

BASTION: Bayesian Adaptive Seasonality and Trend decomposition Incorporating Outliers and Noise

Jason Cho, David S. Matteson

January 18, 2025

Abstract

We introduce BASTION, Bayesian Adaptive Seasonality and Trend Decomposition Incorporating Outliers and Noise, a flexible Bayesian framework for decomposing time series into trend and multiple seasonality components. BASTION provides three key advantages over existing time series decomposition methods: locally adaptive estimation of trend and seasonality for improved accuracy, explicit modeling of outliers and time-varying volatility for robustness, and rigorous uncertainty quantification via credible intervals. Using global-local shrinkage priors, we derive the model and propose an efficient Gibbs sampling scheme for posterior inference. Simulation studies demonstrate BASTION’s superiority over existing methods, such as TBATS, STR, and MSTL, in both accuracy and uncertainty quantification. We further showcase its practical utility through applications to real-world datasets, including electricity demand, bike rentals, and airline passenger data, where BASTION captures complex dynamics and adapts to local changes while accounting for irregular components such as outliers and heteroskedasticity, thereby providing a more nuanced and interpretable decomposition of the data.

1 Introduction

Time-series decomposition is a powerful analytical method used to break down a univariate time-series into its constituent components, such as trend and seasonality. Decomposition reveals deeper insights into underlying long or medium-term trend that may not be immediately apparent. One of the most prominent applications of the method is in economic reporting by government agencies, where it is used to adjust for seasonality when publishing key macroeconomic indicators such as the

Consumer Price Index (CPI) (of Labor Statistics [1977]) and unemployment rates (Dagum [1979]). By separating long-term trends from short-term seasonal fluctuations, decomposition allows for clearer interpretation of economic conditions.

Beyond economics, time-series decomposition is widely used across disciplines, including climate science (Grieser et al. [2002], Zhou et al. [2015]), epidemiology (Abeku et al. [2002], Ke et al. [2016]), and business and management (Gardner Jr and Diaz-Saiz [2002], Zarnowitz and Ozyildirim [2006], Rosselló and Sansó [2017]), where seasonal effects play a crucial role in data analysis. Additionally, time-series decomposition is being incorporated into state-of-the-art time-series forecasting models. Notable examples include TBATS by Alysha M. De Livera and Snyder [2011], ETSformer by Woo et al. [2022] Prophet by Taylor and Letham [2018], and DeepFS by Jiang et al. [2022].

Various decomposition methods have been developed over the past 70 years. Based on the “ratio to moving average” technique by Macaulay [1972], the U.S. Census Bureau introduced the first computerized decomposition method called Census II by Shiskin [1957]. The method was followed by successors like X-11 (Shiskin et al. [1965]), X-11-ARIMA (Dagum et al. [1980]), X-12-ARIMA (David F. Findley and Chen [1998]), and the most recent version, X-13-SEATS (Findley [2005]). Another classical and widely used method is STL (Seasonal-Trend decomposition using LOESS) by Cleveland [1990], which iteratively uses Locally Weighted Scatterplot Smoothing (LOWESS) to extract seasonality from trend.

Existing decomposition methods, though widely used, exhibit several limitations: (1) inability to accommodate multiple seasonal patterns, (2) difficulty in adapting to abrupt changes in the trend, (3) absence of uncertainty quantification, (4) inability to handle heteroskedastic noise, and (5) lack of robustness to outliers. While addressing some of these challenges, no existing model fully resolves all of them. For example, Multiple STL Bandara et al. [2021], an extension of STL, can handle multiple seasonalities but lacks both uncertainty quantification and adaptability to sudden trend shifts. Seasonality-Trend decomposition Dokumentov and Hyndman [2022] supports complex seasonal patterns and includes uncertainty quantification, yet it remains limited in handling abrupt changes. Robust STL Wen et al. [2019] manages multiple seasonal patterns, abrupt changes, and outliers but does not provide uncertainty quantification. Furthermore, none of these methods explicitly address heteroskedastic noise, which is found in various datasets across disciplines.

In this paper, we present Bayesian Adaptive Seasonality Trend decomposition Incorporating

Outlier and Noise (BASTION), a novel, flexible, fully Bayesian framework for time-series decomposition. A key innovation in BASTION is the use of a global-local shrinkage prior, a class of continuous shrinkage prior distribution commonly applied in high-dimensional regression (Carvalho et al. [2010], Bhadra et al. [2015]). Recently, this prior has been extended to time-series applications, such as Bayesian smoothing (Kowal et al. [2019], Schafer and Matteson [2023], Huber and Pfarrhofer [2021], Cho and Matteson [2024]), time-series regression (Cadonna et al. [2020]), and changepoint detection (Wu et al. [2024], Banerjee [2022]). The global-local shrinkage prior excels at reducing noise while preserving significant signals, enabling BASTION to produce locally adaptive and robust yet smooth estimates of trend and seasonality. Moreover, since BASTION is fully Bayesian, it offers uncertainty quantification for each component of the decomposition.

BASTION shares foundational similarities with STR Dokumentov and Hyndman [2022], as both treat decomposition as a high-dimensional regression task and can incorporate time-varying covariates to capture complex effects in the data. However, BASTION’s Bayesian framework introduces additional advantages not available in frequentist approaches. For instance, BASTION can include extra components beyond trend and seasonality, such as an additive outlier term inspired by the ABCO model Wu et al. [2024], which enhances robustness to anomalies. Furthermore, BASTION addresses heteroskedasticity by modeling time-varying volatility through stochastic volatility model Hull and White [1987]. This flexibility allows BASTION to effectively handle seasonality, trend, outliers, and heteroskedastic noise within a cohesive, Bayesian framework.

The paper is organized as follows: Section 2 introduces the full model and its parameters. Appendix A provides the derivation of the Gibbs sampling scheme for posterior inference. Section 3 compares BASTION with existing models, including TBATS (Alysha M. De Livera and Snyder [2011]), MSTL (Bandara et al. [2021]), and STR (Dokumentov and Hyndman [2022]), using a variety of simulated time series. Section 4 presents empirical analyses of three datasets: monthly airline passengers in the U.S., daily electricity demand in New York state, and daily bike rentals in Washington D.C.

2 Method

2.1 Single Seasonality and Trend Decomposition with BASTION

Similar to existing decomposition methods, BASTION employs an additive observation equation structure to decompose a length N time-series $\{y_t\}_{t=1}^N$, into a trend $\{T_t\}_{t=1}^N$, and a seasonality component $\{S_t\}_{t=1}^N$ with k being the length of the season.

$$y_t = T_t + S_t + R_t \quad [R_t | \sigma_y] \stackrel{iid}{\sim} N(0, \sigma_y^2), \quad (1)$$

Different decomposition methods employ a variety of smoothing algorithms to extract trend and seasonal components from a noisy time series. As their names suggest, STL and MSTL rely on LOESS, while STR achieves smooth estimates of trend and seasonality by imposing L_1 or L_2 penalties on the second and seasonal differences of the respective components, similar to the Hodrick-Prescott filter by Hodrick and Prescott [1997] or the L_1 trend filter by Kim et al. [2009], Tibshirani [2014]. Meanwhile, methods developed by the Census Bureau, such as X-12, employ iterated moving average filters to achieve smoothing.

BASTION adopts a similar approach to STR by simultaneously penalizing both the second and seasonal differences. However, as a Bayesian method, BASTION incorporates these constraints by using global-local shrinkage priors, which provide a probabilistic framework for adaptively smoothing the data. The application of global-local shrinkage priors to trend filtering in time series was pioneered in Kowal et al. [2019] with its Bayesian Trend Filter (BTF), which applies the Dynamic Shrinkage Process, an extension of horseshoe prior Carvalho et al. [2010], to the first or second differences of the mean component. This method demonstrated improved performance over traditional smoothing algorithms such as the L_1 trend filter by Kim et al. [2009], Tibshirani [2014] and smoothing splines by Schoenberg [1964], particularly in terms of accuracy and uncertainty quantification.

The defining characteristic of global-local shrinkage priors is their use of a global parameter τ , which controls overall shrinkage, paired with a time-varying local parameter λ_t that captures local variation. There are various choices for these priors, including the horseshoe prior Carvalho et al. [2010], horseshoe plus Bhadra et al. [2015], dynamic shrinkage processes Kowal et al. [2019], the regularized horseshoe Piironen and Vehtari [2017], and the triple Gamma prior Cadonna et al.

[2020]. These priors differ in their degree of shrinkage and computational efficiency. Considering the large number of parameters to be estimated for time-series decomposition task, we chose the horseshoe prior for BASTION, as it requires the least number of parameters to be estimated compared to other options.

BASTION extends the BTF to time-series decomposition by penalizing the second differencing of the trend component and the seasonal differencing of the seasonal components. This simultaneous penalization allows both trend and seasonality to adapt to local changes while maintaining smoothness. For the trend component, the prior distributions are specified as below,

$$\begin{aligned}
[T_t | \sigma_y, \lambda_{T,t}] &\sim N(0, \sigma_y^2 \lambda_{T,t}^2) && \forall t \in \{1, 2\}, \\
[\Delta^2 T_t | \sigma_y, \tau_T, \lambda_{T,t}] &\sim N(0, \sigma_y^2 \tau_T^2 \lambda_{T,t}^2) && t \geq 3, \\
[\tau_T] &\sim C^+(0, 1), && [\lambda_{T,t}] \stackrel{iid}{\sim} C^+(0, 1),
\end{aligned}$$

and for the seasonality component,

$$\begin{aligned}
[S_1] &= 0, && [S_2 | \sigma_y, \lambda_{S,t}] \sim N(0, \sigma_y^2 \lambda_{S,t}^2), \\
[\Delta^2 S_t | \sigma_y, \tau_S, \lambda_S] &\sim N(0, \sigma_y^2 \tau_S^2 \lambda_{S,t}^2) && \forall t \in \{3, \dots, k\}, \\
[(1 - B)^k S_t | \sigma_y, \tau_S, \lambda_{S,t}] &\sim N(0, \sigma_y^2 \tau_S^2 \lambda_{S,t}^2) && t \geq (k + 1), \\
[\tau_S] &\sim C^+(0, 1), && [\lambda_{S,t}] \stackrel{iid}{\sim} C^+(0, 1).
\end{aligned}$$

For identifiability, $S_1^k = 0$. For $t \in \{3, k\}$, priors are specified on the second differencing to induce smoothness. For $t \geq k$, priors on the seasonal differencing is imposed. By applying global-local shrinkage priors to both the trend and seasonality, BASTION provides locally adaptive yet smooth estimate of both the trend and seasonality component.

2.2 Multiple Seasonalities

Naturally, the framework described in Section 2.1 to account for multiple seasonalities. Let's consider p seasonality terms with k_i representing the length of the cycle for each seasonality represented by $i = 1, \dots, P$. We have the following observation equations for the trend and multiple

seasonalities decomposition:

$$y_t = T_t + \sum_{i=1}^P S_{i,t} + R_t \quad [R_t | \sigma_y] \stackrel{iid}{\sim} N(0, \sigma_y^2),$$

and similar to the single seasonality model described in Section 2.1, the horseshoe prior are assumed on the seasonal differencing of $\{S_{i,t}\}_{t=1}^N$ for $i = \{1, \dots, P\}$.

2.3 Additive Outlier

In addition to providing a locally adaptive estimate of the trend, a key feature of BASTION is its explicit modeling of additive outliers. Existing methods typically handle outliers by pre-adjusting the observed series or by applying smoothing methods that are robust to outliers, rather than directly addressing them within the decomposition framework. For instance, the X-12-ARIMA model uses an ARIMA approach for pre-adjustment, while the Robust STL method applies bilateral filtering to denoise the series before decomposition. The STR addresses outliers by imposing an L_1 penalty instead of L_2 , thereby making the trend and seasonality estimates more robust. In contrast, MSTL and STL do not address outliers at all.

BASTION, however, is the first decomposition method to explicitly model outliers through the use of an extreme shrinkage-inducing prior inspired by the approach in Wu et al. [2024]. Let $\{\zeta_t\}_{t=1}^N$ represent the outlier process. To capture large deviations at isolated time points, BASTION applies the horseshoe+ prior by Bhadra et al. [2015] to ζ_t . By nesting two levels of half-Cauchy distributions, the horseshoe+ prior provides more aggressive shrinkage, allowing larger deviations from zero at a fewer number of locations when compared to the regular horseshoe prior. Specifically,

$$\begin{aligned} [\zeta_t | \sigma_y, \lambda_{\zeta,t}] &\sim N(0, \sigma_y^2 \lambda_{\zeta,t}^2) & [\lambda_{\zeta,t} | \tau_{\zeta}, \xi_{\zeta,t}] &\sim C^+(0, \tau_{\zeta} \xi_{\zeta,t}), \\ [\tau_{\zeta}] &\sim C^+(0, 1), & [\xi_{\zeta,t}] &\stackrel{iid}{\sim} C^+(0, 1). \end{aligned}$$

The outlier term ζ_t is then incorporated into the observation equation, as described in previous sections. By explicitly modeling outliers, BASTION enables these anomalous points to be isolated for further analysis. Outliers often signify unique or noteworthy events at specific time points, and by pinpointing their locations, one can conduct additional analyses to better understand the nature and causes of these anomalies.

