DS 1018 Report- Modeling and Analyzing Hourly Bike-Sharing Demand: Forecasting with ARMA and LDS, and Regime Analysis with HMM

Eric Han
New York University

Haojie Yin
New York University

Abstract

Hourly bike-sharing demand exhibits strong non-stationarity driven by daily and weekly seasonal patterns, posing challenges for short-term forecasting. Using the Capital Bikeshare dataset from 2011–2012, we study probabilistic time-series models for short-term demand prediction. Deterministic seasonality is removed using hour-based and hour-by-weekday averaging to obtain residual series that better satisfy stationarity assumptions, which are assessed via autocorrelation analysis and the Augmented Dickey–Fuller test. We fit and compare ARMA models, linear dynamical systems estimated via Kalman filtering, and hidden Markov models on the residual series. Residual forecasts are combined with seasonal components to recover demand predictions, and HMM latent states are analyzed as an exploratory tool to characterize potential regime structure in residual demand dynamics.

1 Introduction

Accurate short-term forecasting of bike-sharing demand is important for system operation and resource allocation. However, hourly bike rental counts often exhibit strong non-stationarity driven by deterministic daily and weekly seasonal patterns, such as commuting peaks on weekdays. These patterns violate the stationarity assumptions underlying many classical time-series models and complicate short-term demand prediction.

In this project, we study hourly bike-sharing demand from the Capital Bikeshare system and investigate probabilistic time-series models for short-term forecasting. We first remove deterministic seasonality using hour-based and hour-by-weekday averaging to obtain residual series that better satisfy stationarity assumptions. We then apply and compare ARMA models, linear dynamical systems estimated via Kalman filtering, and hidden Markov models on the residual series. Forecasts are generated on the residual scale and combined with seasonal components to recover demand predictions in the original scale. In addition to forecasting performance, we examine whether latent states inferred by HMMs provide insight into regime changes potentially related to external factors such as weather.

2 Data Description

The dataset Fanaee-T [2013] contains hourly records of bike rental activity from the Capital Bikeshare system over the years 2011 and 2012. The primary variable of interest is `cnt`, which represents the total number of bikes rented during each hour (bike rental count).

Each observation is indexed by date and hour, and includes calendar-related variables such as weekday and holiday indicators, as well as environmental features including weather condition and humidity. In total, the dataset contains 17,379 hourly observations with 13 recorded features.

3 Experimental Setup

3.1 Experimental Setup for Time-Series Forecasting

3.1.1 Data Partitioning

For ARMA and LDS predictive modeling, the dataset is split into a training set (80%) and a test set (20%) using a chronological split, preserving the temporal ordering of observations.

3.1.2 Target Representation and Seasonality Removal

We consider three representations of the target variable (bike rental counts) for time-series forecasting:

- **Raw count**: the original bike rental count series without preprocessing.
- **Seasonality removal (hour only)**: the average bike rental count for each hour of the day, estimated from the training data, is subtracted from the raw bike rental count series.
- **Seasonality removal (hour $\times$ weekday)**: the average bike rental count for each hour–weekday combination, estimated from the training data, is subtracted from the raw bike rental count series.

The latter two representations remove deterministic seasonal structure while preserving stochastic temporal dynamics, yielding residual series that better satisfy the stationarity assumptions of ARMA models. After seasonal mean removal, the resulting residual series is used as the target variable for model fitting and forecasting.

3.1.3 Stationarity Analysis via ACF and PACF

To assess the effectiveness of seasonality removal, we examine the autocorrelation function (ACF) and partial autocorrelation function (PACF) under each preprocessing strategy. The raw bike rental count series (Appendix 1) exhibits strong, persistent autocorrelation with repeating peaks and no stable decay, indicating pronounced nonstationarity driven by daily and weekly seasonal effects.

After removing the hourly seasonal mean, the ACF (Appendix 2) becomes irregular and noisy. Although the daily cycle is partially reduced, substantial weekly structure remains, suggesting that hour-only removal is insufficient to isolate stochastic temporal dynamics.

