Rolling Robust Regime Detection and Explaining Transitions through Decision Trees

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Yiting Liu [1,2]

Under the supervision of Prof. Dr. Ali Hirsa and Miao Wang

[1]Bern University of Applied Sciences [2]University of Twente

R2-RD: Robust Rolling Regime Detection

Identifying evolving macroeconomic regimes is critical for understanding structural shifts in financial systems. Traditional regime detection methods, such as fixed-state HMMs, assume a static number of latent states. **R2-RD** (Robust Rolling Regime Detection) lifts this constraint by dynamically updating both model parameters and regime structure over time.

Design Overview:

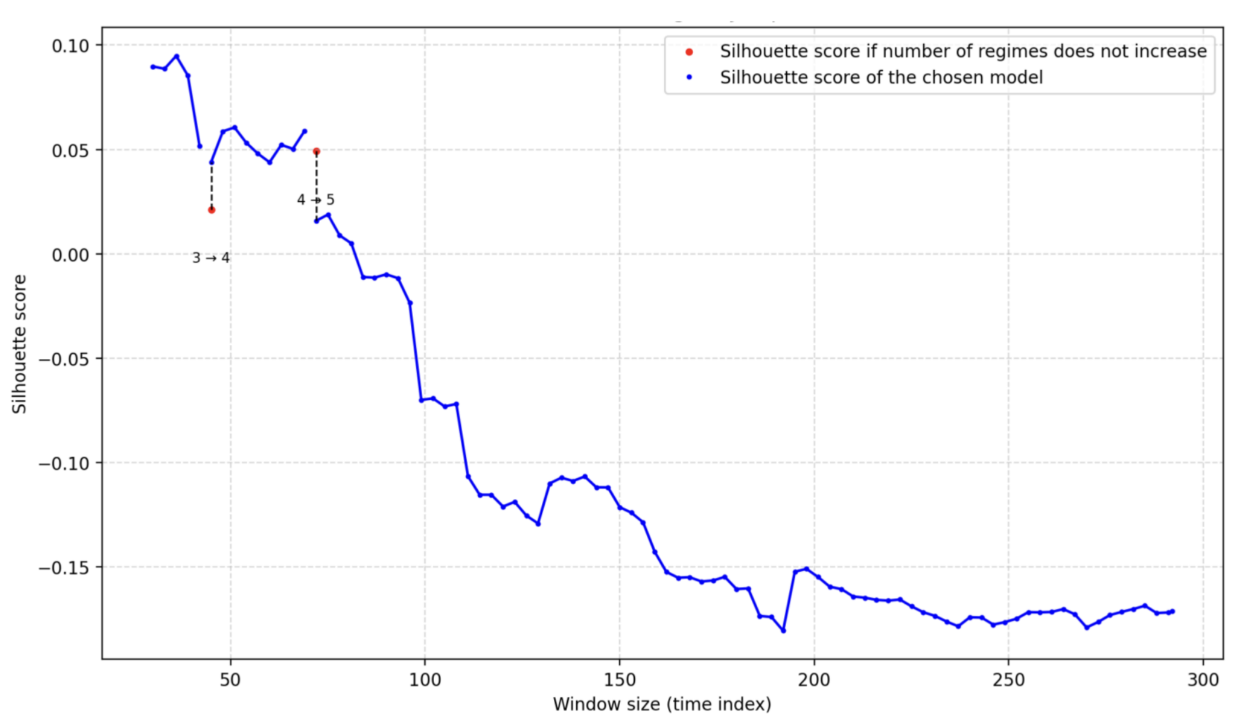
- **Rolling HMM retraining**: At each time step, a Gaussian HMM is re-estimated on a growing window of recent principal components.
- **Model pairwise comparison**: Two models are fitted at every step:
  - Model $\mathcal{I}$ with current regime count K
  - Model $\mathcal{J}$ with $K+1$ regimes (candidate expansion)
- **Label alignment problem**: Since HMM regimes are unordered, we use the *Hungarian algorithm* to align new regime labels with prior assignments, minimizing assignment cost c.
- **Expansion rule**:

