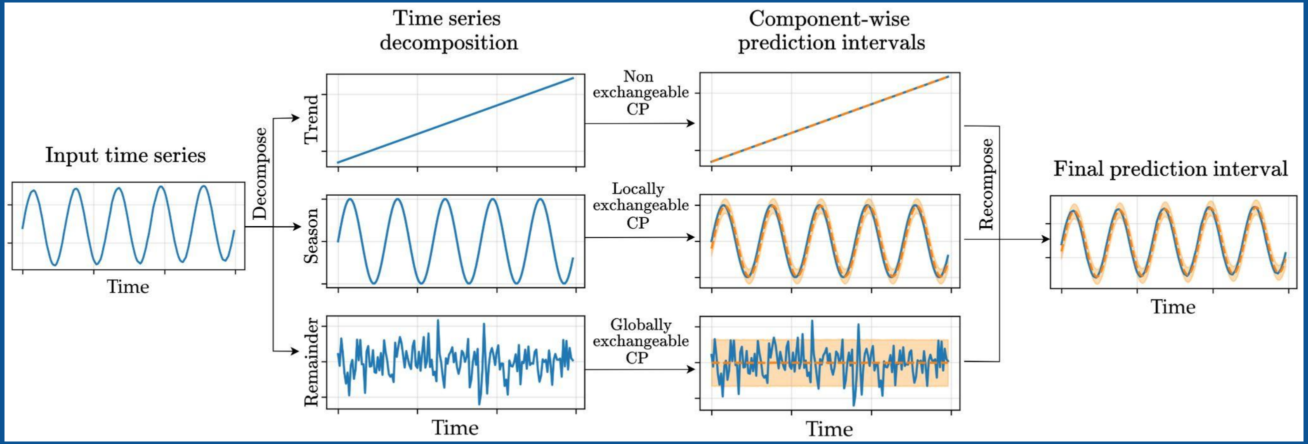
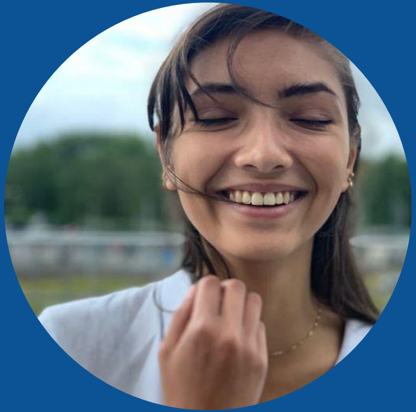
Conformal time series decomposition with component-wise exchangeability



Who are we?

Derck Prinzhorn
MSc AI student

Thijmen Nijdam
MSc AI student



Putri van der Linden
PhD candidate

Alexander Timans
PhD candidate

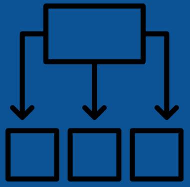
Outline

1. Introduction
2. Background
3. Method
4. Experiments
5. Conclusion

Introduction

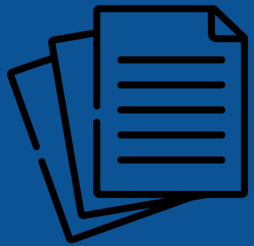
Motivation

non-exchangeable



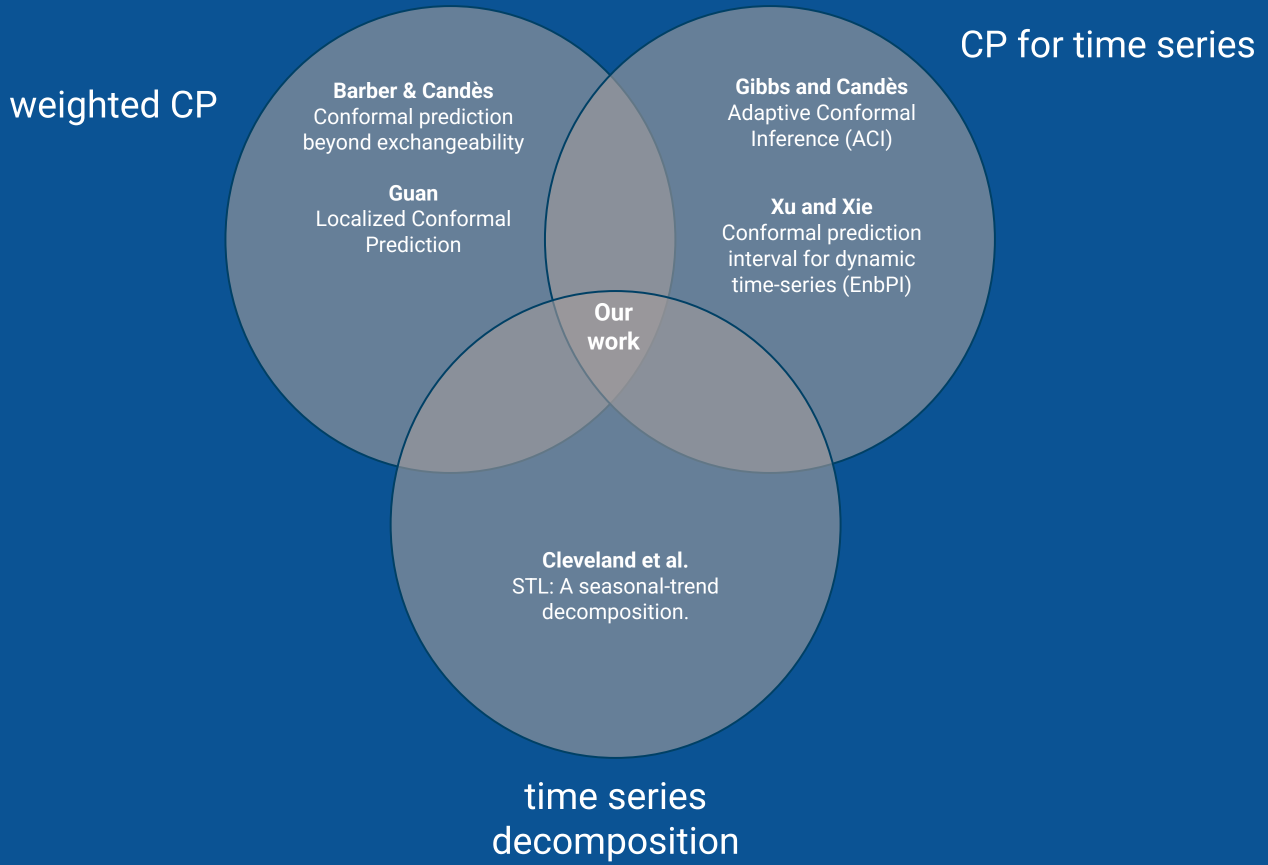
TSD allows for analysis at different scales

