August 3-7, 2025
KDD2025

CrossLinear: Plug-and-Play Cross-Correlation Embedding for Time Series Forecasting with Exogenous Variables

Hello!

Pengfei ZHOU, Yunlong LIU, Junli LIANG, Qi SONG* & Xiangyang LI

University of Science and Technology of China

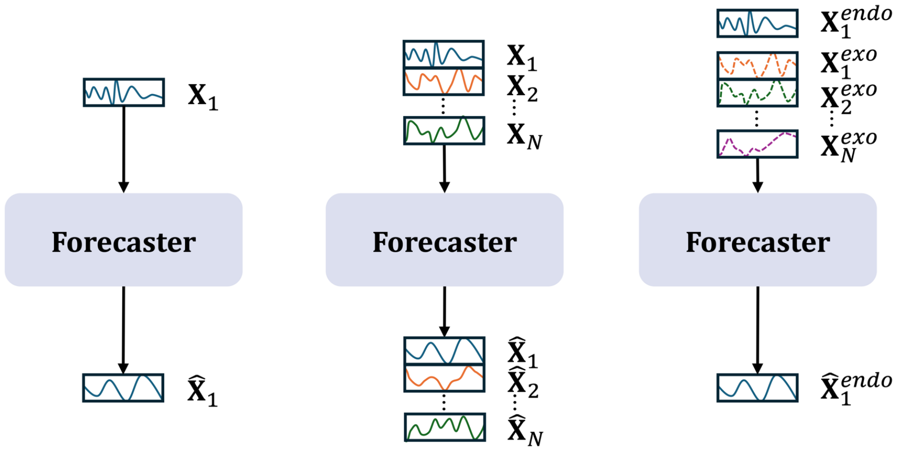
Email: pengfeizhou@mail.ustc.edu.cn (Pengfei ZHOU)

Github link: https://github.com/mumiao2000/CrossLinear
OR scan this →

Background:
What & Why is Forecasting with Exogenous Variables?

Two traditional forecasting paradigms:
- Univariate forecasting (one-to-one);
- Multivariate forecasting (many-to-many).

A practical example:
Traffic volume is often influenced by
- Holidays;
- Weather conditions;
- Other *exogenous* factors.



(a) Univariate (b) Multivariate (c) Ours

A critical emerging paradigm:
- *Forecasting with exogenous variables* (many-to-one);
- Has aroused interest in both statistics area and machine learning area.
  - ARIMAX, TimeXer, ChronosX, *etc*.

Background:
Existing Methods & Problems

Start

Background

Insight

Methodology

Results

Conclusion

End

Channel (*aka*. variable) independent or dependent?

- **Channel independent:**
  - *Implicit* modeling of variable dependencies;
  - Computationally *efficient*;
- **Channel dependent:**
  - *Explicit* modeling of variable dependencies;
  - Mathematically *powerful*.

Problems:

- **Channel independent:**
  - *Performance ceiling* due to no explicit variable dependency modeling;
- **Channel dependent:**
  - *Overfitting tendency* due to a mismatch between modeling complexity and dataset sparsity.

Insight:
Direct & Time-Invariant Variable Dependencies

Existing problems:

- Variable dependencies are complex, and datasets are sparse.

A simple and straightforward approach to modeling variable dependencies is needed.

Intuitive assumption:

- Only *direct* and *time-invariant* variable dependencies are critical.
- **Direct:**
  - Exo A directly affects Endo (STRONG);
  - Exo A affects Exo B which in turn affects Endo (WEAK).
- **Time-invariant:**
  - Exo A affects Endo uniformly across various time points (STRONG);
  - Exo A affects Endo differently across various time points (WEAK).

Practical solution:

- 1D convolution is perfect for capturing direct (layer = 1) and time-invariant (kernel = 3) variable dependencies.

