Enhancing Time Series Forecasting Models under Concept Drift by Data-centric Online Ensembling

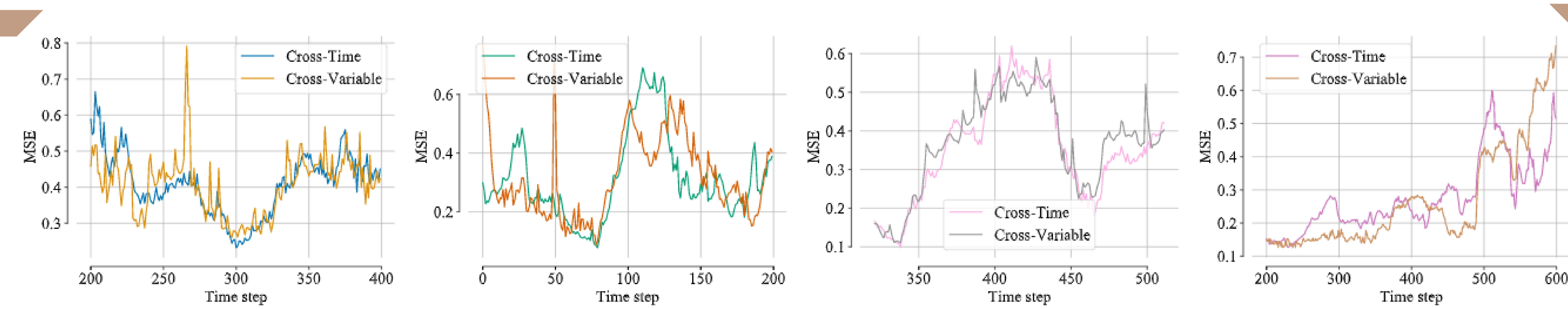
Yi-Fan Zhang, Qingsong Wen, Xue Wang, Weiqi Chen, Liang Sun, Zhang Zhang, Liang Wang, Rong Jin, Tieniu Tan

Online Time Series Forecasting

1. **Concept drift:** Concepts in the real world are often dynamic and can change over time, which is especially true for scenarios like weather prediction and customer preferences.
2. **Online time series forecasting** is a widely used technique in real-world due to the sequential nature of the data and the frequent drift of concepts. In this approach, the learning process takes place over a sequence of rounds, where the model receives a look-back window and predicts the forecast window. The true values are then revealed to improve the model's performance in the next rounds.

Existing Models

1. **Cross-time dependency** predicting each variable independently, which is crucial for the robustness of models under concept drift.
2. **Cross-variable dependency:** for a specific variable, information from associated series in other variables may improve forecasting results.



(a) ETTh2. (b) ETTm1. (c) WTH. (d) ECL.

Figure 1. **A motivating example for online ensembling**, where the reported metric is MSE and forecast horizon length is set to 48 during online adaptation. Cross-Time refers to a TCN backbone that assumes independence among covariates and only models the temporal dependence, and cross-variable refers to a TCN backbone that takes into cross-variable dependence.