2.4 Heteroskedastic Noise

BASTION explicitly makes a distributional assumption on the remainder term, $R_t \sim N(0, \sigma_y^2)$. While this assumption is commonly applied, it may be overly restrictive in cases where the remainder exhibits heteroskedasticity, as it could lead to inaccurate estimates of both the trend and seasonal components. Heteroskedastic noise, characterized by changing volatility over time, is observed in various academic disciplines such as finance Black and Scholes [1973], Hull and White [1987], Taylor [2008], Melino and Turnbull [1990], epidemiology Achcar et al. [2020], Sarkar and Chatterjee [2017], geophysics Mariani et al. [2018], Wang et al. [2005], and environmental science Tippett [2014], Modarres and Ouarda [2014], Hor et al. [2006]. Despite its prevalence, no existing decomposition methods explicitly account for time-varying volatility within their frameworks.

To address this gap, BASTION introduces a stochastic volatility (SV) model to capture heteroskedasticity in the remainder term. In addition to the overall variance parameter, σ_y , we introduce a time-varying variance component $\{\nu_t\}_{t=1}^N$, which follows a first order stochastic volatility model by Kim et al. [1998]. This allows the model to adapt to changes in volatility over time. Specifically, the process is defined as follows:

$$\begin{aligned} [R_t | \sigma_y, \nu_t] &\sim N(0, \sigma_y^2 \nu_t^2) \\ \log(\nu_t^2) &= \mu + \phi(\log(\nu_{t-1}^2) - \mu) + \sigma_\nu \epsilon_t \\ [\epsilon_t] &\stackrel{iid}{\sim} N(0, 1). \end{aligned}$$

Here, μ represents the mean log-variance, ϕ is a persistence parameter that controls how strongly current volatility depends on past values, and σ_ν dictates the variability of the log-variance process. By incorporating a time-varying volatility model, BASTION provides a more flexible decomposition that adapts to heteroskedastic structures in the data. This extension allows for more accurate trend and seasonality estimation, particularly in data with substantial fluctuations in residual variance.

2.5 Additional Covariates

Additional covariates can often help explain variation in a time series. For example, Stock and W.Watson [2003] demonstrates that certain asset prices can be effective predictors of inflation, while Hamilton [1983] uses oil prices to explain macroeconomic fluctuations. In the STR model,

Dokumentov and Hyndman [2022] incorporates temperature as an additional covariate to account for variation in electricity demand, beyond the trend and seasonal components. The additive framework of BASTION naturally facilitates the inclusion of regression terms, allowing the inclusion of time-varying covariates. Given k time varying covariates $\{x_{1,t}\}_{t=1}^N \dots \{x_{k,t}\}_{t=1}^N$, we have the following trend and multiple seasonality decomposition with time varying covariates:

$$y_t = T_t + \sum_{i=1}^P S_{i,t} + \sum_{j=1}^k \beta_j x_{j,t} + R_t \quad [R_t | \sigma_y] \stackrel{iid}{\sim} N(0, \sigma_y^2)$$

$$[\beta_j | \eta_j] \sim N(0, \eta_j^2) \quad [\eta_j] \sim \eta_j^{-2} d\eta_j^2, j = 1 \dots k.$$

In BASTION, as described in Sections 2.1 and 2.2, seasonality is modeled through a seasonal differencing operator, which may not fully capture irregular seasonal effects. The regression framework additionally allows for the incorporation of holiday effects or other complex, non-regular seasonal patterns that fixed seasonal differencing might miss, thereby enhancing the model’s flexibility in capturing nuanced temporal variations.

2.6 Full Model

In order to conduct posterior inference on the estimated parameters, Markov Chain Monte Carlo is used. Note that other than the parameter ϕ , the persistence parameter for the SV model in the remainder term, the full conditional distributions can be derived analytically as they are all conditionally conjugate priors, which then Gibbs sampling may be used. For sampling the parameters related to the horseshoe prior, parameter expanded Gibbs sampling scheme by Makalic and Schmidt [2016] is used. For all the state variables, $\{T_t\}$, $\{S_t\}$, and $\{\zeta_t\}$, which are all conditionally Gaussian, efficient multivariate Gaussian sampler by Rue [2001] is used. All Without a Loop (AWOL) sampler by Kastner and Frühwirth-Schnatter [2014] for sampling parameters is for sampling the SV model in the remainder term. For clarity, the parameterization and the prior

distributions of BASTION as described in Section 2 is described below:

Observation Equation

$$y_t = T_t + \sum_{i=1}^P S_{i,t}^{k_i} + \sum_{j=1}^k \beta_j x_{j,t} + R_t, \quad [R_t | \sigma_y, \nu_t] \stackrel{iid}{\sim} N(0, \sigma_y^2 \nu_t^2)$$

Trend

$$\begin{aligned} [T_t | \sigma_y, \lambda_{T,t}] &\sim N(0, \sigma_y^2 \lambda_{T,t}^2) && \forall t \in \{1, 2\}, \\ [\Delta^2 T_t | \sigma_y, \tau_T, \lambda_{T,t}] &\sim N(0, \sigma_y^2 \tau_T^2 \lambda_{T,t}^2) && t \geq 3, \\ [\tau_T] &\sim C^+(0, 1), && [\lambda_{T,t}] \stackrel{iid}{\sim} C^+(0, 1). \end{aligned}$$

Multiple Seasonality

$$\begin{aligned} [S_{i,1}] &= 0, && [S_{i,2} | \sigma_y, \lambda_{S_{i,2}}] \sim N(0, \sigma_y^2 \lambda_{S_{i,2}}^2), \\ [\Delta^2 S_{i,t} | \sigma_y, \tau_{S_i}, \lambda_{S_{i,t}}] &\sim N(0, \sigma_y^2 \tau_{S_i}^2 \lambda_{S_{i,t}}^2) && \forall t \in \{3, \dots, k_i\}, \\ [(1-B)^{k_i} S_{i,t} | \sigma_y, \tau_{S_i}, \lambda_{S_{i,t}}] &\sim N(0, \sigma_y^2 \tau_{S_i}^2 \lambda_{S_{i,t}}^2) && \forall t \geq (k_i + 1), \\ [\tau_{S_i}] &\sim C^+(0, 1), && [\lambda_{S_{i,t}}] \stackrel{iid}{\sim} C^+(0, 1). \end{aligned}$$

Outliers

$$\begin{aligned} [\zeta_t | \sigma_y, \lambda_{\zeta,t}] &\sim N(0, \sigma_y^2 \lambda_{\zeta,t}^2), && [\lambda_{\zeta,t} | \tau_\zeta, \xi_{\zeta,t}] \sim C^+(0, \tau_\zeta \xi_{\zeta,t}) \\ [\tau_\zeta] &\sim C^+(0, 1), && [\xi_{\zeta,t}] \stackrel{iid}{\sim} C^+(0, 1). \end{aligned}$$

Remainder

$$\begin{aligned} [\sigma_y^2] &\sim \sigma_y^{-2} d\sigma_y^2 \\ \log(\nu_t^2) &= \mu + \phi(\log(\nu_{t-1}^2) - \mu) + \sigma_\nu \epsilon_t, && [\epsilon_t] \stackrel{iid}{\sim} N(0, 1) \\ [\mu] &\sim N(0, 100), && [\phi] \sim \text{Beta}(5, 1.5) \\ [\sigma_\nu^2] &\sim IG(1/2, 1/2) \end{aligned}$$

Regression Coefficients

$$[\beta_j | \eta_j] \sim N(0, \sigma_y^2 \eta_j^2), \quad [\eta_j] \sim \eta_j^{-2} d\eta_j^2, j = 1 \dots m$$

Let's first define $\mathbf{Y} = [y_1, \dots, y_N]'$ and similarly for the variables $\mathbf{T}, \mathbf{S}_i, \boldsymbol{\zeta}, \boldsymbol{\nu}, \boldsymbol{\beta}, \mathbf{X}$. The full conditional posterior distributions for Gibbs sampling are derived in the Appendix A.

3 Simulation Study

3.1 Set Up

	Trend	Seasonality	Seasonal Length	Remainder	Outlier
DGS 1	Piecewise Linear	Fourier Pairs	12, 40	Constant	FALSE
DGS 2	Linear	Piecewise Constant	40	Constant	FALSE
DGS 3	Quadratic	Fourier Pair	50	Stochastic Volatility	FALSE
DGS 4	Piecewise Linear	Piecewise Constant & Fourier Pair	12, 40	Stochastic Volatility	TRUE
DGS 5	Quadratic	Fourier Pair	50	Constant	FALSE

Table 1: Descriptions of data generating schemes (DGS) with regards to their trend, seasonality, seasonal lengths, remainder, and outliers. The sample size of each time-series is 500, and each DGS is replicated 1000 times with random coefficients for both the trend and seasonality components.

In this section, BASTION is compared against existing multiple seasonalities decomposition models: MSTL by Bandara et al. [2021], STR by Dokumentov and Hyndman [2022], and TBATS by Alysha M. De Livera and Snyder [2011] across various simulation scenarios described in Table 1. Exact simulation schemes are detailed in Appendix B. The comparisons are made in terms of their ability to accurately extract trend, seasonality, and signal (trend and seasonality combined), by measuring the mean squared error (MSE). As a statistical model, BASTION allows for uncertainty quantification via credible region. STR, a Frequentist counterpart, is the only other decomposition method with uncertainty quantification. Empirical coverages based on BASTION’s credible interval and STR’s confidence intervals are compared to measure each method’s accuracy of its uncertainty quantification.

Due to the incorporation of global-local shrinkage prior, BASTION excels at capturing abrupt changes in trend and seasonality components, offering a significant advantage in scenarios with structural breaks or rapid changes. In addition, its Bayesian framework allows for the explicit models additive outliers and heteroskedastic noise, overlooked by existing methods, enabling more accurate and robust estimation, particularly in complex data environments. Consequently, BASTION is expected to outperform existing models in DGS 1 through 4 due to their distinct characteristics: trend discontinuities (Figure 1a), seasonal discontinuities (Figure 1b), heteroskedastic noise (Figure 1c), or a combination of these with additive outliers (Figure 1d). DGS 5, on the other

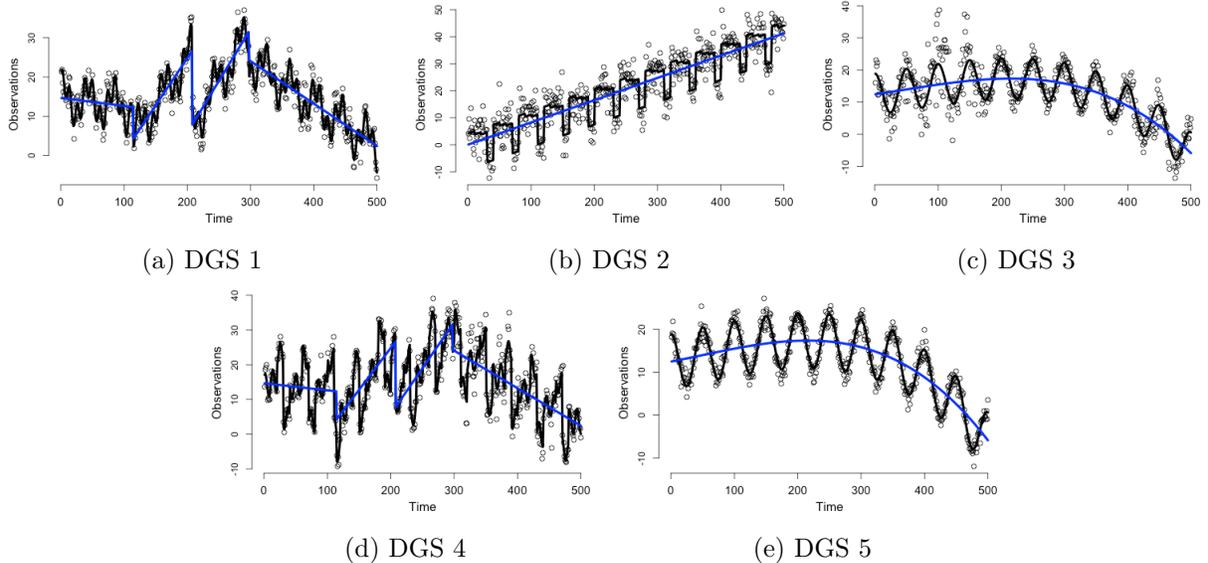


Figure 1: Synthetic time series generated by adding a trend component, T_t , seasonal components, S_t , and remainders R_t based on the descriptions in Table 1. T_t are drawn in blue and $T_t + S_t$ are drawn in black. Figures represent one replication of each DGS.

hand, replicates the simulation study design used by Dokumentov and Hyndman [2022] (Figure 1e), providing a benchmark for fair comparison against existing methods.

All models are implemented in R (R Core Team [2024]). TBATS and MSTL are implemented using the `forecast` package (Hyndman et al. [2024]), and STR is implemented using the `str` package (Dokumentov and Hyndman [2024]). For BASTION, we extend the efficient sampling approach for the horseshoe prior from Makalic and Schmidt [2016], the Gaussian state-space model framework from Rue [2001], and the stochastic volatility model from Kastner [2016].

