

A THEORETICAL ANALYSIS OF ReLD

As shown in Figure 3 of main paper, we can observe LD values oscillating within a certain range similarly to original time series. We discuss the following points about this observation.

1. When the time series is sampled from a periodic function, LD is also a periodic function.
2. When the time series is sampled from a bounded periodic function, LD is bounded.

First, we can easily prove that LD is also periodic when the time series is sampled from a periodic function, which the model should learn from data (i.e., normal states).

Theorem 1. *If f is a periodic function that satisfies $f(t) = f(t + p)$,*

$$LD(a, a + L) = \frac{m(a) - m(a + L)}{\sqrt{\frac{s(a)}{N} + \frac{s(a+L)}{N}}} \quad (3)$$

is also a periodic function with period p , where $m(a) = \frac{1}{N} \sum_{t \in I(a)} f(t)$, $s(a) = \frac{1}{N} \sum_{t \in I(a)} (f(t) - m(a))^2$, and $I(a) = \{a + \frac{L}{N} \cdot i\}_{i=0}^{N-1}$ for range $[a, a + L]$ and sampling interval L/N .

To prove Theorem 1, we prove and use Proposition 1 and 2 with respect to the mean and the variance of the periodic function.

Proposition 1. *If f is a periodic function that satisfies $f(t) = f(t + p)$, $m(a) = \frac{1}{N} \sum_{t \in I(a)} f(t)$ is also a periodic function with period p where $I(a) = \{a + \frac{L}{N} \cdot i\}_{i=0}^{N-1}$ for range $[a, a + L]$ and sampling interval L/N .*

Proposition 2. *If f is a periodic function that satisfies $f(t) = f(t + p)$, $s(a) = \frac{1}{N} \sum_{t \in I(a)} (f(t) - m(a))^2$ is also a periodic function with period p where $I(a) = \{a + \frac{L}{N} \cdot i\}_{i=0}^{N-1}$ for range $[a, a + L]$ and sampling interval L/N .*

By proposition 1 and 2, we prove Theorem 1 as follows:

$$\begin{aligned} LD(a + p, a + p + L) &= \frac{m(a + p) - m(a + p + L)}{\sqrt{\frac{s(a+p)}{N} + \frac{s(a+p+L)}{N}}} \\ &= \frac{m(a) - m(a + L)}{\sqrt{\frac{s(a)}{N} + \frac{s(a+L)}{N}}} \\ &= LD(a, a + L) \end{aligned}$$

Regarding the bound of LD, if we define LD by ignoring the variances (i.e., $LD(a, a + L) = m(a) - m(a + L)$), we can obtain the bound $2 \cdot B$ given that f is bounded function that satisfies $|f(t)| \leq B$ for all t .

However, since we use the variance of input and output, LD can diverge when both variances of input sequence and output sequence are equal to zero. There are two cases when the variance equals to zero.

1. f is a constant function.
2. The window size L is $N \cdot p$ for data points, which are sampled from a periodic function f with the period p and sampling interval L/N .

In the first case, since the time series dataset has a constant target value, the prediction also remains as the constant value, leading to a trivial solution. In the second case, the variance is no longer zero if the window size L is adjusted.

In practice, we use epsilon ϵ as a numerical stabilizer to solve the case where variances are zero as shown in Equation 1. Note that we do not use these bounds as thresholds in the proposed method.

Proofs for the Proposition 1 and 2 are as follows.

Proposition 1. *If f is periodic function that satisfy $f(t) = f(t + p)$, $m(a) = \frac{1}{N} \sum_{t \in I(a)} f(t)$ is also periodic function with period p where $I(a) = \{a + \frac{L}{N} \cdot i\}_{i=0}^{N-1}$ for range $[a, a + L]$ and sampling interval L/N .*

Proof:

$$\begin{aligned} m(a + p) &= \frac{1}{N} \sum_{t \in I(a+p)} f(t) \\ &= \frac{1}{N} \left\{ f(a + p) + f(a + p + \frac{L}{N}) + \cdots + f(a + p + L - \frac{L}{N}) \right\} \\ &= \frac{1}{N} \left\{ f(a) + f(a + \frac{L}{N}) + \cdots + f(a + L - \frac{L}{N}) \right\} \\ &= m(a) \end{aligned}$$

Proposition 2. *If f is periodic function that satisfy $f(t) = f(t + p)$, $s(a) = \frac{1}{N} \sum_{t \in I(a)} (f(t) - m(a))^2$ is also periodic function with period p where $I(a) = \{a + \frac{L}{N} \cdot i\}_{i=0}^{N-1}$ for range $[a, a + L]$ and sampling interval L/N .*

Proof:

$$\begin{aligned} s(a + p) &= \frac{1}{N} \sum_{t \in I(a+p)} \{f(t) - m(a + p)\}^2 \\ &= \frac{1}{N} \left\{ (f(a + p) - m(a + p))^2 + (f(a + p + \frac{L}{N}) - m(a + p))^2 \right. \\ &\quad \left. + \cdots + (f(a + p + L - \frac{L}{N}) - m(a + p))^2 \right\} \\ &= \frac{1}{N} \left\{ (f(a) - m(a))^2 + (f(a + \frac{L}{N}) - m(a))^2 + \cdots + (f(a + L - \frac{L}{N}) - m(a))^2 \right\} \\ &= s(a) \end{aligned}$$

B IN-DEPTH ANALYSIS ON ReLD

This section provides various analysis for our reweighting framework.

B.1 ABRUPT CHANGE WITH EXTERNAL VARIABLES

We assumed that an abrupt change can be caused by unobserved and external events as we mentioned in Section 1 and Section 5. If the abrupt change can be predicted using an external variable, down-weighting the loss of the abrupt change would get in the way of learning such correlation for the model. However, utilizing additional variables without thorough verification causes the model to learn a spurious correlation between variables, which worsens the generalization ability. Moreover, some abrupt changes have unknown causes (e.g., sensor malfunction), which cannot be addressed by simply collecting external variables. In fact, as shown in the Table 5, we observed that training baseline models with external variables (i.e., Multivariate to Univariate denoted as Mul2Uni setting) rather shows lower performance than training those with the target time series only (i.e., Univariate to Univariate denoted as Uni2Uni setting). These results indicate that simply adding covariates does not guarantee performance gains. Note that multivariate forecasting we mentioned in main paper is multivariate to multivariate setting (i.e., Mul2Mul), which is different with Mul2Uni setting. In addition, applying our method on Mul2Uni outperformed the Uni2Uni in several cases (see Autoformer 96/96 and 336/168 of Table 5).

B.2 ReLD ON REPEATED CHANGES

To further understand our ReLD, we present a rectangular time series as a special case, which generally includes a large shift during a short period of time and shows increasing amplitude (see 1st row

Table 5: Comparison between Mul2Uni and Uni2Uni forecasting on ETTm1 dataset.

Setting	I \rightarrow O	Pyraformer			Autoformer		
		96 \rightarrow 96	336 \rightarrow 168	336 \rightarrow 336	96 \rightarrow 96	336 \rightarrow 168	336 \rightarrow 336
Uni2Uni		0.0821 \pm 0.0289	0.1286 \pm 0.0346	0.1941 \pm 0.0523	0.0577 \pm 0.0081	0.0881 \pm 0.0284	0.0903 \pm 0.0096
Uni2Uni + Ours		0.0576 \pm 0.0079	0.1218 \pm 0.0395	0.1843 \pm 0.0507	0.0522 \pm 0.0035	0.0723 \pm 0.0068	0.0847 \pm 0.0084
Mul2Uni		0.1757 \pm 0.0372	0.2926 \pm 0.0778	0.5920 \pm 0.0591	0.0619 \pm 0.0090	0.0799 \pm 0.0188	0.1367 \pm 0.0392
Mul2Uni + Ours		0.1137 \pm 0.0295	0.2984 \pm 0.1246	0.5533 \pm 0.0761	0.0496 \pm 0.0021	0.0674 \pm 0.0071	0.1193 \pm 0.0267

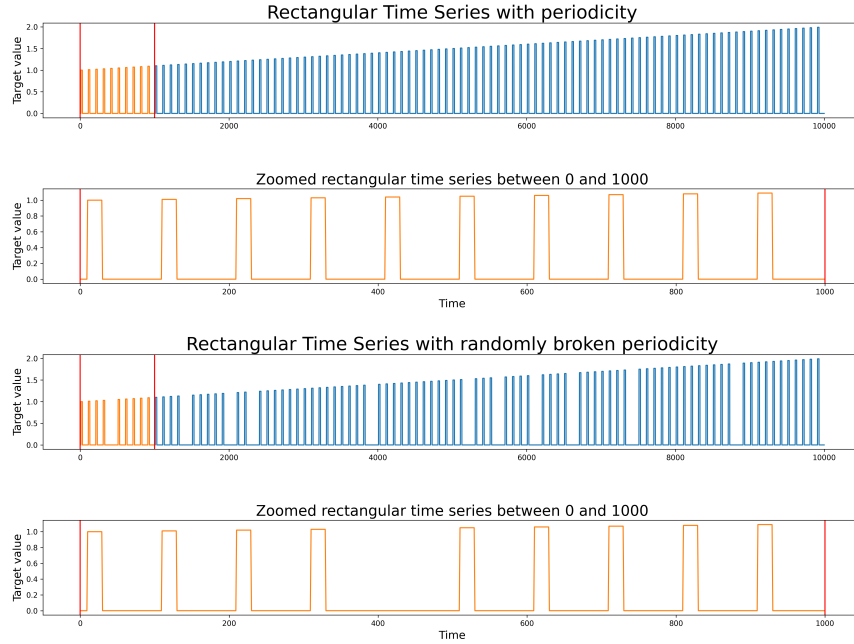


Figure 6: Two rectangular time series which include large shift in a short time. As in the first row, if a rectangular pattern with a large change in time exists several times, ReLD learns it normally without down-weighting it. However, if the periodicity is broken in the third row, ReLD mitigates impact of anomaly pattern in training phase.

of Figure 6). Although this series includes large shifts, we do not regard those as abrupt changes defined in our paper since rectangular patterns are repeated (i.e., seasonal component). Also, the increasing amplitude (i.e., trend component) is considered one of the trend types. Since we calculate LD by sliding the window, the increasing amplitude does not change the LD values. For example, the LD value of the window, which of size is large enough to cover period, has a value less than 0 (greater than 0 if the amplitude decreases) regardless of time. In this case, since the LD values of all windows are similar, they will be given the same weights. Therefore, even if our method is applied, we would observe more or less the same performance as shown in Rect-Normal dataset of Table 7.