relaxed assumptions

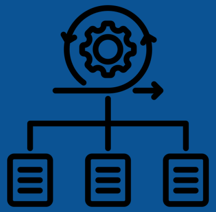


exchangeability regimes

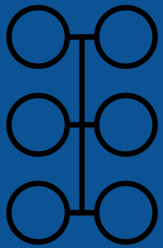
Earlier work



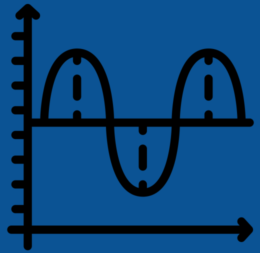
Contributions



TSD pipeline for conformal time series forecasting



conceptually relating TS components to notions of exchangeability



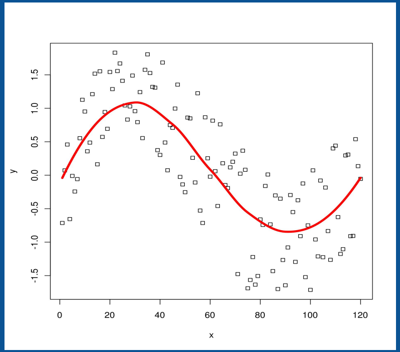
weighting strategies for seasonal component

Background

Background

Time series decomposition

STL: Seasonal-Trend decomposition based on LOESS

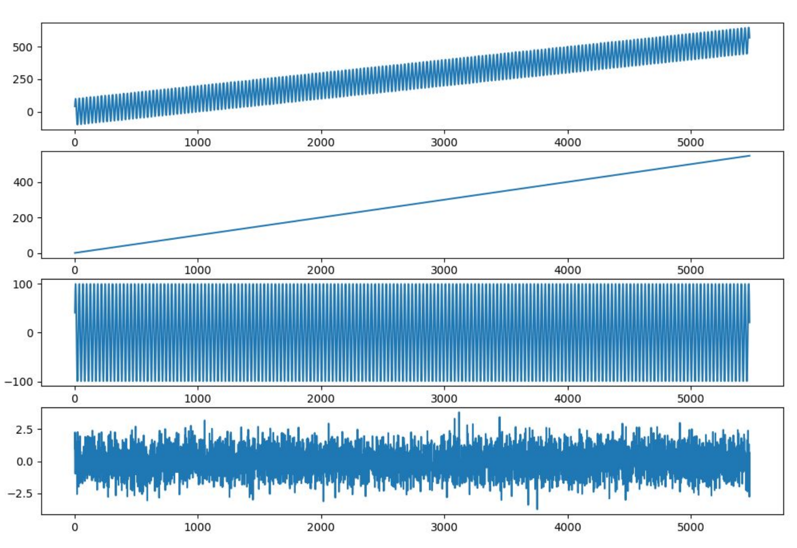


$$Y_t = T_t \circ S_t \circ R_t$$

Trend

Season

Remainder



Background

Exchangeability regimes

Global exchangeability

Definition 1 (Exchangeability) *A sequence of random variables $X_1, \ldots, X_n$ is exchangeable if for any permutation $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ with $n \geq 1$ we have that*

$$P(X_1, \ldots, X_n) = P(X_{\pi(1)}, \ldots, X_{\pi(n)}).$$

Local exchangeability

Weighted CP

$$\hat{F}_w(S) = \sum_{i=1}^{n} \tilde{w}_i \, \delta_{s_i} + \tilde{w}_{n+1} \, \delta_{+\infty}$$

Definition 2 (Local exchangeability) *Consider a sequence of random variables $X_1, \ldots, X_n$ and denote its index set by $\mathcal{I} := \{1, \ldots, n\}$. For any subset $\mathcal{I}' := \{1, \ldots, k\} \subset \mathcal{I}$ the associated random variables $X_1, \ldots, X_k$ are locally (approximately) exchangeable if for any permutation $\tilde{\pi} : \mathcal{I}' \to \mathcal{I}'$ with $|\mathcal{I}'| \geq 1$ we have that $P(X_1, \ldots, X_k) \simeq P(X_{\tilde{\pi}(1)}, \ldots, X_{\tilde{\pi}(k)})$.*

Lack of exchangeability

Adaptive Conformal Inference

$$\alpha_{t+1} = \alpha_t + \gamma \left(\alpha - \mathbb{1}[Y_t \notin \hat{C}(X_t)]\right)$$