3.2 Results

Table 2 summarizes the performance of BASTION compared to MSTL, TBATS, and STR in terms of accuracy (measured by Mean Squared Error, MSE) and uncertainty quantification (measured by empirical coverage). BASTION consistently achieved the lowest or near-lowest MSE across all five simulation scenarios for trend, seasonality, and their combination (signal), highlighting its precision in decomposition. In terms of uncertainty quantification, BASTION demonstrated superior performance, with empirical coverage consistently exceeding the nominal 95% across all simulations. In contrast, STR frequently under-covers, with empirical coverage as low as 62% in

	Mean Squared Error				Empirical Coverage	
	TBATS	MSTL	STR	BASTION	STR	BASTION
DGS 1						
Signal	13.878	11.760	10.829	0.376	0.799	0.998
Trend	3.624	13.227	13.785	0.581	0.616	0.970
Seasonality	11.672	2.712	2.575	0.536	0.815	0.999
DGS 2						
Signal	1.411	1.214	0.7020	0.3922	0.869	0.989
Trend	0.621	0.184	0.0878	0.0579	0.716	0.976
Seasonality	0.838	1.032	0.6156	0.3412	0.793	0.956
DGS 3						
Signal	11.307	3.041	0.706	0.439	0.940	0.995
Trend	10.509	0.364	0.274	0.293	0.812	0.929
Seasonality	0.900	2.683	0.444	0.278	0.847	0.981
DGS 4						
Signal	11.829	11.430	20.548	2.877	0.679	0.981
Trend	11.111	12.328	13.431	5.210	0.668	0.939
Seasonality	5.364	3.045	11.358	2.562	0.623	0.999
DGS 5						
Signal	0.682	0.3457	0.1087	0.0946	0.951	0.998
Trend	0.598	0.0432	0.0387	0.0650	0.829	0.943
Seasonality	0.093	0.3040	0.0731	0.0756	0.848	0.983

Table 2: Mean Squared Error (MSE) and Empirical Coverage (EC) for Trend, Seasonality, and their combined component, denoted as Signal. Each data generating scheme (DGS), described in Table 1, is replicated 1000 times. The EC is calculated using 95% confidence intervals for STR and 95% credible intervals for BASTION.

some cases and achieving nominal coverage only for specific scenarios. These results confirm that BASTION provides not only accurate decomposition but also reliable uncertainty quantification, outperforming existing methods.

Superior results of BASTION on DGS 1 and DGS 2 highlight its ability to accurately estimate both trend and seasonality components in the presence of abrupt changes. Figure 2 displays one replication of DGS 1, where discontinuities in the true trend occur around $t = 100, 200,$ and 300 . The estimates by MSTL (Figure 2b) exhibit excessive smoothness at these points, failing to adapt to local changes. While STR (Figure 2d) is more locally adaptive than MSTL, it still misses the level shift at $t = 300$. TBATS (Figure 2c) captures the discontinuities but suffers from overfitting, resulting in noisy estimates. In contrast, BASTION (Figure 2e) strikes a balance: it produces a smooth overall trend while effectively adapting to sharp changes, offering the most accurate

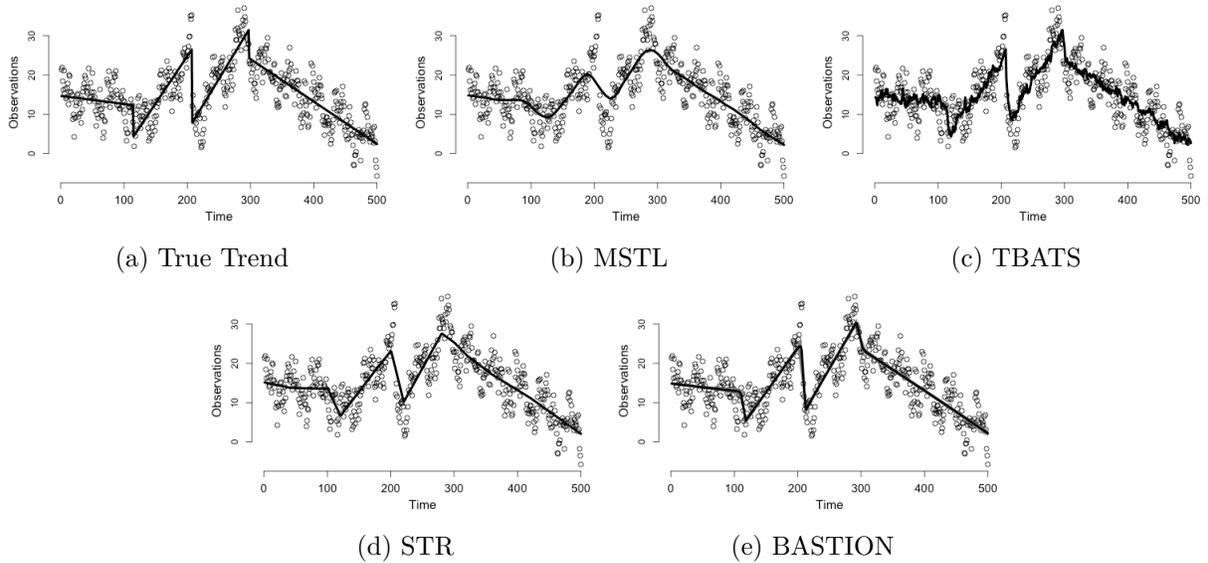


Figure 2: One replication of synthetic time series generated from the data generating scheme 1 (DGS 1). The true trend alongside estimates produced by MSTL, TBATS, STR, and BASTION are drawn in black line from panel (a) through (e) respectively.

representation of the underlying process. In DGS 2, the seasonality consists of repeating piecewise constant functions with sharp discontinuities. BASTION again achieves the lowest MSE, followed by STR, MSTL, and TBATS, further confirming its superior ability to handle abrupt changes in the data.

DGS 3 and DGS 4 highlight BASTION’s ability to explicitly model both heteroskedastic noise and additive outliers. Specifically, DGS 3 exhibits heteroskedastic noise following a stochastic volatility (SV) model with lag order 1 (Figure 3a). BASTION not only accurately captures the underlying trend and seasonality but also estimates the time-varying variance, as shown in Figure 3c. In DGS 4, the data contain seasonality with sharp discontinuities as well as large outliers, which are denoted by red dashed lines in Figure 3b. BASTION successfully identifies and explicitly measures these outliers while also providing accurate estimates of the trend and seasonality, as depicted in Figure 3d.

Even in the absence of abrupt changes, outliers, or stochastic volatility, as seen in DGS 5, BASTION continues to provide competitive results. While STR achieves slightly better accuracy in this smooth scenario with a slowly changing mean and seasonality modeled using Fourier pairs, as shown in Dokumentov and Hyndman [2022], BASTION still produces reasonable and reliable

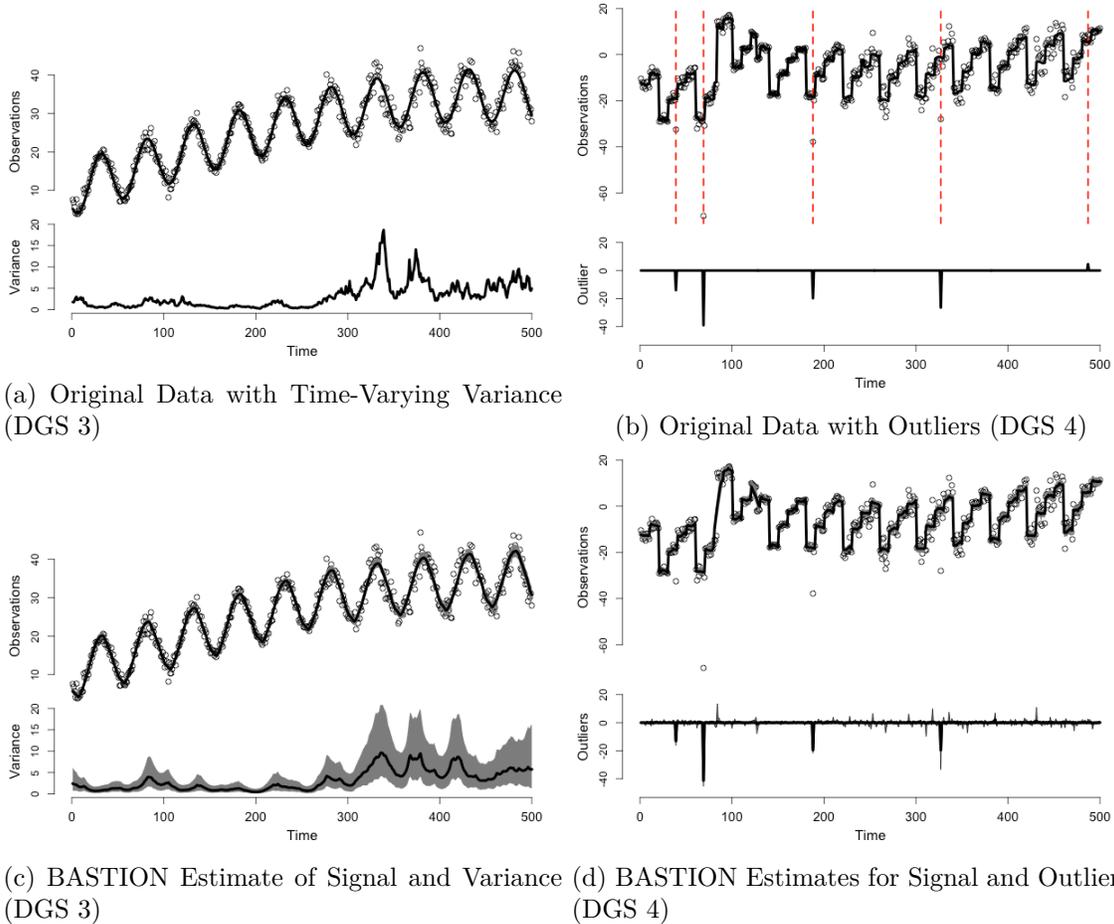


Figure 3: Comparison of sample paths from DGS 3 and DGS 4 with corresponding BASTION model estimates. Panels (a) and (b) display the original data: DGS 3 highlights time-varying variance, while DGS 4 includes outliers marked by red dashed lines. Panel (c) shows BASTION’s estimates of trend, seasonality, and time-varying variance with 95% credible intervals shaded in grey. Panel (d) presents BASTION’s estimates of trend, seasonality, and outliers, with outlier estimates shown below and 95% credible intervals shaded in grey.

estimates. This demonstrates that BASTION performs well not only in challenging scenarios with sharp changes or noise complexities but also in relatively smooth and well-behaved settings.

Overall, the results across all simulation scenarios (DGS 1-5) demonstrate BASTION’s superior performance in both accuracy and uncertainty quantification compared to existing decomposition methods. In terms of accuracy, BASTION consistently achieves the lowest or near-lowest Mean Squared Error (MSE) across trend, seasonality, and their combination (signal), outperforming MSTL, TBATS, and STR. This is particularly evident in scenarios with abrupt changes, time-varying variance, and outliers, where BASTION balances smoothness with adaptiveness to sharp

transitions.

In terms of uncertainty quantification, BASTION’s empirical coverage consistently exceeds the nominal 95% credible intervals across all simulations. In contrast, STR underperforms, with its confidence intervals often falling well below the nominal level, particularly in scenarios with more complex noise structures or outliers. These results highlight BASTION’s ability to provide reliable and accurate estimates of the underlying time series components while offering a robust framework for uncertainty quantification.

4 Empirical Data Analysis

In this section, we apply BASTION to three real-world datasets that exhibit complex seasonality patterns, abrupt changes, heteroskedastic noise or combination of these characteristics. The first dataset consists of monthly U.S. airline traffic from 2003 to 2023, obtained from Kaggle (Yan [2023]), which highlights sharp declines in passenger volume during significant disruptions such as the COVID-19 pandemic. The second dataset records the daily number of bike rentals from the Capital Bikeshare system serving Washington, D.C., providing an example of yearly patterns, as well as demand spikes and heteroskedastic noise (Bikeshare [2024]). Finally, we analyze the average daily electricity demand (in Megawatts per hour) for the state of New York, sourced from the New York Independent System Operator (NYISO) via the U.S. Energy Information Administration [EIA]. This dataset captures complex long-term trend, with complex seasonal patterns and heteroskedastic noise.

4.1 U.S. Airline Traffic Data

We analyze the monthly international airline traffic data from 2003 to 2023, as shown in Figure 4a. The seasonality component (Figure 4c) reveals a clear yearly cycle. Specifically, passenger numbers are low at the beginning of the year, followed by a steady increase, peaking in July and August, likely due to the summer vacation season. This is followed by a sharp decline during September to November, before experiencing a moderate rise again in December and January, coinciding with the end-of-year holiday season.