  $$\text{Accept model } \mathcal{J} \text{ if } c_0 > c_1 \quad \text{and} \quad c_0 > \bar{c}$$

  where c_0 and c_1 are label assignment costs for $\mathcal{I}$ and $\mathcal{J}$ respectively, and $\bar{c}$ is the 80th percentile of historical c values.
- **Stabilization**: Once a new regime is accepted, we reset the counter and allow the system to stabilize before considering further expansion.

Compared to standard HMMs, R2-RD introduces:

- Adaptive regime complexity
- Robust state alignment across time
- A principled expansion policy guided by both fit and historical separation cost



R2-PCA and Regime Visualization

We apply a tailored dimensionality reduction method, **R2-PCA**, to extract key latent components from macroeconomic indicators. This method stabilizes the feature space used for regime detection while enhancing visualization clarity.

PC Trajectories with Regime Overlay: The first figure shows PC1–PC5 time series with background colors denoting detected regimes. Transitions often coincide with directional shifts or volatility bursts in PC1.

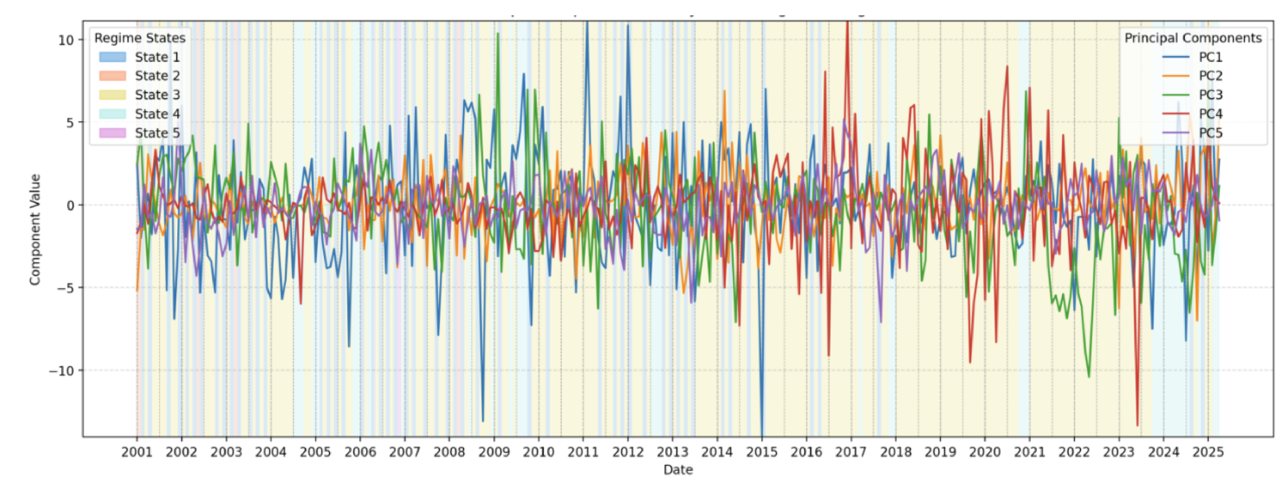


Figure 1. *

PC1–PC3 series with regime-colored background.

PC1–PC2 Regime Distributions: The second plot displays all time points in the PC1–PC2 plane. Gaussian contours fitted per regime illustrate spatial separation and compactness.

