LOCALLY CONNECTED ECHO STATE NETWORKS FOR TIME SERIES FORECASTING

Filip Matzner - filip.matzner@mff.cuni.cz
František Mráz - frantisek.mraz@mff.cuni.cz

Charles University, Faculty of Mathematics and Physics

ICLR 2025

RESEARCH OBJECTIVE

A Breath of Fresh Air

Recurrent model significantly different from the industry standard

State of the Art

Results competitive with deep feed-forward models

Adress two limitations of conventional ESN approaches

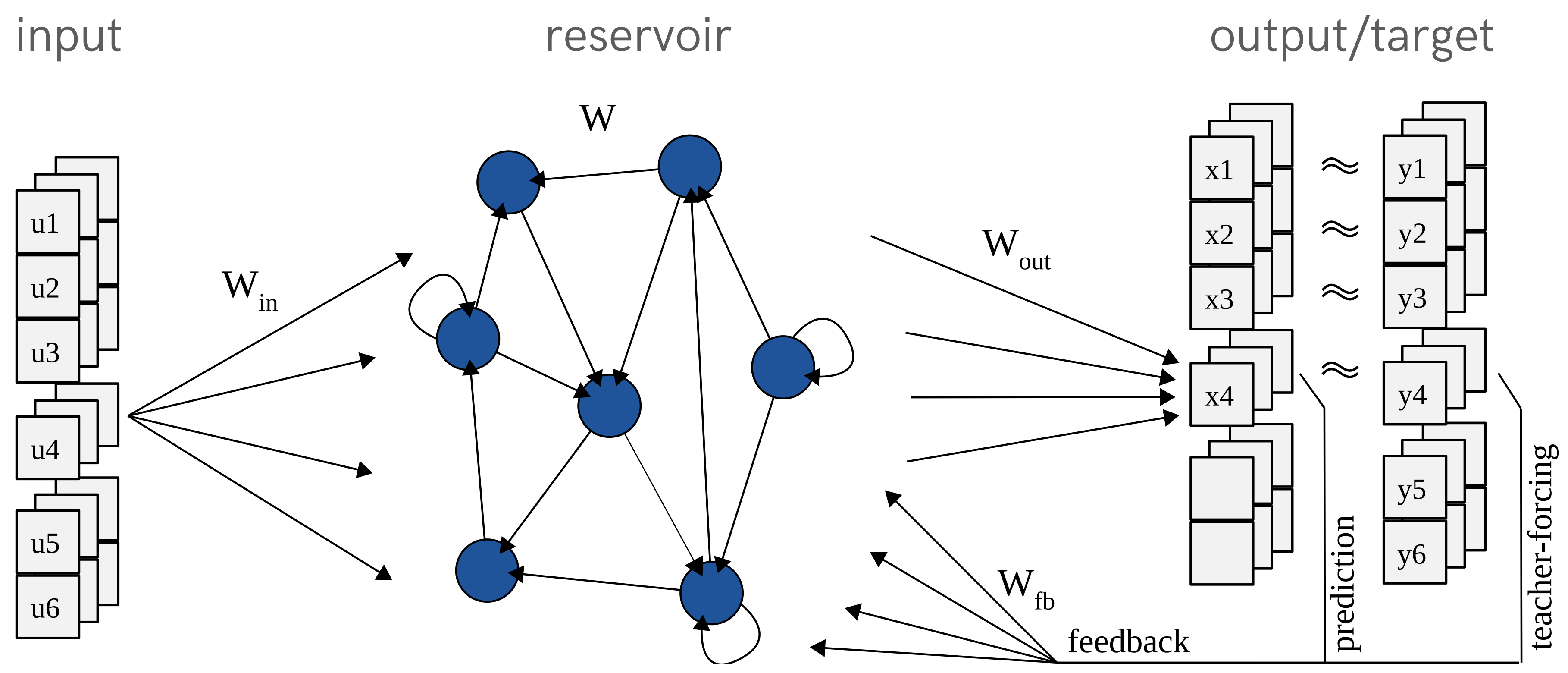
- Quadratic time and space complexity of ESN step
- Trade-off between memory capacity and network stability

ECHO STATE NETWORKS

A gradient-free model with a large random recurrent reservoir and a simple linear readout layer trained by linear regression.

The random reservoir represents a biologically plausible approach proving that a **sufficient number of random features makes the readout of any derived output easy**.