Additionally, we conducted experiments by removing rectangles randomly from the dataset (see 3rd row of Figure 6). This can be considered abrupt changes (e.g., broken periodicity). We observe that our ReLD brings performance gain in such cases (see Rect-Broken of Table 7). This again demonstrates that our proposed method promotes the model to be robust to abrupt changes.

B.3 IMPACT OF THE IN-OUTPUT RATIO

We conducted an experiment by fixing the output length and changing the input length from 48 to 720 to explore the performance change according to the I/O ratio. We conducted experiments on the three datasets, ETTh1, ETTh2, and ETTm1. Applying our method brings consistent performance improvements, although there exists different performance gains depending on the input lengths.

Table 6: Rectangular Time Series with the increasing amplitude and the randomly broken periodicity.

Models		Pyraformer		Autoformer		Informer	
MSE		base	our	base	our	base	our
Rect-Normal	96	0.2405 \pm 0.0211	0.2468 \pm 0.0243	0.9748 \pm 0.4805	0.9131 \pm 0.4100	0.5409 \pm 0.0280	0.5480 \pm 0.0316
	168	0.2614 \pm 0.0115	0.2652 \pm 0.0102	1.4711 \pm 0.9902	1.7147 \pm 0.7723	1.3573 \pm 0.1377	1.3556 \pm 0.1068
	336	0.3179 \pm 0.0075	0.3193 \pm 0.0062	0.5271 \pm 0.2047	0.4492 \pm 0.1558	1.2945 \pm 0.0874	1.2764 \pm 0.0891
	720	0.4034 \pm 0.0084	0.4088 \pm 0.0089	2.7076 \pm 0.4390	2.6447 \pm 0.6276	1.6796 \pm 0.0450	1.7306 \pm 0.0614
	Imp.	1.46%		-1.72%		0.71%	
Rect-Broken	96	0.4028 \pm 0.0434	0.2343 \pm 0.0061	0.9883 \pm 0.1790	1.1399 \pm 0.4871	0.5781 \pm 0.0558	0.3912 \pm 0.0463
	168	0.3261 \pm 0.0138	0.3020 \pm 0.0174	1.7361 \pm 0.8130	1.2914 \pm 0.3595	0.5622 \pm 0.0386	0.5227 \pm 0.0178
	336	0.2548 \pm 0.0112	0.2579 \pm 0.0155	0.7678 \pm 0.1225	0.5648 \pm 0.1603	0.5144 \pm 0.0394	0.4988 \pm 0.0494
	720	0.3098 \pm 0.0617	0.3037 \pm 0.0427	2.0861 \pm 0.1504	1.8670 \pm 0.3713	0.5936 \pm 0.0473	0.5569 \pm 0.0390
	Imp.	-12.49%		-11.80%		-12.14%	

Table 7: Impact of the ratio I/O on multivariate time series forecasting.

Models		Pyraformer		Autoformer		Informer		Imp.
Output-96	Input	base	our	base	our	base	our	
ETTh1	48	0.6314 \pm 0.0371	0.5166 \pm 0.0104	0.4748 \pm 0.0328	0.4675 \pm 0.0504	1.0632 \pm 0.2707	0.8125 \pm 0.0679	-23.58%
	96	0.6453 \pm 0.0583	0.5345 \pm 0.0073	0.4531 \pm 0.0282	0.4452 \pm 0.0153	0.9075 \pm 0.0479	0.8476 \pm 0.0532	-6.6%
	168	0.6330 \pm 0.0241	0.5604 \pm 0.0145	0.4477 \pm 0.0247	0.4594 \pm 0.0426	0.8997 \pm 0.0738	0.7928 \pm 0.0698	-11.88%
	336	0.7195 \pm 0.0206	0.6310 \pm 0.0293	0.4667 \pm 0.0240	0.4826 \pm 0.0188	1.1695 \pm 0.2012	1.0384 \pm 0.1830	-11.21%
	720	0.7290 \pm 0.0757	0.6540 \pm 0.0191	0.6354 \pm 0.0386	0.5057 \pm 0.0673	1.6608 \pm 0.1402	1.3866 \pm 0.1341	-16.51%
ETTh2	48	1.5411 \pm 0.1880	1.0982 \pm 0.1702	0.3637 \pm 0.0091	0.3394 \pm 0.0051	1.7225 \pm 0.1508	1.1461 \pm 0.0743	-22.96%
	96	1.6090 \pm 0.0866	1.1733 \pm 0.2271	0.3731 \pm 0.0294	0.3464 \pm 0.0102	3.4245 \pm 0.4814	2.4505 \pm 0.4804	-20.89%
	168	1.7787 \pm 0.2003	1.3081 \pm 0.2461	0.4414 \pm 0.0271	0.3833 \pm 0.0069	5.6370 \pm 0.8005	2.9705 \pm 0.4835	-28.97%
	336	1.7924 \pm 0.2872	1.5560 \pm 0.1676	0.4897 \pm 0.0565	0.4174 \pm 0.0517	6.2992 \pm 0.9310	3.9496 \pm 0.8050	-21.75%
	720	2.0959 \pm 0.1960	1.9368 \pm 0.2612	0.6769 \pm 0.1552	0.4701 \pm 0.0869	9.1387 \pm 2.0638	6.9792 \pm 1.5690	-20.59%
ETTm1	48	0.5559 \pm 0.0225	0.4922 \pm 0.0134	0.5673 \pm 0.0542	0.5182 \pm 0.0396	0.6389 \pm 0.0270	0.5925 \pm 0.0348	-9.13%
	96	0.5364 \pm 0.0318	0.4713 \pm 0.0299	0.5128 \pm 0.0635	0.4545 \pm 0.0410	0.6438 \pm 0.0596	0.5367 \pm 0.0439	-13.38%
	168	0.5015 \pm 0.0431	0.4174 \pm 0.0176	0.4987 \pm 0.0241	0.4603 \pm 0.0636	0.6907 \pm 0.0578	0.5677 \pm 0.0281	-14.09%
	336	0.4876 \pm 0.0325	0.4316 \pm 0.0171	0.5374 \pm 0.0361	0.5053 \pm 0.0462	0.8487 \pm 0.0578	0.6078 \pm 0.0541	-15.28%
	720	0.4841 \pm 0.0381	0.4546 \pm 0.0224	0.5799 \pm 0.0914	0.4988 \pm 0.0384	1.0951 \pm 0.1741	0.8007 \pm 0.1602	-15.65%

Table 8: The processing time of ReLD and training time of Autoformer during 1 epoch.

Dataset	# of Windows	Window size (I + O)	# of Series	ReLD Preprocessing time (a) (seconds)	Training time (b) (seconds per epoch)	Ratio (a) / (a) + (b)
ETTh1	8449	192 (96 + 96)	7	0.18	38.12	0.47%
ETTh2	8449	192 (96 + 96)	7	0.17	39.11	0.43%
ETTm1	34369	192 (96 + 96)	7	0.69	154.68	0.44%
ETTm2	34369	192 (96 + 96)	7	0.68	160.31	0.42%
Weather-hour	5093	1056 (336 + 720)	21	0.63	121.19	0.52%
Pump	9610	672 (336 + 336)	35	1.22	125.30	0.96%
ECL	17741	672 (336 + 336)	1	0.08	94.67	0.08%
Traffic	11225	1056 (336 + 720)	1	0.07	99.05	0.07%

B.4 COMPUTATIONAL COST OF ReLD

Our reweighting framework requires a marginal amount of additional computational cost of calculating the weights for all input-output sequences before training. As shown in Table 8, the cost of calculating the weights on datasets with multiple settings is less than 1% of the time it takes to train with the dataset during one epoch. The absolute time was mostly less than 1 second.

C COMPARISON WITH OTHER METHODS

C.1 COMPARISON WITH SMOOTHING AND OUTLIER FILTERING

We compared our proposed method with 1) smoothing and 2) outlier filtering which are expected to perform well with drastic changes (e.g., fluke) in time series datasets. Smoothing techniques are used to remove nosiness and reduce outliers, allowing meaningful temporal patterns to stand out. Conventional methods include moving average (MA) smoothing as follows:

$$s_t = \frac{(x_{t-k+1} + x_{t-k+2} + \dots + x_t)}{k} \quad (4)$$

where s_t is the smoothed observation at t and x_t is the original observation. The other method is exponential (EMA) smoothing calculated by Equation as follows:

$$s_t = \alpha \cdot x_t + (1 - \alpha) \cdot s_{t-1} \quad (5)$$

where $\alpha \in (0, 1)$. We smoothed the training time series and train forecasting models. To use outlier filtering method for forecasting task, a simple way to detect outliers is to assume that the target value follows a Gaussian and remove values that exceed a certain range of values. We train forecasters after removing outliers which exceed a certain value.