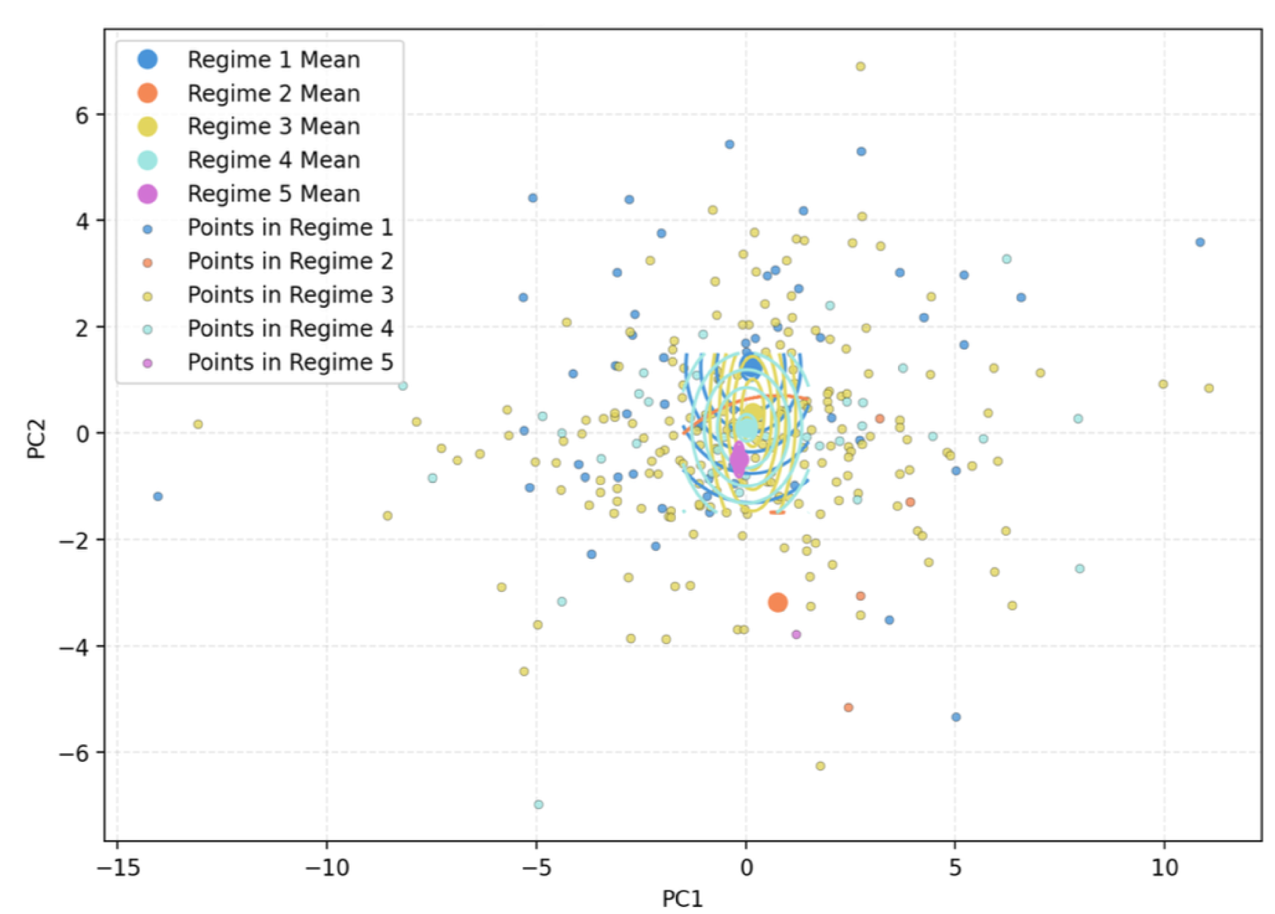


Figure 2. *

Regime clusters in PC1–PC2 space via 2D Gaussian contours.

Interpreting Regime Transitions with Decision Trees

To analyze how regime transitions unfold, we restructure the dataset to reflect regime-to-regime dynamics. For each time point t, we extract the economic indicators as features, and use the regime label at $t+1$ as the prediction target. Samples are grouped by their regime at time t, so each decision tree is trained on data from a specific starting regime. This produces multiple regime-specific trees that learn how macroeconomic signals drive transitions. As shown in the figure below, the resulting tree provides interpretable rules mapping current indicators to next-period regime outcomes.

