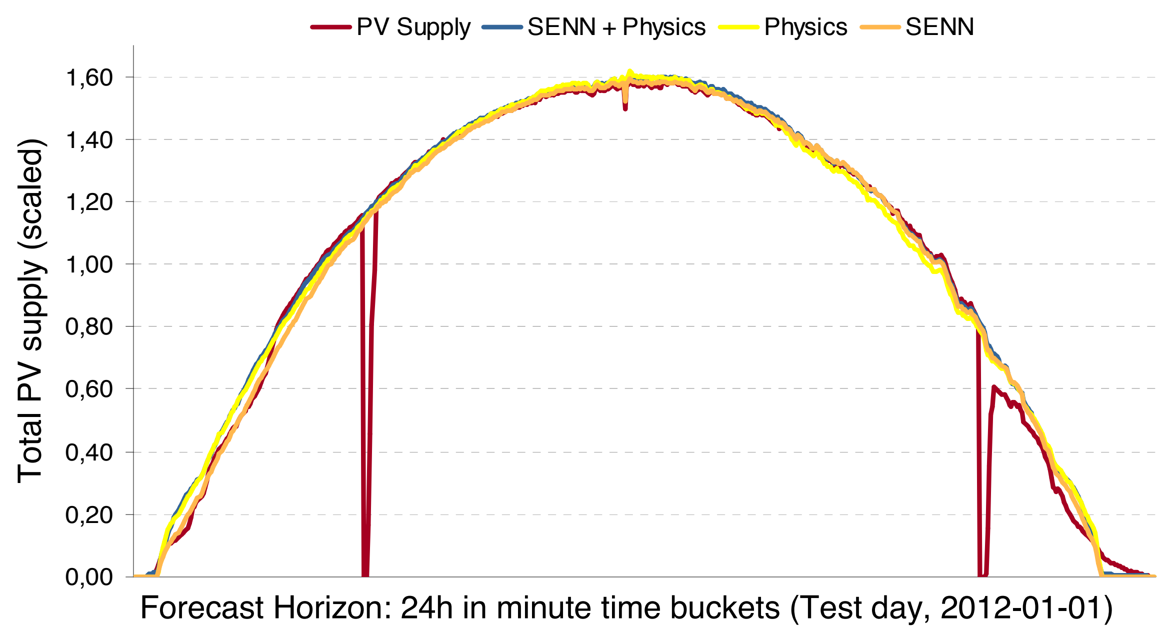
- ***constant over time*** (= linear relationship)
- ***monotone*** (input can be used in 1dim. control)
- ***non-monotone*** (only multi-dim control possible)
- ***~ zero*** (input useless in modeling and control)