After removing one more weekday "dimension", the hour $\times$ weekday–removed series (Appendix 3) displays a smoother and more stable autocorrelation structure, with gradual decay and consistent oscillatory patterns at 24-hour intervals that do not grow over time. This behavior is characteristic of stationary processes rather than nonstationarity. The corresponding PACF shows smaller and rapidly diminishing partial correlations, indicating reduced residual dependence.

Overall, removing the interaction between hour and weekday most effectively eliminates deterministic seasonal structure while preserving stable stochastic dynamics, making it the most suitable representation for ARMA and LDS forecasting models.

3.2 Experimental Setup for HMM

To analyze potential latent regime structure beyond deterministic seasonality, we apply a hidden Markov model (HMM) to the seasonally adjusted residual series. Specifically, we use the residuals obtained after removing hour-by-weekday seasonal means, which exhibit reduced deterministic structure and are more suitable for regime analysis.

We model the residual series using a Gaussian HMM with two latent states. The observations consist of one-dimensional continuous residual values, and each latent state is associated with a Gaussian emission distribution. Model parameters, including transition probabilities and emission parameters, are estimated via maximum likelihood using the Expectation–Maximization (EM) algorithm as implemented in the `hmmlearn` library. The HMM is trained only on the residuals from the training set to avoid information leakage.

Unlike ARMA and LDS models, the HMM is not used as a primary forecasting model. Instead, it is employed for exploratory regime analysis, with the goal of examining whether the inferred latent states correspond to meaningful changes in demand dynamics potentially related to external factors such as weather or calendar effects.

4 Forecasting Models Fitting and Prediction

ARMA Model: Model orders for ARMA are selected based on ACF and PACF plots of the preprocessed series. Across all three input representations (raw, hour-only removal, and hour $\times$ weekday removal), the PACF consistently exhibits a strong spike at lag 1 and a second significant spike at lag 2, followed by rapid decay within the confidence bounds, as shown in ACF at Appendix 3. This pattern suggests an AR(2) component.

The ACF shows a slow and oscillatory decay without a sharp cutoff at any specific lag. Such behavior does not indicate pure a moving-average models. However, the presence of weak residual correlation at short lags suggests that a small moving-average component may help capture short-memory shock effects not fully explained by the AR terms.

Based on these diagnostics, we fit AR(2) and ARMA(2,1) models for all three data representations with corresponding target values. Then we generate one-step-ahead predictions in a rolling way on the test set.

Linear Dynamical System (LDS): Model parameters are estimated on the training data using Expectation–Maximization (EM) algorithm, and rolling one-step-ahead forecasts are generated via Kalman filtering on the test set.

5 Results and Analysis

5.1 Forecasting Results and Analysis

5.1.1 Effect of Target Representation on Forecasting Performance

Since reconstruction from residuals to bike rental counts involves adding a deterministic seasonal mean, MAE and RMSE remain unchanged after reconstruction, allowing direct comparison across target representations.

Across both MAE and RMSE, forecasting MAE and RMSE improves consistently as the target representation progresses from raw bike rental counts to hour-only residuals and further to hour $\times$ weekday residuals (Appendix B). Among all models and representations, ARMA(2,1) achieves the lowest error and is consistently the best-performing model within each representation.

5.1.2 Residual-Based Forecasting and Reconstruction Limitations

For both seasonality removal representations, residual predictions from all three models remain well aligned across the test period. However, toward later timestamps, the reconstructed bike rental counts deviate noticeably from the observed values, as illustrated in Appendix 5 and Appendix 4.

Since reconstructed bike rental counts are obtained by adding predicted residuals to the seasonal mean estimated from the training set, this divergence cannot be attributed to residual prediction error. Instead, it reflects a mismatch in the seasonal mean component. This is confirmed by Appendix 6, where the heatmap of seasonal mean differences (test minus training) shows higher seasonal means in the test set across many hour $\times$ weekday combinations. The largest increases, reaching approximately 250 additional rentals, occur around 8 AM and 5 PM on Thursdays, indicating substantial growth in peak commute demand during the testing period.

Because ARMA models assume a fixed seasonal structure learned from the training data, they cannot adapt to such shifts in the seasonal mean. This limitation explains the observed reconstruction bias and highlights a fundamental constraint of residual-based forecasting when seasonal patterns evolve over time. The same behavior is observed for both ARMA and LDS models under seasonality removal (Appendix 5).