C.2 COMPARISON WITH ERROR-BASED REWEIGHTING

As we mentioned in the main paper, we observed that abrupt changes significantly contribute to the total loss in the training phase. In this situation, we can simply reweight a loss of sample that have large error while considering the sample including abrupt change. Reweighting inversely to the error may down-weight the loss of the abrupt change without additional LD calculation. In the main paper, we presented two error-based methods: Focal-R and filp Focal-R. Focal-R loss is calculated as $\sigma(\beta |e_i|)^\gamma L_i$ where e_i is error of i -th sample, L_i is loss of i -th sample, and $\sigma(\cdot)$ is sigmoid function. β and γ are hyperparameters. In case of filp Focal-R, β is negative to flip the sigmoid function along the y axis. Additionally, we provide L2 error-based reweighting results, namely invL2 which is written as $\frac{L_i}{e_i + \epsilon}$. In case of invL2, as the model forecasts accurately and thus the error of the normal states is close to zero, the parameter moves with larger steps by up-weighted loss. Table 9 shows the performance in the case of reweighting inversely to the error of each window.

C.3 VARIANTS FOR LOCAL DISCREPANCY

We propose the *Local Discrepancy* (LD) based on the statistics formulated by a statistical test, Welch’s t-test Welch (1938), in order to measure how two adjacent in-output sequences, \mathcal{X}_t and \mathcal{Y}_t , are different from each other. There may exist other metrics to measure the local discrepancy such

Table 9: Comparison with error-based reweighting (invL2) in the multivariate forecasting (Top) and in the univariate forecasting (Bottom) using ETTh1 dataset.

Multivariate I / O	Pyraformer			Autoformer			Informer		
	base	invL2	ReLD	base	invL2	ReLD	base	invL2	ReLD
96 / 96	0.6453 ± 0.0583	0.6083 ± 0.0149	0.5345 ± 0.0073	0.4422 ± 0.0242	0.4458 ± 0.0212	0.4438 ± 0.0143	0.9084 ± 0.0485	0.8506 ± 0.0280	0.8031 ± 0.0317
336 / 168	0.8644 ± 0.0905	0.7842 ± 0.0250	0.7415 ± 0.0399	0.5042 ± 0.0515	0.4772 ± 0.0144	0.4906 ± 0.0263	1.3720 ± 0.2422	1.2150 ± 0.1333	0.8858 ± 0.0258
336 / 336	0.9328 ± 0.0341	0.9643 ± 0.0404	0.8895 ± 0.0548	0.5694 ± 0.1115	0.5450 ± 0.0886	0.5110 ± 0.0990	1.3425 ± 0.0725	1.2857 ± 0.0710	0.9850 ± 0.0308
336 / 720	0.9843 ± 0.0213	1.0003 ± 0.0228	0.9781 ± 0.0196	0.5348 ± 0.0212	0.5589 ± 0.0613	0.5207 ± 0.0106	1.3933 ± 0.0892	1.3735 ± 0.0386	1.1994 ± 0.0597
Imp.	-	-2.50%	-9.17%	-	-1.08%	-3.81%	-	-5.86%	-21.89%

Univariate I / O	Pyraformer			Autoformer			Informer		
	base	invL2	ReLD	base	invL2	ReLD	base	invL2	ReLD
96 / 96	0.2074 ± 0.0728	0.1928 ± 0.0365	0.1831 ± 0.0533	0.0859 ± 0.0063	0.0861 ± 0.0031	0.0841 ± 0.0067	0.1203 ± 0.0730	0.1132 ± 0.0441	0.1020 ± 0.0472
336 / 168	0.1819 ± 0.0257	0.1750 ± 0.0581	0.1725 ± 0.0406	0.1077 ± 0.0130	0.0949 ± 0.0109	0.0999 ± 0.0072	0.0862 ± 0.0292	0.0946 ± 0.0244	0.0848 ± 0.0297
336 / 336	0.1716 ± 0.0597	0.1853 ± 0.0622	0.1649 ± 0.0426	0.1055 ± 0.0219	0.1135 ± 0.0198	0.1008 ± 0.0157	0.0862 ± 0.0025	0.0870 ± 0.0084	0.0897 ± 0.0165
336 / 720	0.1974 ± 0.0415	0.1746 ± 0.0312	0.1667 ± 0.0298	0.1352 ± 0.0207	0.1251 ± 0.0110	0.1244 ± 0.0270	0.2025 ± 0.0961	0.1800 ± 0.0945	0.1550 ± 0.0237
Imp.	-	-3.60%	-9.09%	-	-2.88%	-5.45%	-	-1.59%	-9.06%

as multivariate t -statistic Hotelling (1992) (i.e., Hotelling’s t -squared statistic) and stationarity tests (e.g., Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests Kwiatkowski et al. (1992)). We also report the performance of our reweighting framework using a different metric other than t -statistics for measuring the local discrepancy in Table 10. Hotelling’s t -squared statistic is a generalization of Student’s t -statistic that is used in multivariate hypothesis testing. We can naturally utilize t -squared statistic as LD for multivariate forecasting (i.e., $\mathbf{s} \in \mathbb{R}^m$ and $m > 1$) as follows:

$$\text{LocalDis}(\mathcal{X}_t, \mathcal{Y}_t) = \frac{I \cdot O}{I + O} (\bar{\mathcal{X}}_t - \bar{\mathcal{Y}}_t)' \hat{\Sigma}^{-1} (\bar{\mathcal{X}}_t - \bar{\mathcal{Y}}_t) := v_t^2 \quad (6)$$

where the mean and covariance are defined as follows:

$$\bar{\mathcal{X}}_t = \frac{1}{I} \sum_{i=1}^I \mathbf{s}_{t-i}, \quad \bar{\mathcal{Y}}_t = \frac{1}{O} \sum_{i=0}^{O-1} \mathbf{s}_{t+i}, \quad \hat{\Sigma} = \frac{(I-1) \hat{\Sigma}_{\bar{\mathcal{X}}} + (O-1) \hat{\Sigma}_{\bar{\mathcal{Y}}}}{I + O - 2},$$

$$\hat{\Sigma}_{\bar{\mathcal{X}}} = \frac{1}{I-1} \sum_{i=1}^I (\mathbf{s}_{t-i} - \bar{\mathcal{X}}_t) (\mathbf{s}_{t-i} - \bar{\mathcal{X}}_t)', \quad \hat{\Sigma}_{\bar{\mathcal{Y}}} = \frac{1}{O-1} \sum_{i=1}^{O-1} (\mathbf{s}_{t+i} - \bar{\mathcal{Y}}_t) (\mathbf{s}_{t+i} - \bar{\mathcal{Y}}_t)'.$$

We can interpret the time-series data in terms of stochastic processes. KPSS tests are used for testing a null hypothesis that an observable time series is stationary around a deterministic trend (i.e., trend-stationary) against the alternative of a unit root. When the given time series is trend stationary, the KPSS statistic has small value, which is close to zero. Thus, to measure the degree of abruptness of a change in a given period of time, we leverage the KPSS statistic as LD:

$$\text{LocalDis}(\text{concat}(\mathcal{X}_t, \mathcal{Y}_t)) = \frac{1}{(I+O)^2} \cdot \sum_{i=-I}^{O-1} \frac{\mathcal{E}_{t+i}^2}{\hat{\sigma}^2} := v_t \quad (7)$$

where \mathcal{E}_t is partial sum of the residuals and $\hat{\sigma}^2$ is the estimate of the long-run variance of the residuals as follows:

$$\mathcal{E}_k = \sum_{i=1}^t e_i, \quad e = (e_{t-I}, e_{t-I+1}, \dots, e_{O-1})$$

where e means OLS residuals when regressing the concated in-output sequence (i.e., $\text{concat}(\mathcal{X}_t, \mathcal{Y}_t)$). We observe that our reweighting framework consistently outperforms the ones without our framework regardless of the statistics used for measuring the local discrepancy. While we empirically confirmed that using t -statistic is more suitable for LD compared to KPSS or t -Squared statistic, such result demonstrates that our framework can be used with any statistics measure the user deems appropriate.

Table 10: Ablation study on variants of local discrepancy used in our reweighting framework. We compare models which uses 1) KPSS, 2) t -squared, and 3) t -statistic. The t -statistic shows more consistent and superior results compared to other statistics in the multivariate setting.

Dataset		ETTh1			ETTh2			ETTm1			Imp.
Model	Predict-O	96	168	336	96	168	336	96	168	336	
Pyraformer		0.645	0.864	0.933	1.609	5.014	4.356	0.536	0.563	0.697	-
Pyraformer + KPSS		0.554	0.782	0.909	1.482	4.590	5.327	0.470	0.527	0.604	-5.84%
Pyraformer + t -squared		0.640	0.809	0.898	1.440	3.112	3.912	0.490	0.557	0.632	-9.84%
Pyraformer + t -statistic		0.534	0.742	0.889	1.173	3.976	3.281	0.471	0.506	0.573	-16.51%
Autoformer		0.442	0.504	0.569	0.386	0.439	0.494	0.524	0.534	0.561	-
Autoformer + KPSS		0.446	0.528	0.486	0.358	0.436	0.516	0.456	0.538	0.513	-3.68%
Autoformer + t -squared		0.454	0.521	0.515	0.357	0.403	0.436	0.503	0.548	0.512	-4.60%
Autoformer + t -statistic		0.444	0.491	0.511	0.351	0.413	0.424	0.455	0.500	0.514	-7.74%
Informer		0.908	1.372	1.343	3.400	5.796	3.901	0.640	1.224	1.390	-
Informer + KPSS		0.850	1.215	1.215	3.050	5.593	4.202	0.535	0.844	1.087	-11.41%
Informer + t -squared		0.871	1.262	1.234	2.796	4.393	3.419	0.594	0.992	1.195	-12.76%
Informer + t -statistic		0.856	1.113	1.151	2.462	4.723	3.788	0.543	0.751	1.008	-18.81%

D OUR FRAMEWORK DETAILS

D.1 IMPLEMENTATION DETAILS

We include 12 baselines to validate our ReLD. All models were implemented with PyTorch. As for recent models (*i.e.*, FEDformer², Pyraformer³, Autoformer⁴, and Informer⁵), we used the official code released by the original authors, rather than implementing from scratch. For a fair comparison between ReLD and the existing framework, we set the same hyperparameters found in each work. We trained all models from scratch to 10 epochs. To assign weights to all training samples in ReLD, the LD is computed only once before training, and it takes only a negligible amount of time compared to the training time. Most models, which leverage a generative decoding, take an average of less than an hour to train on a TITAN-Xp GPU except for LSTMa which uses auto-regressive decoding. The source code is available in the following address: <https://tinyurl.com/iclr3913>.