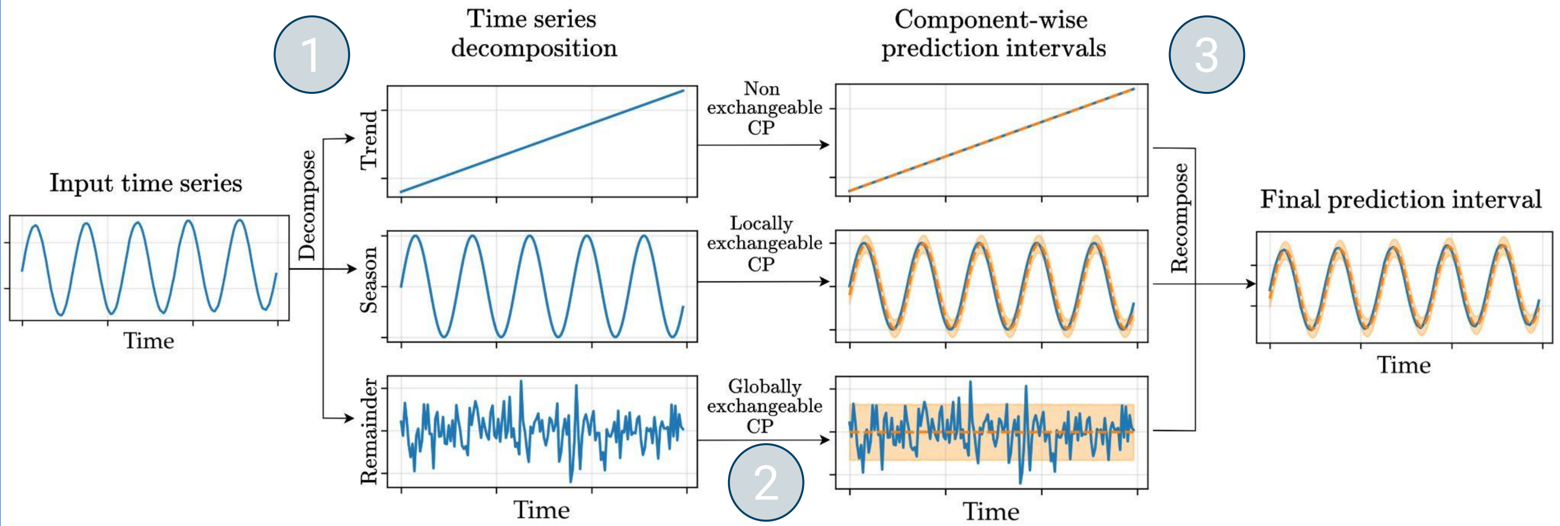
Ensemble batch prediction intervals

$$S_{t+1} := (S_t \backslash \{s_1\}) \cup \{s_{t+1}\}$$

Method

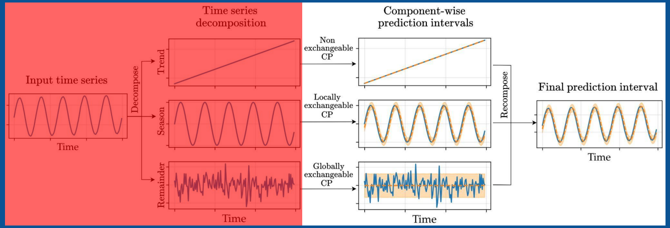
Method

Leveraging exchangeability regimes for time series components



Method

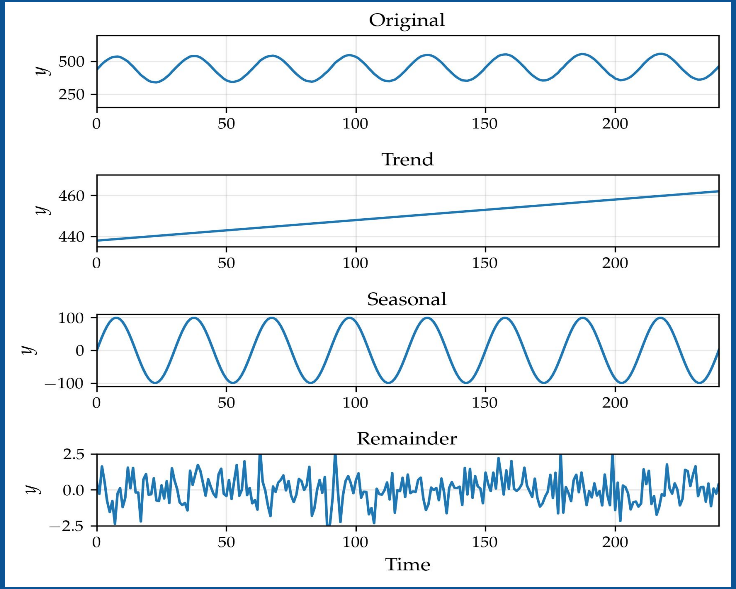
1) Time series decomposition



Trend

Season

Remainder



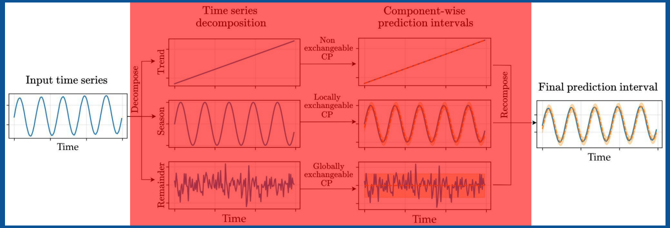
Non-exchangeable

Local exch.

Global exch.

Method

2) Component-wise prediction intervals



Trend component
EnbPI and ACI

Remainder component
CV+

Seasonal component (our contribution)

BinaryPoint

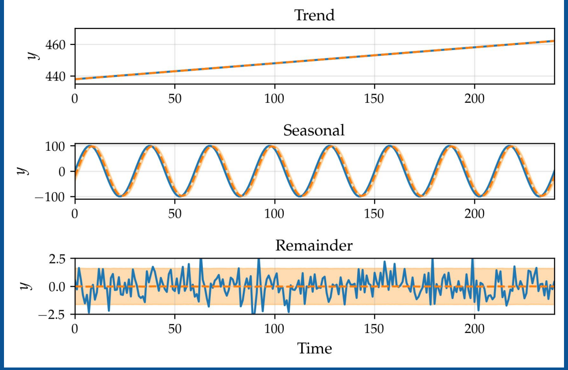
$$\mathcal{I}_{cal}^{\text{BP}} := \{i \in \mathcal{I}_{cal} : i \bmod \tau = \tau_{n+1}\}$$

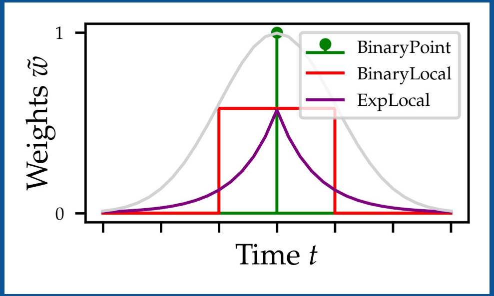
BinaryLocal

Local neighbourhood around τ_{n+1}

ExpLocal

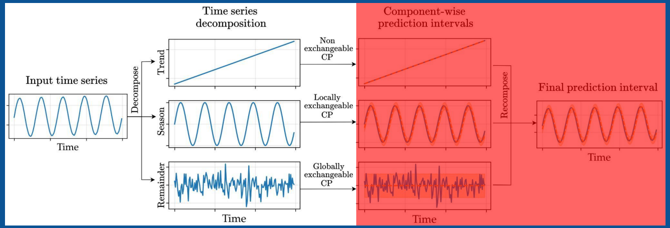
$$\lambda e^{-\lambda \cdot \Delta\tau_i} \qquad \Delta\tau_i = |\tau_i - \tau_{n+1}|$$





Method

3) Prediction interval (PI) recomposition



$$\hat{L}(X_{n+1}) = \hat{L}_T(X_{n+1}) + \hat{L}_S(X_{n+1}) + \hat{L}_R(X_{n+1})$$
$$\hat{U}(X_{n+1}) = \hat{U}_T(X_{n+1}) + \hat{U}_S(X_{n+1}) + \hat{U}_R(X_{n+1})$$

