Bayesian Machine Learning for Forecasting Climate Policy Uncertainty

Donia Besher

Sorbonne University Abu Dhabi

December 15, 2024

Relevance of the Problem

- **Definition:** Unpredictability in climate-related legislative actions.

Relevance of the Problem

- **Definition:** Unpredictability in climate-related legislative actions.
- **Impacts:**
  - Climate policy uncertainty shocks lead to economic downturns.
  - Delayed investment in green technologies.
  - High uncertainty correlates with renewable energy market volatility.

Relevance of the Problem

- **Definition:** Unpredictability in climate-related legislative actions.
- **Impacts:**
  - Climate policy uncertainty shocks lead to economic downturns.
  - Delayed investment in green technologies.
  - High uncertainty correlates with renewable energy market volatility.
- **Objective:** Forecast climate policy uncertainty to provide insights for policymakers and industries.

SORBONNE UNIVERSITY ABU DHABI

Relevance of the Problem

- **Definition:** Unpredictability in climate-related legislative actions.
- **Impacts:**
  - Climate policy uncertainty shocks lead to economic downturns.
  - Delayed investment in green technologies.
  - High uncertainty correlates with renewable energy market volatility.
- **Objective:** Forecast climate policy uncertainty to provide insights for policymakers and industries.
- **Methodology:** Employ the BSTS model and explore the interplay between climate policy uncertainty and macroeconomic variables.

Relevance of the Problem

- **Definition:** Unpredictability in climate-related legislative actions.
- **Impacts:**
  - Climate policy uncertainty shocks lead to economic downturns.
  - Delayed investment in green technologies.
  - High uncertainty correlates with renewable energy market volatility.
- **Objective:** Forecast climate policy uncertainty to provide insights for policymakers and industries.
- **Methodology:** Employ the BSTS model and explore the interplay between climate policy uncertainty and macroeconomic variables.
- **Justification:**
  - Incorporates prior knowledge and updates with observed data.
  - Filters irrelevant variables for robust forecasts.

SORBONNE UNIVERSITY ABU DHABI

DETAILED OUTLINE

What is a Time Series?

Definition

A time series is denoted as $\{X_t : t = 1, 2, 3, \ldots, T\}$, where t represents time and X_t is the observed value at time t.

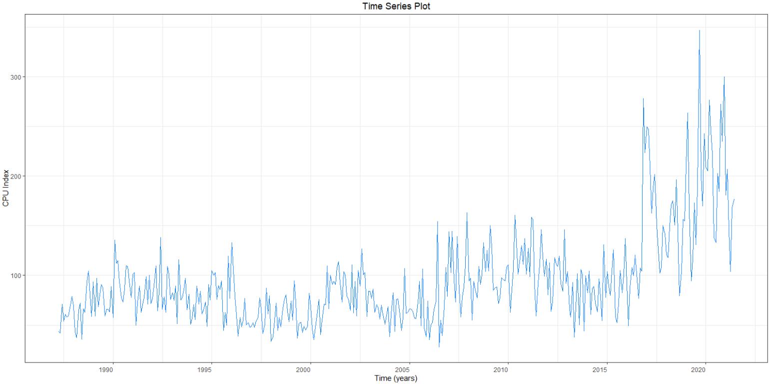


Figure: The time series plot of the climate policy uncertainty series.

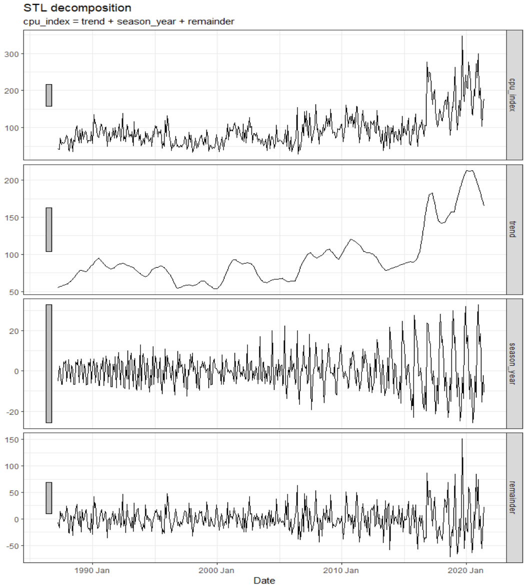
What are the components of a time series?