D.2 PSEUDO CODE FOR ReLD

Algorithm 1 ReLD: Reweighting framework based on Local Discrepancy Density

Require: Training set $\mathcal{D} = \{(\mathcal{X}_t, \mathcal{Y}_t)\}_{t=1}^N$, bin size Δb , symmetric kernel distribution $k(v, v')$

Compute Local Discrepancy $LD(\mathcal{X}_t, \mathcal{Y}_t) = \frac{\bar{\mathcal{X}}_t - \bar{\mathcal{Y}}_t}{\sqrt{\frac{s_{\mathcal{X}_t}^2}{T} + \frac{s_{\mathcal{Y}_t}^2}{O} + \epsilon}} := v_t$

Compute the empirical label density distribution $p(v)$ based on Δb and \mathcal{D}

Compute the effective label density distribution $\tilde{p}(v') := \int_{\mathcal{V}} k(v, v') p(v) dv$

for all $(\mathcal{X}_t, \mathcal{Y}_t, v_t) \in \mathcal{D}$ **do**

 Assign weight for each sample as $w_t \propto c \cdot \tilde{p}(v_t)$ (constant c as scaling factor)

end for

for all number of training iterations **do**

 Sample a mini-batch $\{(\mathcal{X}_b, \mathcal{Y}_b, w_b)\}_{b=1}^B$ from \mathcal{D}

 Forward $\{\mathcal{X}_b\}_{b=1}^B$ and get corresponding predictions $\{\hat{\mathcal{Y}}_b\}_{b=1}^B$

 Do one training step using the weighted loss $\frac{1}{B} \sum_{b=1}^B \mathcal{L}_{w_b}(\hat{\mathcal{Y}}_b, \mathcal{Y}_b)$

end for

²<https://github.com/MAZiqing/FEDformer>

³<https://github.com/alipay/Pyraformer>

⁴<https://github.com/thuml/Autoformer>

⁵<https://github.com/zhouhaoyi/Informer2020>

Table 11: Performance change according to the number of bins.

Dataset		ETTh1	ETTh2	
Model	# bins	336 \rightarrow 336	96 \rightarrow 96	Imp.
Autoformer	-	0.5694 ± 0.1115	0.3859 ± 0.0260	-
Autoformer + ReLD	40	0.5245 ± 0.1543	0.3529 ± 0.0262	-8.55%
	120	0.4907 ± 0.0337	0.3455 ± 0.0229	-10.47%
	200	0.4903 ± 0.0610	0.3501 ± 0.0168	-9.28%
	300	0.4881 ± 0.0413	0.3472 ± 0.0072	-10.03%
	500	0.5130 ± 0.0529	0.3485 ± 0.0142	-9.69%

Table 12: Performance change according to the KDE kernel types.

Dataset		ETTh1	ETTh2	
Model	KDE kernel	336 \rightarrow 336	96 \rightarrow 96	Imp.
Autoformer	-	0.5694 ± 0.1115	0.3859 ± 0.0260	-
Autoformer + ReLD	Gaussian	0.4903 ± 0.0610	0.3501 ± 0.0168	-9.28%
	Triangle	0.4792 ± 0.0171	0.3453 ± 0.0232	-10.52%
	Laplace	0.4786 ± 0.0380	0.3496 ± 0.0087	-9.41%

Table 13: Performance changes according to the KDE kernel size.

Dataset		ETTh1	ETTh2	
Model	KDE kernel size	336 \rightarrow 336	96 \rightarrow 96	Imp.
Autoformer	-	0.5694 ± 0.1115	0.3859 ± 0.0260	-
Autoformer + ReLD	5	0.4903 ± 0.0610	0.3501 ± 0.0168	-9.28%
	10	0.4896 ± 0.0728	0.3466 ± 0.0018	-10.18%
	15	0.4836 ± 0.0189	0.3567 ± 0.0293	-7.57%
	20	0.4841 ± 0.0444	0.3482 ± 0.0383	-9.77%
	25	0.4835 ± 0.0317	0.3490 ± 0.0166	-9.56%

Table 14: Performance change according to the KDE kernel sigma.

Dataset		ETTh1	ETTh2	
Model	KDE kernel sigma	336 \rightarrow 336	96 \rightarrow 96	Imp.
Autoformer	-	0.5694 ± 0.1115	0.3859 ± 0.0260	-
Autoformer + ReLD	1	0.5545 ± 0.1916	0.3521 ± 0.0189	-8.76%
	2	0.4903 ± 0.0610	0.3501 ± 0.0168	-9.28%
	4	0.5218 ± 0.1632	0.3586 ± 0.0567	-7.07%
	8	0.4767 ± 0.0307	0.3435 ± 0.0172	-10.99%
	16	0.4836 ± 0.0318	0.3494 ± 0.0310	-9.46%

We illustrate the pseudo code of the ReLD in Algorithm 1.

D.3 HYPERPARAMETER SENSITIVITY

We used KDE to smooth the LD distribution. Related parameters include the bin size that determines how many sections continuous LD is divided into, KDE’s kernel type, kernel size and kernel sigma. In our experiment, we set the bin size to 200, kernel type to Gaussian, and kernel size and sigma to 5 and 2, respectively, as default parameters. We conducted experiments on ETTh1 and ETTh2 to observe the variance of performance according to each parameter. As shown in Table 11, Table 12,

Table 13, and Table 14, we observe that our proposed method is robust to the hyper-parameters while showing consistent performance improvements.

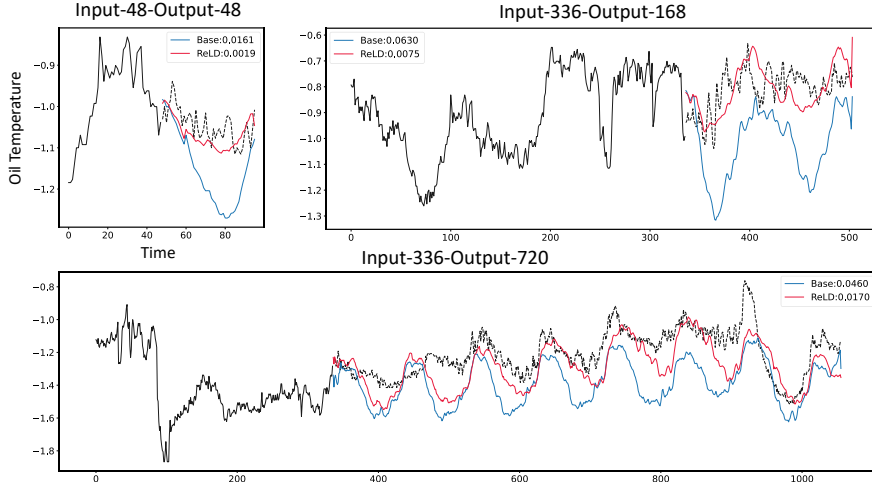


Figure 7: Forecasting results of Autoformer trained on ETTm1 with three different length settings: Input-48-Output-48, Input-336-Output-168, and Input-336-Output-720. The blue line indicates the forecasting results of the baselines without our ReLD and the red line indicates those with our ReLD.

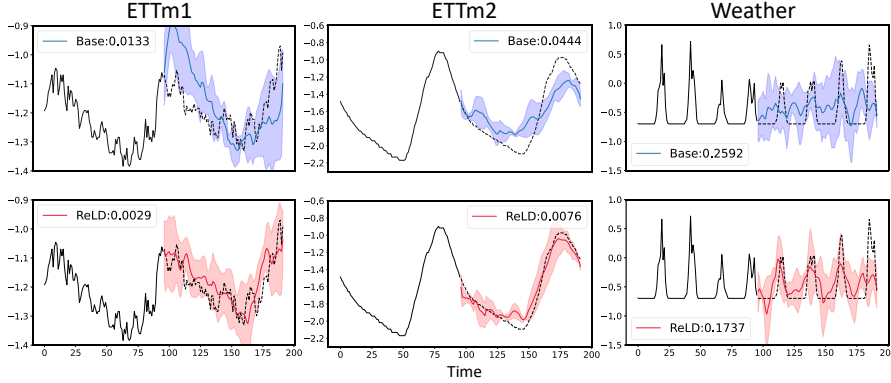


Figure 8: Forecasting results of Autoformer on three datasets: ETTm1, ETTm2, and Weather. The first row shows the forecasting results of the baseline without our ReLD and the second row shows those with our ReLD.

E QUALITATIVE RESULTS

This section visualizes the forecasting results using three criteria: in-output length (Figure 7), dataset (Figure 8), and model architecture (Figure 9). All samples are from the test set of each dataset. The solid black line denotes the input series and the dotted black line denotes the ground truth series that a model should predict. For a reliable comparison, we plot the averaged forecasting results of the independent models trained from different random initializations. The shaded part of the forecasting result indicates the forecasting variation at a given time stamp. In Figure 7, we only report the mean of forecasting results without the forecasting variation for better clarity.