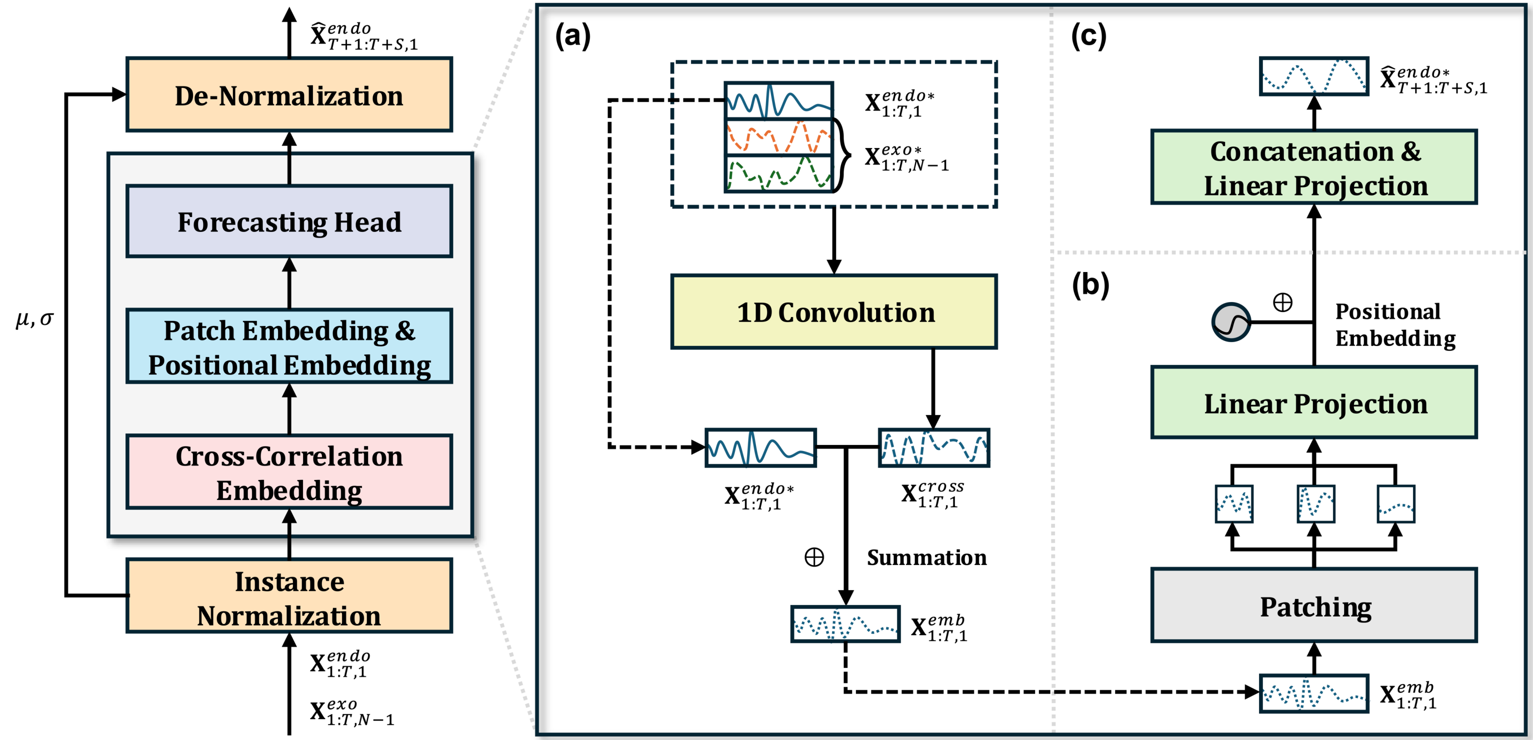
Methodology: Variable & Temporal Dependencies Modeling

Backbone of CrossLinear:

(a) Cross-correlation embedding: to capture *variable* dependencies;

(b) Patch embedding: to extract *short-term temporal* dependencies,

(c) Forecasting head: to model *long-term temporal* dependencies.



August 3-7, 2025
KDD2025

Results:
Main Results

Start

Background

Insight

Methodology

Results

Conclusion

End

Forecasting with exogenous variables:
- Top rankings in *30 cases* for MSE and *29 cases* for MAE;

Multivariate forecasting:
- Top rankings in *31 cases* for MSE and *28 cases* for MAE.