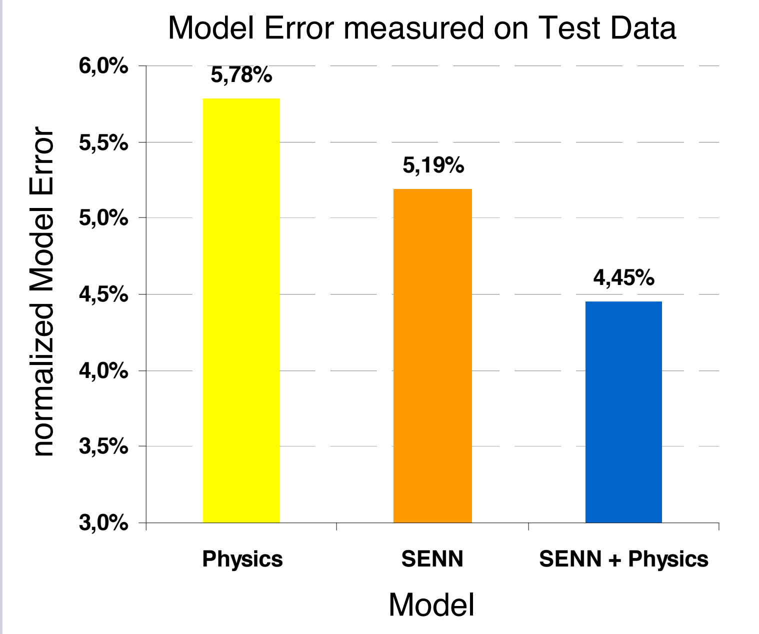
Neural Forecasting of Wind Power Supply



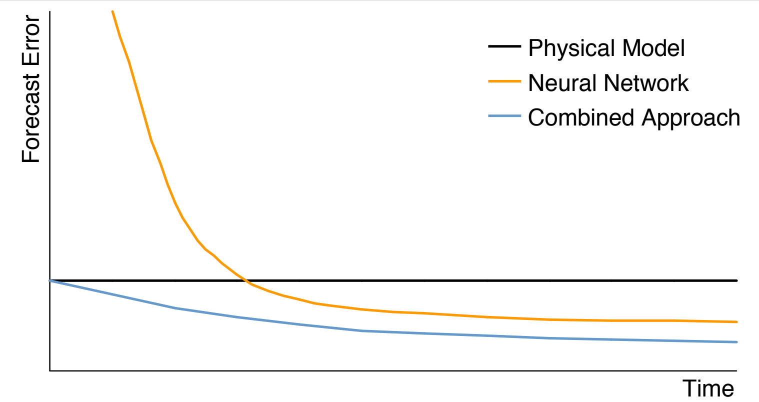
- **Task:** Forecast the wind energy supply of an entire wind field over the next 24 hours
- **Solution:** Deep Neural Network (DNN) with 10 hidden layers. Inputs are wind speed and direction
- **Results:** DNN (**RMSD: 7.20%**) outperform the analytical Jensen model (benchmark, **RMSD: 10.22%**)

Forecasting of Solar Power Supply



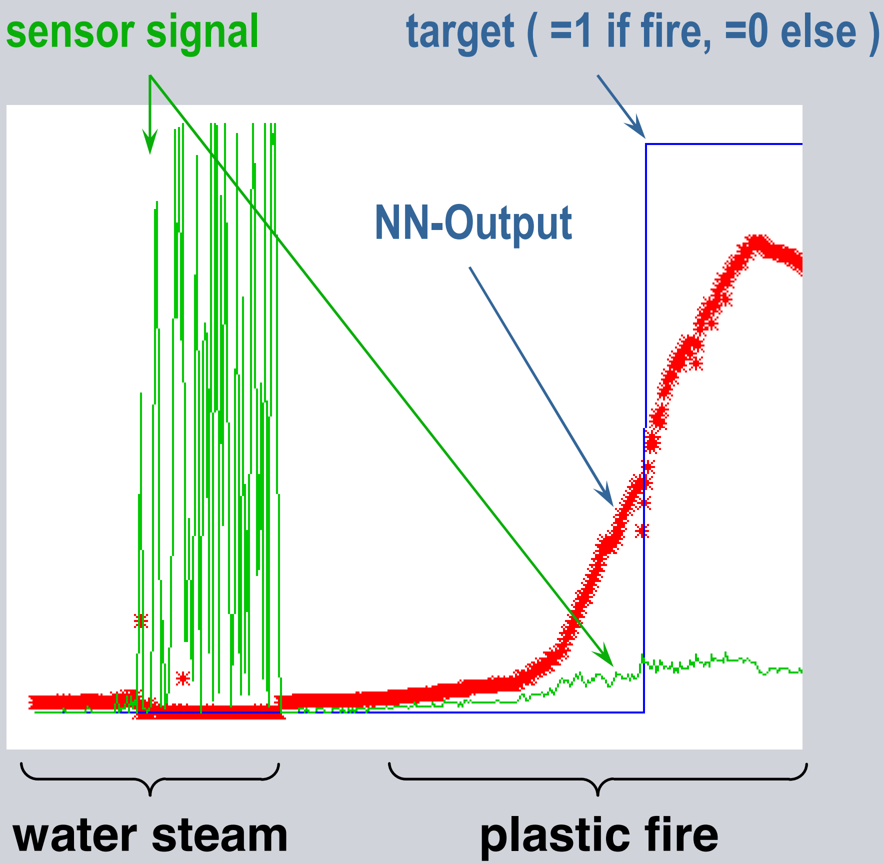


- Forecast the energy supply of a PV plant with a neural network (SENN) based on weather forecasts and/or a physics based model
- Data set: 2011-09-08 to 2012-01-23. Minute data (200k data points). 40k data points are randomly selected as test data.
- Performance of a purely data-driven model (SENN) is comparable to the physics based model
- A hybrid model (SENN & physics) improves the forecast accuracy



SIEMENS

Optimizing Sensors with Neural Networks



(time series of 2 different scenarios)

Improve Optical Smoke Detectors:

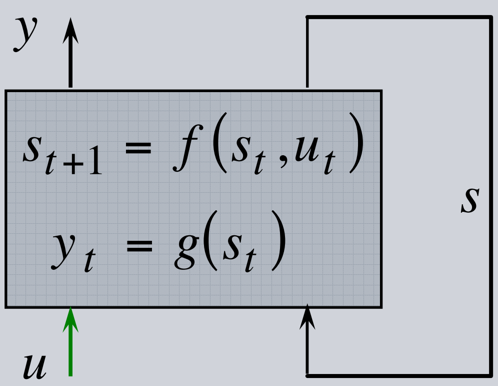
- Smoke-poor fires are hard to detect
- Avoid false alarms caused by e.g. steam, welding, exhaust fumes

