This document provides a comprehensive comparison and detailed mathematical formulation for three Bayesian regime-switching models: `RS.R`, `BMA_corrected.R`, and `BHS_TP_corrected.R`. These models simulate time series data with regime switching and employ Bayesian hierarchical methods to estimate regime-specific parameters. Key differences among these models arise in their approach to model averaging, the inclusion of dynamic weights for each regime, and their diagnostic performance based on leave-one-out cross-validation (LOO).

1 Data Simulation Process

Each model generates synthetic time series data for $N = 10$ individuals over $T = 50$ time points, with observations generated from one of four regimes. Each regime has unique mean μ_k, standard deviation σ_k, and transition probabilities that govern regime changes over time.

1.1 Regime Transitions

For each individual i and time step t, a regime indicator $r_{i,t} \in \{1, 2, 3, 4\}$ is used to track the active regime. Transitions between regimes follow a predefined transition matrix $\mathbf{T}$, with T_{ij} denoting the probability of transitioning from regime i to regime j:

$$\mathbf{T} = \begin{bmatrix} 0.8 & 0.05 & 0.1 & 0.05 \\ 0.1 & 0.8 & 0.05 & 0.05 \\ 0.1 & 0.05 & 0.8 & 0.05 \\ 0.1 & 0.05 & 0.1 & 0.75 \end{bmatrix}. \tag{1}$$

1.2 Observation Generation for Each Regime

At each time step, the observation $y_{i,t}$ for individual i depends on the current regime $r_{i,t}$. Each regime has its own mean μ_k and standard deviation σ_k, and observations are adjusted by autoregressive (AR) or moving-average (MA) components as follows:

- Define the target mean $\mu^*_{i,t}$, incorporating smoothing based on the previous observation $y_{i,t-1}$:

  $$\mu^*_{i,t} = \lambda \mu_{r_{i,t}} + (1 - \lambda) y_{i,t-1}, \tag{2}$$

  where λ is a smoothing parameter for each regime.

- Generate $y_{i,t}$ according to the current regime $r_{i,t}$:

$$y_{i,t} \sim \mathcal{N}((1 - \varphi) \cdot \mu^*_{i,t} + \varphi \cdot y_{i,t-1}, \sigma_1), \quad \text{if } r_{i,t} = 1, \tag{3}$$
$$y_{i,t} \sim \mathcal{N}((1 - \psi) \cdot \mu^*_{i,t} + \psi \cdot y_{i,t-1}), \sigma_2), \quad \text{if } r_{i,t} = 2, \tag{4}$$
$$y_{i,t} \sim \mathcal{N}(\mu^*_{i,t} + \theta \cdot (y_{i,t-1} - \mu_{i,t-1}), \sigma_3), \quad \text{if } r_{i,t} = 3, \tag{5}$$
$$y_{i,t} \sim \mathcal{N}(\mu^*_{i,t}, \sigma_4), \quad \text{if } r_{i,t} = 4. \tag{6}$$

Here, φ, ψ, and θ are the AR/MA coefficients specific to regimes 1, 2, and 3, respectively. Regime 4 has no AR or MA component.

2 Hierarchical Bayesian Models

The models `RS.R`, `BMA_corrected.R`, and `BHS_TP_corrected.R` estimate the parameters for each regime, including means μ_k, standard deviations σ_k, and AR/MA coefficients. They differ in how they handle model averaging and regime weighting.

2.1 `RS.R` - Regime Switching Model

This model uses a single regime-switching approach without explicit model averaging across regimes. For each observation, $y_{i,t}$ is generated directly based on the current regime. The model's priors are:

$$\mu_k \sim \mathcal{N}(0, 5), \tag{7}$$

$$\sigma_k \sim \text{Cauchy}(0, 2), \tag{8}$$

$$\lambda \sim \text{Beta}(2, 2), \tag{9}$$

$$\varphi \sim \mathcal{N}(0.7, 0.3), \quad \psi \sim \mathcal{N}(-0.7, 0.3), \quad \theta \sim \mathcal{N}(0, 0.5). \tag{10}$$

2.2 `BMA_corrected.R` - Bayesian Model Averaging

The `BMA_corrected.R` model introduces fixed Bayesian model averaging by assigning static weights w_j across regimes, applied to observations at each time step:

$$y_{i,t} \sim \mathcal{N}\left(\sum_{j=1}^{4} w_j y_{i,t}^{(j)}, \sigma_{r_{i,t}}\right), \tag{11}$$

where $y_{i,t}^{(j)}$ is the regime-specific prediction and $\sum_{j=1}^{4} w_j = 1$. The weights w_j are drawn from a Dirichlet prior.

2.3 `BHS_TP_corrected.R` - Bayesian Hierarchical Stacking with Transition Probabilities

In the `BHS_TP_corrected.R` model, a dynamic weighting mechanism is applied, where weights $\mathbf{w}_t$ adjust at each time step according to transition probabilities:

$$\mathbf{w}_t = \mathbf{T}\mathbf{w}_{t-1}. \tag{12}$$

The transition matrix $\mathbf{T}$ follows a Dirichlet prior with concentration parameters that encourage diagonal stability. The predictive distribution for $y_{i,t}$ becomes:

$$y_{i,t} \sim \mathcal{N}\left(\sum_{j=1}^{4} w_t^{(j)} y_{i,t}^{(j)}, \sigma_{r_{i,t}}\right). \tag{13}$$

3 Results and Visualizations

3.1 Time Series with Regime Switching for Each Individual

The following figure displays time series data for each of the 10 individuals, highlighting regime switching over 50 time points. Each color corresponds to a different regime:

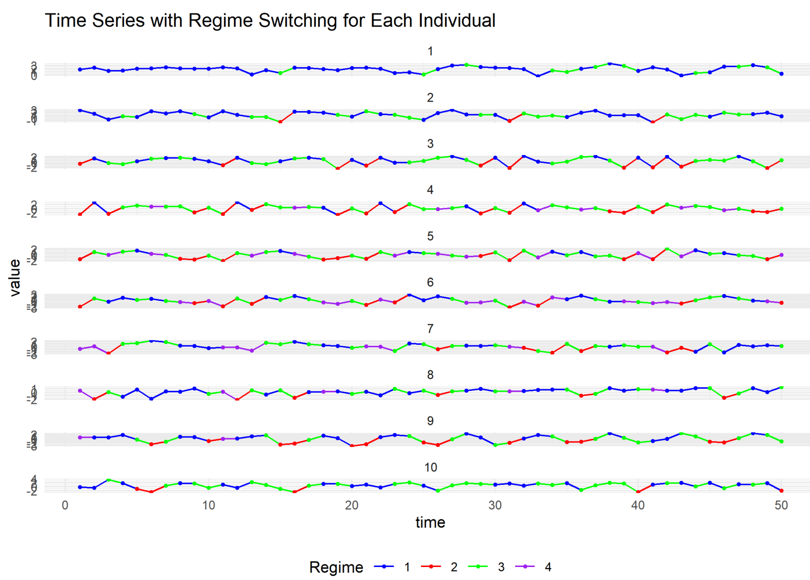

Figure 1: Time series with regime switching for each individual. Each line represents an individual's time series data, with colors indicating different regimes.

3.2 Frequency of Each Regime

This figure shows the frequency of each regime across all time points and individuals. It provides insight into the prevalence of each regime in the simulated data:

4 Model Comparison with Leave-One-Out Cross-Validation (LOO)

In terms of LOOIC, the `RS.R` model demonstrates superior predictive accuracy with the lowest LOOIC value (802.4) among the models. This suggests that `RS.R` provides a better out-of-sample predictive fit. However, `BHS_TP_corrected.R` with dynamic weights (especially in set 1) offers greater flexibility and may

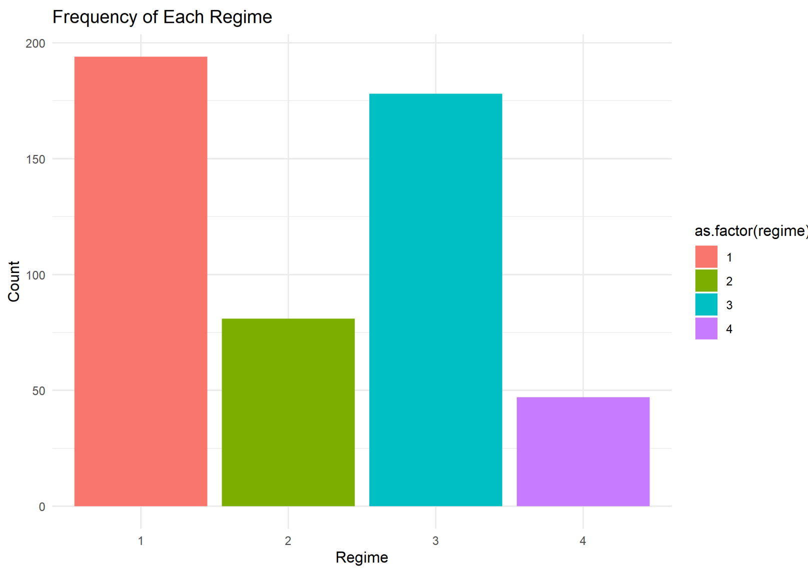

Figure 2: Frequency distribution of each regime across all individuals and time points.

Table 1: Summary of LOO Metrics for Each Model

Model	elpd_loo	SE (elpd_loo)	p_loo	LOOIC
`RS.R`	−401.2	16.2	10.8	802.4
`BMA_corrected.R`	−557.1	29.2	5.5	1114.1
`BHS_TP_corrected.R` (set 1)	−557.4	29.5	7.6	1114.8
`BHS_TP_corrected.R` (set 2)	−558.0	29.6	8.7	1116.0

better accommodate more complex, non-stationary data scenarios where regime-switching models benefit from adaptive weighting across regimes.

5 Pareto-k Diagnostics

For all three models, the majority of Pareto-k estimates fall in the "good" range ($k < 0.5$), indicating that the diagnostics are stable across observations. Specifically: `-` Both `RS.R` and `BMA_corrected.R` show 100% of their Pareto-k values in the "good" range, suggesting very stable diagnostics. `- BHS_TP_corrected.R` shows slight variation between set 1 and set 2: `-` In `set 1`, 99.4% of Pareto-k values are "good," with a small proportion (0.6%) in the "bad" range ($0.7 \leq k < 1$). `-` In `set 2`, all Pareto-k values are in the "good" range ($k < 0.5$), indicating optimal stability.

Table 2: Pareto-k Diagnostic Values

Model	**Good** ($k < 0.5$)	**OK** ($0.5 \leq k < 0.7$)	**Bad** ($0.7 \leq k < 1$)	**Very Bad** (k
`RS.R`	490 (100%)	0	0	0
`BMA_corrected.R`	490 (100%)	0	0	0
`BHS_TP_corrected.R` (set 1)	487 (99.4%)	0	3 (0.6%)	0
`BHS_TP_corrected.R` (set 2)	490 (100%)	0	0	0