As shown in Figure 7, our ReLD demonstrated enhanced forecasting results in both short-term and long-term settings. We observe that applying ReLD significantly reduces the MSE loss regardless of datasets (see Figure 8) and model architectures (Figure 9). For example, by applying ReLD on Weather dataset (see Figure 8), the prediction variations (red-shaded regions) are fitted to the fluctuations of the target times series which was underfitted without applying ReLD (blue-shaded regions).

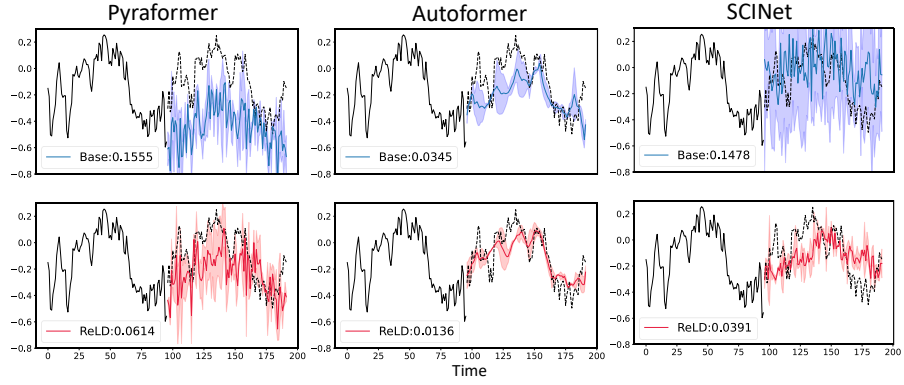


Figure 9: Forecasting results of the recent three models on the same sample in HULL series of ETTm2. The first row shows the forecasting results of the baselines without our ReLD and the second row shows those with our ReLD.

F FULL BENCHMARK ON THE REAL-WORLD DATASETS

Models	FEDformer			Pyraformer			Autoformer			Informer			LogTrans			Reformer			Transformer			LSTNet			LSTMa			TCN		
	MSE	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our			
ETTm1	48	0.3479 ±0.0017	0.3483 ±0.0017	0.5354 ±0.0157	0.4438 ±0.0139	0.4160 ±0.0216	0.3842 ±0.0181	0.6872 ±0.0736	0.5734 ±0.0409	0.7247 ±0.0764	0.5902 ±0.0473	0.7393 ±0.0260	0.6409 ±0.0173	0.6615 ±0.0767	0.5443 ±0.0370	0.6272 ±0.0115	0.5586 ±0.0069	0.6938 ±0.0447	0.5809 ±0.0334	0.5655 ±0.0287	0.4862 ±0.0151									
	96	0.3774 ±0.0018	0.3743 ±0.0016	0.6453 ±0.0584	0.5345 ±0.0074	0.4422 ±0.0243	0.4438 ±0.0143	0.9084 ±0.0486	0.8561 ±0.0522	0.6584 ±0.0574	0.5772 ±0.0307	0.8031 ±0.0317	0.7306 ±0.0216	0.8652 ±0.1242	0.7160 ±0.0289	0.7605 ±0.0365	0.6816 ±0.0354	0.8290 ±0.1048	0.6613 ±0.0551	0.7628 ±0.0429	0.6794 ±0.0152									
	168	0.4446 ±0.0164	0.4367 ±0.0153	0.8644 ±0.0905	0.7415 ±0.0399	0.5042 ±0.0515	0.4906 ±0.0264	1.3720 ±0.2422	1.1125 ±0.0813	1.0497 ±0.1452	0.9355 ±0.0647	0.8858 ±0.0259	0.7677 ±0.0172	1.0911 ±0.1563	0.8628 ±0.0823	0.9276 ±0.0890	0.7904 ±0.0403	1.0522 ±0.0564	0.8622 ±0.0960	0.9406 ±0.0313	0.9319 ±0.0327									
	336	0.4479 ±0.0075	0.4453 ±0.0059	0.9328 ±0.0341	0.8895 ±0.0548	0.5694 ±0.1115	0.5110 ±0.0990	1.3425 ±0.0725	1.1507 ±0.0305	1.0618 ±0.2052	0.9874 ±0.1435	0.9850 ±0.0308	0.8481 ±0.0130	1.1712 ±0.0721	1.0429 ±0.0560	1.0125 ±0.0758	0.9586 ±0.0449	1.1677 ±0.0875	0.9672 ±0.1548	1.1093 ±0.0555	1.1120 ±0.0625									
	720	0.5088 ±0.0131	0.5064 ±0.0192	0.9843 ±0.0214	0.9781 ±0.0196	0.5348 ±0.0213	0.5207 ±0.0107	1.3933 ±0.0893	1.3912 ±0.0444	1.0621 ±0.1705	1.1995 ±0.2349	1.1994 ±0.0597	1.0481 ±0.0449	1.1264 ±0.0765	1.0834 ±0.0754	1.1154 ±0.0656	1.0965 ±0.0391	1.4609 ±0.1134	1.3927 ±0.1274	1.1604 ±0.0653	1.1484 ±0.0246									
Imp.		-0.71%		-10.75%		-4.57%		-11.13%		-7.17%		-12.44%		-14.13%		-8.62%		-15.28%		-5.33%										
ETTm2	48	0.2539 ±0.0004	0.2569 ±0.0010	0.7756 ±0.0910	0.6186 ±0.0369	0.2984 ±0.0072	0.2815 ±0.0061	0.9560 ±0.1512	0.6438 ±0.0916	1.1329 ±0.0825	0.6747 ±0.1191	0.8518 ±0.0350	0.6777 ±0.0400	1.1950 ±0.3054	0.6794 ±0.0608	0.8930 ±0.0805	0.6791 ±0.0480	0.8816 ±0.0832	0.5750 ±0.0213	0.9044 ±0.1185	0.6387 ±0.1442									
	96	0.3420 ±0.0024	0.3304 ±0.0033	1.6090 ±0.0866	1.1733 ±0.2271	0.3859 ±0.0261	0.3514 ±0.0124	3.4000 ±0.5346	2.4616 ±0.4278	2.7501 ±0.6731	2.1171 ±0.3310	1.8490 ±0.2863	1.3124 ±0.1231	2.0744 ±0.1893	1.7618 ±0.3022	1.4426 ±0.0179	1.1167 ±0.0990	1.6695 ±0.5567	0.9905 ±0.1973	1.7404 ±0.4478	1.5293 ±0.1764									
	168	0.4021 ±0.0096	0.3837 ±0.0069	5.0138 ±1.0231	3.9761 ±0.6641	0.4394 ±0.0927	0.4131 ±0.0343	5.7957 ±0.8701	4.7229 ±0.5132	2.2557 ±0.3390	2.2054 ±0.5046	3.1301 ±0.3670	2.2042 ±0.2261	3.9472 ±0.8010	4.0993 ±0.8544	3.3118 ±0.2903	2.8871 ±0.2880	3.9065 ±1.1970	2.5133 ±0.8563	2.4321 ±0.1782	2.4058 ±0.1341									
	336	0.4006 ±0.0108	0.3842 ±0.0071	4.3559 ±0.3957	3.2814 ±0.5145	0.4942 ±0.0307	0.4239 ±0.0338	3.9011 ±0.6699	3.7882 ±0.7034	4.4043 ±0.5353	3.6032 ±0.2468	2.9343 ±0.3673	2.2374 ±0.1497	2.8095 ±0.5978	2.9328 ±0.3583	3.9730 ±0.3591	3.0537 ±0.2037	2.9757 ±0.4939	2.1797 ±0.3570	2.9652 ±0.0794	2.9037 ±0.1516									
	720	0.4520 ±0.0363	0.4429 ±0.0315	4.2013 ±0.5338	3.4771 ±0.2490	0.4502 ±0.0414	0.4209 ±0.0232	4.0175 ±0.5374	3.9606 ±0.3356	2.4391 ±0.2176	2.7142 ±0.5075	2.6307 ±0.2001	2.6276 ±0.2424	2.6471 ±0.5341	3.0069 ±0.4458	3.7246 ±0.1502	2.9905 ±0.3913	3.2982 ±0.4699	2.6797 ±0.2820	3.1513 ±0.0452	3.1424 ±0.0870									
Imp.		-2.58%		-21.98%		-8.26%		-16.62%		-14.52%		-20.58%		-7.28%		-20.44%		-31.32%		-8.99%										
ETTm1	48	0.5146 ±0.0091	0.5190 ±0.0124	0.5632 ±0.0285	0.5564 ±0.0443	0.5487 ±0.0865	0.4924 ±0.0329	0.6332 ±0.0150	0.5984 ±0.0266	0.5800 ±0.0225	0.5696 ±0.0186	0.7703 ±0.0320	0.7276 ±0.0077	0.6136 ±0.0164	0.5882 ±0.0439	0.6415 ±0.0463	0.6136 ±0.0256	0.6744 ±0.0253	0.6478 ±0.0336	0.6768 ±0.0620	0.6478 ±0.0606									
	96	0.3594 ±0.0035	0.3580 ±0.0033	0.5364 ±0.0319	0.4713 ±0.0300	0.5239 ±0.0851	0.4547 ±0.0476	0.6404 ±0.0570	0.5429 ±0.0368	0.7070 ±0.0470	0.5969 ±0.0310	0.7773 ±0.0698	0.6409 ±0.0582	0.5787 ±0.0672	0.5157 ±0.0293	0.5484 ±0.0358	0.5365 ±0.0251	0.7054 ±0.0793	0.5920 ±0.0838	0.6758 ±0.0333	0.5936 ±0.0259									
	168	0.3850 ±0.0259	0.3790 ±0.0138	0.5631 ±0.0737	0.5056 ±0.0346	0.5341 ±0.0366	0.5001 ±0.0300	1.2238 ±0.1911	0.7507 ±0.0776	0.8866 ±0.1458	0.7534 ±0.0856	0.8398 ±0.0346	0.6889 ±0.0220	0.7928 ±0.0628	0.7268 ±0.0320	0.6323 ±0.0282	0.5774 ±0.0338	0.8710 ±0.0883	0.6477 ±0.0661	0.9382 ±0.0412	0.9128 ±0.0288									
	336	0.4033 ±0.0107	0.3955 ±0.0090	0.6970 ±0.0855	0.5729 ±0.0707	0.5613 ±0.0217	0.5145 ±0.0214	1.3896 ±0.1481	1.0085 ±0.0877	0.9685 ±0.0582	0.8368 ±0.0570	0.9874 ±0.0519	0.8946 ±0.0184	0.9579 ±0.1637	0.9262 ±0.1313	0.7985 ±0.1058	0.6858 ±0.0390	1.1254 ±0.1963	0.6810 ±0.0222	1.1476 ±0.0318	1.1258 ±0.0166									
	720	0.5010 ±0.0456	0.4805 ±0.0297	0.9037 ±0.1019	0.6822 ±0.0624	0.5601 ±0.0578	0.5278 ±0.0301	1.3329 ±0.1013	1.0780 ±0.0966	1.1222 ±0.1009	0.9226 ±0.0750	1.1216 ±0.0227	1.0027 ±0.0421	1.0858 ±0.1066	1.0393 ±0.0785	0.9248 ±0.0516	0.8168 ±0.0886	0.9782 ±0.1870	0.8280 ±0.1288	1.2768 ±0.0243	1.2381 ±0.0397									
Imp.		-1.42%		-13.17%		-8.79%		-21.19%		-12.75%		-12.21%		-6.19%		-8.20%		-20.10%		-4.82%										
ETTm2	48	0.1421 ±0.0019	0.1386 ±0.0003	0.2806 ±0.0215	0.2200 ±0.0283	0.1708 ±0.0051	0.1689 ±0.0023	0.2953 ±0.0350	0.2143 ±0.0188	0.3191 ±0.0620	0.2394 ±0.0261	0.4635 ±0.0969	0.3293 ±0.0322	0.3317 ±0.1093	0.2794 ±0.0382	0.2923 ±0.0328	0.2826 ±0.0174	0.2736 ±0.0374	0.2494 ±0.0509	0.3619 ±0.0255	0.2535 ±0.0233									
	96	0.1886 ±0.0008	0.1844 ±0.0012	0.3709 ±0.1396	0.2480 ±0.0164	0.2934 ±0.0691	0.2213 ±0.0126	0.4451 ±0.0519	0.2860 ±0.0221	0.5480 ±0.0984	0.3844 ±0.0594	0.7433 ±0.0529	0.4491 ±0.0230	0.4636 ±0.0845	0.2993 ±0.0384	0.3426 ±0.0377	0.3807 ±0.0391	0.3807 ±0.0377	0.2798 ±0.0165	0.3839 ±0.0607	0.3537 ±0.0067									
	168	0.3429 ±0.0240	0.2791 ±0.0017	0.5655 ±0.0840	0.5515 ±0.0586	0.3095 ±0.0130	0.2838 ±0.0087	2.2831 ±0.1434	1.4531 ±0.1662	1.4953 ±0.2619	1.1491 ±0.1603	1.2078 ±0.1790	0.8359 ±0.0428	1.7806 ±0.2845	1.3650 ±0.2038	0.9504 ±0.0953	0.8304 ±0.0878	0.6010 ±0.2799	0.6010 ±0.0853	1.8678 ±0.1214	1.9559 ±0.0890									
	336	0.3381 ±0.0060	0.3191 ±0.0050	1.6013 ±0.3115	1.3296 ±0.2605	0.5084 ±0.1483	0.3642 ±0.0491	2.4795 ±0.1344	1.7639 ±0.1074	3.4383 ±0.4704	2.6429 ±0.3334	2.2391 ±0.2275	1.4248 ±0.0750	1.9661 ±0.2138	1.5664 ±0.0497	1.6096 ±0.2870	1.0193 ±0.1413	1.4788 ±0.1412	0.7454 ±0.1184	2.7687 ±0.1485	2.7734 ±0.0594									
	720	0.4322 ±0.0130	0.3939 ±0.0094	5.4764 ±0.9723	5.0372 ±0.6450	0.5022 ±0.0937	0.4110 ±0.0286	6.5797 ±0.7368	5.7770 ±0.6946	5.6465 ±0.3868	6.1251 ±0.9705	3.0679 ±0.1061	2.8266 ±0.0524	4.5449 ±0.5983	4.1617 ±1.2639	6.1303 ±1.4973	4.4494 ±1.3402	3.0833 ±0.3581	2.3814 ±0.1011	3.2036 ±0.1119	3.1873 ±0.0779									
Imp.		-7.56%		-16.44%		-16.10%		-28.12%		-18.53%		-28.71%		-20.66%		-20.54%		-31.34%		-11.25%										