Modeling with Neural Networks:

- **Input:** 3 inputs extracted from one sensor over 40 time steps. 50 fire scenarios.
- **Network Design:** Feedforward network including local and global modeling.
- **Network Learning:** Stochastic learning including cleaning noise
- **Result:** Neural networks was able to classify all patterns correctly

SIEMENS

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



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

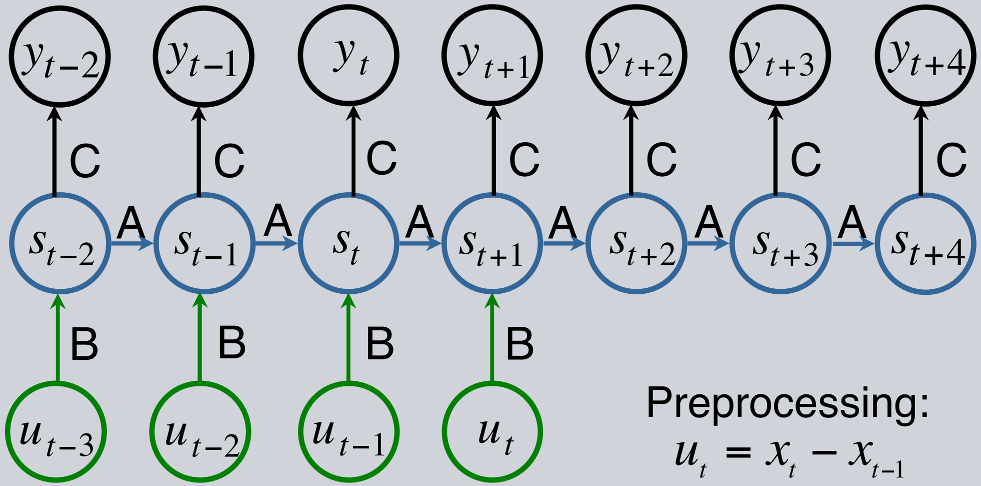
$$y_t = Cs_t \quad \text{output equation}$$

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

Finite unfolding in time transforms time into a spatial architecture. We assume, that x_t=const in the future.

The analysis of open systems by RNNs allows a decomposition of its autonomous & external driven subsystems.

Long-term predictability depends on a strong autonomous subsystem.



Preprocessing:
$u_t = x_t - x_{t-1}$

Modeling Dynamical Systems with Error Correction Neural Networks (ECNN)

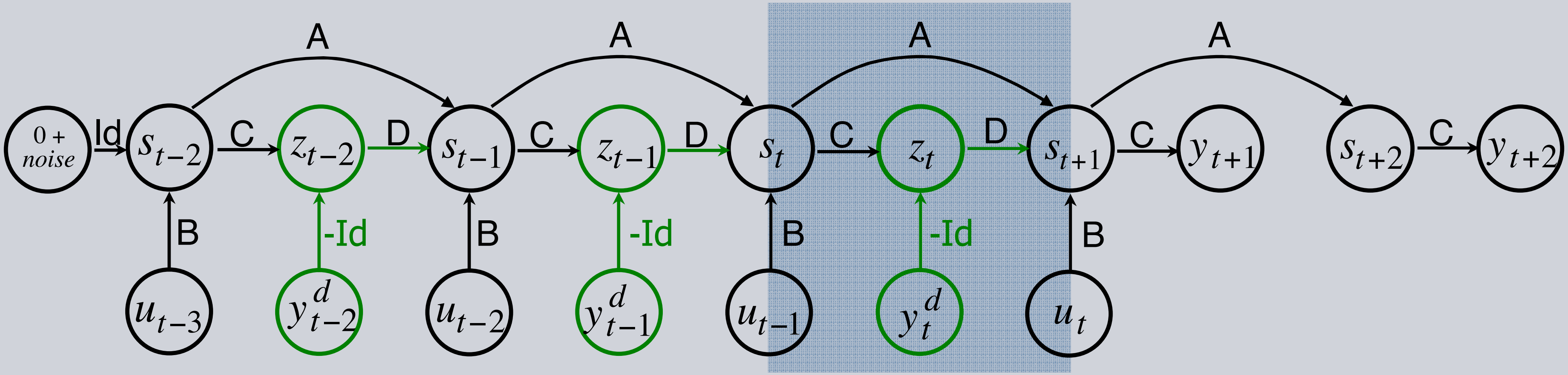
An error correction system considers the forecast error in present time as a reaction on unknown external information.

In order to correct the forecasting this error is used as an additional input, which substitutes the unknown external information.