- **Trend** (T_t): Long-term progression (increase or decrease) in the data.
- **Seasonal** (S_t): Recurring patterns within fixed time periods.
- **Cyclic** (C_t): Irregular oscillations not confined to a specific period.
- **Noise** (R_t): Random variation or errors in the data.

What are the components of a time series?

- **Trend** (T_t): Long-term progression (increase or decrease) in the data.
- **Seasonal** (S_t): Recurring patterns within fixed time periods.
- **Cyclic** (C_t): Irregular oscillations not confined to a specific period.
- **Noise** (R_t): Random variation or errors in the data.

- **Decomposition**: $y_t = T_t + S_t + C_t + R_t$



What are the characteristics of a time series?

- **Stationarity:** Statistical properties such as mean and variance remain constant over time, which can be visually examined through the ACF and PACF plots.

What are the characteristics of a time series?

- **Stationarity:** Statistical properties such as mean and variance remain constant over time, which can be visually examined through the ACF and PACF plots.
- **Long-Range Dependence:** Indicates the persistence of correlation between observations across long time horizons. It is measured by the Hurst Exponent:

$$H = \begin{cases} 0.5 & \text{Random walk (no memory)} \\ H > 0.5 & \text{Persistence (long memory)} \\ H < 0.5 & \text{Anti-persistent (mean-reverting)} \end{cases}$$

DETAILED OUTLINE

SORBONNE UNIVERSITY ABU DHABI

Local Level Model

Definition

Let y_t denote observation t in a real-valued time series and μ_t denote the unobserved state.

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, \qquad & \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \mu_{t+1} &= \mu_t + u_t, \qquad & u_t \sim N(0, \sigma_u^2) \end{aligned} \tag{1}$$

Local Linear Trend Model

Definition

Let y_t denote observation t, μ_t denote the unobserved state and δ_t denote the slope.

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\
\mu_t &= \mu_{t-1} + \delta_{t-1} + u_t, & u_t &\sim N(0, \sigma_u^2) \\
\delta_t &= \delta_{t-1} + v_t, & v_t &\sim N(0, \sigma_v^2)
\end{aligned} \tag{2}
$$

SORBONNE UNIVERSITY ABU DHABI

Local Trend with Seasonality

Definition

Let y_t denote observation t, μ_t denote the local linear trend and τ_t denote the seasonal component (set of S dummy variables with coefficients constrained to have zero expectation over a full cycle of S seasons.).

$$\begin{aligned} y_t &= \mu_t + \tau_t + \varepsilon_t, && \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \\ \tau_t &= -\sum_{s=1}^{S-1} \tau_{t-s} + w_t, && w_t \sim N(0, \sigma_w^2) \end{aligned} \tag{3}$$

SORBONNE UNIVERSITY ABU DHABI

Local Trend with Seasonality and Regression

Definition

Let y_t denote observation t, μ_t denote the local linear trend, τ_t denote the seasonal component and $\beta_t^T x_t$ denote the regression component.

$$y_t = \mu_t + \tau_t + \beta_t^T x_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{4}$$

General Form

Definition

Let y_t denote observation t and α_t denote the vector of latent state variables.

$$\begin{aligned} y_t &= Z_t^T \alpha_t + \varepsilon_t, & \varepsilon_t &\sim N(0, H_t) \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, & \eta_t &\sim N(0, Q_t) \end{aligned} \tag{5}$$

where $\eta_t = (u_t, v_t, w_t)$ contains independent components of Gaussian random noise, T_t is a square transition matrix, R_t allows working with a full rank variance matrix Q_t (constant diagonal matrix with diagonal elements σ_u^2, σ_v^2 and σ_w^2) and we consider H_t to be a positive scalar σ_ε^2.