Loose lower bound on nominal coverage of $1 - 3\alpha$

Experiments

Experimental Setup

Evaluation metrics

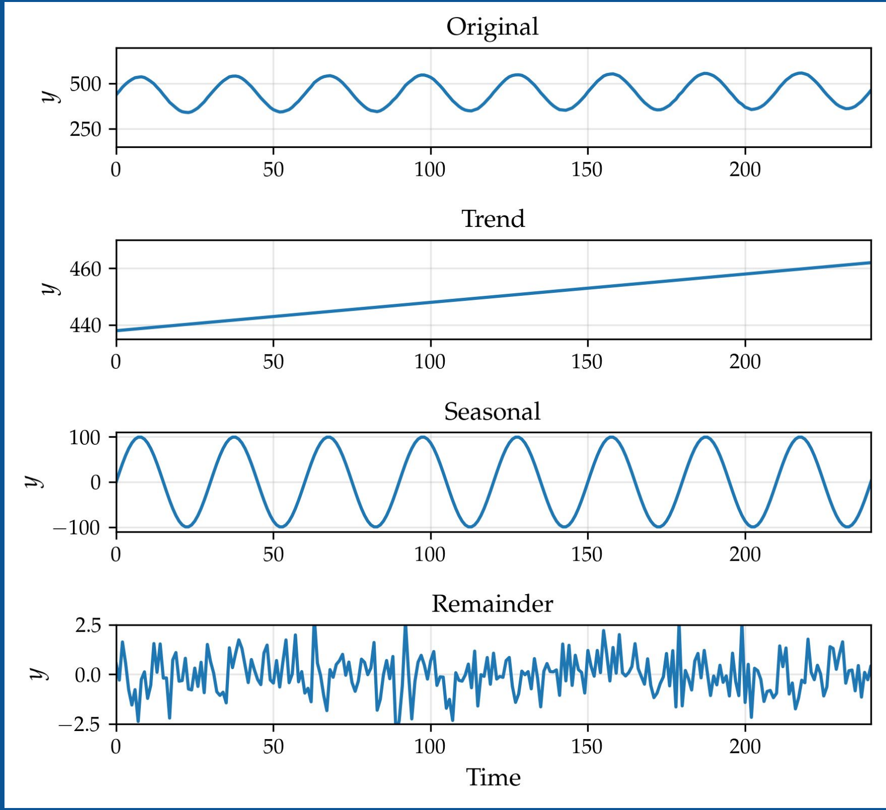
$$\mathrm{PICP} = \frac{1}{n_t} \sum_{j=1}^{n_t} 1 \left[Y_j \in \hat{C}(X_j) \right]$$

$$\mathrm{PIAW} = \frac{1}{n_t} \sum_{j=1}^{n_t} \left| \hat{C}(X_j) \right|$$

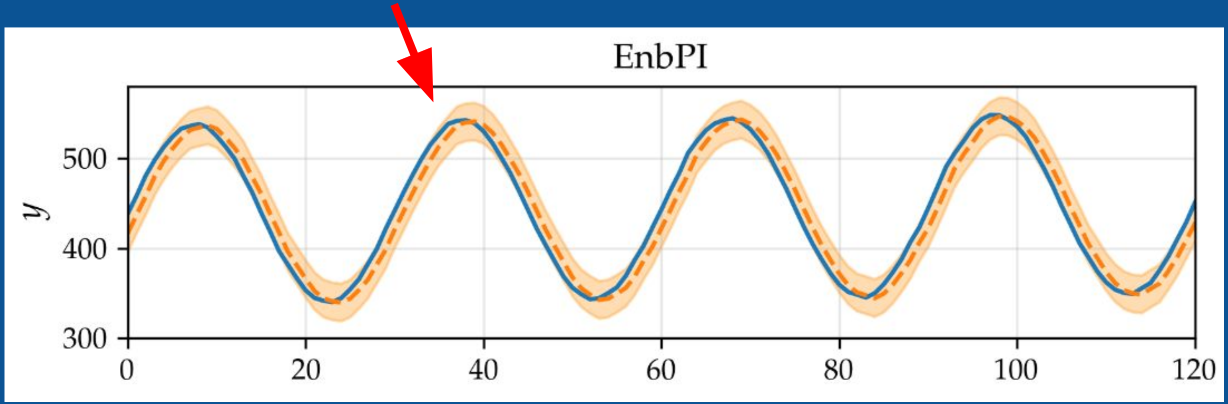
Synthetic data

$$T_t = 0.1t, \quad S_t = 100 \cdot \sin\left(2\pi \cdot \frac{t}{30}\right),$$

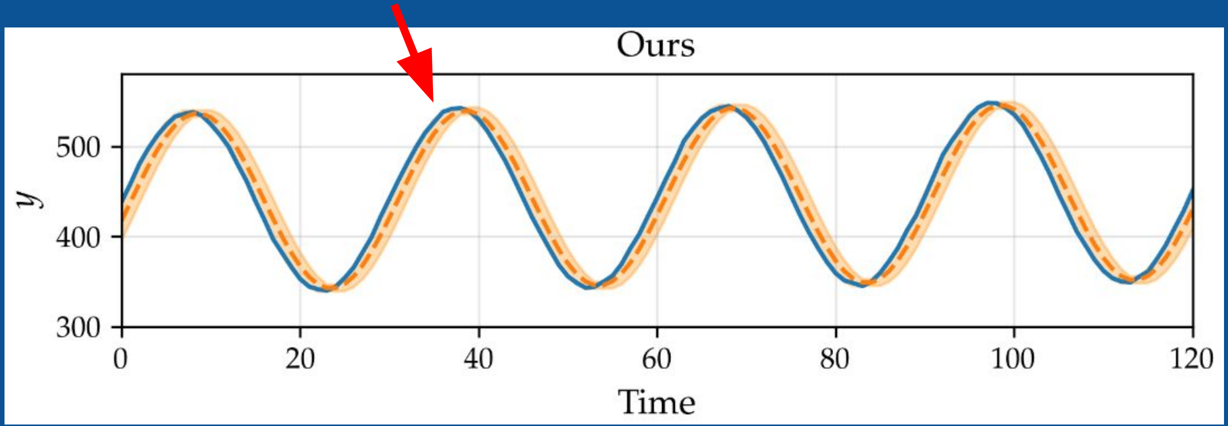
$$R_t \sim \mathcal{N}(0, 1), \quad t = 1, \ldots, T$$