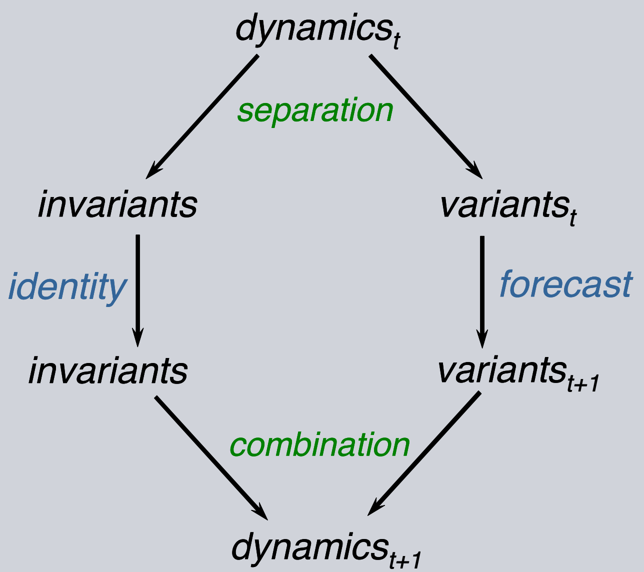
$$s_{t+1} = f\left(s_t, u_t, \left(y_t - y_t^d\right)\right)$$

$$y_t = g(s_t)$$

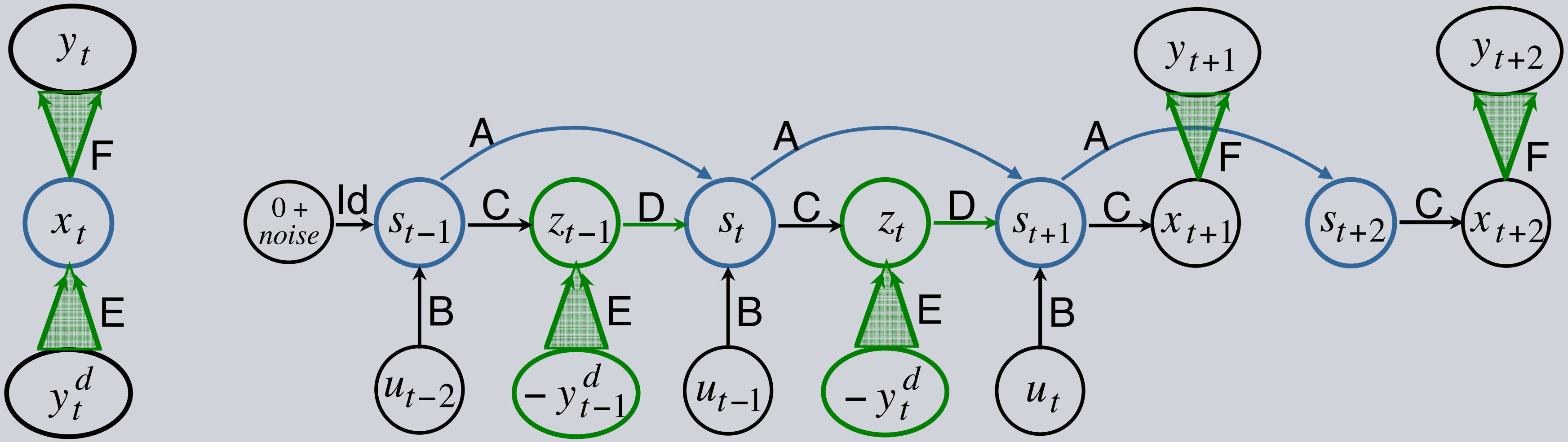
$$\sum_{t=1}^{T} \left(y_t - y_t^d\right)^2 \rightarrow \min_{f,g}$$



Combining Variance - Invariance Separation with ECNN



- The bottleneck autoassociator solves the variance - invariance decomposition.
- The Error Correction Neural Network solves the transformed temporal problem.
- The sub-networks are implicitly coupled by shared weights.

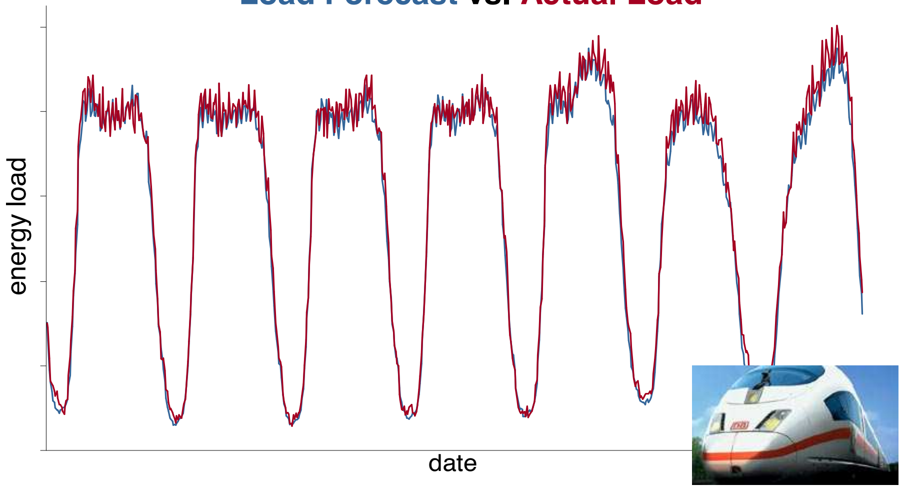


![SIEMENS]