Over time, the magnitude of the seasonal pattern shows a gradual increase from 2003 to 2015

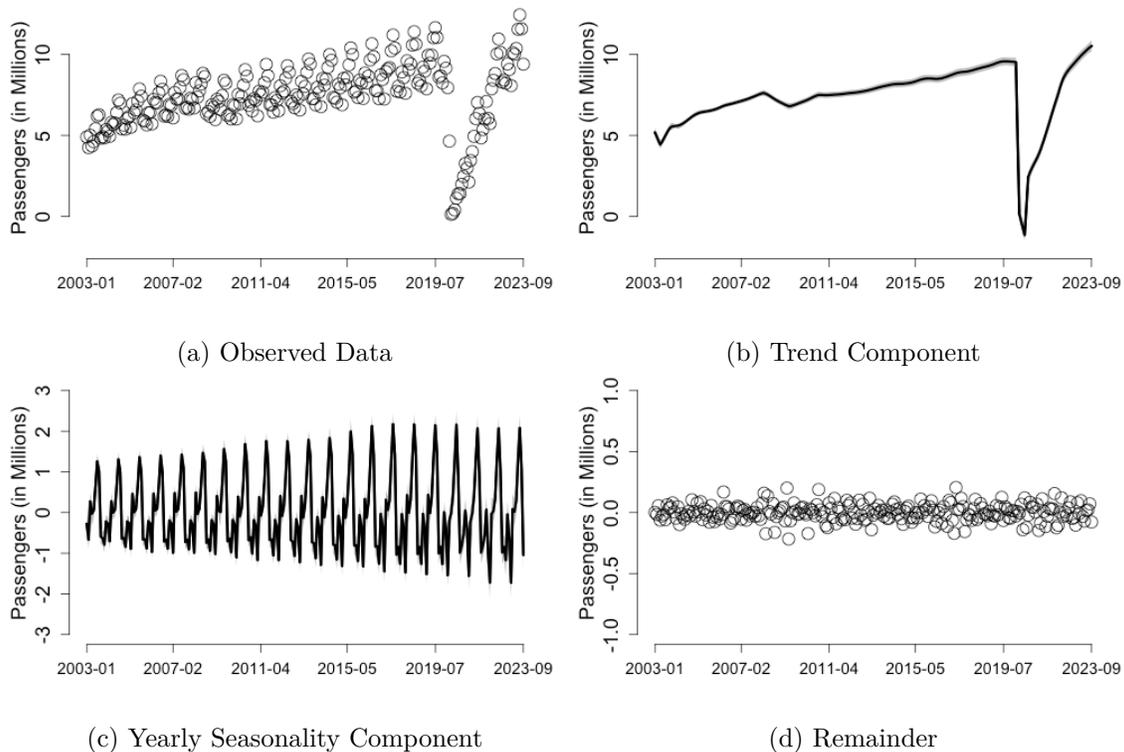


Figure 4: Monthly international airline traffic in the U.S from 2003 to 2023 (Figure 4a), and its Trend (Figure 4b) Seasonality (Figure 4c), and remainder (Figure 4d) decomposition based on the proposed model, BASTION. 95% credible regions are generated and shaded in grey for the trend and seasonality components.

but stabilizes between 2015 and 2023. The trend component (Figure 4b) highlights a significant drop in the late 2019, corresponding to the COVID-19 pandemic and resulting travel restrictions. Finally, the remainder term (Figure 4d) does not exhibit any substantial patterns or structural changes over the observed time horizon, suggesting that most variability is accounted for by the trend and seasonality components.

The most notable feature of this dataset is the sharp decline in the number of airline passengers during late 2019 and early 2020, corresponding to the travel ban imposed in response to the COVID-19 outbreak. This event introduces a clear discontinuity in the trend, similar to the behavior observed in DGS 1. Figure 9 presents the trend estimates produced by TBATS, MSTL, STR, and our proposed model, BASTION. BASTION provides a smooth yet adaptive trend estimate, accurately capturing the abrupt drop caused by the travel ban. In contrast, MSTL and STR produce oversmoothed estimates, failing to identify the sharp structural break and instead treating the drop as a gradual change. TBATS, while capable of detecting the discontinuity, results in a

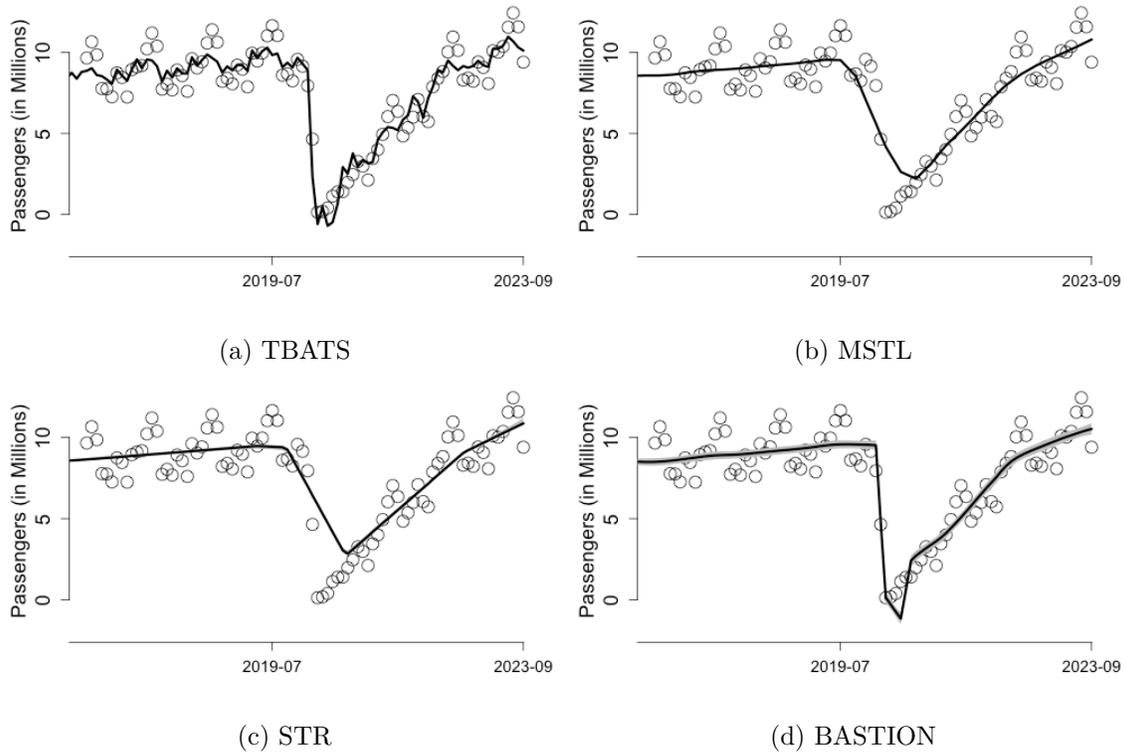


Figure 5: Trend estimates for monthly international airline traffic in the U.S. from 2016 to 2023, highlighting the discontinuity caused by the COVID-19 travel ban in late 2019 and early 2020. Figures (a) - (d) display estimates from TBATS, MSTL, STR, and BASTION, respectively.

noisier estimate that lacks smoothness. These results demonstrate BASTION’s ability to maintain a smooth trend while effectively adapting to abrupt changes in the series.

4.2 New York Daily Average Electricity Demand

The daily average electricity demand data for New York, spanning from July 1, 2015, to June 30, 2024, reveals clear seasonal patterns. As shown in Figure 6a(a), electricity demand peaks during the summer months due to the widespread use of air conditioning and rises again in winter, reflecting heating needs. In contrast, demand decreases during the spring and fall, when milder temperatures reduce the reliance on cooling or heating. The 30-day rolling standard deviation, illustrated in Figure 6b, highlights significant heteroskedasticity, with variability in demand itself exhibiting a seasonal pattern. Variability is higher during the summer and winter months, coinciding with periods of peak demand, while it remains lower during the spring and fall, when electricity usage is more stable. Over the long term, no significant trend is observed, though there is a slight overall

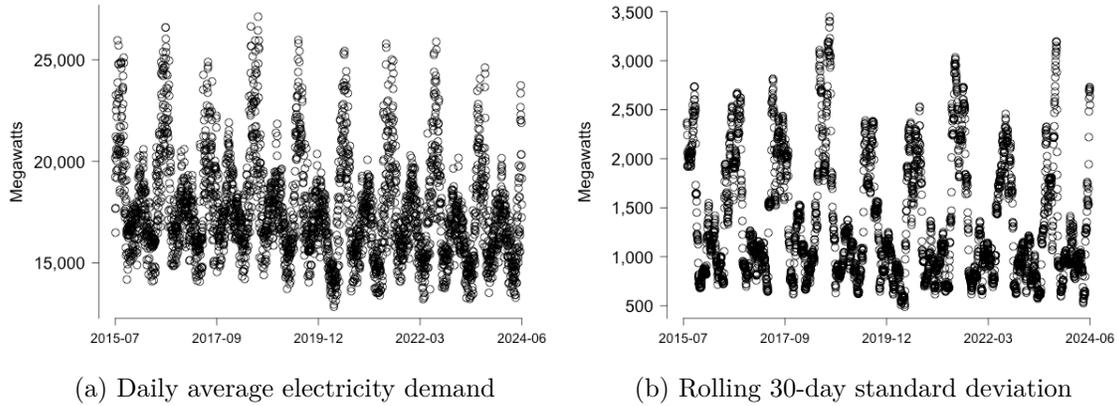


Figure 6: Figure 6a shows daily average electricity demand (in megawatts) in New York State from 2015-07-01 to 2024-06-30, displaying seasonal patterns associated with weather-driven demand fluctuations. Figure 6b shows rolling 30-day standard deviation of electricity demand. Data are sourced from the New York Independent System Operator (NYISO) through the U.S. Energy Information Administration.

decline in electricity demand, potentially reflecting the impact of improved energy efficiency over the years.

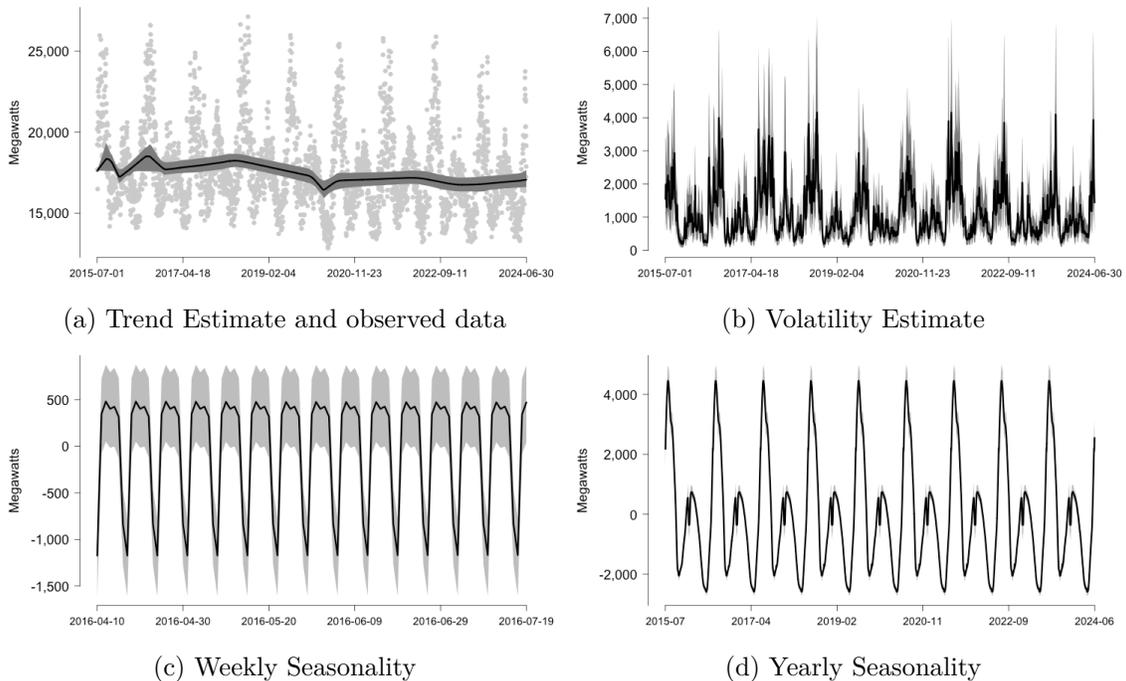


Figure 7: Decomposition of the trend (7a), volatility (7b), weekly seasonality (7c), and yearly seasonality (7d) based on BASTION for daily average electricity demand from 2015-07-01 to 2024-06-30. 95% credible regions are drawn in dark grey.

By decomposing the noisy observed time series into multiple components, BASTION facilitates

a more detailed analysis of the data, as shown in Figure 7. Figure 7a illustrates the estimated trend and its 95% credible region overlaid on the observed data. After accounting for seasonal effects, the average daily electricity demand remains relatively stable over time, with a noticeable dip in late 2019, likely reflecting the impact of COVID-19-related shutdowns.

BASTION’s estimate of time-varying volatility, depicted in Figure 6b, highlights not only the heteroskedasticity of the noise term but also the presence of seasonality in the volatility itself. Specifically, higher volatility is observed during the winter months, while lower volatility occurs during the summer. This seasonal pattern in volatility aligns closely with the yearly seasonality pattern shown in Figure 7d. Notably, BASTION is unique among decomposition methods in its ability to directly model volatility, offering valuable insights into the dynamics of the data that other models cannot provide.

The seasonality components extracted by BASTION also reveal patterns that are not immediately evident in the raw data. Both weekly and yearly seasonalities are modeled. While the weekly pattern is difficult to discern from the observed data, Figure 7c clearly demonstrates it: electricity demand remains relatively steady during weekdays and decreases on weekends, corroborating common patterns of electricity usage.

The differences in the yearly seasonality estimates produced by BASTION compared to existing models are studied in detailed, as shown in Figure 8. All four methods accurately capture the broad seasonal patterns typical of the East Coast, with peak electricity demand during the summer months of July and August due to increased cooling needs. This is followed by a steady decline in the fall and a moderate rise during the winter, reflecting increased heating requirements.

One interesting feature visible in the BASTION and MSTL estimates is a small dip in demand lasting approximately 30 days, beginning in mid-December and ending in mid-January. This drop coincides with the holiday season, during which many businesses and industries scale back operations or temporarily close. The reduction in commercial and industrial activity significantly decreases electricity consumption, which is a primary contributor to overall demand.