Spike and Slab Prior

Spike

Spike $p(\gamma)$: prior for probability of regression coefficients being zero

Spike and Slab Prior

Spike

Spike $p(\gamma)$: prior for probability of regression coefficients being zero

$$\gamma_k = \begin{cases} 1 & , \quad \beta_k \neq 0 \\ 0 & , \quad \beta_k = 0 \end{cases}$$

Spike and Slab Prior

Spike

Spike $p(\gamma)$: prior for probability of regression coefficients being zero

$$\gamma_k = \begin{cases} 1 & , \quad \beta_k \neq 0 \\ 0 & , \quad \beta_k = 0 \end{cases}$$

$$\gamma \sim \prod_{k=1}^{K} \pi_k^{\gamma_k} (1 - \pi_k)^{1-\gamma_k} \tag{6}$$

Spike and Slab Prior

Spike

Spike $p(\gamma)$: prior for probability of regression coefficients being zero

$$\gamma_k = \begin{cases} 1 & , \quad \beta_k \neq 0 \\ 0 & , \quad \beta_k = 0 \end{cases}$$

$$\gamma \sim \prod_{k=1}^{K} \pi_k^{\gamma_k} (1 - \pi_k)^{1-\gamma_k} \tag{6}$$

Equation (6) is simplified by assuming all the π_k are the same value $\pi = \frac{p}{K}$ where K is the dimension of x_t and p is the number of non-zero predictors.

Spike and Slab Prior

Slab

Slab: prior on the nonzero regression coefficients and variance

- Ω^{-1}: symmetric matrix
- Ω_{γ}^{-1}: submatrix corresponding to $\gamma_k = 1$

Spike and Slab Prior

Slab

Slab: prior on the nonzero regression coefficients and variance

- Ω^{-1}: symmetric matrix
- Ω_γ^{-1}: submatrix corresponding to $\gamma_k = 1$

Conjugate prior:

$$\beta_\gamma | \sigma_\varepsilon^2, \gamma \sim N(b_\gamma, \sigma_\varepsilon^2 (\Omega_\gamma^{-1})^{-1}) \quad \frac{1}{\sigma_\varepsilon^2} | \gamma \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{ss}{2}\right) \tag{7}$$

Spike and Slab Prior

Slab

Slab: prior on the nonzero regression coefficients and variance

- Ω^{-1}: symmetric matrix
- Ω_γ^{-1}: submatrix corresponding to $\gamma_k = 1$

Conjugate prior:

$$\beta_\gamma|\sigma_\varepsilon^2, \gamma \sim N(b_\gamma, \sigma_\varepsilon^2(\Omega_\gamma^{-1})^{-1}) \quad \frac{1}{\sigma_\varepsilon^2}|\gamma \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{ss}{2}\right) \tag{7}$$

Default values for equation (7):

- vector of prior means: $b = 0$
- prior sum of squares: ss expected R^2 from the regression
- prior sample size: ν number of observations worth of weight
- $\frac{ss}{\nu} = (1 - R^2)s_y^2$ where s_y^2 is the marginal standard deviation of the response
- $\Omega^{-1} = \frac{\kappa(\omega X^T X + (1-\omega)\text{diag}(X^T X))}{n}$, κ is the number of observations worth of weight on the prior mean b, $\omega = \frac{1}{2}$ and $\kappa = 1$

Spike and Slab Prior

Full Prior

$$p(\beta, \sigma_\varepsilon^2, \gamma) = p(\beta_\gamma | \sigma_\varepsilon^2, \gamma) p(\sigma_\varepsilon^2 | \gamma) p(\gamma) \tag{8}$$

Posterior Distribution

Conditional Posterior of β_γ and σ_ε^2 given γ

Let $y_t^* = y_t - {Z_t^*}^T \alpha_t$, where ${Z_t^*}^T$ is the observation matrix from equation (5) with $\beta^T x_t$ set to zero and $\mathbf{y}^* = y_{1:n}^*$.

Posterior Distribution

Conditional Posterior of β_γ and σ_ε^2 given γ

Let $y_t^* = y_t - Z_t^{*T}\alpha_t$, where Z_t^{*T} is the observation matrix from equation (5) with $\beta^T x_t$ set to zero and $\mathbf{y}^* = y_{1:n}^*$.

$$\beta_\gamma|\sigma_\varepsilon^2, \gamma, \mathbf{y}^* \sim N\left(\tilde{\beta}_\gamma, \sigma_\varepsilon^2 (V_\gamma^{-1})^{-1}\right), \quad \frac{1}{\sigma_\varepsilon^2}|\gamma, \mathbf{y}^* \sim \text{Gamma}\left(\frac{N}{2}, \frac{SS_\gamma}{2}\right) \tag{9}$$

Posterior Distribution

Conditional Posterior of β_γ and σ_ε^2 given γ

Let $y_t^* = y_t - Z_t^{*T} \alpha_t$, where Z_t^{*T} is the observation matrix from equation (5) with $\beta^T x_t$ set to zero and $\mathbf{y}^* = y_{1:n}^*$.