Effective Load Forecasting with Recurrent Neural Networks

Load Forecasting for Electrical Energy

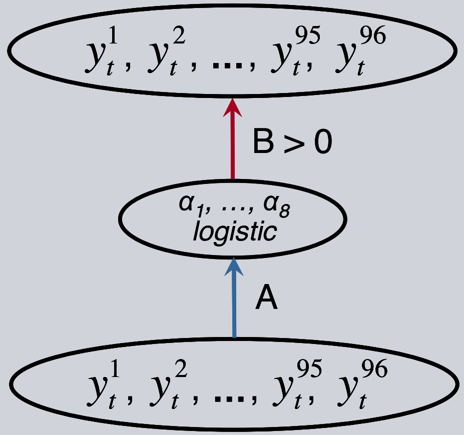
Load Forecast vs. **Actual Load**



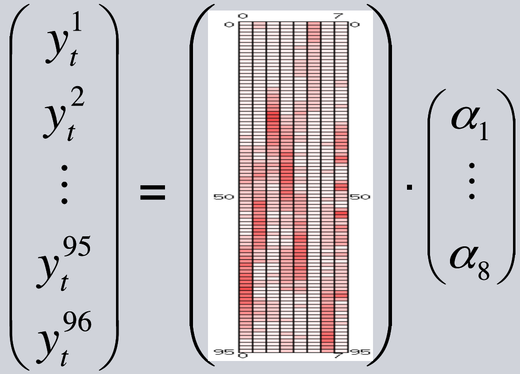
Accuracy of load forecast is 97,95% (Benchmark 95%)

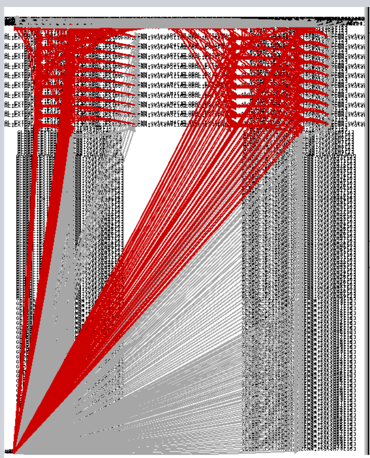
- **Task:** Predict the upcoming energy demand on a 15 min. time grid up to 5 days ahead.
- **Difficulty:** Incorporate the impact of external influences on the energy demand.