Synthetic data



$$\mathrm{PICP} = 0.89$$
$$\mathrm{PIAW} = 41.7$$

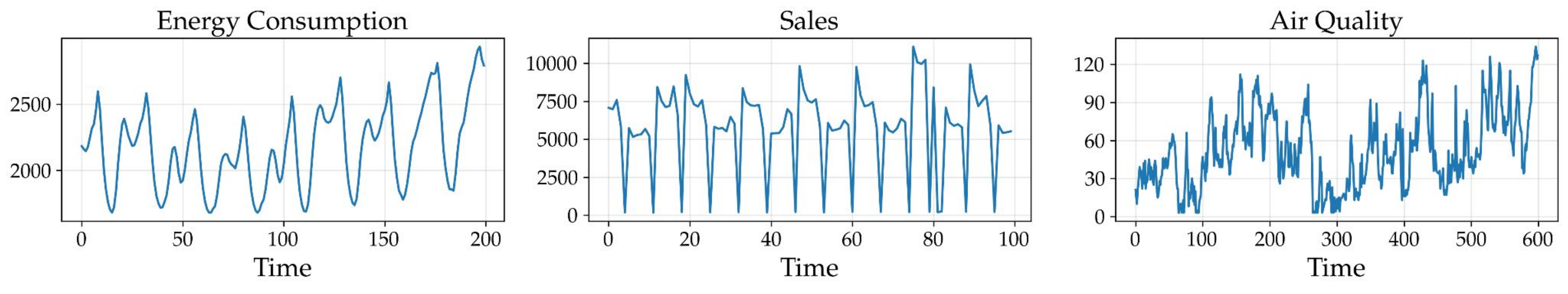


$$\mathrm{PICP} = 0.95$$
$$\mathrm{PIAW} = 29.8$$

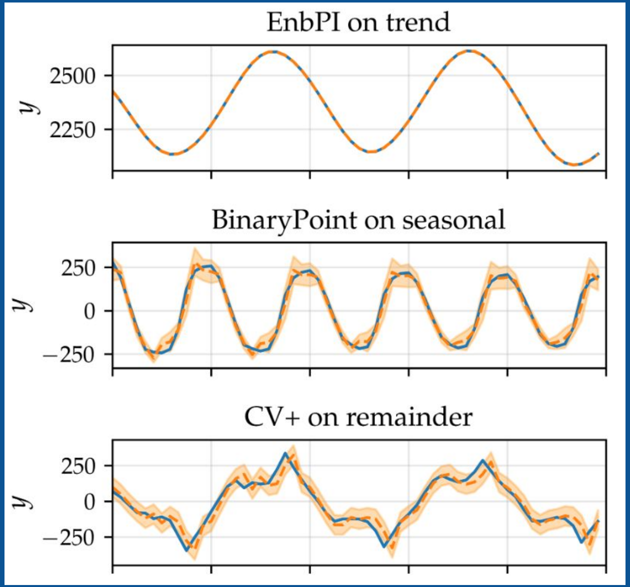
Synthetic data

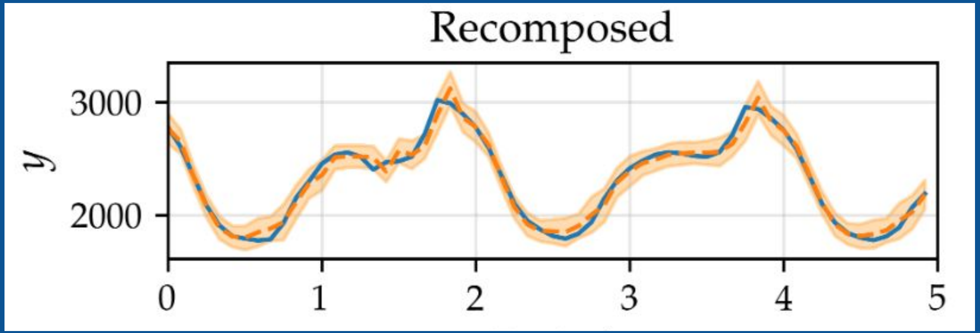
Method	Decomposed	Linear Reg.		MLP Reg.		Gradient Boosting	
		PICP	*PIAW* (↓)	*PICP*	*PIAW* (↓)	*PICP*	*PIAW* (↓)
EnbPI	✗	0.889	41.694	0.887	41.626	0.897	11.922
EnbPI	✓	0.985	44.013	0.985	44.197	0.987	12.536
ACI	✗	0.898	42.065	0.898	42.168	0.900	11.991
ACI	✓	0.985	43.995	0.986	44.237	0.994	554.298
BinaryPoint (Ours)	✓	0.946	**29.755**	0.947	**29.771**	0.947	**9.524**
BinaryLocal (Ours)	✓	0.994	41.323	0.994	41.347	0.990	11.942
ExpLocal (Ours)	✓	0.994	44.086	0.993	44.009	0.990	12.510