$$\beta_\gamma | \sigma_\varepsilon^2, \gamma, \mathbf{y}^* \sim N\left(\tilde{\beta}_\gamma, \sigma_\varepsilon^2 (V_\gamma^{-1})^{-1}\right), \quad \frac{1}{\sigma_\varepsilon^2} | \gamma, \mathbf{y}^* \sim \text{Gamma}\left(\frac{N}{2}, \frac{SS_\gamma}{2}\right) \tag{9}$$

where the sufficient statistics can be written:

- $V_\gamma^{-1} = (X^T X)_\gamma + \Omega_\gamma^{-1}$
- $\tilde{\beta}_\gamma = (V_\gamma^{-1})^{-1} (X_\gamma^T \mathbf{y}^* + \Omega_\gamma^{-1} b_\gamma)$
- $N = \nu + n$
- $SS_\gamma = ss + \mathbf{y}^{*T} \mathbf{y}^* + b_\gamma^T \Omega_\gamma^{-1} b_\gamma - \tilde{\beta}_\gamma^T V_\gamma^{-1} \tilde{\beta}_\gamma$

Posterior Distribution

The marginal posterior of γ

By conjugacy, we can marginalize over β_γ and σ_ε^2 to obtain:

Posterior Distribution

The marginal posterior of γ

By conjugacy, we can marginalize over β_γ and σ_ε^2 to obtain:

$$\gamma|\mathbf{y}^* \sim C(\mathbf{y}^*)\frac{\left|\Omega_\gamma^{-1}\right|^{\frac{1}{2}} p(\gamma)}{\left|V_\gamma^{-1}\right|^{\frac{1}{2}} SS_\gamma^{\frac{N}{2}-1}} \tag{10}$$

Posterior Distribution

The marginal posterior of γ

By conjugacy, we can marginalize over β_γ and σ_ε^2 to obtain:

$$\gamma|\mathbf{y}^* \sim C(\mathbf{y}^*) \frac{\left|\Omega_\gamma^{-1}\right|^{\frac{1}{2}} p(\gamma)}{\left|V_\gamma^{-1}\right|^{\frac{1}{2}} SS_\gamma^{\frac{N}{2}-1}} \tag{10}$$

where $C(\mathbf{y}^*)$ is a normalizing constant that depends on $\mathbf{y}^*$ but not on γ, MCMC does not require it to be computed explicitly.

Markov Chain Monte Carlo

Key Concepts in Markov Chain Monte Carlo

- Class of algorithms used for sampling from a probability distribution based on constructing a Markov chain that has the desired distribution as its stationary distribution.

Markov Chain Monte Carlo

Key Concepts in Markov Chain Monte Carlo

- Class of algorithms used for sampling from a probability distribution based on constructing a Markov chain that has the desired distribution as its stationary distribution.
- Common MCMC Algorithms:
  - Metropolis-Hastings Algorithm: generates a candidate sample and accepts or rejects it based on a calculated acceptance probability. The acceptance probability ensures that the stationary distribution of the Markov chain is the target distribution.
  - Gibbs Sampling: sampling is performed for each variable conditional on the current values of all other variables, useful for high-dimensional distributions that can be broken down into simpler conditional distributions.

SORBONNE UNIVERSITY ABU DHABI

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b|y_{1:T}) \quad \forall b \in \{1,\ldots,B\}$

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b|y_{1:T}) \quad \forall b \in \{1,\ldots,B\}$
2. **Iterative Sampling:**

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b|y_{1:T}) \quad \forall b \in \{1,\ldots,B\}$
2. **Iterative Sampling:**
   - $\forall i \, \forall b, \quad$ draw $\theta_b^c \sim q_b(\theta_b|y_{1:T})$

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b | y_{1:T}) \quad \forall b \in \{1, \ldots, B\}$
2. **Iterative Sampling:**
   - $\forall i \, \forall b, \quad$ draw $\theta_b^c \sim q_b(\theta_b | y_{1:T})$
   - Calculate the acceptance ratio

$$r = \frac{p_b^*(\theta_b^c | \theta_1^{(i)}, \ldots, \theta_{b-1}^{(i)}, \theta_{b+1}^{(i-1)}, \ldots, \theta_B^{(i-1)}, y_{1:T}) \times q_b(\theta_b^{(i-1)} | y_{1:T})}{p_b^*(\theta_{(b}^{(i-1)} | \theta_1^{(i)}, \ldots, \theta_{b-1}^{(i)}, \theta_{b+1}^{(i-1)}, \ldots, \theta_B^{(i-1)}, y_{1:T}) \times q_b(\theta_b^c | y_{1:T})}$$