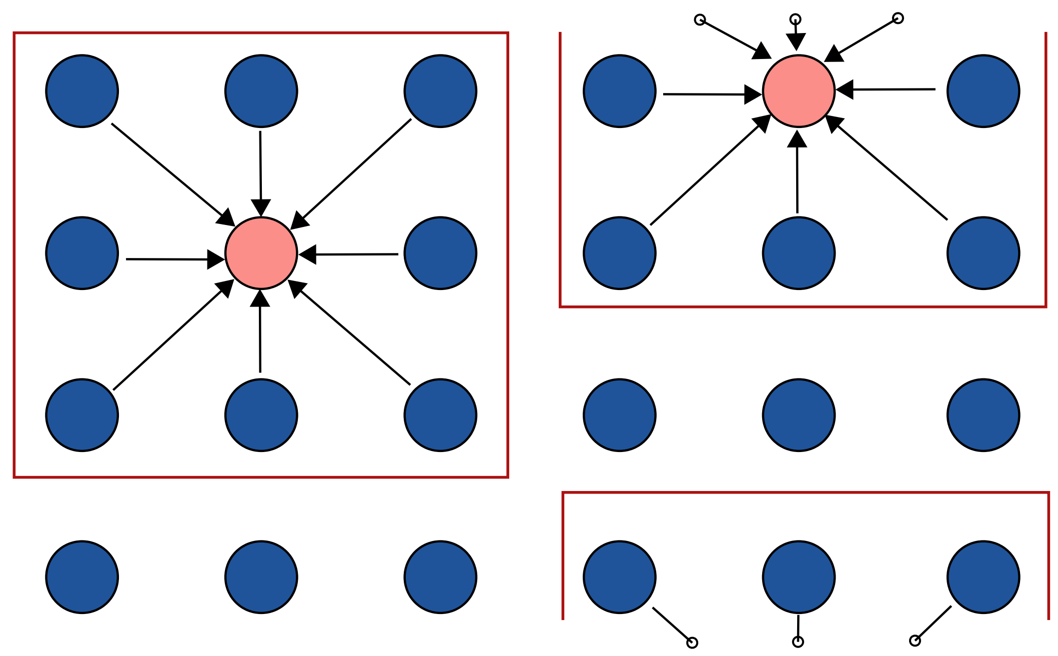
It is necessary to properly tune the hyperparameters (we used the CMA-ES framework from Matzner [2022])



