Coverage Testing in Non-Stationary Time Series Forecasts

This repository accompanies the following **submitted manuscript**:

Title:
Testing Marginal and Conditional Coverage in Non-Stationary Time Series Forecasts through Value-at-Risk Backtesting
Author: Konrad Retzlaff
This manuscript was submitted to CoPA 2025 (Symposium on Conformal and Probabilistic Prediction with Applications) on May 17, 2025. As CoPA imposes no exclusivity requirement, this submission does not pose any copyright or concurrent submission conflict.

Project Summary

This project introduces a **statistical testing framework** for evaluating **Conformal Prediction (CP)** under **non-stationary time series**. Inspired by Value-at-Risk (VaR) backtesting, we assess CP methods using **formal hypothesis tests** for marginal validity, temporal independence, and conditional coverage.

We evaluate CP methods on:

- **Synthetic data** with changepoints and drift *(Barber et al., 2023)*
- **Electricity price forecasts** (ELEC2 dataset from Kaggle)
- **S&P 100 stock forecasts** (retrieved from Yahoo Finance API)

Folder Structure

```
data/
│
├── ELEC2/
│   ├── elec_simulation_CP.csv
│   ├── elec_simulation_nexCP_LS.csv
│   └── elec_simulation_nexCP_WLS.csv
│
├── Financial Timeseries/
│   ├── pred_calib.csv
│   ├── pred_test.csv
│   ├── quantile_calib.csv
│   ├── quantile_test.csv
│   └── Wide_S_P_100_forecasts.csv
│
└── Synthetic Data/
    ├── simulation_CP_LS.csv
    ├── simulation_nexCP_LS.csv
    └── simulation_nexCP_WLS.csv
```

Repository Files

File	Description
`ConformalPredictor.py`	Implements Split CP, CQR, and ACI
`TestingFramework.py`	Contains all 8 formal backtests
`Synthetic_data.ipynb`	Runs CP+LS, NexCP+LS, NexCP+WLS on synthetic data (Barber et al.)
`Elec_data.ipynb`	Applies all 3 Barber methods to the ELEC2 dataset
`Modeltrainer.ipynb`	Trains LightGBM model on S&P 100 stock return data
`Calib_Analysis.ipynb`	Applies Split CP, CQR, ACI and all tests to financial data
`requirements.txt`	Python packages needed to reproduce results

CP Methods Used

Synthetic & ELEC2 (Barber et al., 2023)

- **CP+LS**: Full Conformal Prediction with Least Squares
- **NexCP+LS**: Non-exchangeable CP with exponential weights
- **NexCP+WLS**: Non-exchangeable CP with Weighted Least Squares

Financial Time Series (S&P 100)

- **Split CP**
- **Conformalized Quantile Regression (CQR)**
- **Adaptive Conformal Inference (ACI)**

Statistical Tests (8 total, grouped in 4 categories)

Category	Purpose	Included Tests
1. Marginal Coverage	Is average coverage close to target rate α?	• Binomial Test (Lower) • Binomial Test (Upper) • Binomial Test (Two-sided) • First Geometric Test (for parameter (a))
2. Violation Independence	Are violations temporally independent?	• Second Geometric Test (for parameter (b))
3. Conditional Coverage	Are violations predictable from inputs?	• Joint Geometric Test (for (a) and (b)) • Dynamic Binary Test
4. Interval Score Comparison	Are intervals both valid and efficient?	• Comparative Interval-Score Test (Diebold-Mariano)

Installation

```
git clone https://github.com/KonradRtz/Coverage-Testing-in-Non-stationary-Time-
Series.git
cd Coverage-Testing-in-Non-stationary-Time-Series
pip install -r requirements.txt
```

Dependencies include:
`numpy`, `pandas`, `scipy`, `matplotlib`, `scikit-learn`, `lightgbm`

How to Run

Synthetic Data

Run `Synthetic_data.ipynb` to evaluate CP+LS, NexCP+LS, and NexCP+WLS.
Default predictions are included in `data/Synthetic Data/`.

> **Alternative**: You may regenerate synthetic predictions using the official implementation by Barber et al.:
> https://rinafb.github.io/code/nonexchangeable_conformal.zip
> Convert the output to:
>
> ```
> method, time, true_y, lower_bound, upper_bound, violation
> ```

Electricity Forecasts (ELEC2)

Run `Elec_data.ipynb` using files in `data/ELEC2/`.
You may optionally regenerate predictions using the same external code and formatting as above.

Financial Forecasts

1. Train forecasting model in `Modeltrainer.ipynb`
2. Evaluate with Split CP, CQR, and ACI in `Calib_Analysis.ipynb` using Yahoo Finance-based data

---

Output

Results include full statistical test reports, p-values, and interval score comparisons per method.

References

- **Barber, Candès, Ramdas, Tibshirani (2023)**
  Conformal Prediction Beyond Exchangeability

arXiv:2202.13415

- **Romano, Patterson, Candès (2019)**
  Conformalized Quantile Regression
  arXiv:1905.03222

- **Gibbs & Candès (2021)**
  Adaptive Conformal Inference
  arXiv:2106.00170

- **ELEC2 Dataset**
  https://www.kaggle.com/datasets/yashsharan/the-elec2-dataset

---

Contact

kretzlaff.student@maastrichtuniversity.nl

arXiv:2202.13415

- **Romano, Patterson, Candès (2019)**
  Conformalized Quantile Regression
  arXiv:1905.03222