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b|y_{1:T}) \quad \forall b \in \{1,\ldots,B\}$
2. **Iterative Sampling:**
   - $\forall i\, \forall b, \quad$ draw $\theta_b^c \sim q_b(\theta_b|y_{1:T})$
   - Calculate the acceptance ratio

$$r = \frac{p_b^*(\theta_b^c|\theta_1^{(i)},\ldots,\theta_{b-1}^{(i)},\theta_{b+1}^{(i-1)},\ldots,\theta_B^{(i-1)},y_{1:T}) \times q_b(\theta_b^{(i-1)}|y_{1:T})}{p_b^*(\theta_{(b}^{(i-1)}|\theta_1^{(i)},\ldots,\theta_{b-1}^{(i)},\theta_{b+1}^{(i-1)},\ldots,\theta_B^{(i-1)},y_{1:T}) \times q_b(\theta_b^c|y_{1:T})}$$

   - Accept the new value θ_c with probability $\min(1,r)$, otherwise retain the current value $\theta_b^{(i)}$.

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b|y_{1:T}) \quad \forall b \in \{1,\ldots,B\}$

2. **Iterative Sampling:**
   - $\forall i \, \forall b, \quad$ draw $\theta_b^c \sim q_b(\theta_b|y_{1:T})$
   - Calculate the acceptance ratio

$$r = \frac{p_b^*(\theta_b^c|\theta_1^{(i)},\ldots,\theta_{b-1}^{(i)},\theta_{b+1}^{(i-1)},\ldots,\theta_B^{(i-1)},y_{1:T}) \times q_b(\theta_b^{(i-1)}|y_{1:T})}{p_b^*(\theta_{(b}^{(i-1)}|\theta_1^{(i)},\ldots,\theta_{b-1}^{(i)},\theta_{b+1}^{(i-1)},\ldots,\theta_B^{(i-1)},y_{1:T}) \times q_b(\theta_b^c|y_{1:T})}$$

   - Accept the new value θ_c with probability $\min(1,r)$, otherwise retain the current value $\theta_b^{(i)}$.

3. **Burn-in Period:** Discard the initial portion of the chain.

Steps of Metropolis-Hastings

1. **Initialization:** Specify an initial value $\theta^{(0)}$ and partition the parameter set into B mutually exclusive blocks, and a proposal distribution $q_b(\theta_b|y_{1:T}) \quad \forall b \in \{1,\ldots,B\}$

2. **Iterative Sampling:**
   - $\forall i \, \forall b, \quad$ draw $\theta_b^c \sim q_b(\theta_b|y_{1:T})$
   - Calculate the acceptance ratio

$$r = \frac{p_b^*(\theta_b^c|\theta_1^{(i)},\ldots,\theta_{b-1}^{(i)},\theta_{b+1}^{(i-1)},\ldots,\theta_B^{(i-1)},y_{1:T}) \times q_b(\theta_b^{(i-1)}|y_{1:T})}{p_b^*(\theta_{(b}^{(i-1)}|\theta_1^{(i)},\ldots,\theta_{b-1}^{(i)},\theta_{b+1}^{(i-1)},\ldots,\theta_B^{(i-1)},y_{1:T}) \times q_b(\theta_b^c|y_{1:T})}$$

   - Accept the new value θ_c with probability $\min(1,r)$, otherwise retain the current value $\theta_b^{(i)}$.

3. **Burn-in Period:** Discard the initial portion of the chain.

4. **Convergence:**Ensure that the markov chain has converged to the target distribution $p(\theta|y_{1:T})$.