BASTION stands out for its ability to detect such subtle changes while maintaining a smooth and accurate estimate of the seasonal component. This characteristic highlights BASTION’s capability to adapt to abrupt changes in seasonality patterns, a strength that is also demonstrated in DGS 2 of the simulation study. These findings reinforce the utility of BASTION in capturing nuanced

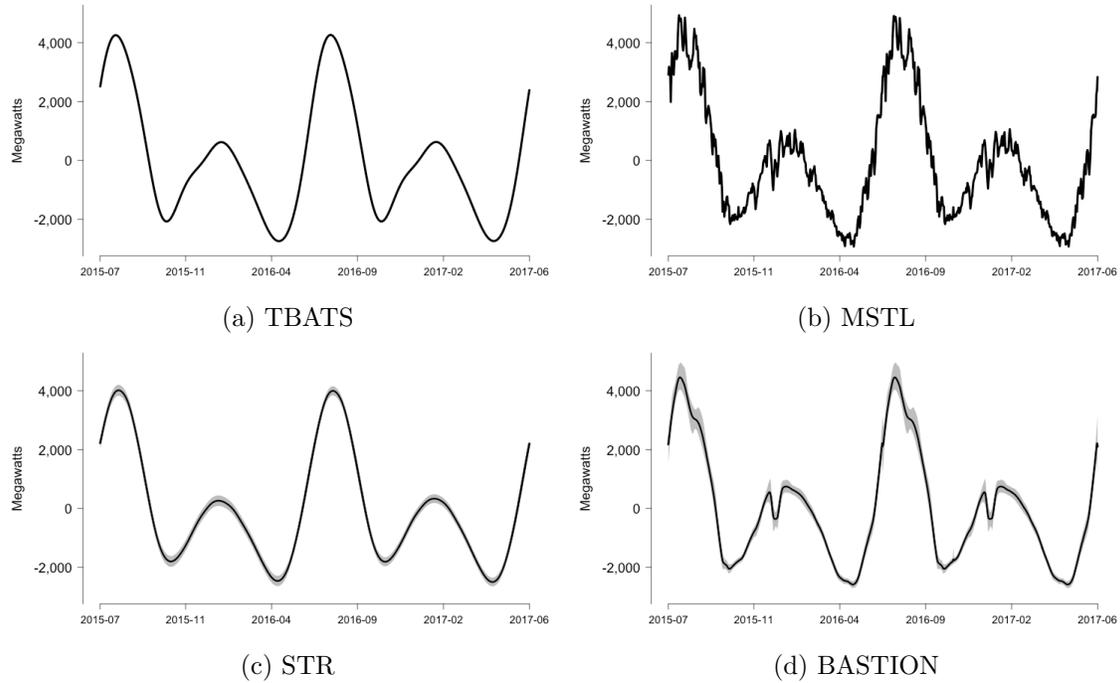


Figure 8: Yearly seasonality estimates for daily average electricity demand in New York from July 2015 to June 2017. Figures (a) - (d) display estimates from TBATS, MSTL, STR, and BASTION, respectively.

seasonal variations that other methods may miss.

4.3 Daily Bike Rentals via Capital Bikeshare System

We analyze daily bike rentals from the Capital Bikeshare system, a company providing bike rentals in the Washington, D.C., and surrounding areas. The studied period is between 2012-01-01 and 2017-12-31. The dataset (Bikeshare [2024]), categorizes bike rentals into two groups: members and non-members. This analysis focuses on non-member users, as their rental patterns exhibit more discernible seasonality patterns (Figure 9a).

The data’s most notable feature is its yearly seasonality. Bike rentals increase starting in January, peak during the summer months (June to August), and decline toward December, reflecting seasonal demand driven by weather. A similar seasonal pattern appears in the 30-day rolling standard deviation (Figure 9b), indicating that rental activity becomes more variable during high-demand periods. The BASTION’s estimate of the volatility (Figure 9d) also highlights increased volatility during the high demand period, also capturing the seasonal pattern present in the data.

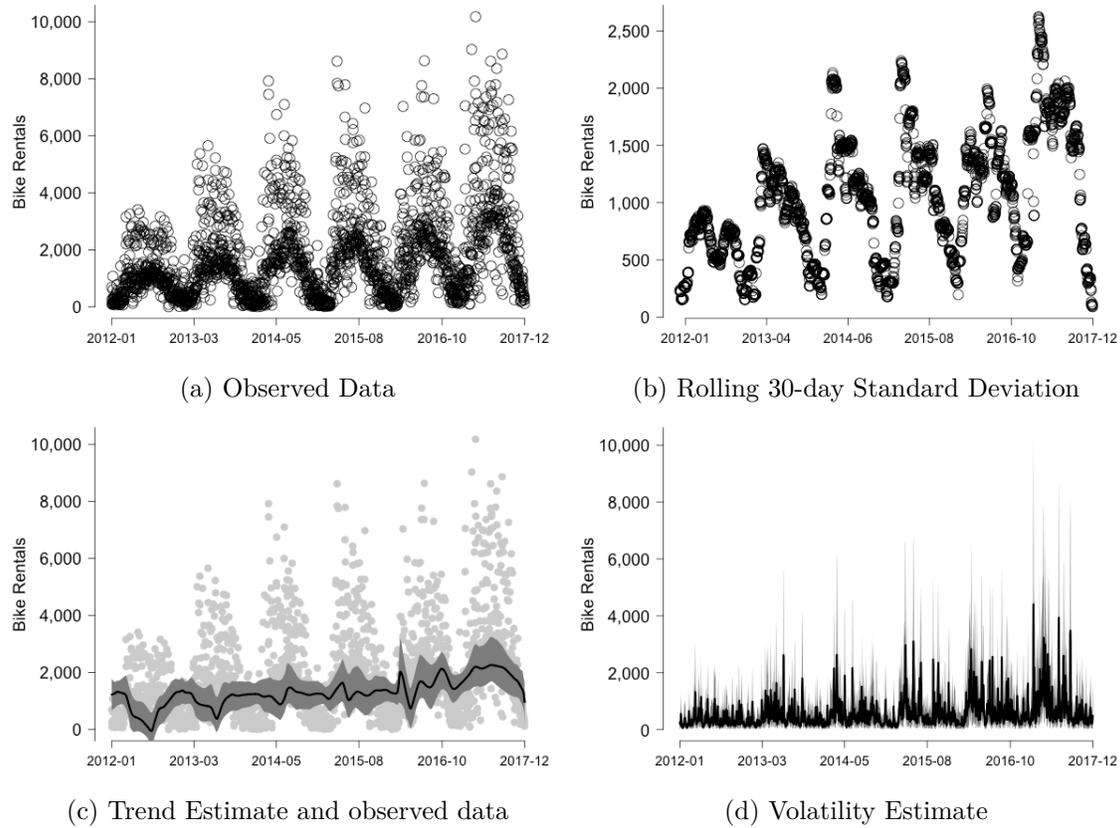


Figure 9: Daily total number of bike rentals via Capital Bikeshare system by non-members from 2012-01-01 to 2017-12-31 (9a), its rolling 30-day standard deviation (9b), BASTION estimate of the trend (7a) and volatility (7b). 95% credible regions are drawn in grey.

In contrast to the strong seasonal patterns, no clear long-term trend is apparent in the observed data (Figure 9a). The BASTION trend estimate (Figure 9c) suggests a slight overall increase in bike rentals, but the trend is not pronounced and exhibits localized fluctuations. These fluctuations, while present in the trend estimate, are accompanied by relatively large uncertainty bands, indicating that they are likely attributable to noise rather than meaningful long-term changes in demand.

Figure 10 presents the weekly and yearly seasonality components estimated using MSTL, STR, and BASTION. The TBATS model is excluded from the analysis as it fails to extract a meaningful seasonality component. The weekly seasonal components reveal a consistent pattern across all three models, showing higher bike rentals during weekends and lower, steady demand on weekdays. However, the long-term patterns over multiple years differ across models. STR (Figure 10b) estimates a steady weekly pattern that declines in 2015 before rebounding in 2016. In contrast, MSTL (Fig-

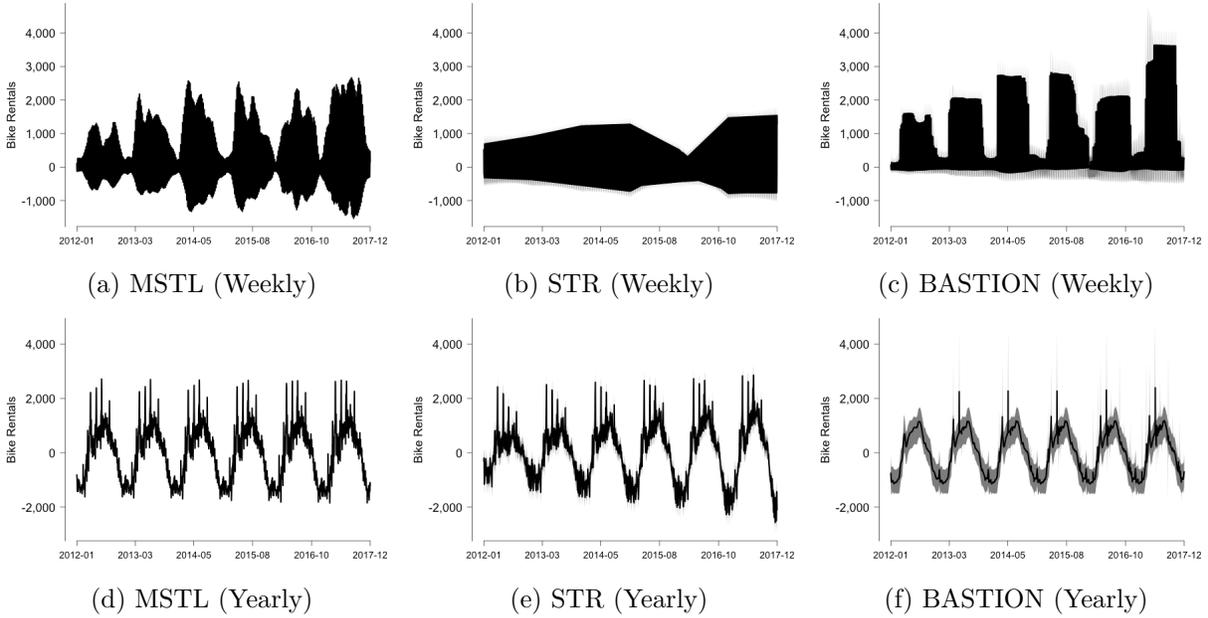


Figure 10: Weekly (a-c) and yearly (d-f) seasonal components of daily non-member bike rentals from the Capital Bikeshare system from 2012-01-01 to 2017-12-31, as estimated by MSTL, STR, and BASTION. The weekly panels highlight short-term periodic patterns, while the yearly panels capture long-term seasonal variations across the study period.

ure 10a and BASTION (Figure 10c produce more dynamic weekly patterns, with smaller seasonal cycles in winter and larger cycles in summer.

The estimates from MSTL and BASTION align more closely with the nature of the data, which focuses on non-member rentals. Non-members are likely tourists or occasional users, whose activity is highly seasonal. Tourists are less active during colder months and more active during warmer months, particularly in a city like Washington, D.C., where winter weather often discourages biking. The dynamic patterns captured by MSTL and BASTION with smaller weekend peaks in winter and larger ones in summer reflect these seasonal trends and align well with real-world factors influencing bike rental activity.

All three models produced consistent yearly components, aligning with the observed patterns in Figure 9a, where bike demand peaks during the summer months and steadily declines toward winter, reflecting seasonal trends. However, the estimates also reveal local outliers during the summer months, with varying degrees of intensity across the models. Notably, BASTION, which incorporates an explicit outlier term, produces the smoothest estimates with the fewest outliers, highlighting its ability to account for irregularities in the data more effectively.

5 Conclusion

In this paper, we introduced BASTION (Bayesian Adaptive Seasonality and Trend Decomposition Incorporating Outliers and Noise), a novel and flexible Bayesian framework for decomposing time series into trend, multiple seasonalities with explicit modeling of both outliers and volatility. BASTION’s key contributions lie in its ability to provide locally adaptive estimates of both trend and seasonality, explicit modeling of outliers and volatility, and rigorous uncertainty quantification through credible intervals. By leveraging global-local shrinkage priors, we derived an efficient Gibbs sampling scheme for posterior inference, enabling robust and accurate decomposition even in complex time-series datasets.

Through simulation studies, BASTION demonstrated superior performance compared to existing models such as TBATS, STR, and MSTL. It provides not only more accurate estimate of the trend and seasonality components but also more reliable uncertainty quantification, as evidenced by its robust nominal coverage in various scenarios. Strengths of BASTION is also shown via analysis of data including airline passenger, electricity demand, and bike rentals. BASTION provides insights that aligns with the underlying characteristics of the data and offered a more nuanced and comprehensive understanding of the data compared to traditional methods.

Building upon this work, future research could extend BASTION to incorporate covariate-driven seasonalities and trends, allowing for more nuanced modeling of time-series data influenced by external factors. Additionally, BASTION’s ability to effectively extract trend and seasonality from noisy time-series opens up opportunities for its application in diverse fields such as finance, environmental science, and epidemiology, where understanding complex temporal dynamics is critical.

References

- Tarekegn A. Abeku, Sake J. De Vlas, Gerard Borsboom, Awash Teklehaimanot, Asnakew Kebede, Dereje Olana, Gerrit J. Van Oortmarssen, and J. D. F. Habbema. Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of ethiopia: a simple seasonal adjustment method performs best. *Tropical Medicine & International Health*, 7(10):851–857, 2002. doi: <https://doi.org/10.1046/j.1365-3156.2002.00924.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-3156.2002.00924.x>.
- Jorge Alberto Achcar, Ricardo Puziol de Oliveira, and Emerson Barili. Use of stochastic volatility models in epidemiological data: Application to a dengue time series in s√lo paulo city, brazil. *Journal of Biostatistics and Epidemiology*, 6(1):19–29, Oct. 2020. doi: 10.18502/jbe.v6i1.4755. URL <https://jbe.tums.ac.ir/index.php/jbe/article/view/338>.