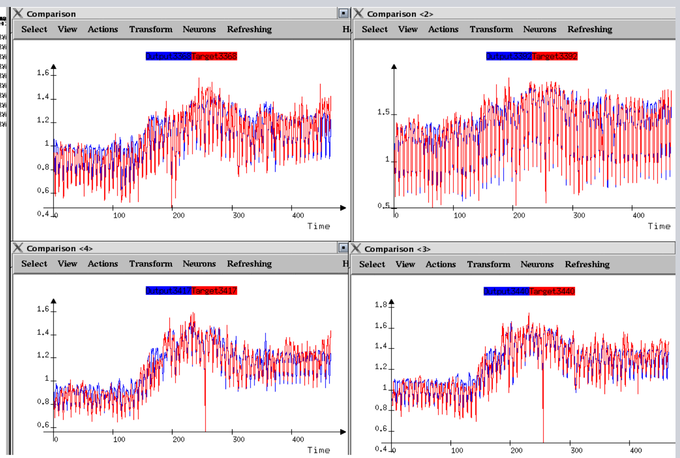
Compression of Electrical Load Curves



Daily load is represented by a superposition of intraday sub-load curves

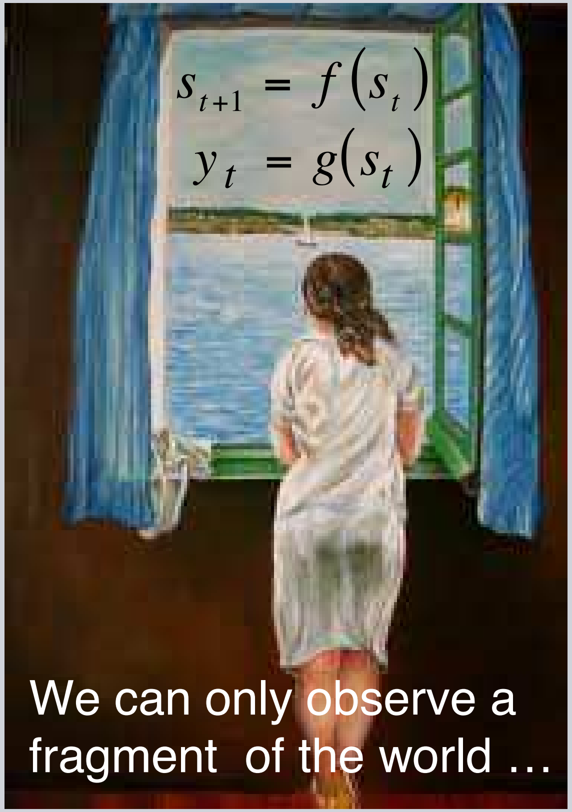






 Intelligent Systems & Control

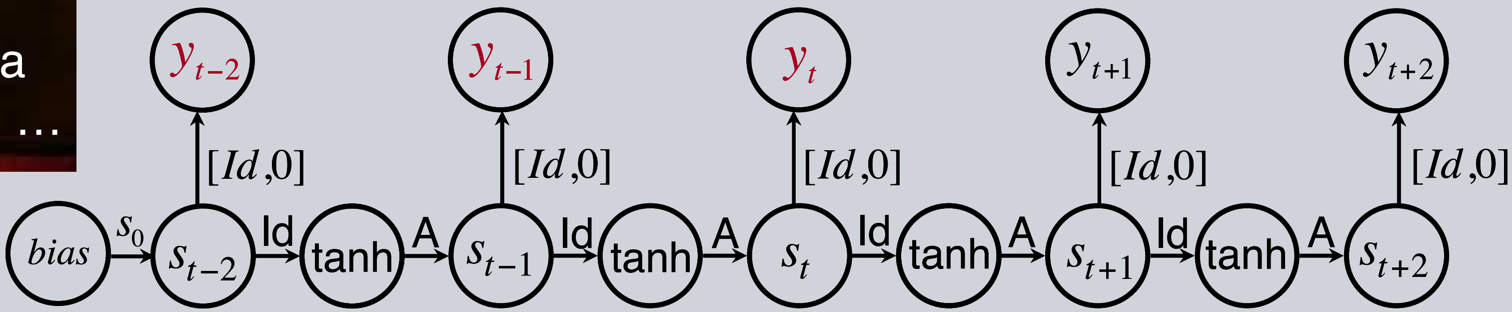
Modeling Closed Dynamical Systems with Recurrent Neural Networks



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

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

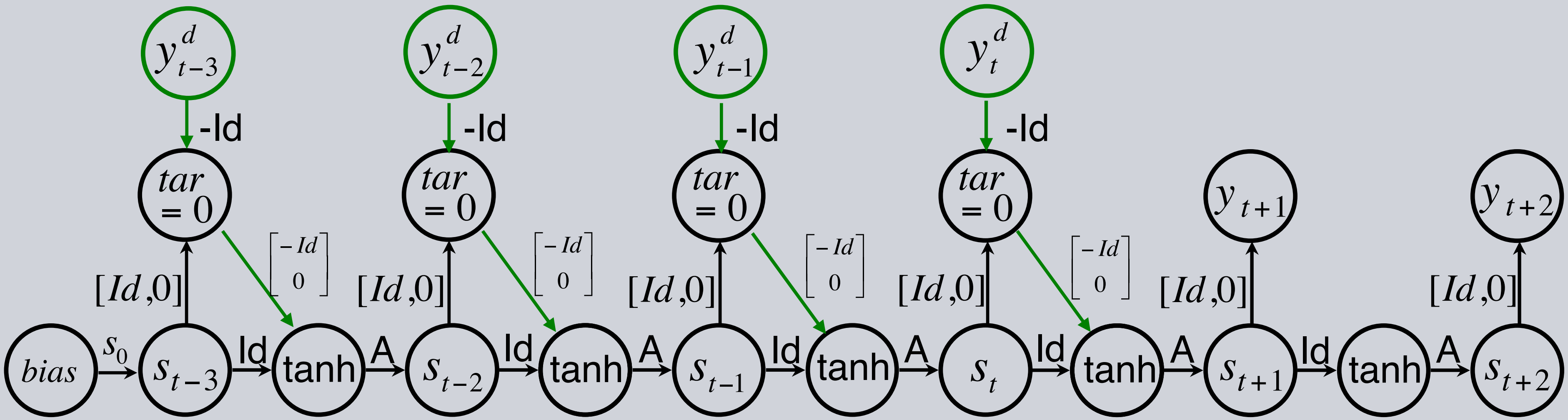
$$\sum_{t=1}^{T} \left(y_t - y_t^d\right)^2 \rightarrow \min_{A, s_0}$$ identification