DETAILED OUTLINE

Autoregressive Integrated Moving Average

The ARIMA model is a widely used statistical method for time series forecasting consisting of three key components:

- **Autoregression (AR):** $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \xi_t$
- **Differencing (I):** $B^d(y_t) = y_{t-d}$
- **Moving Average (MA):** $y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$

$$\left(1 - \sum_{i=1}^{p} \phi_i B^i\right) (1-B)^d y_t = \left(1 + \sum_{j=1}^{q} \theta_j B^j\right) \varepsilon_t \tag{11}$$

where p is the number of AR terms, d is the order of differencing, q is the number of MA terms, and ξ_t represents white noise.

Autoregressive Fractionally Integrated Moving Average

ARFIMA Model

An extension of the ARIMA framework by incorporating fractional differencing, making it suitable for time series with long memory characterized by slowly decaying autocorrelations.

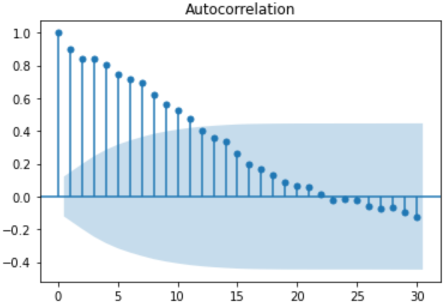


Illustration of Long Memory

$$\left(1-\sum_{i=1}^{p}\phi_i B^i\right)(1-B)^d y_t = \left(1+\sum_{j=1}^{q}\theta_j B^j\right)\varepsilon_t, \quad d \in (0, 0.5)$$

Fractional Differencing

Fractional differencing series ($p = q = 0$):

$$(1-B)^d y_t = \left(\sum_{j\geq 0} \pi_j B^j\right) y_t = \left(\sum_{j\geq 0} \pi_j y_{t-j}\right) = w_t$$

Fractional Differencing

Fractional differencing series ($p = q = 0$):

$$(1-B)^d y_t = \left(\sum_{j\geq 0} \pi_j B^j\right) y_t = \left(\sum_{j\geq 0} \pi_j y_{t-j}\right) = w_t$$

Taylor's Expansion:

$$(1-B)^d = 1 - dB + \frac{d(d-1)}{2} B^2 + o(B^2) = \left(\sum_{j\geq 0} \pi_j B^j\right); \quad \pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$$

Fractional Differencing

Fractional differencing series ($p = q = 0$):

$$(1-B)^d y_t = \left(\sum_{j\geq 0} \pi_j B^j\right) y_t = \left(\sum_{j\geq 0} \pi_j y_{t-j}\right) = w_t$$

Taylor's Expansion:

$$(1-B)^d = 1 - dB + \frac{d(d-1)}{2}B^2 + o(B^2) = \left(\sum_{j\geq 0} \pi_j B^j\right); \quad \pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$$

π_j can be represented recursively as:

$$\pi_{j+1} = \frac{(j-d)\pi_j}{j+1}, \quad \pi_0 = 1.$$

Fractional Differencing

Fractional differencing series ($p = q = 0$):

$$(1-B)^d y_t = \left(\sum_{j\geq 0} \pi_j B^j\right) y_t = \left(\sum_{j\geq 0} \pi_j y_{t-j}\right) = w_t$$

Taylor's Expansion:

$$(1-B)^d = 1 - dB + \frac{d(d-1)}{2} B^2 + o(B^2) = \left(\sum_{j\geq 0} \pi_j B^j\right); \quad \pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$$

π_j can be represented recursively as:

$$\pi_{j+1} = \frac{(j-d)\pi_j}{j+1}, \quad \pi_0 = 1.$$

The fractional differencing parameter d is estimated by minimizing the residual sum of squares: $\sum w_t^2(d)$, $w_t = (1-B)^d y_t$, ensuring that d is data-dependent and adjusting for the long memory.

DETAILED OUTLINE

SORBONNE UNIVERSITY ABU DHABI

Characteristics of the Time Series

The global characteristics of the train (April 1987 to June 2021) climate policy uncertainly series:

Skewness	Kurtosis	Linearity	Seasonality	Stationarity	Long Range Dependence
1.863	4.531	Nonlinear	Seasonal	Nonstationary	0.780

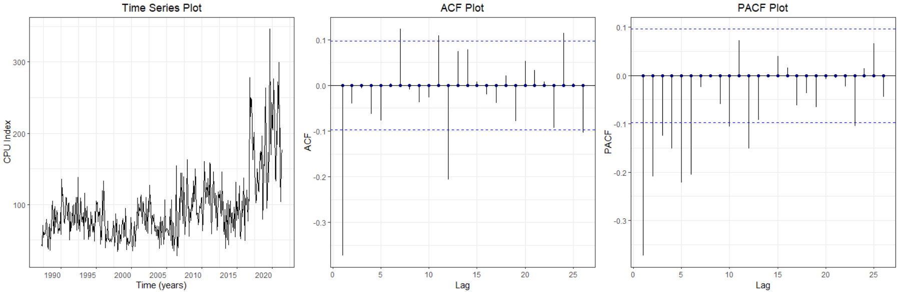


Figure: The upward trend indicates increased climate-related policy uncertainty.