- Rob J. Hyndman Alysha M. De Livera and Ralph D. Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):1513–1527, 2011. doi: 10.1198/jasa.2011.tm09771. URL <https://doi.org/10.1198/jasa.2011.tm09771>.
- Kasun Bandara, Rob J Hyndman, and Christoph Bergmeir. Mstl: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns, 2021.
- Sayantana Banerjee. Horseshoe shrinkage methods for bayesian fusion estimation. *Computational Statistics & Data Analysis*, 174:107450, 2022. ISSN 0167-9473. doi: <https://doi.org/10.1016/j.csda.2022.107450>. URL <https://www.sciencedirect.com/science/article/pii/S0167947322000305>.
- Anindya Bhadra, Jyotishka Datta, Nicholas G. Polson, and Brandon Willard. The horseshoe+ estimator of ultra-sparse signals, 2015.
- Capital Bikeshare. Capital bikeshare system data, 2024. URL <https://capitalbikeshare.com/system-data>. Accessed: 2024-06-16.
- Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. ISSN 00223808, 1537534X. URL <http://www.jstor.org/stable/1831029>.
- Annalisa Cadonna, Sylvia Frühwirth-Schnatter, and Peter Knaus. Triple the gamma—a unifying shrinkage prior for variance and variable selection in sparse state space and tvp models. *Econometrics*, 8(2):20, 2020.
- CARLOS M. CARVALHO, NICHOLAS G. POLSON, and JAMES G. SCOTT. The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480, 2010. ISSN 00063444, 14643510. URL <http://www.jstor.org/stable/25734098>.
- Jason B. Cho and David S. Matteson. Locally adaptive random walk stochastic volatility, 2024. URL <https://arxiv.org/abs/2408.11315>.
- R. B. Cleveland. Stl : A seasonal-trend decomposition procedure based on loess. 1990. URL <https://api.semanticscholar.org/CorpusID:64570714>.
- E.B. Dagum, Statistics Canada. Seasonal Adjustment, and Time Series Analysis Staff. *The X-II-ARIMA Seasonal Adjustment Method*. Cat. 12-564 E Occasional. Statistics Canada, Seasonal Adjustment and Time Series Staff, 1980. URL <https://books.google.com/books?id=PWN6Mcnbc3EC>.
- Estela Bee Dagum. *On the seasonal adjustment of economic time series aggregates: A case study of the unemployment rate*. Number 31. National Commission on Employment and Unemployment Statistics, 1979.
- William R. Bell Mark C. Otto David F. Findley, Brian C. Monsell and Bor-Chung Chen. New capabilities and methods of the x-12-arma seasonal-adjustment program. *Journal of Business & Economic Statistics*, 16(2):127–152, 1998. doi: 10.1080/07350015.1998.10524743. URL <https://doi.org/10.1080/07350015.1998.10524743>.
- Alexander Dokumentov and Rob Hyndman. *stR: STR Decomposition of Time Series*, 2024. URL <https://pkg.robjhyndman.com/stR/>. R package version 0.7.
- Alexander Dokumentov and Rob J. Hyndman. Str: Seasonal-trend decomposition using regression. *INFORMS Journal on Data Science*, 1(1):50–62, 2022. doi: 10.1287/ijds.2021.0004. URL <https://doi.org/10.1287/ijds.2021.0004>.
- U.S. Energy Information Administration (EIA). New york independent system operator electricity demand data, 2024. URL https://www.eia.gov/electricity/gridmonitor/dashboard/electric_overview/regional/REG-NY. Accessed: 2024-06-16.

- David F Findley. Some recent developments and directions in seasonal adjustment. *Journal of official statistics*, 21(2):343, 2005.
- Everette S Gardner Jr and Joaquin Diaz-Saiz. Seasonal adjustment of inventory demand series: a case study. *International Journal of Forecasting*, 18(1):117–123, 2002.
- J. Grieser, S. Trömel, and C.-D. Schönwiese. Statistical time series decomposition into significant components and application to european temperature. *Theoretical and Applied Climatology*, 71(3):171–183, Feb 2002. ISSN 1434-4483. doi: 10.1007/s007040200003. URL <https://doi.org/10.1007/s007040200003>.
- James D. Hamilton. Oil and the macroeconomy since world war ii. *Journal of Political Economy*, 91(2):228–248, 1983. ISSN 00223808, 1537534X. URL <http://www.jstor.org/stable/1832055>.
- Robert J. Hodrick and Edward C. Prescott. Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16, 1997. ISSN 00222879, 15384616. URL <http://www.jstor.org/stable/2953682>.
- Ching-Lai Hor, Simon J Watson, and Shanti Majithia. Daily load forecasting and maximum demand estimation using arima and garch. In *2006 International conference on probabilistic methods applied to power systems*, pages 1–6. IEEE, 2006.
- Florian Huber and Michael Pfarrhofer. Dynamic shrinkage in time-varying parameter stochastic volatility in mean models. *Journal of Applied Econometrics*, 36(2):262–270, 2021. doi: <https://doi.org/10.1002/jae.2804>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2804>.
- JOHN Hull and ALAN White. The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300, 1987. doi: <https://doi.org/10.1111/j.1540-6261.1987.tb02568.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1987.tb02568.x>.
- Rob Hyndman, George Athanasopoulos, Christoph Bergmeir, Gabriel Caceres, Leanne Chhay, Mitchell O’Hara-Wild, Fotios Petropoulos, Slava Razbash, Earo Wang, and Farah Yasmeen. *forecast: Forecasting functions for time series and linear models*, 2024. URL <https://pkg.robjhyndman.com/forecast/>. R package version 8.23.0.
- Song Jiang, Tahin Syed, Xuan Zhu, Joshua Levy, Boris Aronchik, and Yizhou Sun. Bridging self-attention and time series decomposition for periodic forecasting. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, CIKM ’22*, page 3202–3211, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392365. doi: 10.1145/3511808.3557077. URL <https://doi.org/10.1145/3511808.3557077>.
- Gregor Kastner. Dealing with Stochastic Volatility in Time Series Using the R Package *stochvol*. *Journal of Statistical Software*, 69(i05), 2016. doi: <http://hdl.handle.net/10.1002/jss/jsts/v69i05>. URL <https://ideas.repec.org/a/jss/jsts/v69i05.html>.
- Gregor Kastner and Sylvia Frühwirth-Schnatter. Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis*, 76:408–423, aug 2014. doi: 10.1016/j.csda.2013.01.002. URL <https://doi.org/10.1016/j.csda.2013.01.002>.
- Guibao Ke, Yao Hu, Xin Huang, Xuan Peng, Min Lei, Chaoli Huang, Li Gu, Ping Xian, and Dehua Yang. Epidemiological analysis of hemorrhagic fever with renal syndrome in china with the seasonal-trend decomposition method and the exponential smoothing model. *Scientific reports*, 6(1):39350, 2016.
- Sangjoon Kim, Neil Shephard, and Siddhartha Chib. Stochastic volatility: Likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393, 1998. ISSN 00346527, 1467937X. URL <http://www.jstor.org/stable/2566931>.

- Seung-Jean Kim, Kwangmoo Koh, Stephen Boyd, and Dimitry Gorinevsky. l_1 trend filtering. *SIAM Rev.*, 51(2):339–360, 2009. ISSN 0036-1445,1095-7200. doi: 10.1137/070690274. URL <https://doi.org/10.1137/070690274>.
- Daniel R. Kowal, David S. Matteson, and David Ruppert. Dynamic Shrinkage Processes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 81(4):781–804, 05 2019. ISSN 1369-7412. doi: 10.1111/rssb.12325. URL <https://doi.org/10.1111/rssb.12325>.
- F.R. Macaulay. *The Smoothing of Time Series, By Frederick R. Macaulay*. Publication, no. 19. National Bureau of Economic Research, 1972. URL https://books.google.com/books?id=_6pnnQEACAAJ.
- Enes Makalic and Daniel F. Schmidt. A simple sampler for the horseshoe estimator. *IEEE Signal Processing Letters*, 23(1):179–182, 2016. doi: 10.1109/LSP.2015.2503725.
- Maria C Mariani, Md Al Masum Bhuiyan, Osei K Tweneboah, Hector Gonzalez-Huizar, and Ionut Florescu. Volatility models applied to geophysics and high frequency financial market data. *Physica A: Statistical Mechanics and its Applications*, 503:304–321, 2018.
- Angelo Melino and Stuart M. Turnbull. Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45(1-2):239–265, 1990. URL <https://EconPapers.repec.org/RePEc:eee:econom:v:45:y:1990:i:1-2:p:239-265>.
- Reza Modarres and Taha BMJ Ouarda. Modeling the relationship between climate oscillations and drought by a multivariate garch model. *Water Resources Research*, 50(1):601–618, 2014.
- United States. Bureau of Labor Statistics. *The Consumer Price Index: Concepts and Content Over the Years*. Report (United States. Bureau of Labor Statistics). The Bureau, 1977. URL <https://books.google.com/books?id=sPDxroxMTi0C>.
- Yasuhiro Omori, Siddhartha Chib, Neil Shephard, and Jouchi Nakajima. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449, 2007. ISSN 0304-4076. doi: <https://doi.org/10.1016/j.jeconom.2006.07.008>. URL <https://www.sciencedirect.com/science/article/pii/S0304407606001436>.
- Juho Piironen and Aki Vehtari. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, 11(2):5018 – 5051, 2017. doi: 10.1214/17-EJS1337SI. URL <https://doi.org/10.1214/17-EJS1337SI>.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2024. URL <https://www.R-project.org/>.
- Jaume Rosselló and Andreu Sansó. Yearly, monthly and weekly seasonality of tourism demand: A decomposition analysis. *Tourism Management*, 60:379–389, 2017.
- Havard Rue. Fast sampling of gaussian markov random fields. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63(2):325–338, 2001. ISSN 13697412, 14679868. URL <http://www.jstor.org/stable/2680602>.
- RR Sarkar and C Chatterjee. Application of different time series models on epidemiological data-comparison and predictions for malaria prevalence. *SM J. Biom. Biostat*, 2(4):1022, 2017.
- Toryn L. J. Schafer and David S. Matteson. Locally adaptive shrinkage priors for trends and breaks in count time series, 2023.
- I. J. Schoenberg. Spline functions and the problem of graduation. *Proceedings of the National Academy of Sciences of the United States of America*, 52(4):947–950, 1964. ISSN 00278424, 10916490. URL <http://www.jstor.org/stable/72390>.
- J. Shiskin, A.H. Young, J.C. Musgrave, and United States. Bureau of the Census. *The X-11 Variant of the Census Method II Seasonal Adjustment Program*. Technical paper. U.S. Government Printing Office, 1965. URL <https://books.google.com/books?id=BFIfiGmatUoC>.

- Julius Shiskin. Appendix a: Computational steps and sample tables. In *Electronic Computers and Business Indicators*, pages 30–39. National Bureau of Economic Research, Inc, 1957. URL <https://EconPapers.repec.org/RePEc:nbr:nberch:2730>.
- James H. Stock and Mark W. Watson. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3):788–829, September 2003. doi: 10.1257/002205103322436197. URL <https://www.aeaweb.org/articles?id=10.1257/002205103322436197>.
- Sean J. Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018. doi: 10.1080/00031305.2017.1380080. URL <https://doi.org/10.1080/00031305.2017.1380080>.
- S.J. Taylor. *Modelling Financial Time Series*. G - Reference, Information and Interdisciplinary Subjects Series. World Scientific, 2008. ISBN 9789812770844. URL <https://books.google.com/books?id=KQ5pDQAAQBAJ>.
- Ryan J. Tibshirani. Adaptive piecewise polynomial estimation via trend filtering. *The Annals of Statistics*, 42(1), February 2014. ISSN 0090-5364. doi: 10.1214/13-aos1189. URL <http://dx.doi.org/10.1214/13-AOS1189>.
- Michael K Tippett. Changing volatility of us annual tornado reports. *Geophysical Research Letters*, 41(19):6956–6961, 2014.
- W Wang, PHAJ M Van Gelder, JK Vrijling, and J Ma. Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes. *Nonlinear processes in Geophysics*, 12(1): 55–66, 2005.
- Qingsong Wen, Jingkun Gao, Xiaomin Song, Liang Sun, Huan Xu, and Shenghuo Zhu. Robuststl: A robust seasonal-trend decomposition algorithm for long time series. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 5409–5416, 2019.
- Gerald Woo, Chenghao Liu, Doyen Sahoo, Akshat Kumar, and Steven Hoi. Etsformer: Exponential smoothing transformers for time-series forecasting, 2022. URL <https://arxiv.org/abs/2202.01381>.
- Haoxuan Wu, Toryn L. J. Schafer, and David S. Matteson. Trend and variance adaptive bayesian changepoint analysis and local outlier scoring. *Journal of Business & Economic Statistics*, 0(0): 1–12, 2024. doi: 10.1080/07350015.2024.2362269. URL <https://doi.org/10.1080/07350015.2024.2362269>.
- Yuxian Yan. U.s. airline traffic data (2003-2023), 2023. URL <https://www.kaggle.com/datasets/yxian/u-s-airline-traffic-data>. Accessed: 2024-06-16.
- Victor Zarnowitz and Ataman Ozyildirim. Time series decomposition and measurement of business cycles, trends and growth cycles. *Journal of Monetary Economics*, 53(7):1717–1739, 2006. ISSN 0304-3932. doi: <https://doi.org/10.1016/j.jmoneco.2005.03.015>. URL <https://www.sciencedirect.com/science/article/pii/S0304393206000596>.
- Jing Zhou, Zhongyao Liang, Yong Liu, Huaicheng Guo, Dan He, and Lei Zhao. Six-decade temporal change and seasonal decomposition of climate variables in lake dianchi watershed (china): stable trend or abrupt shift? *Theoretical and Applied Climatology*, 119:181–191, 2015.