Real-life data

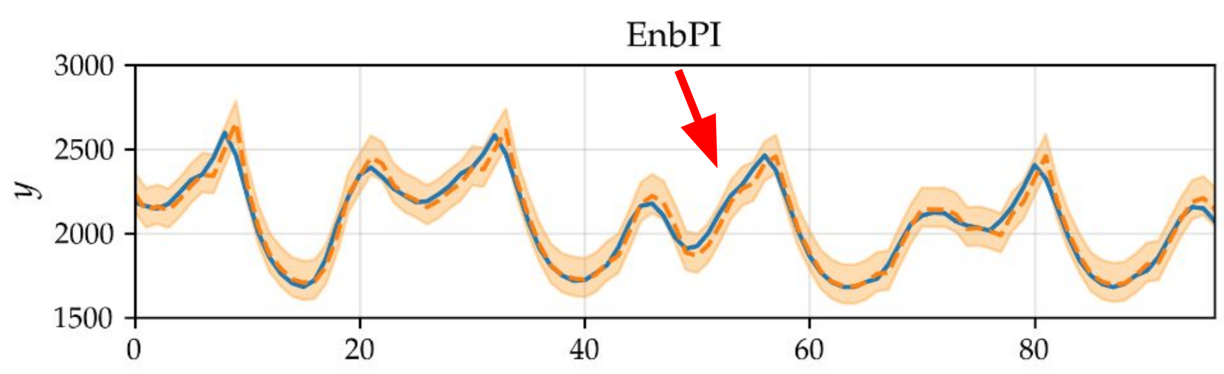


Energy consumption



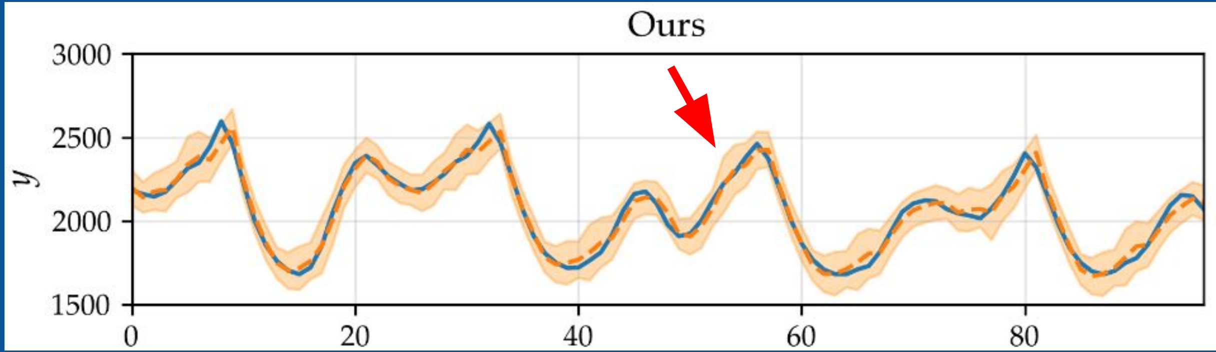


Energy consumption



$\mathrm{PICP} = 0.91$

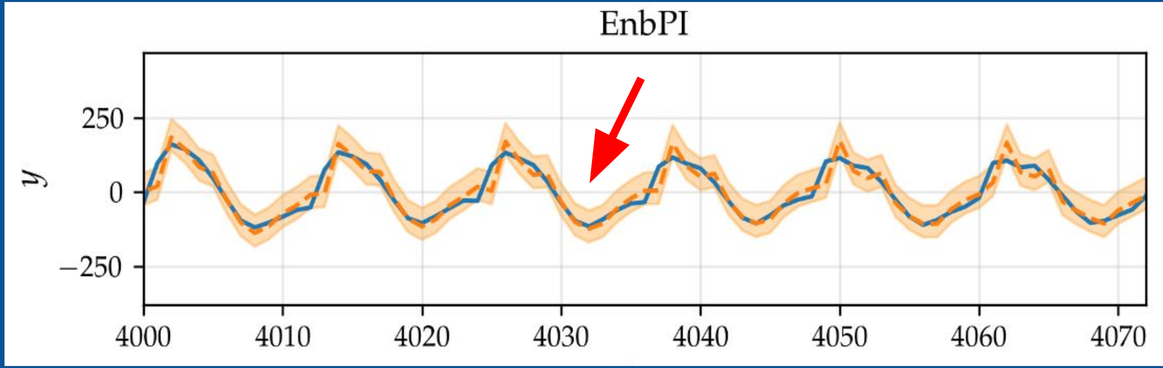
$\mathrm{PIAW} = 223$



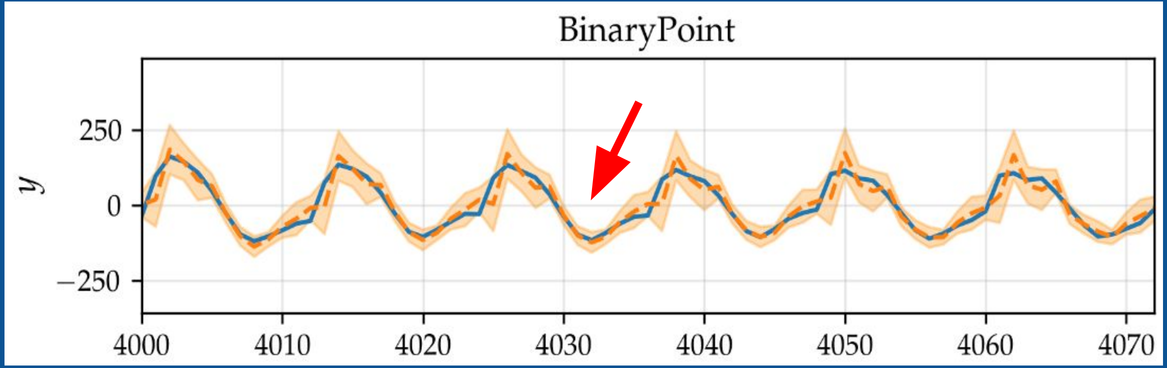
$\mathrm{PICP} = 0.97$

$\mathrm{PIAW} = 233$

Energy consumption - Seasonal component

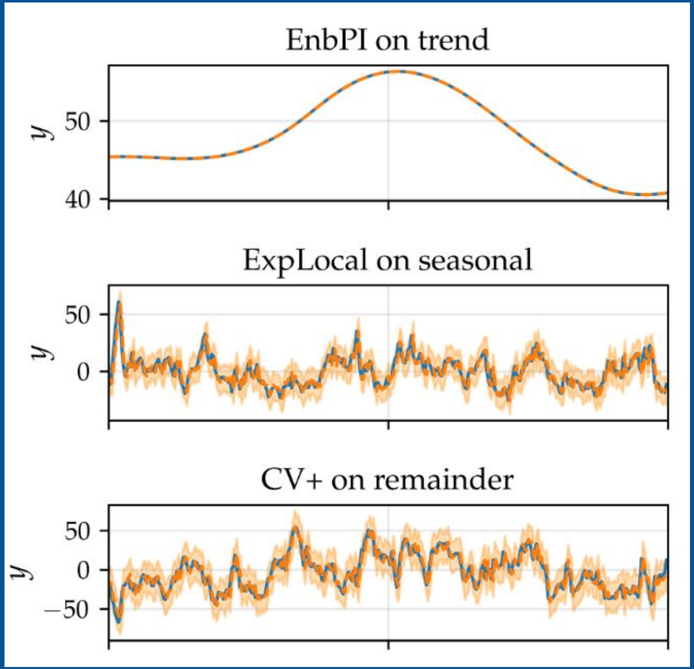


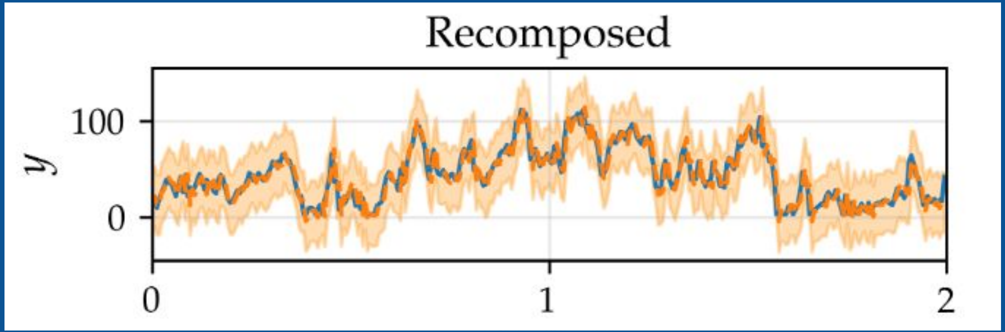
$\mathrm{PICP} = 0.92$
$\mathrm{PIAW} = 108$



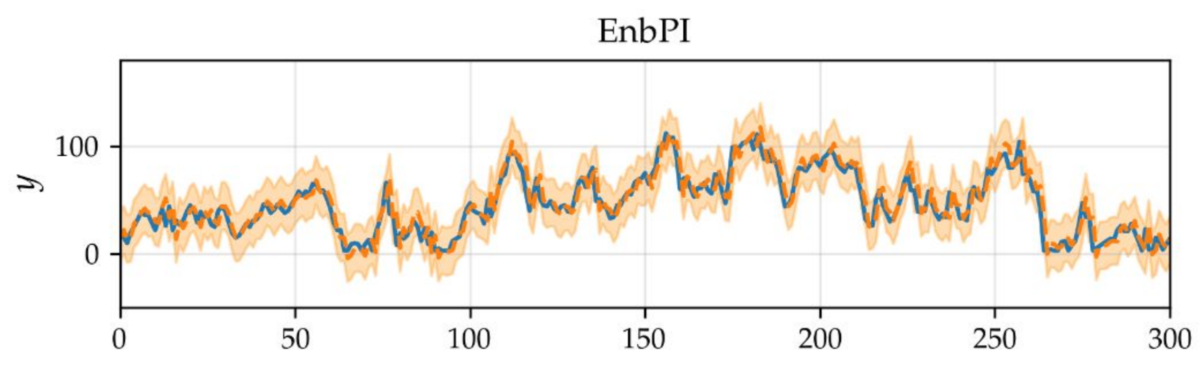
$\mathrm{PICP} = 0.91$
$\mathrm{PIAW} = 98$