5.1.3 Comparison Between ARMA and LDS Models

Under the same seasonality removal strategy, LDS consistently exhibits larger MAE and RMSE than ARMA, even though its predictions appear visually smoother and better aligned with the overall trend. This apparent contradiction arises from fundamental differences in how the two models respond to short-term fluctuations.

LDS predictions are generated through Kalman filtering, which inherently performs smoothing and denoising by pulling predictions toward a latent state that evolves gradually over time. So LDS adapts conservatively to sudden changes and typically requires multiple consecutive deviations to interpret them as meaningful trend shifts. However, in bike rental data, sharp peaks and drops are frequent and often carry important signals, and the delayed response of LDS to these abrupt changes leads to larger pointwise errors that disproportionately increase MAE and RMSE.

In contrast, ARMA models rely directly on recent observations to generate forecasts, making them more reactive to sudden local changes. While this reactivity introduces higher variance in predictions, it allows ARMA to track short-term fluctuations more closely, resulting in lower pointwise error despite noisier behavior. This suggests that bike rental demand remains highly volatile even after seasonality removal.

We also evaluate whether training LDS directly on raw bike rental counts improves performance by allowing the model to adapt to evolving seasonal levels. Although this approach eliminates reconstruction bias caused by shifted seasonal means, it results in even larger MAE and RMSE (Appendix B). Because raw counts have substantially larger magnitudes than residuals, mispredictions are more heavily penalized, and the same smoothing behavior of LDS leads to amplified errors. Consequently, despite better aligning capability, LDS remains less effective than residual-based ARMA for this forecasting task.

5.2 Regime Analysis with Hidden Markov Models

After completing residual-based time-series forecasting, we analyze whether the seasonally adjusted demand residuals exhibit distinct latent regimes.

5.2.1 Separation of Demand Regimes

We first verify whether the inferred HMM states correspond to meaningfully different demand behaviors. As shown in Appendix 7, the two states differ primarily in the dispersion of residual demand. State 0 exhibits a tightly concentrated distribution centered near zero, whereas state 1 displays substantially higher variance and heavier tails. This indicates that the HMM separates periods of relatively stable demand from periods with elevated volatility.

This distinction is also reflected in the distribution of original bike rental counts across states (Appendix 8). Observations assigned to state 1 consistently show higher demand levels and greater variability, confirming that the inferred states represent structurally different demand regimes.

5.2.2 State-Level Summary and Weather Patterns

To explore whether the inferred demand regimes are associated with systematic differences in environmental conditions, we summarize the average demand and weather variables by state in Appendix D. Across both the training and test sets, the low-demand regime (state 0) is associated with higher average humidity, while the high-demand regime (state 1) corresponds to lower average humidity.

These consistent differences motivate a closer examination of the relationship between humidity and the inferred demand regimes. However, the magnitude of the weather differences is modest compared to the separation observed in demand levels.

5.2.3 Distributional Analysis of Humidity and Residual Demand

To assess whether humidity meaningfully differentiates the inferred regimes, we examine the full distribution of humidity values conditioned on HMM state. As shown in Appendix 9, although the low-demand state tends to exhibit higher average humidity, the humidity distributions overlap substantially across states. This overlap suggests that humidity alone does not sharply distinguish the inferred demand regimes.

We further visualize the joint relationship between residual demand and humidity in Appendix 10. While residual demand variability differs clearly across states, no strong monotonic or sharply separable relationship between humidity and residual demand is observed. Instead, humidity varies gradually within each regime.

5.2.4 Interpretation

Taken together, these results indicate that the inferred HMM states are primarily distinguished by differences in the magnitude and variability of demand fluctuations, rather than by distinct shifts in associated weather conditions. Although humidity exhibits consistent average differences across regimes, the substantial overlap in its distribution suggests that weather variables play at most a secondary role in defining regime boundaries.

6 Conclusion

Residual-based target representations substantially improve forecasting performance compared to using raw bike rental counts. In particular, removing deterministic seasonality via hour $\times$ weekday mean removal consistently reduces MAE and RMSE across all models, with ARMA(2,1) achieving the best overall performance.