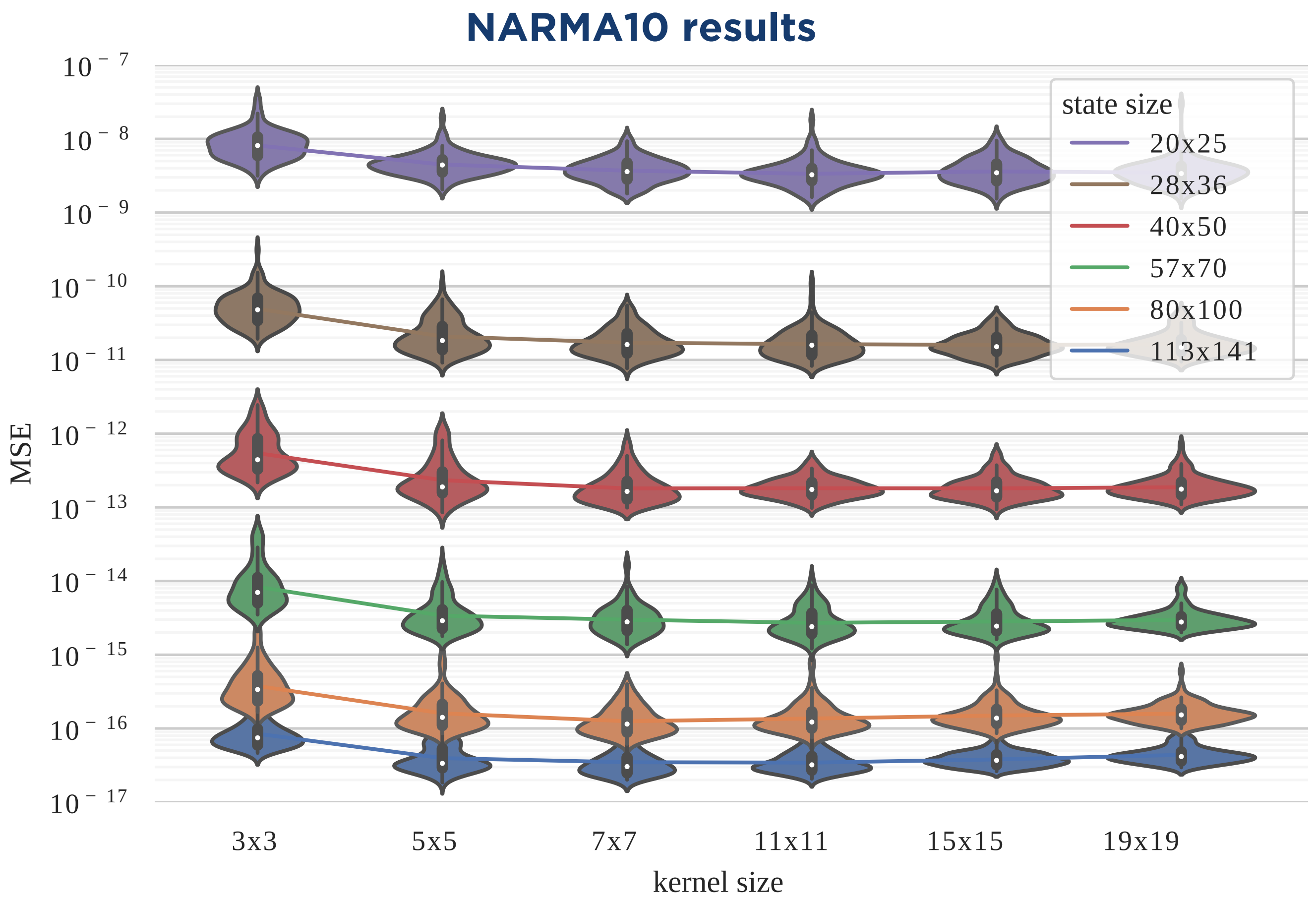
$$z(t) = Wa(t-1) + W_{in}u(t) + W_{fb}x(t-1) + \mu_b \,,$$
$$a(t) = (1-\gamma)a(t-1) + \tanh(z(t)) \,,$$
$$x(t) = W_{out}a(t) \,,$$

Matzner, Filip. Hyperparameter Tuning in Echo State Networks. In: GECCO '22: Proceedings of the Genetic and Evolutionary Computation Conference.

LOCALLY CONNECTED RESERVOIR



Limit the length of the connections to only reach the closest K neurons



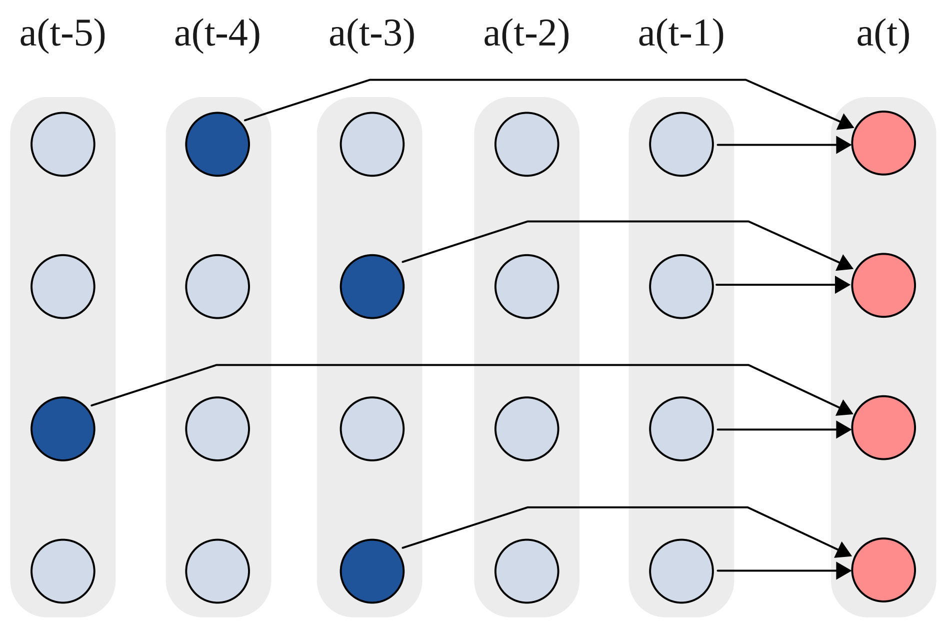
Reduced time and memory complexity of ESN step from O(N^2) to O(N*K) while retaining accuracy