Air Quality

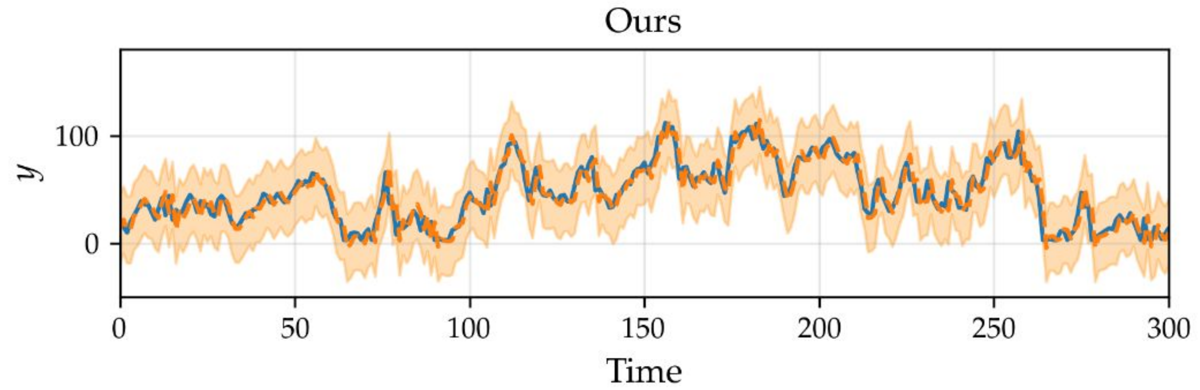

EnbPI on trend
y
50
40
ExpLocal on seasonal
y
50
0
CV+ on remainder
y
50
0
−50


Recomposed
y
100
0
0
1
2

Air Quality



$$\text{PICP} = 0.90$$
$$\text{PIAW} = 43$$



$$\text{PICP} = 0.95$$
$$\text{PIAW} = 60.9$$

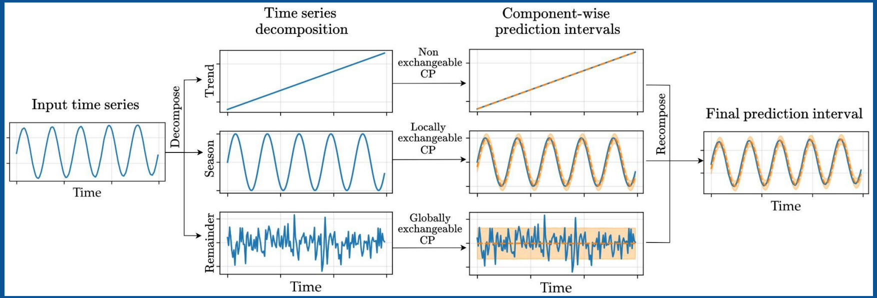
Conclusion

- Effectiveness of method heavily depends on:
  - Complexity of TS
  - Quality of decomposition
- Recomposing predictions results in overcoverage
- Promising results on individual components

Future work

In each step of the pipeline improvements can be made

- More advanced TSD algorithms
  - Multiple STL (MSTL) (Bandara et al., 2021)
  - RobustSTL (Wen et al., 2019)
  - OnlineSTL (Mishra et al., 2022)
- More sophisticated methods for aggregating prediction intervals
  - (Yang et al., 2024)



Questions?

Appendix: coverage guarantee

$$
\begin{aligned}
&\mathbb{P}(Y_{n+1} \in \hat{C}(X_{n+1})) \\
&\quad = \mathbb{P}(T_{n+1} \in \hat{C}_T(X_{n+1}) \wedge S_{n+1} \in \hat{C}_S(X_{n+1}) \wedge R_{n+1} \in \hat{C}_R(X_{n+1})) \\
&\quad = 1 - \mathbb{P}(T_{n+1} \notin \hat{C}_T(X_{n+1}) \vee S_{n+1} \notin \hat{C}_S(X_{n+1}) \vee R_{n+1} \notin \hat{C}_R(X_{n+1})) \\
&\quad \geq 1 - \Big[\underbrace{\mathbb{P}(T_{n+1} \notin \hat{C}_T(X_{n+1}))}_{\leq \alpha_T} + \underbrace{\mathbb{P}(S_{n+1} \notin \hat{C}_S(X_{n+1}))}_{\leq \alpha_S} + \underbrace{\mathbb{P}(R_{n+1} \notin \hat{C}_R(X_{n+1}))}_{\leq \alpha_R}\Big] \\
&\quad \geq 1 - 3\,\alpha.
\end{aligned}
$$

Appendix

B.1. Synthetic data.

Conformal algorithm			Linear Reg.		MLP		Gradient Boost.	
Trend	Season	Remainder	PICP	PIAW	PICP	PIAW	PICP	PIAW
Original (EnbPI)			0.889	41.694	0.887	41.626	0.897	11.922
Original (ACI)			0.898	42.065	0.898	42.168	0.9	11.991
Original (CV+)			0.879	41.645	0.917	42.41	0.904	12.031
EnbPI	EnbPI	EnbPI	0.985	44.013	0.985	44.197	0.987	12.536
EnbPI	BinaryPoint	EnbPI	0.944	29.764	**0.946**	**29.768**	**0.948**	**9.494**
EnbPI	BinaryLocal	EnbPI	0.994	41.332	0.994	41.344	0.99	11.913
EnbPI	ExpLocal	EnbPI	0.994	44.096	0.993	44.005	0.99	12.481
EnbPI	EnbPI	CV+	0.985	44.003	0.985	44.201	0.988	12.566
EnbPI	BinaryPoint	CV+	**0.946**	**29.755**	0.947	29.771	0.947	9.524
EnbPI	BinaryLocal	CV+	0.994	41.323	0.994	41.347	0.99	11.942
EnbPI	ExpLocal	CV+	0.994	44.086	0.993	44.009	0.99	12.510
ACI	ACI	ACI	0.985	43.995	0.986	44.237	0.994	554.298
ACI	BinaryPoint	ACI	0.946	29.793	0.946	29.822	0.969	551.304
ACI	BinaryLocal	ACI	0.995	41.361	0.995	41.398	0.996	553.722
ACI	ExpLocal	ACI	0.994	44.125	0.994	44.06	0.996	554.29
ACI	ACI	CV+	0.984	43.956	0.985	44.186	0.993	554.321
ACI	BinaryPoint	CV+	**0.946**	**29.755**	0.947	29.771	0.968	551.326
ACI	BinaryLocal	CV+	0.994	41.323	0.994	41.347	0.995	553.744
ACI	ExpLocal	CV+	0.994	44.086	0.993	44.009	0.995	554.312

Table 3: Detailed results for synthetic data.

Appendix

B.2. Real-world data: San Diego energy consumption.

This dataset records energy usage in the city of San Diego, collected over a period of five years. The response consists of hourly energy consumption measurements in Watt hours. Available time-dependent features include seasonal indicators, public holidays, temperature, and air conditioning use, among others.