Models	FEDformer			Pyraformer			Autoformer			Informer			LogTrans			Reformer			Transformer			LSTNet			LSTMa			TCN			
	base	our	MSE	base	our	MSE	base	our	MSE	base	our	MSE	base	our	MSE	base	our	MSE	base	our	MSE	base	our	MSE	base	our	MSE	base	our		
Pump	0.5246	0.5227	48	0.7242	0.7151	0.5767	0.5439	0.8648	0.7934	0.7886	0.7129	0.8083	0.7282	0.8434	0.6645	0.8234	0.7343	0.7540	0.7155	0.8649	0.8310										
	± 0.0024	± 0.0012		± 0.0184	± 0.0153	± 0.0364	± 0.0042	± 0.0294	± 0.1065	± 0.0569	± 0.0673	± 0.0318	± 0.0155	± 0.0664	± 0.0697	± 0.0376	± 0.0110	± 0.0238	± 0.0295	± 0.0920	± 0.0964										
	0.5197	0.5135	96	0.8484	0.7958	0.5576	0.5381	0.8314	0.8699	0.9948	0.9839	0.8261	0.7605	0.9108	0.8816	1.0155	1.0068	0.8126	0.7660	1.0367	0.9699										
	± 0.0026	± 0.0021		± 0.0309	± 0.0284	± 0.0098	± 0.0192	± 0.0732	± 0.0987	± 0.1275	± 0.1926	± 0.0619	± 0.0332	± 0.0997	± 0.1224	± 0.1262	± 0.1740	± 0.0676	± 0.0374	± 0.1635	± 0.0383										
	0.5502	0.5363	168	0.8512	0.8431	0.5974	0.5814	1.7047	1.5268	1.2820	1.3845	1.0941	0.8555	1.1326	1.1846	1.3271	1.2025	0.9088	0.8161	1.1093	1.0771										
AirQuality	± 0.0061	± 0.0057		± 0.0470	± 0.0320	± 0.0252	± 0.0194	± 0.1863	± 0.1589	± 0.1141	± 0.2210	± 0.2937	± 0.0902	± 0.1874	± 0.1399	± 0.2779	± 0.2055	± 0.1720	± 0.0700	± 0.0529	± 0.0788										
	0.5931	0.5638	336	0.9225	0.9508	0.6615	0.6209	1.6763	1.4923	0.9181	0.9100	0.9658	0.9183	1.1494	1.1369	1.6537	1.2921	0.9340	0.8587	1.5206	1.2081										
	± 0.0525	± 0.0109		± 0.0405	± 0.0520	± 0.0248	± 0.0144	± 0.3531	± 0.2763	± 0.0463	± 0.0816	± 0.0598	± 0.0385	± 0.3206	± 0.2823	± 0.4863	± 0.3871	± 0.1100	± 0.0966	± 0.0914	± 0.1237										
	0.7230	0.5802	720	1.3697	1.2829	0.7066	0.6191	1.7041	1.6986	0.9746	0.9082	1.3278	1.2181	1.5309	1.7408	1.6076	1.6421	1.4645	1.2439	2.0748	1.5455										
	± 0.2988	± 0.0318		± 0.2434	± 0.1263	± 0.0144	± 0.0203	± 0.2126	± 0.0835	± 0.0806	± 0.0435	± 0.0993	± 0.0630	± 0.3202	± 0.3619	± 0.3964	± 0.1623	± 0.4705	± 0.5097	± 0.4642	± 0.1738										
Imp.	-5.75%		-2.34%		-6.08%		-5.07%		-2.08%		-10.57%		-1.44%		-8.16%		-8.83%		-11.87%												
Weather-h	0.7306	0.7447	48	0.8780	0.8579	0.8606	0.8520	0.9094	0.8682	0.9281	0.8010	1.1148	0.9913	0.9718	0.8877	1.0227	0.9738	0.9273	0.8540	0.8829	0.8358										
	± 0.0131	± 0.0254		± 0.0675	± 0.0712	± 0.0191	± 0.0425	± 0.0579	± 0.0610	± 0.1007	± 0.0443	± 0.0509	± 0.0463	± 0.0409	± 0.0707	± 0.0654	± 0.0522	± 0.1077	± 0.0242	± 0.1588	± 0.1005										
	0.8251	0.8166	96	1.1207	1.1122	0.9976	0.9271	1.3526	1.1934	1.0586	1.0653	1.2099	1.1961	1.1338	1.0816	1.1462	1.1413	1.1454	1.0810	1.0260	0.9915										
	± 0.0025	± 0.0027		± 0.1378	± 0.1346	± 0.0667	± 0.0638	± 0.0255	± 0.0788	± 0.0867	± 0.0535	± 0.0452	± 0.0546	± 0.1105	± 0.1563	± 0.0491	± 0.0279	± 0.1172	± 0.1114	± 0.0392	± 0.0360										
	0.8108	0.8080	168	1.1930	1.1149	0.9141	0.9256	1.7961	1.5947	1.8682	1.9269	1.4728	1.3455	1.8278	1.7274	1.2313	1.1560	1.6440	1.3756	1.2456	1.1629										
Imp.	± 0.0141	± 0.0164		± 0.1415	± 0.0557	± 0.0214	± 0.0844	± 0.1993	± 0.1155	± 0.2617	± 0.4315	± 0.1751	± 0.1253	± 0.4175	± 0.2465	± 0.0753	± 0.0128	± 0.2611	± 0.2291	± 0.0523	± 0.0954										
	0.8918	0.8724	336	1.2240	1.2140	0.9525	0.9295	1.7579	1.7060	1.6436	1.4867	1.4731	1.3958	1.4636	1.4199	1.3990	1.3881	1.3524	1.2055	1.3006	1.2844										
	± 0.0066	± 0.0112		± 0.0475	± 0.0475	± 0.0683	± 0.0525	± 0.0597	± 0.0608	± 0.2307	± 0.1592	± 0.1082	± 0.0760	± 0.1249	± 0.0744	± 0.0595	± 0.0429	± 0.2118	± 0.1609	± 0.0564	± 0.0660										
	0.9974	0.9527	720	2.1962	1.9817	1.1018	1.0593	2.9136	2.9846	3.3683	2.8386	1.7234	1.6715	2.4630	2.4953	1.8259	1.9214	2.3104	2.2325	1.4419	1.4256										
	± 0.0302	± 0.0261		± 0.1120	± 0.1065	± 0.0378	± 0.0508	± 0.1185	± 0.2300	± 0.3599	± 0.5806	± 0.0728	± 0.0258	± 0.1439	± 0.2564	± 0.0761	± 0.0689	± 0.2172	± 0.1665	± 0.0792	± 0.0703										
Imp.	-1.22%		-4.04%		-2.62%		-5.61%		-7.04%		-5.82%		-4.09%		-1.37%		-8.82%		-3.54%												
Imp.	0.3376	0.3357	48	0.2922	0.2786	0.3441	0.3425	0.3454	0.2941	0.3593	0.3262	0.3429	0.3127	0.3047	0.2860	0.3181	0.3099	0.3457	0.3245	0.3476	0.3270										
	± 0.0026	± 0.0023		± 0.0047	± 0.0057	± 0.0075	± 0.0064	± 0.0149	± 0.0131	± 0.0263	± 0.0147	± 0.0237	± 0.0113	± 0.0040	± 0.0095	± 0.0078	± 0.0060	± 0.0158	± 0.0071	± 0.0133	± 0.0091										
	0.4031	0.3997	96	0.3933	0.3578	0.4636	0.4455	0.4528	0.4428	0.4471	0.3856	0.5264	0.4163	0.4279	0.3704	0.4142	0.3857	0.4090	0.3873	0.4502	0.4243										
	± 0.0057	± 0.0198		± 0.0095	± 0.0074	± 0.0277	± 0.0215	± 0.0551	± 0.0291	± 0.0229	± 0.0217	± 0.0448	± 0.0186	± 0.0372	± 0.0214	± 0.0118	± 0.0115	± 0.0158	± 0.0131	± 0.0416	± 0.0279										
	0.4578	0.4466	168	0.4207	0.3978	0.5163	0.4907	0.5304	0.4984	0.4706	0.4431	0.6591	0.6726	0.5106	0.4813	0.4641	0.4612	0.4198	0.4164	1.0181	1.0051										
Imp.	± 0.0256	± 0.0154		± 0.0160	± 0.0105	± 0.0232	± 0.0190	± 0.0582	± 0.0593	± 0.0369	± 0.0555	± 0.1473	± 0.1001	± 0.0384	± 0.0218	± 0.0388	± 0.0245	± 0.0168	± 0.0187	± 0.0500	± 0.1028										
	0.5099	0.5162	336	0.4540	0.4401	0.6123	0.5662	0.5921	0.5680	0.5736	0.5854	0.8411	0.7824	0.5051	0.4726	0.4905	0.4730	0.4727	0.4518	1.1468	1.2091										
	± 0.0250	± 0.0197		± 0.0269	± 0.0201	± 0.0374	± 0.0339	± 0.0244	± 0.0204	± 0.1297	± 0.0899	± 0.0633	± 0.0867	± 0.0455	± 0.0457	± 0.0198	± 0.0133	± 0.0159	± 0.0212	± 0.1591	± 0.1111										
	Imp.	-0.65%		-4.21%		-6.79%		-6.69%		-8.66%		-7.94%		-3.41%		-4.17%		-1.88%													

