Sequential Predictive Conformal Inference for Time Series

Chen Xu, Yao Xie

H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Motivation and objective

- Predictive inference beyond point prediction.

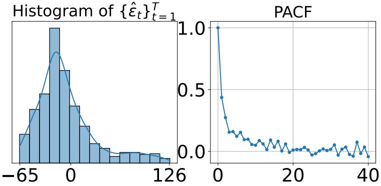
- Conformal prediction (CP) has been gaining popularity (distribution-free and model-free).

Motivation and objective

- Predictive inference beyond point prediction.

- Conformal prediction (CP) has been gaining popularity (distribution-free and model-free).

- Limitation to dependent and sequential data.

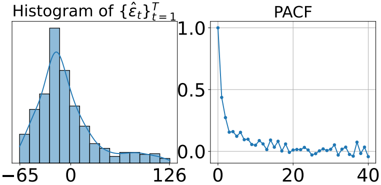


Motivation and objective

- Predictive inference beyond point prediction.

- Conformal prediction (CP) has been gaining popularity (distribution-free and model-free).

- Limitation to dependent and sequential data.



- **Goal:** Sequentially construct CP prediction intervals by utilizing feedback and dependency within time-series data.

Related works

- Traditional CP: regression (Papadopoulos et al., 2007; Barber et al., 2019; Kim et al., 2020) and classification (Romano et al., 2020; Angelopoulos et al., 2021).

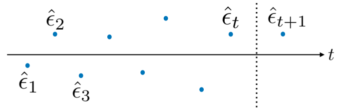
Related works

- Traditional CP: regression (Papadopoulos et al., 2007; Barber et al., 2019; Kim et al., 2020) and classification (Romano et al., 2020; Angelopoulos et al., 2021).

- Sequential CP: leverage feedback during prediction (Xu & Xie, 2021; Gibbs & Candes, 2021; Zaffran et al., 2022; Barber et al., 2022).

$$(\hat{\epsilon}_1, \hat{\epsilon}_2, \ldots, \hat{\epsilon}_t) \overset{d}{=} (\hat{\epsilon}_{\sigma(1)}, \hat{\epsilon}_{\sigma(2)}, \ldots, \hat{\epsilon}_{\sigma(t)})$$

Traditional CP

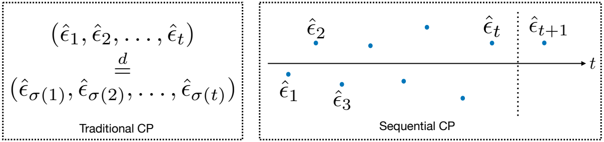


Sequential CP

Related works

- Traditional CP: regression (Papadopoulos et al., 2007; Barber et al., 2019; Kim et al., 2020) and classification (Romano et al., 2020; Angelopoulos et al., 2021).

- Sequential CP: leverage feedback during prediction (Xu & Xie, 2021; Gibbs & Candes, 2021; Zaffran et al., 2022; Barber et al., 2022).



- Probabilistic forecasting via quantile regression: (Wen et al., 2017; Salinas et al., 2020; Lim et al., 2021)

Proposed SPCI

- In essence, SPCI performs time-adaptive re-estimation of residual quantiles by leveraging dependency among residuals.

Proposed `SPCI`

• In essence, `SPCI` performs time-adaptive re-estimation of residual quantiles by leveraging dependency among residuals.

• Data (X_t, Y_t) arrive sequentially, with unknown $Y_t|X_t$. Given T initial samples, we train a point predictor $\hat{f}$ and obtain *prediction* residuals $\hat{\epsilon}$ (Papadopoulos et al., 2007; Xu & Xie, 2021).

Proposed SPCI

- In essence, SPCI performs time-adaptive re-estimation of residual quantiles by leveraging dependency among residuals.

- Data (X_t, Y_t) arrive sequentially, with unknown $Y_t|X_t$. Given T initial samples, we train a point predictor $\hat{f}$ and obtain *prediction* residuals $\hat{\boldsymbol{\epsilon}}$ (Papadopoulos et al., 2007; Xu & Xie, 2021).

- We then fit a quantile regression estimator $\widehat{Q}_t$ on $\hat{\boldsymbol{\epsilon}}$.

- For $\alpha \in (0, 1)$, SPCI interval with feature X_t is

$$\widehat{C}_{t-1}(X_t) = [\hat{f}(X_t) + \widehat{Q}_t(\hat{\beta}), \hat{f}(X_t) + \widehat{Q}_t(1 - \alpha + \hat{\beta})], \quad (1)$$

- Intervals in (1) are sequentially constructed on updated $\hat{\boldsymbol{\epsilon}}$.

Remarks on SPCI

- **Generality**:
  - Most CP methods use the empirical quantile $\widehat{Q}_t$.
  - SPCI can be used with any quantile estimator.