A Full Conditional Distribution for Gibbs Sampling

A.1 Trend

For the likelihood, we have:

$$[\mathbf{Y}|\mathbf{T}, \dots] \sim N\left(\mathbf{T} + \sum_{i=1}^P \mathbf{S}_i + \boldsymbol{\zeta} + \mathbf{X}\boldsymbol{\beta}, \sigma_y^2 \boldsymbol{\nu}^2 \mathbf{I}\right).$$

The prior distribution on \mathbf{T} are imposed on its second differencing. Thus, we may derive the conditional prior distribution of \mathbf{T} by taking the linear combination of $\mathbf{T}^* = [T_1, T_2, \Delta^2 T_3, \dots, \Delta^2 T_N]'$, i.e. $\mathbf{T} = A_T \mathbf{T}^*$, where

$$A_T = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ -1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

$$A_T^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}.$$

Thus, given $\boldsymbol{\sigma}_T^2 := [\sigma_y^2 \lambda_{T,1}^2, \sigma_y^2 \lambda_{T,2}^2, \sigma_y^2 \tau_T^2 \lambda_{T,3}^2, \dots, \sigma_y^2 \tau_T^2 \lambda_{T,N}^2]'$

$$\begin{aligned} [\mathbf{T}|\dots] &\sim N(0, A\boldsymbol{\sigma}_T^2 I A') \\ &\sim N\left(0, \left((A_T^{-1})' \frac{\mathbf{1}}{\boldsymbol{\sigma}_T^2} I A_T^{-1}\right)^{-1}\right). \end{aligned}$$

Let's define $Q_T := (A_T^{-1})' \frac{1}{\sigma_T^2} I A_T^{-1}$. Based on our derivation of A_T^{-1} , we have

$$Q_T = \frac{1}{\sigma_y^2 \tau_T^2} \begin{bmatrix} \left(\frac{\tau_T^2}{\lambda_{T,1}^2} + \frac{1}{\lambda_{T,3}^2} \right) & -\frac{2}{\lambda_{T,3}^2} & -\frac{1}{\lambda_{T,3}^2} & 0 & \dots & \dots & 0 \\ -\frac{2}{\lambda_{T,3}^2} & \left(\frac{\tau_T^2}{\lambda_{T,2}^2} + \frac{4}{\lambda_{T,3}^2} + \frac{1}{\lambda_{T,4}^2} \right) & \frac{-2}{\lambda_{T,3}^2} - \frac{2}{\lambda_{T,4}^2} & -\frac{1}{\lambda_{T,4}^2} & \ddots & \ddots & \vdots \\ -\frac{1}{\lambda_{T,3}^2} & \frac{-2}{\lambda_{T,3}^2} - \frac{2}{\lambda_{T,4}^2} & \left(\frac{1}{\lambda_{T,3}^2} + \frac{4}{\lambda_{T,4}^2} + \frac{1}{\lambda_{T,5}^2} \right) & \frac{-2}{\lambda_{T,4}^2} - \frac{2}{\lambda_{T,5}^2} & -\frac{1}{\lambda_{T,5}^2} & \ddots & \ddots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \left(\frac{1}{\lambda_{T,N-2}^2} + \frac{4}{\lambda_{T,N-1}^2} + \frac{1}{\lambda_{T,N}^2} \right) & \left(\frac{-2}{\lambda_{T,N-1}^2} - \frac{2}{\lambda_{T,N}^2} \right) & -\frac{1}{\lambda_{T,N}^2} \\ \vdots & \ddots & \ddots & \ddots & \left(\frac{-2}{\lambda_{T,N-1}^2} - \frac{2}{\lambda_{T,N}^2} \right) & \left(\frac{1}{\lambda_{T,N-1}^2} + \frac{4}{\lambda_{T,N}^2} \right) & -\frac{2}{\lambda_{T,N}^2} \\ 0 & \dots & \dots & \dots & -\frac{1}{\lambda_{T,N}^2} & -\frac{2}{\lambda_{T,N}^2} & \frac{1}{\lambda_{T,N}^2} \end{bmatrix}$$

Therefore,

$$[\mathbf{T}|\mathbf{y}, \dots] \sim N\left(\left(Q_T + \frac{1}{\sigma_y^2 \nu^2} I \right)^{-1} \left(\frac{\mathbf{Y} - \sum_{i=1}^P S_i - \boldsymbol{\zeta} - \mathbf{X}\boldsymbol{\beta}}{\sigma_y^2 \nu^2} \right), \left(Q_T + \frac{1}{\sigma_y^2 \nu^2} I \right)^{-1} \right).$$

For the global parameter τ_T and the local parameter $\lambda_{T,t}$ both following the half-Cauchy distribution, the parameter expansion as described in Makalic and Schmidt [2016] are used. Thus,

$$\begin{aligned} [\tau_T | \psi_{\tau_T}, \dots] &\sim IG(1/2, 1/\psi_{\tau_T}) & [\psi_{\tau_T}] &\sim IG(1/2, 1) \\ [\lambda_{T,t} | \psi_{\lambda_{T,t}}, \dots] &\sim IG(1/2, 1/\psi_{\lambda_{T,t}}) & [\psi_{\lambda_{T,t}}] &\sim IG(1/2, 1) \end{aligned}$$

Since they are conjugate priors, the conditional prior distributions are as follows:

$$\begin{aligned}
[\tau_T^2 | \mathbf{Y}, \psi_{\tau_T}, \dots] &\sim IG\left(\frac{1}{2} + \frac{N-2}{2}, \frac{1}{\psi_{\tau_T}} + \frac{1}{2\sigma_y^2} \sum_{t=3}^N \left(\frac{\Delta^2 T_t}{\lambda_{T,t}}\right)^2\right). \\
[\psi_{\tau_T} | \mathbf{Y}, \tau_T, \dots] &\sim IG\left(1, 1 + \frac{1}{\tau_T^2}\right). \\
[\lambda_{T,t}^2 | \mathbf{Y}, \psi_{\lambda_{T,t}}, \dots] &\sim IG\left(1, \frac{1}{\psi_{\lambda_{T,t}}} + \frac{1}{2} \left(\frac{T_t}{\sigma_y}\right)^2\right), & t \in \{1, 2\}. \\
&\sim IG\left(1, \frac{1}{\psi_{\lambda_{T,t}}} + \frac{1}{2} \left(\frac{\Delta^2 T_t}{\sigma_y \tau_T}\right)^2\right), & t \geq 3. \\
[\psi_{\lambda_{T,t}} | \mathbf{Y}, \lambda_{T,t}, \dots] &\sim IG\left(1, 1 + \frac{1}{\lambda_{T,t}^2}\right).
\end{aligned}$$

A.2 Seasonality

Let's fix $j \in \{1, \dots, P\}$, and consider the seasonality term $\mathbf{S}_j^{k_j}$. To simplify the notation, let $k := k_j$ to denote the length of the cycle, and $\mathbf{S} := \mathbf{S}_j^{k_j}$. Similar to the derivation of the conditional prior distribution for \mathbf{T} in section A.1, the conditional prior distribution of \mathbf{S} is also a linear combination of $\mathbf{S}^* := [S_2, \Delta^2 S_3, \dots, \Delta^2 S_k, (1-B^k)S_{k+1}, \dots, (1-B^k)S_N]'$. Let's also remind ourselves that $S_1 = 0$, thus we only consider $N-1$ terms. The linear transformation A_S , such that $\mathbf{S} := A_S \mathbf{S}^*$, is defined as a the product of the following two matrices:

$$A_S = \begin{bmatrix} I_{k-1} & 0 \\ B_3 & B_2 \end{bmatrix} \begin{bmatrix} B_1 & 0 \\ 0 & I_{N-k} \end{bmatrix}$$

where each matrix $B_{k,1}$, $B_{k,2}$ has the following structure:

$$B_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 2 & 1 & 0 & \cdots & 0 \\ 3 & 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k-1 & k-2 & k-3 & \cdots & 1 \end{bmatrix} \quad B_2 = \begin{bmatrix} I_k & 0 & 0 & \cdots & 0 \\ I_k & I_k & 0 & \cdots & 0 \\ I_k & I_k & I_k & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_k & I_k & I_k & \cdots & I_k \end{bmatrix} \quad B_3 = \begin{bmatrix} \mathbf{0} \\ I_{k-1} \\ \mathbf{0} \\ I_{k-1} \\ \mathbf{0} \\ \vdots \end{bmatrix}.$$

B_1 is a $(k-1) \times (k-1)$ lower triangle matrix for un-differencing the second differencing operator on the first $k-1$ observations. B_2 is a $(N-k) \times (N-k)$ lower triangle matrix and B_3 is a $(N-k) \times (k-1)$ matrix for un-differencing the seasonal differencing operator on the last $T-k$ observations. Similar to the derivation of Q_T in Section A.1, $Q_S^{-1} := \left((A_S^{-1})' \frac{1}{\sigma_S^2} I A_S^{-1} \right)$, where $\sigma_S^2 := [\sigma_y^2 \lambda_{S,2}^2, \sigma_y^2 \tau_S^2 \lambda_{S,3}^2, \sigma_y^2 \tau_S^2 \lambda_{S,4}^2, \dots, \sigma_y^2 \tau_S^2 \lambda_{S,N}^2]'$. The precision matrix Q_{S^k} has a block-tridiagonal structure with blocks separated by the seasonality k except for the top left $(k-1) \times (k-1)$ submatrix, which has a tridiagonal structure resulting from the second differencing.

The conditional posterior distribution for \mathbf{S} therefore,

$$[\mathbf{S}|\mathbf{y}, \dots] \sim N \left(\left(Q_S + \frac{1}{\sigma_y^2 \nu^2} I \right)^{-1} \left(\frac{\mathbf{Y} - \mathbf{T} - \sum_{i \neq j} \mathbf{S}_i - \boldsymbol{\zeta} - \mathbf{X}\boldsymbol{\beta}}{\sigma_y^2 \nu^2} \right), \left(Q_S + \frac{1}{\sigma_y^2 \nu^2} I \right)^{-1} \right).$$

The conditional posterior distributions of τ_S and λ_S are identical to τ_{T^k} and λ_{T^k} explored in Section A.1. With the parameter expansion of the horseshoe prior:

$$\begin{aligned} [\tau_S^2 | \mathbf{Y}, \psi_{\tau_S}, \dots] &\sim IG \left(\frac{1}{2} + \frac{N-2}{2}, \frac{1}{\psi_{\tau_S}} + \frac{1}{2\sigma_y^2} \left(\sum_{t=3}^k \left(\frac{\Delta^2 S_t}{\lambda_{T,t}} \right)^2 + \sum_{t=(k+1)}^N \left(\frac{(1-B^k)S_t}{\lambda_{T,t}} \right)^2 \right) \right). \\ [\psi_{\tau_S} | \mathbf{Y}, \tau_S, \dots] &\sim IG \left(1, 1 + \frac{1}{\tau_S^2} \right). \\ [\lambda_{S,t}^2 | \mathbf{Y}, \psi_{\lambda_{S,t}}, \dots] &\sim IG \left(1, \frac{1}{\psi_{\lambda_{S,t}}} + \frac{1}{2} \left(\frac{S_t}{\sigma_y} \right)^2 \right), & t = 2, \\ &\sim IG \left(1, \frac{1}{\psi_{\lambda_{S,t}}} + \frac{1}{2} \left(\frac{\Delta^2 S_t}{\sigma_y \tau_S} \right)^2 \right), & t = 3, \dots, k, \\ &\sim IG \left(1, \frac{1}{\psi_{\lambda_{S,t}}} + \frac{1}{2} \left(\frac{(1-B^k)S_t}{\sigma_y \tau_S} \right)^2 \right), & t = (k+1), \dots, N. \\ [\psi_{\lambda_{S,t}} | \mathbf{Y}, \lambda_{S,t}, \dots] &\sim IG \left(1, 1 + \frac{1}{\lambda_{S,t}^2} \right). \end{aligned}$$

A.3 Outliers

For the additive outlier term ζ_t , we use the horseshoe+ prior by Bhadra et al. [2015] to provide more shrinkage than the one induced by the horseshoe prior. The parameter expansion by Makalic

and Schmidt [2016] can be applied to the horseshoe+ prior:

$$\begin{aligned}
[\zeta_t | \sigma_y, \lambda_{\zeta,t}] &\sim N(0, \sigma_y^2 \lambda_{\zeta,t}^2), & [\lambda_{\zeta,t}^2 | \psi_{\lambda_{\zeta,t}}] &\sim IG(1/2, 1/\psi_{\lambda_{\zeta,t}}), \\
[\psi_{\lambda_{\zeta,t}} | \tau_{\zeta}, \xi_{\zeta,t}] &\sim IG(1/2, 1/(\tau_{\zeta}^2 \xi_{\zeta,t}^2)), \\
[\tau_{\zeta}^2 | \psi_{\tau_{\zeta}}] &\sim IG(1/2, 1/\psi_{\tau_{\zeta}}), & [\psi_{\tau_{\zeta}}] &\sim IG(1/2, 1), \\
[\xi_{\zeta,t}^2 | \psi_{\xi_{\zeta,t}}] &\sim IG(1/2, 1/\psi_{\xi_{\zeta,t}}), & [\psi_{\xi_{\zeta,t}}] &\sim IG(1/2, 1).
\end{aligned}$$