Allows substantially larger reservoirs and an efficient GPU implementation

FORCED MEMORY

Each neuron's activation is an affine combination of its immediate past state and its historical state from time $t{-}h$

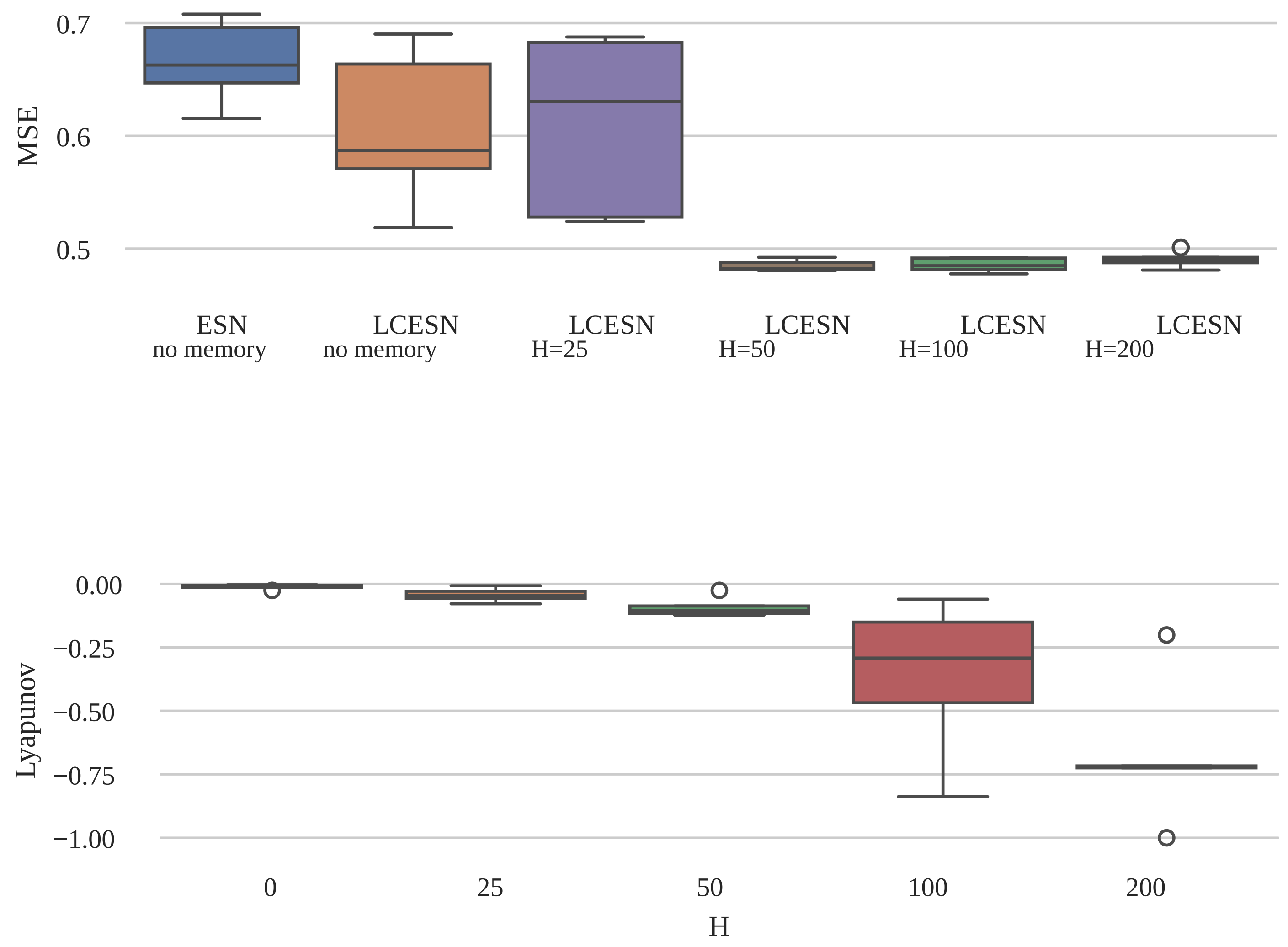
The delay h and the combination coefficients are randomly chosen for each neuron individually and remain fixed