Statistical Significance of the Forecasts: MCB

Multiple Comparisons with the Best: A non-parametric test that ranks forecasters by accuracy and identifies the one with the lowest average rank as the best model.

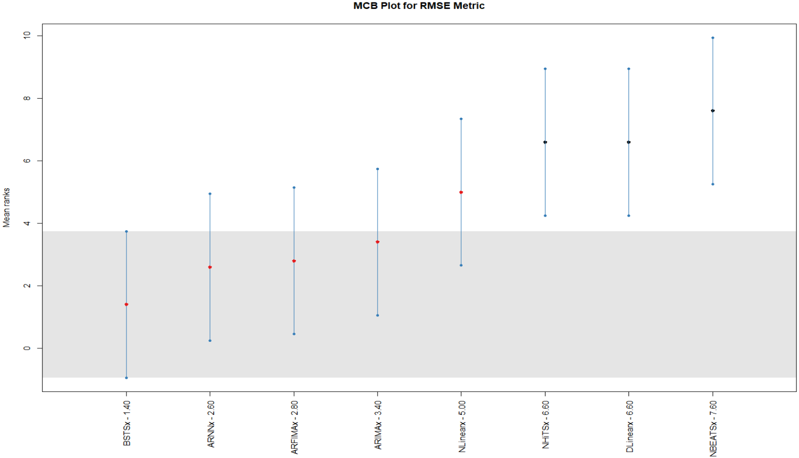


Figure: BSTS achieves superior accuracy due to dynamic adaptation.

Statistical Significance of the Forecasts: MCB

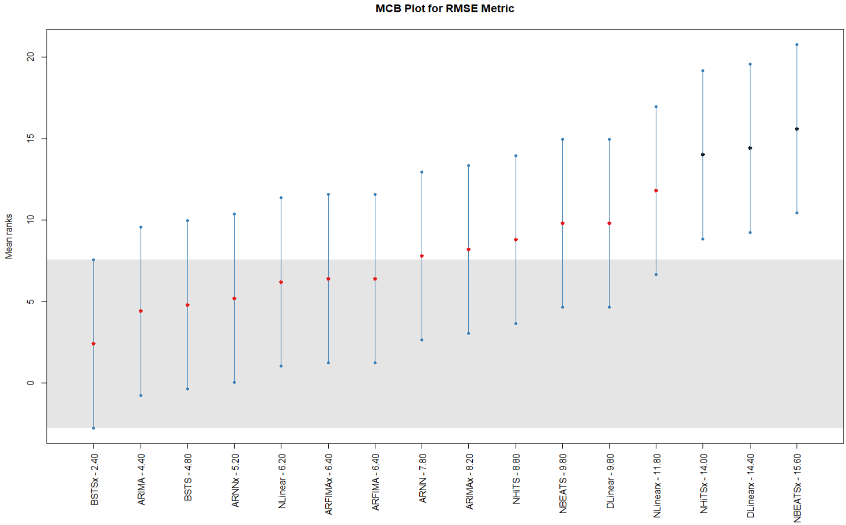


Figure: Inclusion of covariates improves the BSTS model performance.