Models	MSE	FEDformer		Pyraformer		Autoformer		N-BEATS		Informr		LogTrans		Reformer		Transformer		DeepAR		LSTNet		TCN	
		base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our	base	our
ETTm1	48	0.0509 ±0.0010	0.0490 ±0.0008	0.1485 ±0.0149	0.1144 ±0.0079	0.0760 ±0.0134	0.0658 ±0.0057	0.1337 ±0.0290	0.0963 ±0.0230	0.2216 ±0.1100	0.1608 ±0.0463	0.1257 ±0.0474	0.1247 ±0.0249	0.3218 ±0.1346	0.2940 ±0.1004	0.1925 ±0.1114	0.1479 ±0.0983	0.3035 ±0.0400	0.2499 ±0.0196	0.1130 ±0.0030	0.1294 ±0.0051	0.1770 ±0.1064	0.1230 ±0.0264
	96	0.0809 ±0.0014	0.0786 ±0.0007	0.2074 ±0.0078	0.1831 ±0.0053	0.0859 ±0.0064	0.0841 ±0.0067	0.1203 ±0.0472	0.1203 ±0.0472	0.2168 ±0.0425	0.2189 ±0.0443	0.2164 ±0.1065	0.1884 ±0.0804	0.4816 ±0.2308	0.4268 ±0.1818	0.2266 ±0.0960	0.1912 ±0.0881	0.3849 ±0.0529	0.3361 ±0.0394	0.2213 ±0.0112	0.2322 ±0.0068	0.3485 ±0.0808	0.3169 ±0.1107
	168	0.1053 ±0.0069	0.0996 ±0.0071	0.1819 ±0.0057	0.1725 ±0.0046	0.1077 ±0.0131	0.0999 ±0.0131	0.0862 ±0.0292	0.0848 ±0.0292	0.1951 ±0.0514	0.1829 ±0.0567	0.2206 ±0.0313	0.2051 ±0.0293	0.5079 ±0.0215	0.4489 ±0.1202	0.2471 ±0.0475	0.2189 ±0.0437	0.2463 ±0.0177	0.2653 ±0.0144	0.2641 ±0.0890	0.2388 ±0.0443	0.3142 ±0.0449	0.3120 ±0.0170
	336	0.1146 ±0.0075	0.1013 ±0.0059	0.1716 ±0.0057	0.1649 ±0.0042	0.1055 ±0.0220	0.1008 ±0.0157	0.0862 ±0.0025	0.0897 ±0.0025	0.1450 ±0.0181	0.1632 ±0.0483	0.1746 ±0.0408	0.1639 ±0.0440	0.5653 ±0.0168	0.5135 ±0.0152	0.2355 ±0.0201	0.2101 ±0.0190	0.3434 ±0.0390	0.3446 ±0.0333	0.2242 ±0.0302	0.2246 ±0.0302	0.3851 ±0.0668	0.3724 ±0.0472
	720	0.1531 ±0.0365	0.1345 ±0.0224	0.1974 ±0.0416	0.1667 ±0.0299	0.1244 ±0.0207	0.1244 ±0.0207	0.1352 ±0.0962	0.1550 ±0.0962	0.1562 ±0.0705	0.1508 ±0.0676	0.2910 ±0.0748	0.2542 ±0.0434	0.4813 ±0.0808	0.4977 ±0.0808	0.2004 ±0.0533	0.1473 ±0.0324	0.2764 ±0.0283	0.2589 ±0.0291	0.3086 ±0.0237	0.2611 ±0.0348	0.4184 ±0.0466	0.4187 ±0.0376
ETTm2	Imp.	-7.15%	-11.86%	-7.04%	-12.84%	-7.04%	-12.84%	-7.04%	-12.84%	-4.73%	-7.91%	-7.91%	-7.48%	-7.48%	-17.34%	-5.72%	-11.29%	-10.14%	-11.48%	-11.48%	-18.19%	-18.19%	
	48	0.0924 ±0.0003	0.0922 ±0.0003	0.1072 ±0.0008	0.1105 ±0.0035	0.1294 ±0.0074	0.1235 ±0.0070	0.1108 ±0.0195	0.1142 ±0.0202	0.1466 ±0.0027	0.1317 ±0.0067	0.1054 ±0.0041	0.1054 ±0.0041	0.1775 ±0.0077	0.1689 ±0.0054	0.1269 ±0.0098	0.1105 ±0.0098	0.2587 ±0.0234	0.2203 ±0.0075	0.1129 ±0.0051	0.1140 ±0.0046	0.1148 ±0.0052	0.1094 ±0.0094
	96	0.1282 ±0.0013	0.1307 ±0.0008	0.1806 ±0.0058	0.1292 ±0.0074	0.1533 ±0.0038	0.1548 ±0.0073	0.1270 ±0.0076	0.1233 ±0.0076	0.2690 ±0.0109	0.2216 ±0.0364	0.1831 ±0.0364	0.1770 ±0.0481	0.1978 ±0.0070	0.1896 ±0.0047	0.2448 ±0.0617	0.2099 ±0.0534	0.3225 ±0.0196	0.2808 ±0.0289	0.1750 ±0.0056	0.1587 ±0.0044	0.1658 ±0.0052	0.1507 ±0.0049
	168	0.1893 ±0.0170	0.1788 ±0.0092	0.1708 ±0.0133	0.1653 ±0.0140	0.2000 ±0.0096	0.1980 ±0.0119	0.1782 ±0.0188	0.1896 ±0.0242	0.2679 ±0.0164	0.2505 ±0.0214	0.2358 ±0.0268	0.2323 ±0.0332	0.2207 ±0.0054	0.2185 ±0.0099	0.2556 ±0.0416	0.2434 ±0.0333	0.3104 ±0.0599	0.2730 ±0.0234	0.2080 ±0.0189	0.2201 ±0.0249	0.1938 ±0.0177	0.1885 ±0.0143
	336	0.2069 ±0.0241	0.1984 ±0.0098	0.1866 ±0.0170	0.1837 ±0.0102	0.2140 ±0.0189	0.2059 ±0.0091	0.2058 ±0.0254	0.2242 ±0.0304	0.2776 ±0.0322	0.2422 ±0.0140	0.2585 ±0.0403	0.2375 ±0.0164	0.2059 ±0.0031	0.2148 ±0.0042	0.2706 ±0.0189	0.2611 ±0.0111	0.3558 ±0.0590	0.3110 ±0.0534	0.2685 ±0.0281	0.2461 ±0.0360	0.2060 ±0.0138	0.2013 ±0.0085
720	0.2545 ±0.0366	0.2499 ±0.0201	0.1897 ±0.0371	0.1608 ±0.0229	0.2789 ±0.0476	0.2655 ±0.0119	0.3938 ±0.1487	0.3685 ±0.1271	0.3098 ±0.0285	0.2908 ±0.0244	0.2140 ±0.0081	0.2198 ±0.0109	0.2198 ±0.0109	0.1869 ±0.0063	0.1893 ±0.0053	0.2485 ±0.0227	0.2585 ±0.0529	0.2814 ±0.0188	0.3029 ±0.0770	0.3982 ±0.1034	0.3741 ±0.0141	0.2037 ±0.0218	
ETTm1	Imp.	-1.95%	-9.08%	-2.63%	1.81%	-2.63%	1.81%	-2.63%	1.81%	-10.63%	-3.60%	-3.60%	-0.88%	-0.88%	-6.29%	-8.95%	-3.38%	-3.38%	-3.12%	-3.12%	-3.12%	-3.12%	