Although residual forecasts remain accurate throughout the test period, reconstruction of bike rental counts becomes biased when seasonal patterns shift over time. This bias is due to higher seasonal means in the test set, reflecting increased bike rental demand and highlighting a fundamental limitation of residual-based forecasting with fixed seasonal estimates.

Despite producing smoother forecasts, LDS shows higher MAE and RMSE due to its conservative response to abrupt demand fluctuations. In contrast, ARMA models better capture short-term variability and achieve lower pointwise error, making residual-based ARMA the most effective approach for forecasting hourly bike rental demand in this setting.

In addition to forecasting performance, this project explores the structure of demand dynamics through an exploratory regime analysis using Hidden Markov Models. The inferred latent states consistently separate periods of low and high demand and differ primarily in the magnitude and variability of residual demand, suggesting the presence of distinct demand regimes beyond deterministic seasonal effects.

While moderate differences in associated weather variables are observed across regimes, substan-

tial overlap remains, indicating that these latent states are not sharply defined by external environmental conditions. Instead, the HMM results complement the forecasting analysis by providing a descriptive view of intrinsic demand dynamics, rather than being intended as a forecasting model.

References

Hadi Fanaee-T. Bike Sharing. UCI Machine Learning Repository, 2013. DOI: https://doi.org/10.24432/C5W894.

A Seasonality ACF and PACF Plots

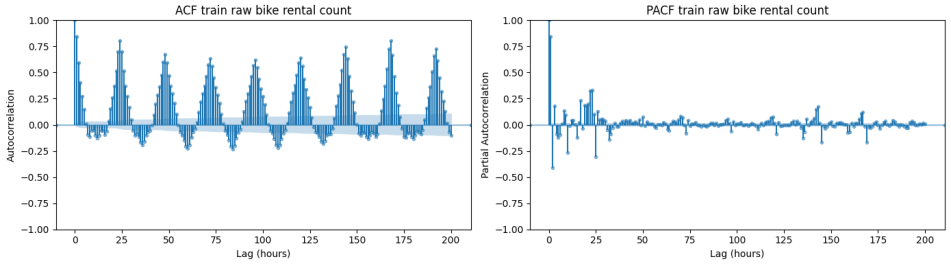

Figure 1: Raw bike rental count series.

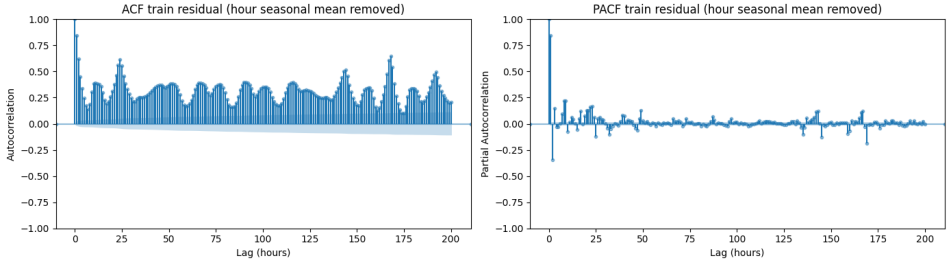

Figure 2: Hour-only seasonality removed.

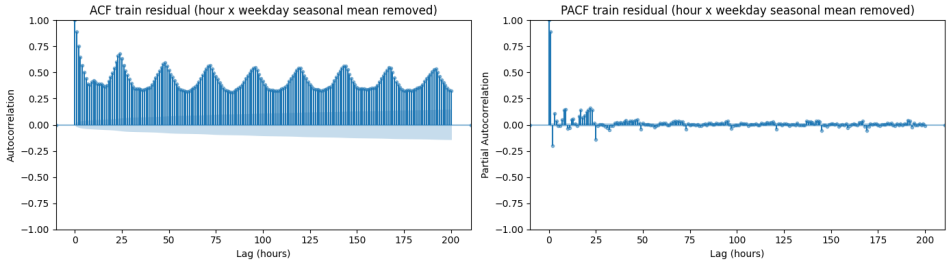

Figure 3: Hour $\times$ weekday seasonality removed.