… 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 Identification of Dynamical Systems in Closed Form

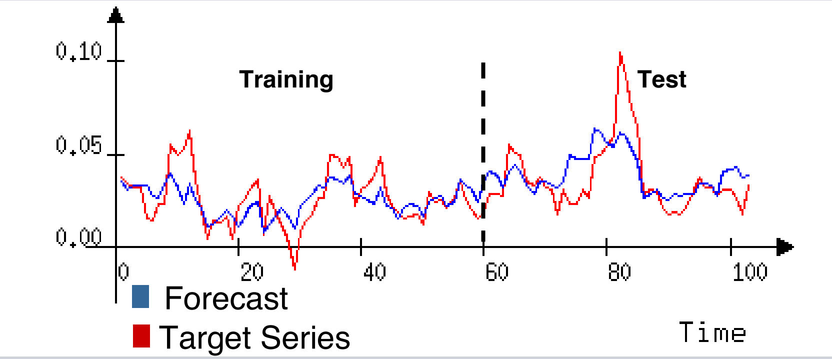
Embed the original architecture into a larger architecture, which is easier to learn. After the training, the extended architecture has to converge to the original model.

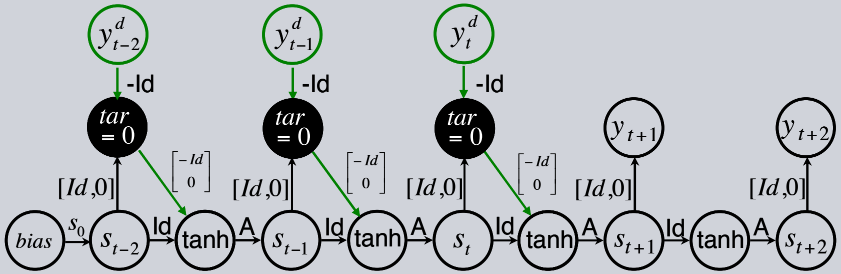


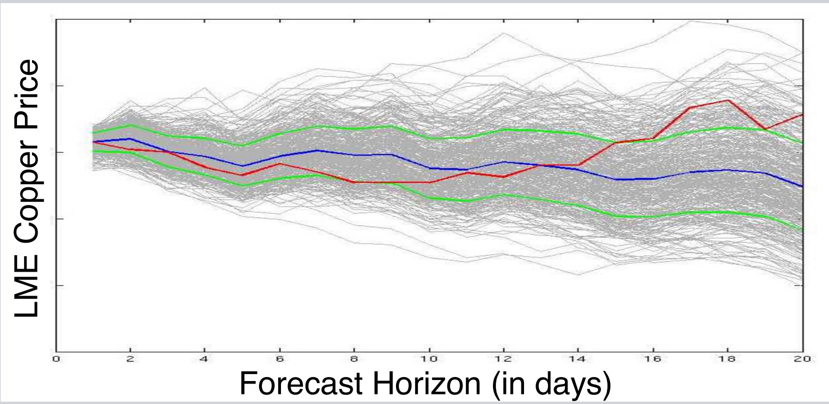
The essential task is NOT to reproduce the past observations, but to identify related hidden variables, which make the dynamics of the observables reasonable.

SIEMENS

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.