	48	0.0230 ±0.0012	0.0221 ±0.0005	0.0288 ±0.0059	0.0255 ±0.0007	0.0357 ±0.0106	0.0328 ±0.0055	0.0306 ±0.0058	0.0295 ±0.0056	0.0428 ±0.0131	0.0356 ±0.0089	0.0375 ±0.0181	0.0332 ±0.0063	0.0784 ±0.0232	0.0699 ±0.0145	0.0315 ±0.0094	0.0317 ±0.0219	0.0843 ±0.0273	0.0501 ±0.0047	0.0304 ±0.0023	0.0287 ±0.0018	0.0342 ±0.0135	0.0327 ±0.0117
	96	0.0359 ±0.0038	0.0350 ±0.0012	0.0821 ±0.0290	0.0576 ±0.0080	0.0577 ±0.0081	0.0522 ±0.0035	0.0851 ±0.0035	0.0851 ±0.0035	0.1051 ±0.0411	0.0831 ±0.0251	0.1051 ±0.0411	0.0831 ±0.0251	0.2280 ±0.0484	0.1608 ±0.0612	0.0888 ±0.0354	0.0569 ±0.0059	0.1790 ±0.0275	0.0849 ±0.0072	0.0622 ±0.0044	0.0512 ±0.0021	0.0841 ±0.0147	0.0740 ±0.0147
	168	0.0830 ±0.0283	0.0662 ±0.0012	0.1286 ±0.0347	0.1218 ±0.0395	0.0881 ±0.0284	0.0723 ±0.0069	0.0771 ±0.0088	0.0701 ±0.0275	0.1808 ±0.0400	0.1576 ±0.0282	0.1924 ±0.0395	0.1612 ±0.0395	0.2840 ±0.0309	0.2227 ±0.0477	0.2211 ±0.0732	0.1960 ±0.0632	0.1999 ±0.0350	0.1985 ±0.0308	0.1243 ±0.0157	0.1223 ±0.0264	0.1945 ±0.0286	0.2129 ±0.0419
	336	0.10808 ±0.0046	0.0758 ±0.0018	0.1941 ±0.0524	0.1384 ±0.0507	0.0903 ±0.0096	0.0847 ±0.0085	0.1336 ±0.0548	0.1248 ±0.0368	0.2631 ±0.0121	0.2273 ±0.0558	0.3041 ±0.0606	0.3041 ±0.0606	0.3394 ±0.0130	0.3039 ±0.0227	0.2739 ±0.0388	0.2469 ±0.1049	0.3821 ±0.0487	0.4323 ±0.0783	0.2973 ±0.0607	0.2629 ±0.0367	0.3286 ±0.0450	0.3506 ±0.0732
720	0.1094 ±0.0136	0.1025 ±0.0253	0.3368 ±0.0928	0.2753 ±0.0563	0.1428 ±0.0618	0.1119 ±0.0270	0.1444 ±0.0342	0.1118 ±0.0342	0.3914 ±0.1449	0.3486 ±0.1133	0.4738 ±0.0667	0.3805 ±0.0369	0.3956 ±0.0255	0.4097 ±0.0291	0.4786 ±0.0737	0.4426 ±0.1325	0.4979 ±0.0795	0.4179 ±0.0847	0.3692 ±0.0291	0.2937 ±0.0219	0.4134 ±0.0892	0.4452 ±0.0289	
ETTm2	Imp.	-7.83%	-13.98%	-12.69%	-12.88%	-12.69%	-12.88%	-12.69%	-12.88%	-15.03%	-21.83%	-21.83%	-13.76%	-13.76%	-12.20%	-19.35%	-11.38%	-11.38%	1.49%	1.49%	1.49%	1.49%	
	48	0.0441 ±0.0013	0.0429 ±0.0002	0.0460 ±0.0015	0.0457 ±0.0016	0.1097 ±0.0106	0.1051 ±0.0107	0.0652 ±0.0063	0.0598 ±0.0047	0.0507 ±0.0024	0.0492 ±0.0009	0.0463 ±0.0018	0.0449 ±0.0019	0.0890 ±0.0029	0.0829 ±0.0021	0.0519 ±0.0054	0.0507 ±0.0061	0.1227 ±0.0176	0.1299 ±0.0088	0.0632 ±0.0009	0.0632 ±0.0009	0.0577 ±0.0084	0.0550 ±0.0041
	96	0.0691 ±0.0030	0.0650 ±0.0009	0.0733 ±0.0050	0.0651 ±0.0042	0.1235 ±0.0108	0.1239 ±0.0120	0.0728 ±0.0085	0.0706 ±0.0047	0.0889 ±0.0095	0.0821 ±0.0025	0.0762 ±0.0068	0.0683 ±0.0047	0.1000 ±0.0046	0.0931 ±0.0131	0.0750 ±0.0091	0.0684 ±0.0053	0.1766 ±0.0124	0.1599 ±0.0072	0.0759 ±0.0024	0.0745 ±0.0052	0.0799 ±0.0084	0.0727 ±0.0050
	168	0.1205 ±0.0174	0.1225 ±0.0038	0.1193 ±0.0282	0.0980 ±0.0166	0.1851 ±0.0400	0.1642 ±0.0296	0.1006 ±0.0127	0.0978 ±0.0127	0.1372 ±0.0104	0.1281 ±0.0116	0.1571 ±0.0230	0.1559 ±0.0096	0.0978 ±0.0082	0.0668 ±0.0068	0.1475 ±0.0135	0.1168 ±0.0106	0.1765 ±0.0166	0.1726 ±0.0166	0.1282 ±0.0132	0.1169 ±0.0152	0.1281 ±0.0139	0.1303 ±0.0092
	336	0.1357 ±0.0085	0.1327 ±0.0038	0.1585 ±0.0129	0.1357 ±0.0310	0.1776 ±0.0158	0.1727 ±0.0109	0.1344 ±0.0130	0.1297 ±0.0130	0.2253 ±0.0169	0.1697 ±0.0195	0.2061 ±0.0107	0.1945 ±0.0311	0.1959 ±0.0133	0.1803 ±0.0180	0.2050 ±0.0249	0.1899 ±0.0232	0.3391 ±0.0429	0.3469 ±0.0784	0.1685 ±0.0158	0.1763 ±0.0174	0.1702 ±0.0215	0.3933 ±0.0458
720	0.1899 ±0.0164	0.1861 ±0.0131	0.1929 ±0.0450	0.2036 ±0.0546	0.2321 ±0.0590	0.2029 ±0.0193	0.1855 ±0.0192	0.1682 ±0.0233	0.2700 ±0.0491	0.2365 ±0.0491	0.2302 ±0.0312	0.2302 ±0.0296	0.2191 ±0.0391	0.2180 ±0.0412	0.2552 ±0.0262	0.2399 ±0.0268	0.3549 ±0.1240	0.2779 ±0.0239	0.2798 ±0.0330	0.2225 ±0.0069	0.2002 ±0.0166	0.1963 ±0.0172	
ECL	Imp.	-2.24%	-8.03%	-6.10%	-5.38%	-6.10%	-5.38%	-6.10%	-5.38%	-10.87%	-3.97%	-3.97%	-9.79%	-9.79%	-9.06%	-5.04%	-9.04%	-9.04%	-3.48%	-3.48%	-3.48%	-3.48%	
	48	0.2942 ±0.0103																					