Aligns with the random nature of ESNs and does not introduce additional hyperparameters that would need to be optimized

Avoids the need to propagate the entire memory through every step and stabilizes the network dynamics

Significantly improves results for benchmarks with long time dependencies, e.g., ETTm1



RESULTS

Competitive results on the longest of the tested datasets
Second place in full relative ranking (all horizons X all datasets)

Models	ESN (2001)		**LCESN** **Ours**		**LCESN-LMS** **Ours**		**LCESN-LR100** **Ours**		**LCESN-LR1** **Ours**		TSMixer (2023)		DLinear (2023)		PatchTST (2023)		iTransformer (2024)		FEDformer (2022)		TimesNet (2023)		RLinear (2023)		Autoformer (2021)	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	0.540	0.527	0.412	0.430	0.364	0.393	0.362	0.392	0.359	0.390	**0.347**	**0.375**	0.403	0.407	0.387	0.400	0.407	0.410	0.382	0.422	0.400	0.406	0.414	0.408	0.515	0.493
ETTm2	0.380	0.424	0.283	0.354	0.275	0.332	**0.257**	0.326	**0.255**	0.326	0.267	**0.322**	0.350	0.401	0.281	0.326	0.288	0.332	0.292	0.343	0.291	0.332	0.286	0.327	0.310	0.357
Weather	0.255	0.298	0.232	0.279	0.236	0.272	**0.221**	**0.264**	**0.221**	**0.264**	0.224	**0.264**	0.265	0.317	0.258	0.280	0.258	0.278	0.310	0.357	0.259	0.286	0.272	0.291	0.335	0.379
Solar	0.268	0.384	**0.202**	0.277	**0.202**	0.268	**0.201**	0.273	**0.201**	0.274	N/A	N/A	0.330	0.401	0.270	0.307	0.233	**0.262**	0.292	0.381	0.301	0.319	0.369	0.355	0.885	0.711
Electricity	0.264	0.351	0.225	0.329	0.197	0.308	0.184	0.294	0.183	0.293	**0.160**	**0.257**	0.212	0.300	0.205	0.290	0.178	0.270	0.207	0.321	0.192	0.295	0.218	0.298	0.214	0.326
Traffic	1.095	0.533	0.882	0.407	0.845	0.402	0.787	0.392	0.782	0.392	**0.408**	0.284	0.624	0.383	0.481	0.304	0.428	**0.282**	0.604	0.372	0.620	0.336	0.626	0.378	0.616	0.384
ETTh1	0.650	0.602	0.818	0.662	0.516	0.510	0.537	0.518	0.541	0.521	**0.412**	**0.428**	0.456	0.452	0.469	0.454	0.454	0.448	0.428	0.454	0.458	0.450	0.446	0.434	0.473	0.477
ETTh2	0.559	0.529	0.438	0.464	0.372	0.419	0.378	0.421	0.381	0.422	**0.355**	0.400	0.559	0.515	0.387	0.407	0.383	0.406	0.388	0.434	0.414	0.427	0.374	**0.398**	0.422	0.443
Exchange	0.703	0.575	0.889	0.661	0.718	0.602	0.656	0.574	0.639	0.566	N/A	N/A	**0.354**	0.414	0.366	0.404	0.360	**0.403**	0.518	0.429	0.416	0.443	0.378	0.418	0.613	0.539
# 1st	0/9	0/9	1/9	0/9	1/9	0/9	3/9	1/9	3/9	1/9	**5/7**	**5/7**	1/9	0/9	0/9	0/9	0/9	3/9	0/9	0/9	0/9	0/9	0/9	1/9	0/9	0/9

CONCLUSION

RNN Refresh

Echo State Networks (ESN) is an overlooked TSF baseline approach significantly different from the industry standard

Addressed Two ESN Limitations

- Local topology allows for much larger networks at the same computational costs
- Forced memory facilitates longer time dependencies and enhances network stability

Locally Connected Echo State Network

We introduced Locally Connected Echo State Network (LCESN) reaching state-of-the-art results

Open Source Implementation

We have released our C++/CUDA implementation of LCESN under a permissive license at our GitHub