SIEMENS

Control of Dynamical Systems with Recurrent Neural Networks

Questions to Solve

System Identification

State Estimation

Controller Design

System Identification, State Estimation & Optimal Control with RNNs

$$s_{\tau+1} = \tanh\left(As_\tau + Bu_\tau\right)$$
$$y_\tau = Cs_\tau$$

system identification

$$\sum_{\tau \le t} \left(y_\tau - y_\tau^d\right)^2 \rightarrow \min_{A,B,C}$$

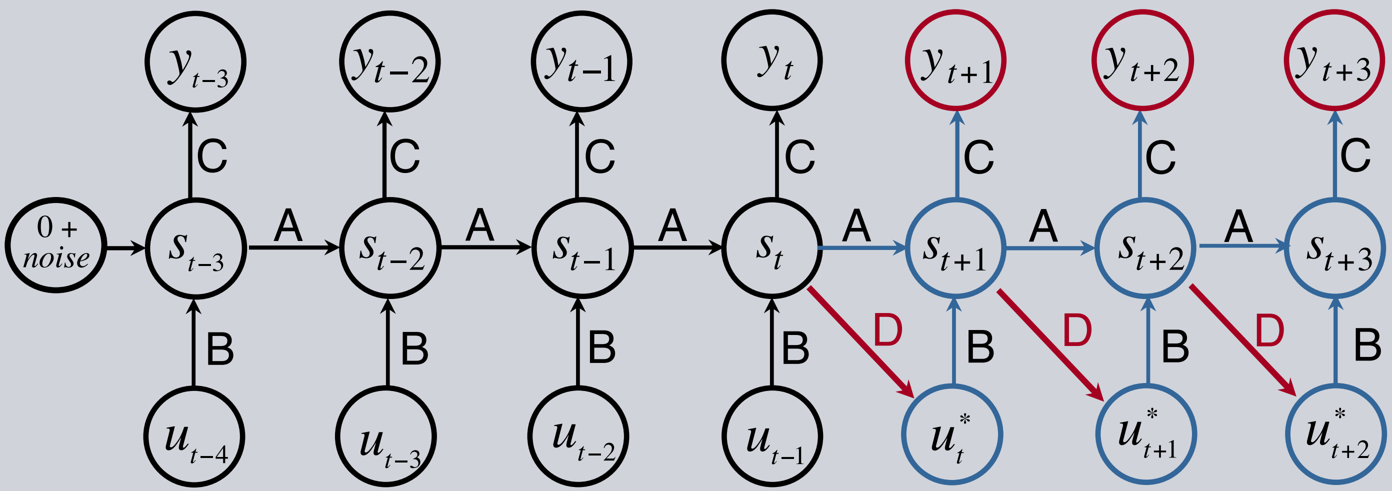
Unfold the RNN into the future, given A,B,C.

Learn a linear feedback controller: $u_\tau^* = u(s_\tau) = Ds_\tau$

$$s_{\tau+1} = \tanh\left(A\, s_\tau + B\, u_\tau(s_\tau)\right)$$
$$y_\tau = C\, s_\tau$$

normative target

$$\sum_{\tau > t} L(y_\tau) \rightarrow \min_{D}$$

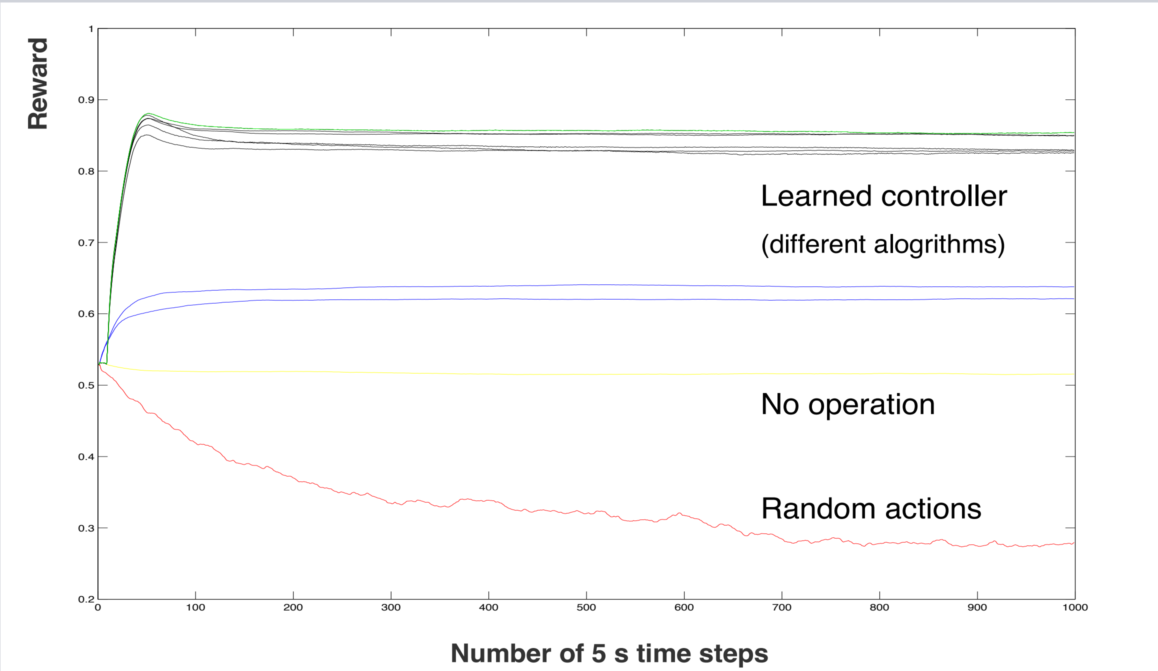


SIEMENS

Reward Maximization with a Learned Controller

Tasks:

System Identification
State Estimation
Controller Design



Challenges:

- High data volume (5000 var/sec)
- Streaming real-time analysis
- Autonomous learning
- Real-time data analytics using 1000 learned models
- Optimization of emissions
- Optimal turbine operation