Model		CrossLinear		TimeXer		iTransformer		MSGNet		SparseTSF		RLinear		PatchTST	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Many-to-one	ECL	**0.323**	**0.405**	0.327	0.408	0.365	0.442	0.387	0.468	0.372	0.431	0.444	0.486	0.394	0.446
	Weather	0.002	0.031	0.002	0.031	0.002	0.031	0.002	0.031	0.002	0.036	**0.002**	**0.029**	0.002	0.031
	ETTh1	**0.072**	**0.208**	0.073	0.209	0.075	0.211	0.076	0.213	0.079	0.224	0.084	0.224	0.078	0.215
	ETTh2	**0.186**	**0.339**	0.189	0.342	0.199	0.352	0.195	0.350	0.194	0.350	0.205	0.356	0.192	0.345
	ETTm1	**0.052**	**0.170**	**0.052**	0.171	0.053	0.175	0.054	0.173	0.053	0.174	0.053	0.173	0.053	0.173
	ETTm2	**0.118**	**0.254**	0.120	0.258	0.127	0.267	0.128	0.270	0.122	0.260	0.122	0.261	0.120	0.258
	Traffic	**0.151**	**0.230**	0.156	0.234	0.161	0.246	0.213	0.313	0.194	0.274	0.324	0.412	0.173	0.253
	EPF	**0.303**	0.266	0.307	**0.265**	0.335	0.289	0.349	0.299	0.351	0.305	0.412	0.339	0.330	0.282
Many-to-many	ECL	0.174	**0.270**	**0.171**	**0.270**	0.178	**0.270**	0.194	0.300	0.213	0.287	0.219	0.298	0.216	0.304
	Weather	**0.238**	**0.269**	0.241	0.271	0.258	0.279	0.249	0.278	0.261	0.293	0.272	0.291	0.259	0.281
	ETTh1	**0.431**	**0.433**	0.437	0.437	0.454	0.447	0.452	0.452	0.439	**0.433**	0.446	0.434	0.469	0.454
	ETTh2	0.368	**0.396**	**0.367**	**0.396**	0.383	0.407	0.396	0.417	0.390	0.411	0.374	0.398	0.387	0.407
	ETTm1	**0.370**	**0.393**	0.382	0.397	0.407	0.410	0.398	0.411	0.403	0.408	0.414	0.407	0.387	0.400
	ETTm2	**0.272**	**0.320**	0.274	0.322	0.288	0.332	0.288	0.330	0.280	0.325	0.286	0.327	0.281	0.326
	Traffic	0.483	0.298	0.466	0.287	**0.428**	**0.282**	0.597	0.346	0.588	0.344	0.626	0.378	0.481	0.304
	EPF	**0.171**	**0.210**	0.175	0.216	0.189	0.227	0.210	0.268	0.202	0.247	0.205	0.258	0.183	0.229

Start

Background

Insight

Methodology

Results

Conclusion

End

Results: Introducing Cross-Correlation Embedding to Other Models

Cross-correlation embedding module is plug-and-play.

We validate its broad applicability by incorporating it to 7 forecasting models:
- **MLP-based:** SparseTSF, RLinear and DLinear;
- **Transformer-based:** PatchTST, Autoformer and iTransformer;
- **GNN-based:** MSGNet.

Up to 15.5% MSE reduction (RLinear on EPF Dataset).

Model		SparseTSF		Rlinear		Dlinear		PatchTST		Autoformer		iTransformer		MSGNet	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ECL	Ori.	0.372	0.431	0.444	0.486	0.393	**0.457**	0.394	0.446	0.495	0.528	0.365	0.442	0.387	0.468
	+ Emb.	**0.350**	**0.420**	**0.406**	**0.473**	**0.383**	0.463	**0.359**	**0.427**	**0.469**	**0.514**	**0.359**	**0.439**	**0.378**	**0.456**
ETTh1	Ori.	0.079	0.224	0.084	0.224	0.116	0.259	0.078	0.215	0.130	0.282	0.075	0.211	**0.076**	0.213
	+ Emb.	**0.078**	**0.221**	**0.080**	**0.218**	**0.113**	**0.256**	**0.077**	**0.214**	**0.114**	**0.264**	**0.074**	**0.210**	**0.076**	**0.211**
EPF	Ori.	0.351	0.305	0.412	0.339	0.366	0.314	0.330	**0.282**	0.453	0.368	**0.335**	**0.289**	0.349	0.299
	+ Emb.	**0.343**	**0.301**	**0.348**	**0.304**	**0.364**	**0.313**	**0.328**	**0.282**	**0.426**	**0.352**	0.345	0.292	**0.313**	**0.273**