Main Intuitions of DoneNet

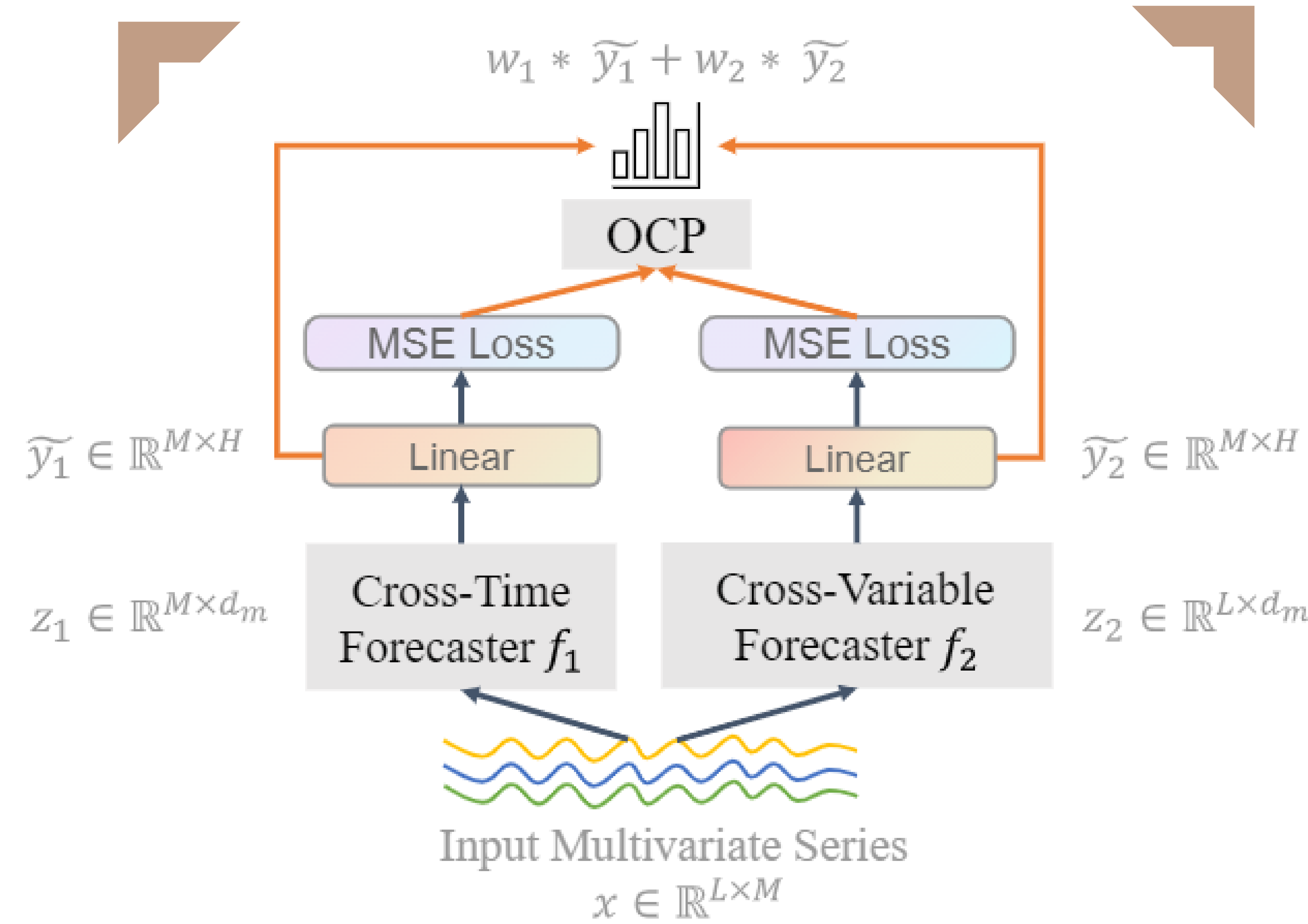
1. A large fluctuation exists in MSE over online adaption, indicating a significant concept drift over time.
2. Neither of these two kinds of models performs consistently better than the other, indicating that neither of the two data assumptions holds true for the entire time series

Main Results

DoneNet reduces the cumulative mean-squared errors (MSE) by 53.1% and mean-absolute errors (MAE) by 34.5%. In particular, the performance gain on challenging dataset ECL is superior, where the MSE is reduced by 59.2% and MAE is reduced by 63.0%

Table 1. **A motivating example for online ensembling**, where the reported metric is MSE and the forecast horizon length is set as 48. Cells are colored from low (red) to medium (white) to high (blue). All methods use the same training and online adaptation strategy.

Dataset	#Variables	Cross-Variable		Cross-Time			Both			Ours
		TCN	FSNet	Time-TCN	DLinear	PatchTST	CrossFormer	TS-Mixer	Fedformer	
ETTh2	7	0.910	0.846	1.307	6.910	2.716	5.772	3.060	1.620	0.609
ETTm1	7	0.250	0.127	0.308	1.120	0.553	0.370	0.660	0.516	0.108
WTH	21	0.348	0.223	0.308	0.541	0.465	0.317	0.482	0.372	0.200
ECL	321	10.800	7.034	5.230	7.388	5.030	94.790	5.764	27.640	2.201



(a) **The overall DoneNet architecture.**

DoneNet

1. **Two-stream forecasters**. The input multivariate time series data is fed into two separate forecasters, a cross-time forecaster and a cross-variable forecaster.
2. **Online convex programming**. Use an improvement of Exponentiated Gradient Descent (EGD) to learn prediction combination weights.
3. **Decoupled training strategy**. Training the OCP block and two-stream forecasters respectively.