B Forecasting Results

	Raw Count		Hour Residual		Hour $\times$ Weekday Residual	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE
AR(2)	141.95	164.78	53.58	82.84	42.05	65.42
ARMA(2,1)	110.78	135.51	53.07	81.80	41.88	64.96
LDS	85.74	130.24	61.74	92.75	45.93	69.68

Table 1: Forecasting performance (MAE and RMSE) for different target representations and models.

C Residual Prediction and Reconstruction Analysis

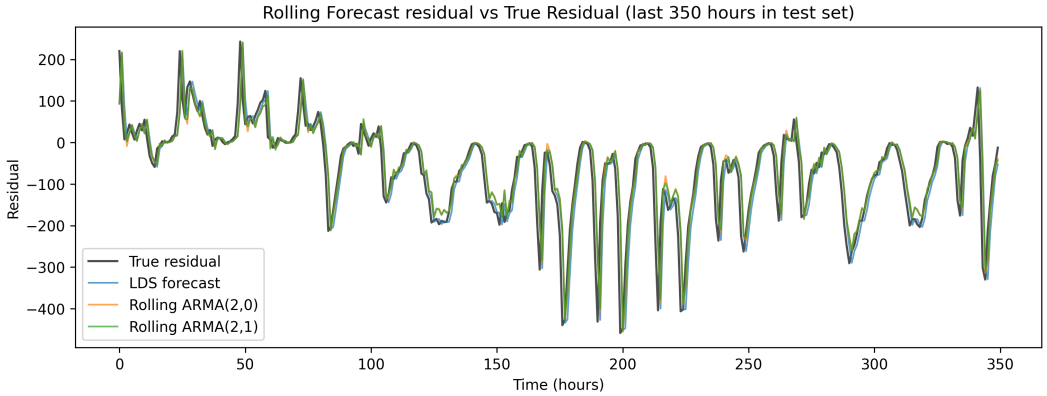

Figure 4: Predicted residuals at the end of the test period. (time $\times$ weekday seasonal mean removal)

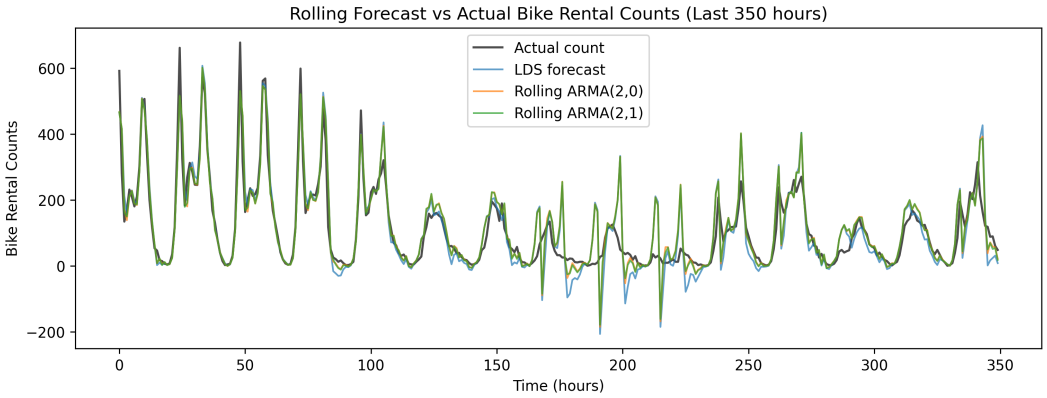

Figure 5: Reconstructed bike rental counts at the end of the test period. (time $\times$ weekday seasonal mean removal)