Significant Macroeconomic Variables

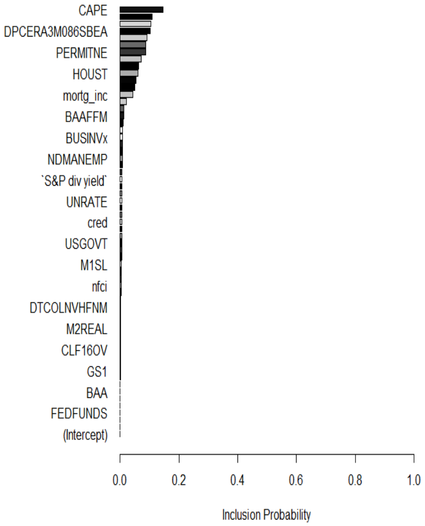
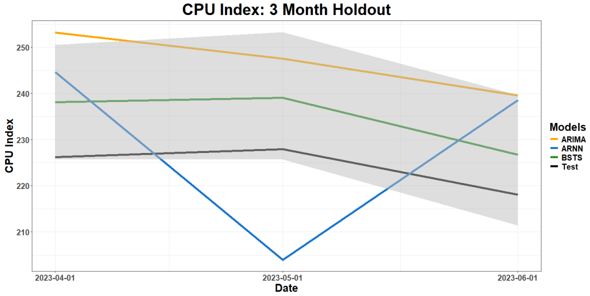


Figure: The inclusion probabilities of the covariates.

- **CAPE:** Cyclically Adjusted Price-to-Earnings ratio, a stock market valuation metric.
- **BCI:** Business Conditions Index, a measure of economic activity.
- **CLI:** Composite Leading Indicator, used to predict economic turning points in business cycles.
- **PERMITNE:** Housing permits in the Northeastern U.S., indicating housing market activity.
- **UEMP15OV:** Unemployment rate for individuals unemployed for 15 weeks or longer.

Credible Intervals

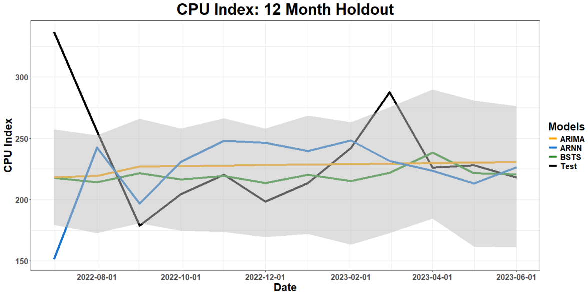


BSTS Forecasts Provide Credible Intervals:

Quantifying uncertainty helps policymakers understand potential risks and confidence levels in forecasting.

BSTS Better Captures Fluctuations:

The BSTS model surpasses ARIMA and ARNN in modeling volatility over long-term horizons and accurately capturing critical trends in the climate policy uncertainty series.



DETAILED OUTLINE

SORBONNE UNIVERSITY ABU DHABI

References

[1] Yang, J., Dong, D., & Liang, C. (2024). Climate policy uncertainty and the U.S. economic cycle. *Technological Forecasting & Social Change*, **202**, 123344.

[2] Liang, C., Umar, M., Ma, F., & Huynh, T. L. D. (2022). Climate policy uncertainty and world renewable energy index volatility forecasting. *Technological Forecasting & Social Change*, **182**, 121810.

[3] Raza, S. A., Khan, K. A., Benkraiem, R., & Guesmi, K. (2023). The importance of climate policy uncertainty in forecasting the green, clean, and sustainable financial markets volatility. *Journal of Sustainable Finance & Investment*. `https://doi.org/10.1016/j.irfa.2023.102984`.

[4] Newell, R. G., Prest, B. C., & Sexton, S. E. (2021). The GDP-temperature relationship: implications for climate change damages. *Journal of Environmental Economics and Management*, **108**, 102445. `https://doi.org/10.1016/j.jeem.2021.102445`.

[5] Akan, T. (2024). The impact of monetary policy on climate change through the mediation of sectoral renewable energy consumption. *Energy Policy*, **192**, 114244. `https://doi.org/10.1016/j.enpol.2024.114244`.

[6] Scott, S. L., & Varian, H. R. (2014). Predicting the Present with Bayesian Structural Time Series. *International Journal of Mathematical Modelling and Numerical Optimisation*, **5**(1-2), 4–23. `https://doi.org/10.1504/IJMMNO.2014.059942`.

[7] Varian, H. R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, **28**(2), 3–28. `https://doi.org/10.1257/jep.28.2.3`.

[8] Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts.

[9] Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, **65**(332), 1509–1526.

[10] Granger, C. W. J., & Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, **1**(1), 15–29.

[11] Koning, A. J., Franses, P. H., Hibon, M., & Stekler, H. O. (2005). The M3 competition: Statistical tests of the results. *International Journal of Forecasting*, *21*(3), 397–409.

SORBONNE UNIVERSITY ABU DHABI