Thus, we have the following conditional posterior distribution:

$$\begin{aligned}
[\boldsymbol{\zeta} | \mathbf{y}, \dots] &\sim N\left(\left(\frac{\boldsymbol{\lambda}_{\zeta}^2}{\boldsymbol{\nu}^2 + \boldsymbol{\lambda}_{\zeta}^2}\right)\left(\mathbf{Y} - \mathbf{T} - \sum_{i=1}^P \mathbf{S}_i - \mathbf{X}\boldsymbol{\beta}\right), \left(\frac{\sigma_y^2 \boldsymbol{\nu}^2 \boldsymbol{\lambda}_{\zeta}^2}{\boldsymbol{\nu}^2 + \boldsymbol{\lambda}_{\zeta}^2}\right)\right), \\
[\lambda_{\zeta,t}^2 | \mathbf{y}, \dots] &\sim IG\left(1, \frac{1}{\psi_{\lambda_{\zeta,t}}} + \frac{1}{2}\left(\frac{\zeta_t}{\sigma_y}\right)^2\right) \\
[\psi_{\lambda_{\zeta,t}} | \mathbf{y}, \dots] &\sim IG\left(1, \frac{1}{\lambda_{\zeta,t}} + \frac{1}{\tau_{\zeta}^2 \xi_{\zeta,t}^2}\right) \\
[\tau_{\zeta}^2 | \mathbf{y}, \dots] &\sim IG\left(\frac{N+1}{2}, \sum_{t=1}^N \left(\frac{1}{\xi_{\zeta,t}^2 \psi_{\lambda_{\zeta,t}}}\right) + \frac{1}{\psi_{\tau_{\zeta}}}\right) \\
[\psi_{\tau_{\zeta}} | \mathbf{y}] &\sim IG\left(1, 1 + \frac{1}{\tau_{\zeta}^2}\right) \\
[\xi_{\zeta,t}^2 | \mathbf{y}, \dots] &\sim IG\left(1, \frac{1}{\tau_{\zeta}^2 \psi_{\lambda_{\zeta,t}}} + \frac{1}{\psi_{\xi_{\zeta,t}}}\right) \\
[\psi_{\xi_{\zeta,t}} | \mathbf{y}] &\sim IG\left(1, 1 + \frac{1}{\xi_{\zeta,t}^2}\right)
\end{aligned}$$

A.4 Remainders

Remainder term also decomposes into two components, the time-invariant σ_y , and the time-varying term ν_t . For the time-invariant term, we have

$$\sigma_y^2 \sim IG\left(\frac{3N + P(N-1) + m}{2}, \frac{1}{2}\left(\left(\frac{\mathbf{Y} - \mathbf{T} - \sum_{i=1}^P \mathbf{S}_i - \boldsymbol{\zeta} - \mathbf{X}\beta}{\boldsymbol{\nu}}\right)^2 + \sum_{t=1}^2 \left(\frac{T_t}{\lambda_{T,t}}\right)^2 + \sum_{t=3}^N \left(\frac{\Delta^2 T_t}{\tau_T \lambda_{T,t}}\right)^2 + \sum_{i=1}^P \left(\left(\frac{S_{i,2}}{\lambda_{S_{i,2}}}\right)^2 + \sum_{t=3}^k \left(\frac{\Delta^2 S_{i,t}}{\tau_{S_i} \lambda_{S_{i,t}}}\right)^2 + \sum_{k+1}^N \left(\frac{(1-B)^{k_i} S_{i,t}}{\tau_{S_i} \lambda_{S_{i,t}}}\right)^2\right) + \sum_{t=1}^N \left(\frac{\zeta_t}{\lambda_{\zeta,t}}\right)^2 + \beta' \left(\frac{1}{\boldsymbol{\eta}^2} \mathbf{I}\right) \beta\right)\right)$$

For the time-varying term $\boldsymbol{\nu}$, we closely follow the sampler proposed by Kastner and Frühwirth-Schnatter [2014]. Define

$$\mathbf{Y}^* = \log\left(\frac{\mathbf{Y} - \mathbf{T} - \sum_{i=1}^P \mathbf{S}_i - \boldsymbol{\zeta} - \mathbf{X}\beta}{\sigma_y}\right)^2$$

$$\mathbf{h} := \log(\boldsymbol{\nu}^2).$$

By transforming the likelihood, we have a linear system with a non-Gaussian error term, which may be approximated by a Gaussian mixture distribution Omori et al. [2007]:

$$\begin{aligned} y_t^* &= h_t + \log(u_t^2), & u_t &\stackrel{iid}{\sim} N(0, 1), \\ y_t^* &\approx h_t + \mu_{j_t} + \sigma_{j_t} u_t, & u_t &\stackrel{iid}{\sim} N(0, 1), \quad j_t \stackrel{iid}{\sim} \text{Categorical}(\boldsymbol{\pi}) \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \sigma_\nu \epsilon & [\epsilon_t] &\stackrel{iid}{\sim} N(0, 1) \end{aligned}$$

With the approximation of the likelihood, the model becomes conditionally Gaussian, allowing efficient sampling of associated parameter \mathbf{h} , μ , ϕ , σ_ν , and the additional parameter \mathbf{j} .

A.5 Regression Coefficients

Prior distributions on the regression coefficients, β , are assumed to be normally distributed with their own variance terms $\boldsymbol{\eta}$ following Jeffreys prior. Define $A = X'X + \frac{1}{\boldsymbol{\eta}} \mathbf{I}$. The conditional

posterior distributions for the related parameters are:

$$\begin{aligned}\boldsymbol{\beta}|\dots &\sim N\left(A^{-1}X^T\left(\mathbf{Y}-\mathbf{T}-\sum_{i=1}^P S_i-\boldsymbol{\zeta}\right),\sigma_y^2A^{-1}\right) \\ \eta_j|\dots &\sim IG\left(\frac{1}{2},\frac{\beta_j^2}{2\sigma_y^2}\right)\end{aligned}\quad j=1,\dots,m.$$

B Simulation Schemes

B.1 Data Generating Scheme 1

Observation Equation

$$Y_t = T_t + S_t^{12} + S_t^{40} + R_t.$$

Trend

$$T_t = \begin{cases} m_1(0.04)(t - b_1) + c_1, & \text{if } 0 \leq t < b_1, \\ m_2(0.04)(t - b_2) + c_2, & \text{if } b_1 \leq t < b_1 + b_2, \\ m_3(0.04)(t - b_3) + c_3, & \text{if } b_1 + b_2 \leq t < b_1 + b_2 + b_3, \\ m_4(0.04)(t - b_4) + c_4, & \text{if } b_1 + b_2 + b_3 \leq t \leq 500, \end{cases}$$

$$[m_1, \dots, m_4] \stackrel{iid}{\sim} U(-20, 20),$$

$$[c_1, \dots, c_4] \stackrel{iid}{\sim} U(-10, 10),$$

$$[b_1, \dots, b_3] \stackrel{iid}{\sim} U(30, 125).$$

Seasonal

$$S_t^{12} = \gamma_1^{12} \sin\left(\frac{2\pi t}{12}\right) + \gamma_2^{12} \cos\left(\frac{2\pi t}{12}\right)$$

$$[\gamma_1^{12}, \gamma_2^{12}] \stackrel{iid}{\sim} N(0, 4^2),$$

$$S_t^{40} = \gamma_1^{40} \sin\left(\frac{2\pi t}{40}\right) + \gamma_2^{40} \cos\left(\frac{2\pi t}{40}\right)$$

$$[\gamma_1^{40}, \gamma_2^{40}] \stackrel{iid}{\sim} N(0, 5^2).$$

Error

$$[R_t] \stackrel{iid}{\sim} N(0, 2^2)$$

B.2 Data Generating Scheme 2

Observation Equation

$$Y_t = T_t + S_t^{40} + R_t.$$

Trend

$$T_t = \frac{mt}{500}$$
$$[m] \sim N(0, 30^2).$$

Seasonal

$$S_t = \begin{cases} \tilde{v}_1, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 1 \leq j \leq 10\}, \\ \tilde{v}_2, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 11 \leq j \leq 20\}, \\ \tilde{v}_3, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 21 \leq j \leq 30\}, \\ \tilde{v}_4, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 31 \leq j \leq 40\}, \end{cases}$$
$$[v_1, \dots, v_4] \stackrel{iid}{\sim} U(-8, 8),$$
$$\tilde{v}_i = v_i - \frac{1}{4} \sum_{j=1}^4 v_j.$$

Error

$$[R_t] \stackrel{iid}{\sim} N\left(0, \frac{m^2}{100}\right).$$

B.3 Data Generating Scheme 3

Observation Equation

$$Y_t = T_t + S_t^{50} + R_t.$$

Trend

$$T_t = b_0 + b_1 \frac{t}{500} + b_2 \left(\frac{t}{500} \right)^2 + b_3 \left(\frac{t}{500} \right)^3$$

$$[b_0] \sim U(-15, 15),$$

$$[b_1, \dots, b_4] \stackrel{iid}{\sim} N(0, 20^2).$$

Seasonal

$$S_t^{50} = \gamma_1^{50} \sin \left(\frac{2\pi t}{50} \right) + \gamma_2^{50} \cos \left(\frac{2\pi t}{50} \right)$$

$$[\gamma_1^{50}, \gamma_2^{50}] \stackrel{iid}{\sim} N(0, 5^2).$$

Error

$$[R_t] \stackrel{iid}{\sim} N(0, \exp(h_t))$$

$$h_t = 2.5 + 0.98(h_{t-1} - 2.5) + 0.2\epsilon_t$$

$$[\epsilon_t] \stackrel{iid}{\sim} N(0, 1)$$

B.4 Data Generating Scheme 4

Observation Equation

$$Y_t = T_t + S_t^{12} + S_t^{50} + O_t + R_t, \quad t = 1, \dots, 500.$$

Trend

$$T_t = \begin{cases} m_1(0.04)(t - b_1) + c_1, & \text{if } 0 \leq t < b_1, \\ m_2(0.04)(t - b_2) + c_2, & \text{if } b_1 \leq t < b_1 + b_2, \\ m_3(0.04)(t - b_3) + c_3, & \text{if } b_1 + b_2 \leq t < b_1 + b_2 + b_3, \\ m_4(0.04)(t - b_4) + c_4, & \text{if } b_1 + b_2 + b_3 \leq t \leq 500, \end{cases}$$

$$[m_1, \dots, m_4] \stackrel{iid}{\sim} U(-20, 20),$$

$$[c_1, \dots, c_4] \stackrel{iid}{\sim} U(-10, 10),$$

$$[b_1, \dots, b_3] \stackrel{iid}{\sim} U(30, 125).$$

Seasonal

$$S_t^{12} = \gamma_1^{12} \sin\left(\frac{2\pi t}{12}\right) + \gamma_2^{12} \cos\left(\frac{2\pi t}{12}\right)$$

$$[\gamma_1^{12}, \gamma_2^{12}] \stackrel{iid}{\sim} N(0, 4^2),$$

$$S_t^{50} = \begin{cases} \tilde{v}_1, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 1 \leq j \leq 10\}, \\ \tilde{v}_2, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 11 \leq j \leq 20\}, \\ \tilde{v}_3, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 21 \leq j \leq 30\}, \\ \tilde{v}_4, & \text{if } t \in \{t \mid t = 40k + j, 0 \leq k \leq 8, 31 \leq j \leq 40\}, \end{cases}$$

$$[v_1, \dots, v_4] \stackrel{iid}{\sim} U(-20, 20),$$

$$\tilde{v}_i = v_i - \frac{1}{4} \sum_{j=1}^4 v_j$$

Outlier

$$O_t = \begin{cases} a_t, & \text{if } t \in \{j_1, j_2, \dots, j_l\}, \\ 0, & \text{if } t \notin \{j_1, j_2, \dots, j_l\}, \end{cases}$$

$$a_t \stackrel{iid}{\sim} N(0, 15^2),$$

$$l \sim Pois(5),$$

$$[j_1, j_2, \dots, j_l] \sim U(1, 500) \text{ without replacement.}$$

Error

$$[R_t] \stackrel{iid}{\sim} N(0, \exp(h_t))$$

$$h_t = 0.8 + 0.98(h_{t-1} - 2.5) + 0.2\epsilon_t$$

$$[\epsilon_t] \stackrel{iid}{\sim} N(0, 1)$$

B.5 Data Generating Scheme 5

Observation Equation

$$Y_t = T_t + S_t^{50} + R_t.$$

Trend

$$T_t = b_0 + b_1 \frac{t}{500} + b_2 \left(\frac{t}{500} \right)^2 + b_3 \left(\frac{t}{500} \right)^3,$$

$$[b_0] \sim U(-15, 15),$$

$$[b_1, \dots, b_4] \stackrel{iid}{\sim} N(0, 20^2).$$

Seasonal

$$S_t^{50} = \gamma_1^{50} \sin \left(\frac{2\pi t}{50} \right) + \gamma_2^{50} \cos \left(\frac{2\pi t}{50} \right),$$

$$[\gamma_1^{50}, \gamma_2^{50}] \stackrel{iid}{\sim} N(0, 5^2).$$

Error

$$[R_t] \stackrel{iid}{\sim} N(0, 1.5^2).$$