    We use quantile random forest (QRF, (Meinshausen, 2006)) in this work to develop theoretical guarantees.

Remarks on `SPCI`

- **Generality**:
  - Most CP methods use the empirical quantile $\widehat{Q}_t$.
  - `SPCI` can be used with any quantile estimator.

We use quantile random forest (QRF, (Meinshausen, 2006)) in this work to develop theoretical guarantees.

- **Differences**:
  - Versus CQR (Romano et al., 2019): fits quantile estimator on the original series.
  - Versus Prob. forecasting methods: we utilize a hybrid approach with theoretical guarantees.

Theoretical results

- Extend guarantee of QRF (Meinshausen, 2006) on i.i.d. data.

Theoretical results

- Extend guarantee of QRF (Meinshausen, 2006) on i.i.d. data.

- We primarily impose constraints on the covariance function over residuals to avoid strong dependency.

Theoretical results

- Extend guarantee of QRF (Meinshausen, 2006) on i.i.d. data.

- We primarily impose constraints on the covariance function over residuals to avoid strong dependency.

- Theorem (informal): We have for any $\alpha \in (0,1)$ that

$$|\mathbb{P}(Y_t \in \widehat{C}_{t-1}(X_t)|X_t) - (1-\alpha)| \xrightarrow{p} 0 \tag{2}$$

- Guaranete (2) is asymptotic, but we show satisfactory marginal coverage performance with finite samples.

Experiments (simulation)

- Non-stationary time-series: $Y_t = f(X_t) + \epsilon_t$, with

$$
\begin{aligned}
f(X_t) &= g(t)h(X_t). \\
g(t) &= \log(t')\sin(2\pi t'/12), t' = \mathrm{mod}(t, 12). \\
h(X_t) &= (|\beta^T X_t| + (\beta^T X_t)^2 + |\beta^T X_t|^3)^{1/4}.
\end{aligned}
\tag{3}
$$

Table: Simulation on non-stationary time-series with $\alpha = 0.1$. We show that SPCI outperforms baselines in terms of interval width without sacrificing valid coverage.

SPCI		EnbPI		AdaptiveCI		NEX-CP	
Coverage	Width	Coverage	Width	Coverage	Width	Coverage	Width
0.92 (2.75e-3)	12.96 (2.56e-2)	0.90 (2.21e-3)	25.41 (4.79e-2)	0.90 (4.12e-3)	28.00 (5.81e-2)	0.93 (3.10e-3)	46.50 (6.29e-2)

Experiments (real data)

- **Data:** time-series related to power generation (e.g., solar and wind) and usage (e.g., electric)
- **Baselines:** three sequential CP methods (Xu & Xie, 2021; Gibbs & Candes, 2021; Barber et al., 2022) and two probabilistic forecasting approaches (Salinas et al., 2020; Lim et al., 2021).

Table: Real time-series with $\alpha = 0.1$. Entries in the bracket indicate standard deviation over three runs. `SPCI` outperforms competitors with a much narrower interval width and does not lose coverage.

	Wind coverage	Wind width	Electric coverage	Electric width	Solar coverage	Solar width
SPCI	0.95 (1.50e-2)	2.65 (1.60e-2)	0.93 (4.79e-3)	0.22 (1.68e-3)	0.91 (1.12e-2)	47.61 (1.33e+0)
EnbPI	0.93 (6.20e-3)	6.38 (3.01e-2)	0.91 (6.84e-4)	0.32 (9.11e-4)	0.88 (4.25e-3)	48.95 (3.38e+0)
AdaptiveCI	0.95 (5.37e-3)	9.34 (3.56e-2)	0.95 (1.81e-3)	0.51 (7.25e-3)	0.96 (1.39e-2)	56.34 (1.15e+0)
NEX-CP	0.96 (8.21e-3)	6.68 (7.73e-2)	0.90 (2.05e-3)	0.45 (2.16e-3)	0.90 (7.73e-3)	102.80 (5.25e+0)
DeepAR	0.95 (5.32e-3)	6.86 (7.86e-3)	0.91 (3.45e-3)	0.62 (2.56e0-3)	0.92 (5.35e-3)	80.23 (4.94e+0)
TFT	0.92 (6.34e-2)	7.56 (5.34e-3)	0.95 (2.34e-2)	0.66 (2.34e-3)	0.93 (2.84e-3)	74.82 (4.23e+0)

Experiments (real data)

- **Data:** time-series related to power generation (e.g., solar and wind) and usage (e.g., electric)
- **Baselines:** three sequential CP methods (Xu & Xie, 2021; Gibbs & Candes, 2021; Barber et al., 2022) and two probabilistic forecasting approaches (Salinas et al., 2020; Lim et al., 2021).