KDD2025 August 3-7, 2025

Results:
Model Analysis

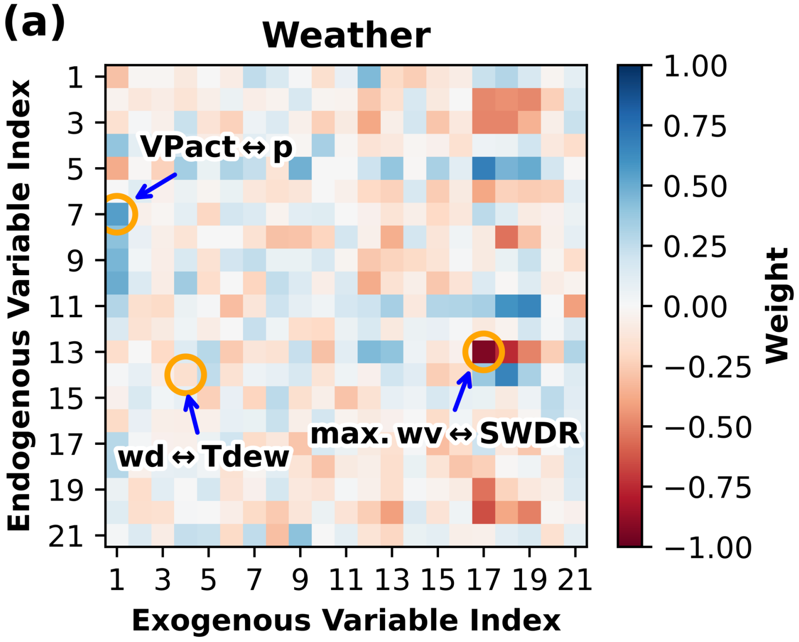
- Start
- Background
- Insight
- Methodology
- **Results**
- Conclusion
- End

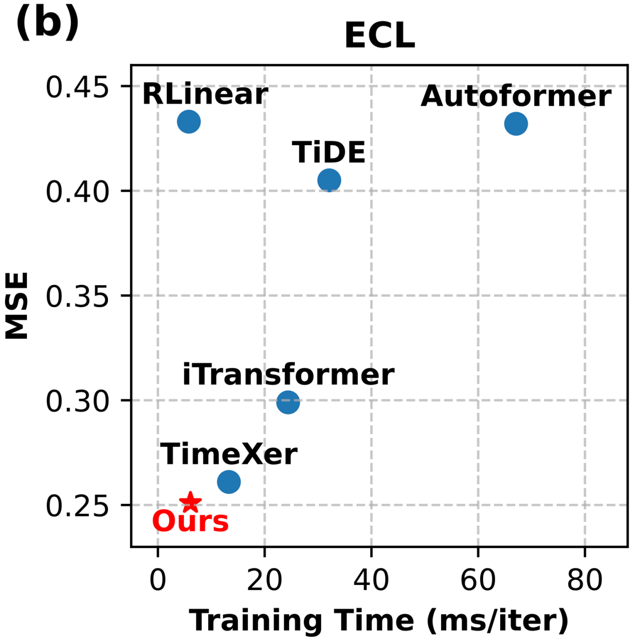
(a) Variate-wise correlation analysis:

- The weight matrix roughly indicates variable dependencies;

(b) Model efficiency analysis:

- Linear computational complexity strikes a balance between efficiency and effectiveness.





Conclusion:
Advantages & Further Work

In this work, we introduce CrossLinear, an advanced model specifically designed to improve time series forecasting with exogenous variables. The core innovation of CrossLinear is a flexible plug-and-play cross-correlation embedding module. Through a series of comprehensive experiments conducted on multiple real-world datasets, we demonstrate that CrossLinear outperforms existing methods in most cases.

Advantages:

- Satisfactory performance;
- Plug-and-play module;
- Versatility across various paradigms.

Further work:

- Exploration of its applicability to non-forecasting tasks;
- Validation on more practical scenarios.

CrossLinear: Plug-and-Play Cross-Correlation Embedding for Time Series Forecasting with Exogenous Variables

Start
Background
Insight
Methodology
Results
Conclusion
End

Thank you!

Pengfei ZHOU, Yunlong LIU, Junli LIANG, Qi SONG* & Xiangyang LI

University of Science and Technology of China

Email: pengfeizhou@mail.ustc.edu.cn (Pengfei ZHOU)

Github link: https://github.com/mumiao2000/CrossLinear
OR scan this →