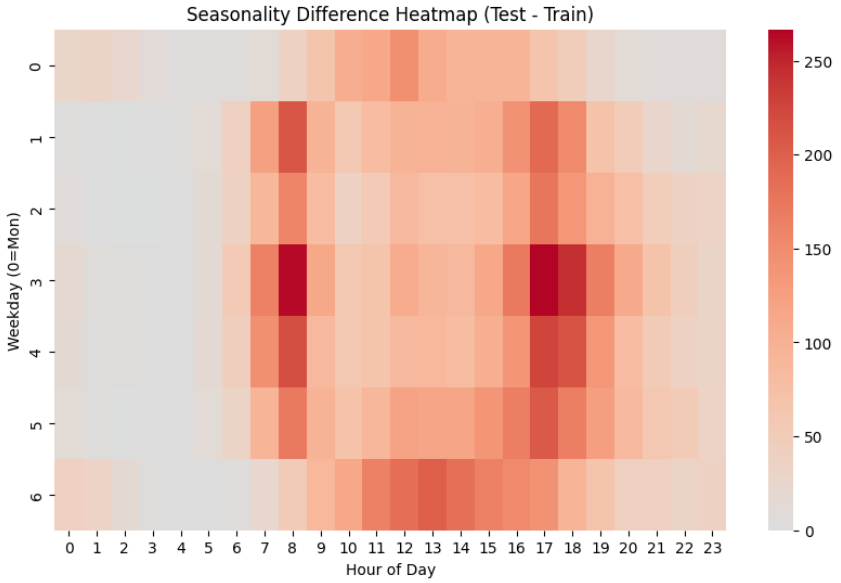

Figure 6: Difference in seasonal mean between test and training sets (test minus train).

D HMM Analysis

Dataset	State	Demand	Temp	Atemp	Humidity	Windspeed
Train	0	105.39	0.494	0.476	0.668	0.176
Train	1	244.52	0.504	0.480	0.578	0.213
Test	0	70.10	0.418	0.406	0.723	0.153
Test	1	345.74	0.526	0.500	0.602	0.186

Table 2: Mean bike rental demand and weather conditions by inferred HMM state.

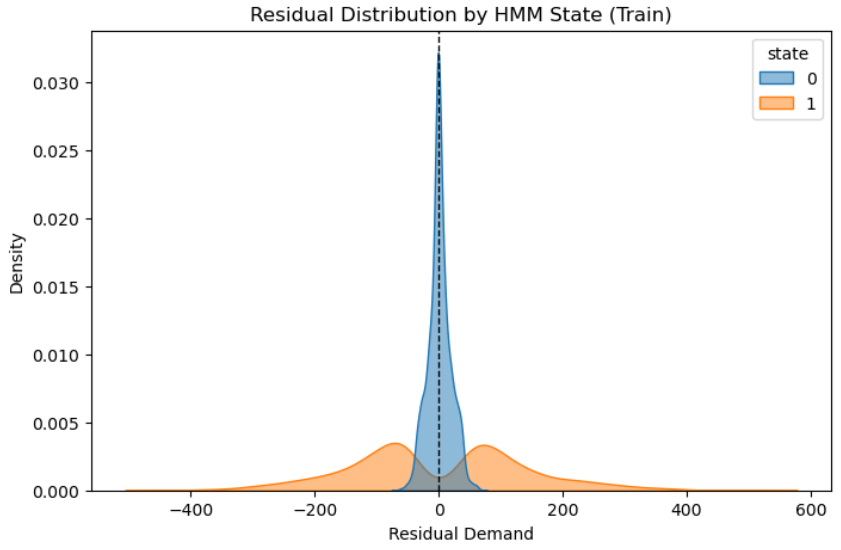

Figure 7: Distribution of residual bike rental demand by inferred HMM state in the training set.