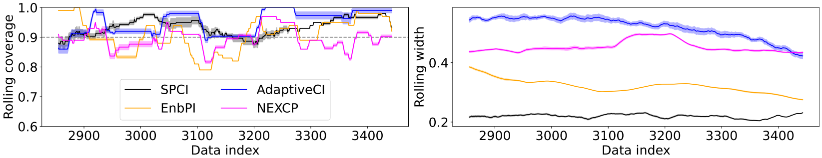
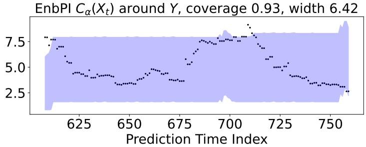


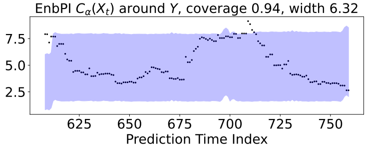
Figure: Rolling coverage and interval width over the electric time-series. `SPCI` in black not only yields valid rolling coverage but also consistently yields the narrowest prediction intervals. Furthermore, the variance of `SPCI` results over trials is also small, as shown by the shaded regions over coverage and width results.

Experiments (real data)

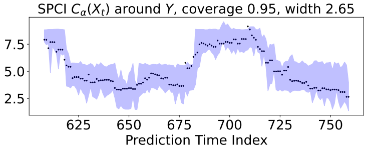
- **Goal:** `SPCI` can also be used for multi-step ahead predictive inference (i.e., construct $\widehat{C}_{t-1}(X_t)$ for multiple t at once).
- **Results:** Compared with `EnbPI` (Xu & Xie, 2021), `SPCI` intervals are much narrower and dynamic, and gets wider as prediction horizon increases to reflect greater predictive uncertainty that naturally exist.



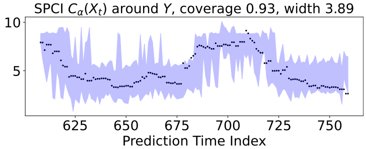
(a) EnbPI, 1 step ahead



(b) EnbPI, 4 steps ahead



(c) SPCI, 1 step ahead



(d) SPCI, 4 step ahead

Figure: Multi-step ahead prediction intervals on wind data.

Summary

• The main novelty of SPCI lies in adaptively re-estimating quantile values of future residuals.

• Against existing sequential CP and probabilistic forecasting approaches, SPCI intervals are much narrower at valid coverage.

• In the future, we will extend the approach to multi-dimensional prediction region construction.

References I

Angelopoulos, A. N., Bates, S., Jordan, M., and Malik, J. Uncertainty sets for image classifiers using conformal prediction. In *International Conference on Learning Representations*, 2021. URL `https://openreview.net/forum?id=eNdiU_DbM9`.

Barber, R. F., Candes, E. J., Ramdas, A., and Tibshirani, R. J. Predictive inference with the jackknife+, 2019.

Barber, R. F., Candes, E. J., Ramdas, A., and Tibshirani, R. J. Conformal prediction beyond exchangeability. *arXiv preprint arXiv:2202.13415*, 2022.

Gibbs, I. and Candes, E. Adaptive conformal inference under distribution shift. *Advances in Neural Information Processing Systems*, 34:1660–1672, 2021.

Kim, B., Xu, C., and Barber, R. F. Predictive inference is free with the jackknife+-after-bootstrap, 2020.

References II

Lim, B., Arık, S. Ö., Loeff, N., and Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4):1748–1764, 2021.

Meinshausen, N. Quantile regression forests. *J. Mach. Learn. Res.*, 7:983–999, 2006.

Papadopoulos, H., Vovk, V., and Gammerman, A. Conformal prediction with neural networks. In *19th IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007)*, volume 2, pp. 388–395, 2007.

Romano, Y., Patterson, E., and Candes, E. Conformalized quantile regression. In *Advances in Neural Information Processing Systems*, pp. 3543–3553, 2019.

Romano, Y., Sesia, M., and Candes, E. Classification with valid and adaptive coverage. *Advances in Neural Information Processing Systems*, 33:3581–3591, 2020.

References III

Salinas, D., Flunkert, V., Gasthaus, J., and Januschowski, T. Deepar: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3):1181–1191, 2020.

Wen, R., Torkkola, K., Narayanaswamy, B. M., and Madeka, D. A multi-horizon quantile recurrent forecaster. In *NeurIPS 2017*, 2017. URL `https://www.amazon.science/publications/a-multi-horizon-quantile-recurrent-forecaster`.

Xu, C. and Xie, Y. Conformal prediction interval for dynamic time-series. In *International Conference on Machine Learning*, pp. 11559–11569. PMLR, 2021.

Zaffran, M., Dieuleveut, A., F'eron, O., Goude, Y., and Josse, J. Adaptive conformal predictions for time series. In *ICML*, 2022.