Conformal algorithm			Linear Reg.		MLP		Gradient Boost.	
Trend	Season	Remainder	PICP	PIAW	PICP	PIAW	PICP	PIAW
	Original (EnbPI)		0.907	222.866	0.865	161.238	0.899	199.334
	Original (ACI)		<u>0.9</u>	<u>218.301</u>	0.9	187.367	0.9	204.007
	Original (CV+)		0.912	227.846	0.898	168.878	<u>0.9</u>	<u>198.872</u>
EnbPI	EnbPI	EnbPI	0.964	243.005	**<u>0.944</u>**	**<u>161.773</u>**	0.966	271.898
EnbPI	BinaryPoint	EnbPI	0.97	232.852	0.961	172.657	0.973	272.751
EnbPI	BinaryLocal	EnbPI	0.968	238.593	0.963	177.651	0.977	281.453
EnbPI	ExpLocal	EnbPI	0.968	238.113	0.963	176.555	0.977	281.194
EnbPI	EnbPI	CV+	0.963	245.122	0.956	170.989	0.967	273.554
EnbPI	BinaryPoint	CV+	0.971	234.969	0.972	181.872	0.975	274.407
EnbPI	BinaryLocal	CV+	0.967	240.71	0.974	186.867	0.977	283.108
EnbPI	ExpLocal	CV+	0.967	240.23	0.974	185.77	0.977	282.849
ACI	ACI	ACI	0.96	229.425	0.966	190.995	0.968	278.927
ACI	BinaryPoint	ACI	**0.969**	**228.552**	0.969	181.361	**0.97**	**270.637**
ACI	BinaryLocal	ACI	0.967	234.293	0.97	186.355	0.974	279.338
ACI	ExpLocal	ACI	0.968	233.814	0.97	185.259	0.973	279.07
ACI	ACI	CV+	0.962	235.855	0.971	191.174	0.972	286.686
ACI	BinaryPoint	CV+	0.971	234.983	0.971	181.54	0.974	278.396
ACI	BinaryLocal	CV+	0.967	240.724	0.973	186.534	0.977	287.097
ACI	ExpLocal	CV+	0.967	240.245	0.973	185.438	0.977	286.838

Table 4: Detailed results for San Diego energy consumption data.

Appendix

B.3. Real-world data: Rossman store sales.

This dataset contains aggregated daily sales from the chain of Rossmann drug stores. The response consists of average daily sales numbers. Available time-dependent features include weekday, average number of customers, and whether holiday indicators.

Conformal algorithm			Linear Reg.		MLP		Gradient Boost.	
Trend	Season	Remainder	PICP	PIAW	PICP	PIAW	PICP	PIAW
	Original (EnbPI)		0.892	1572.343	0.923	1231.569	0.932	5476.084
	Original (ACI)		0.895	1539.925	0.906	1172.582	0.919	4867.123
	Original (CV+)		0.892	1512.231	0.928	1276.386	0.924	5470.683
EnbPI	EnbPI	EnbPI	1.0	5303.825	0.975	4870.155	**0.932**	**6075.189**
EnbPI	BinaryPoint	EnbPI	1.0	6060.792	0.989	6404.697	0.978	8051.558
EnbPI	BinaryLocal	EnbPI	1.0	5441.858	0.984	5553.258	0.973	7547.193
EnbPI	ExpLocal	EnbPI	1.0	5263.612	0.984	5381.6	0.658	7373.728
EnbPI	EnbPI	CV+	1.0	5288.406	0.986	5297.233	0.935	6284.836
EnbPI	BinaryPoint	CV+	1.0	6045.373	1.0	6831.775	0.98	8261.205
EnbPI	BinaryLocal	CV+	1.0	5426.439	0.996	5980.336	0.968	7756.839
EnbPI	ExpLocal	CV+	1.0	5248.193	0.996	5808.679	0.976	7583.375
ACI	ACI	ACI	1.0	5447.39	**0.984**	**4615.531**	0.944	6373.45
ACI	BinaryPoint	ACI	1.0	6015.128	0.993	6313.72	0.977	8003.03
ACI	BinaryLocal	ACI	1.0	5396.194	0.987	5462.281	0.973	7498.665
ACI	ExpLocal	ACI	**1.0**	**5217.948**	0.974	5290.623	0.966	7325.201
ACI	ACI	CV+	1.0	5477.696	0.989	5201.327	0.944	6623.43
ACI	BinaryPoint	CV+	1.0	6045.434	1.0	6899.516	0.98	8253.01
ACI	BinaryLocal	CV+	1.0	5426.5	0.996	6048.076	0.968	7748.645
ACI	ExpLocal	CV+	1.0	5248.254	0.996	5876.419	0.968	7575.181

Table 5: Detailed results for Rossman store sales data.

Appendix

B.4. Real-world data: Beijing air quality.

This dataset consists of hourly readings of air pollutants and meteorological data from the city of Beijing. The primary response is the concentration of PM2.5 pollutants in the air. Available time-dependent features include various pollutant concentrations (*e.g.*, SO2, NO2, CO, O3); air temperature and pressure; or rainfall, wind direction and speed.

Conformal algorithm			**Linear Reg.**		**MLP**		**Gradient Boost.**	
Trend	Season	Remainder	PICP	PIAW	PICP	PIAW	PICP	PIAW
	Original (EnbPI)		0.902	42.918	0.892	40.098	0.905	44.889
	Original (ACI)		0.901	43.549	0.901	44.733	0.901	46.435
	Original (CV+)		0.908	44.499	0.915	44.398	0.913	46.909
EnbPI	EnbPI	EnbPI	0.954	62.721	0.955	64.756	0.951	64.74
EnbPI	BinaryPoint	EnbPI	0.956	65.643	0.959	68.696	0.952	68.127
EnbPI	BinaryLocal	EnbPI	0.953	61.345	0.956	64.376	0.948	63.489
EnbPI	ExpLocal	EnbPI	**0.953**	**60.869**	**0.956**	**63.69**	**0.964**	**62.882**
EnbPI	EnbPI	CV+	0.956	64.092	0.962	67.586	0.954	66.537
EnbPI	BinaryPoint	CV+	0.958	67.014	0.963	71.526	0.954	69.924
EnbPI	BinaryLocal	CV+	0.955	62.716	0.96	67.206	0.951	65.286
EnbPI	ExpLocal	CV+	0.954	62.24	0.96	66.52	0.950	64.68
ACI	ACI	ACI	0.958	63.445	0.961	67.267	0.957	74.507
ACI	BinaryPoint	ACI	0.962	66.195	0.965	70.108	0.96	76.922
ACI	BinaryLocal	ACI	0.957	61.897	0.961	65.789	0.958	72.284
ACI	ExpLocal	ACI	0.957	61.421	0.961	65.103	0.976	71.678
ACI	ACI	CV+	0.958	64.264	0.965	68.233	0.961	74.603
ACI	BinaryPoint	CV+	0.958	67.014	0.963	71.074	0.956	77.019
ACI	BinaryLocal	CV+	0.955	62.716	0.96	66.755	0.953	72.381
ACI	ExpLocal	CV+	0.954	62.24	0.959	66.068	0.953	71.774

Table 6: Detailed results for Beijing air quality data.