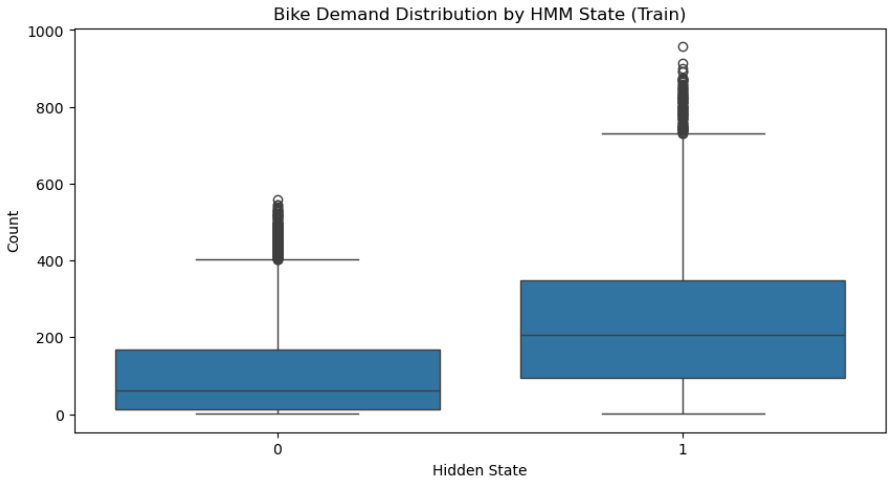

Figure 8: Demand distributions conditioned on inferred HMM states in the training set.

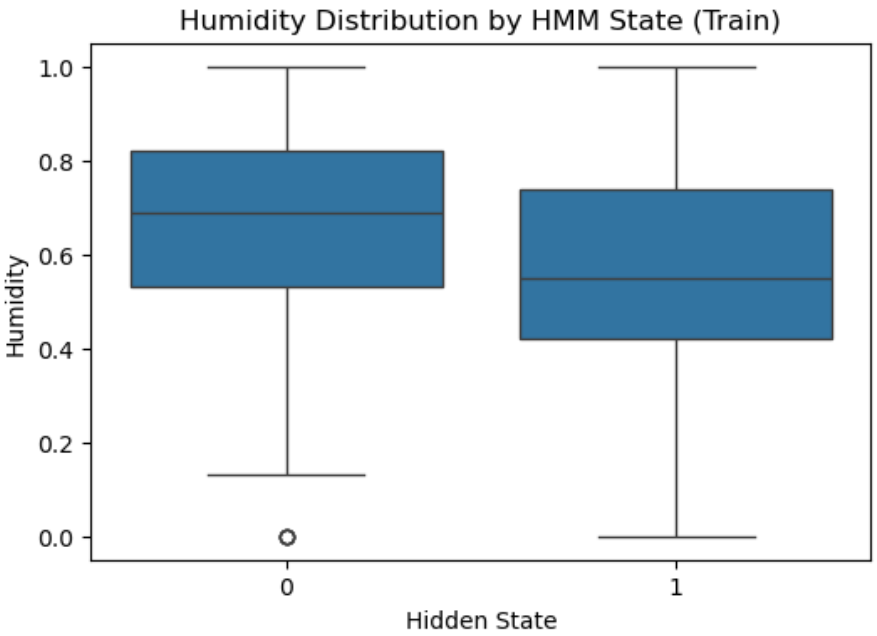

Figure 9: Humidity distributions conditioned on inferred HMM states in the training set.

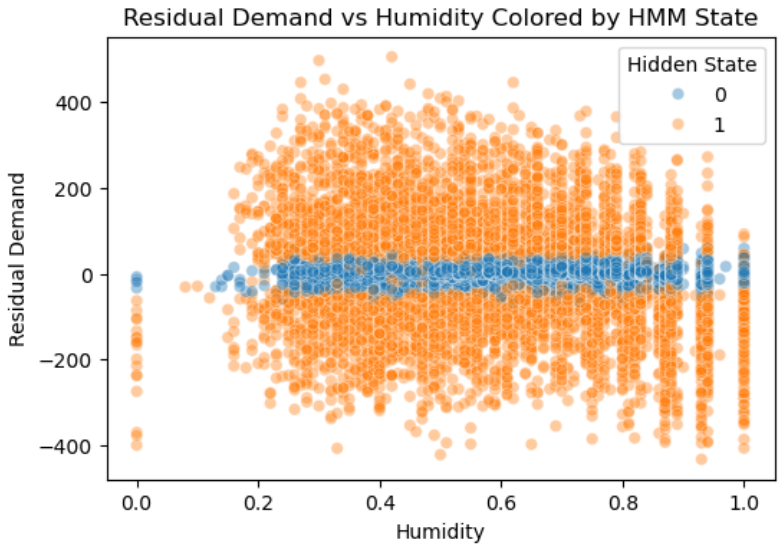

Figure 10: Residual demand versus humidity colored by inferred HMM state in the training set.