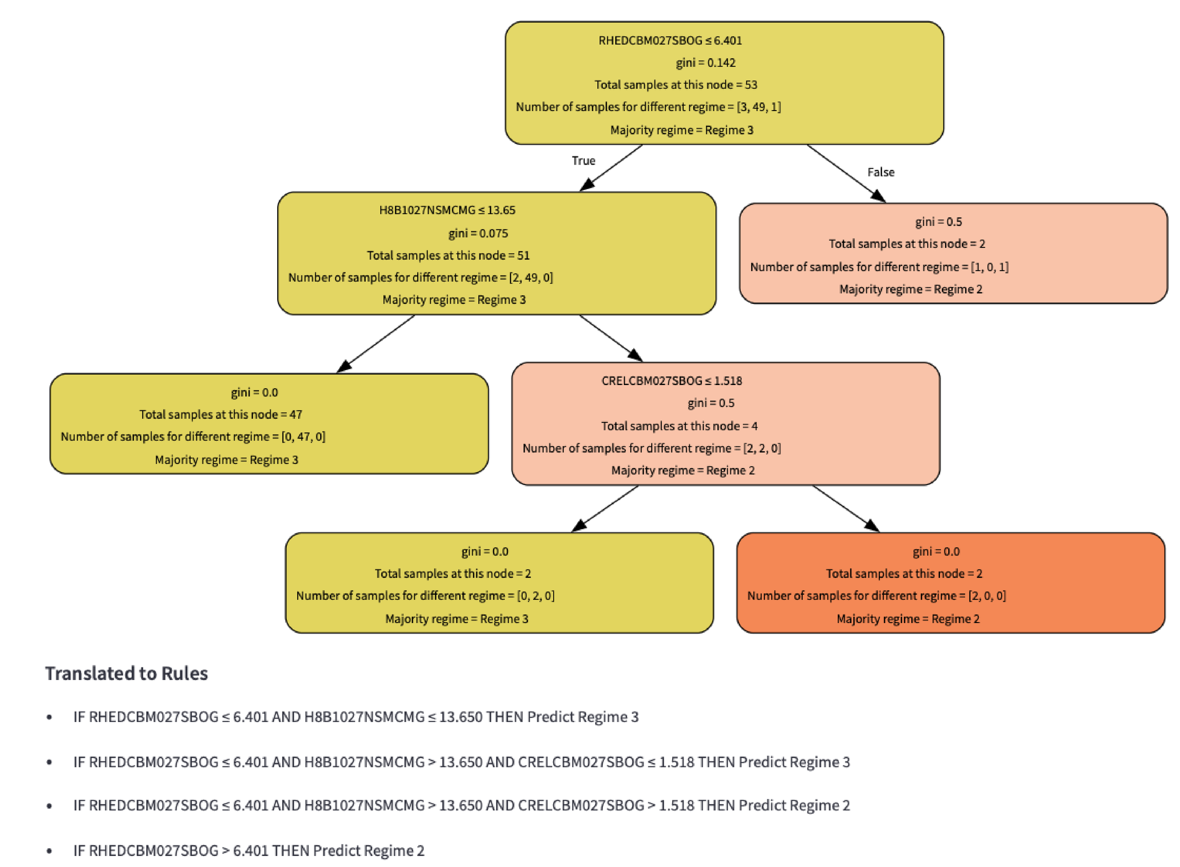


Figure 3. *

Decision tree predicting next regime for samples starting in Regime 1.

To further interpret model behavior, we compute SHAP values for each split node. This highlights the most influential features driving transitions (e.g., interest rates, policy shocks) and reveals how decision boundaries differ across regimes.

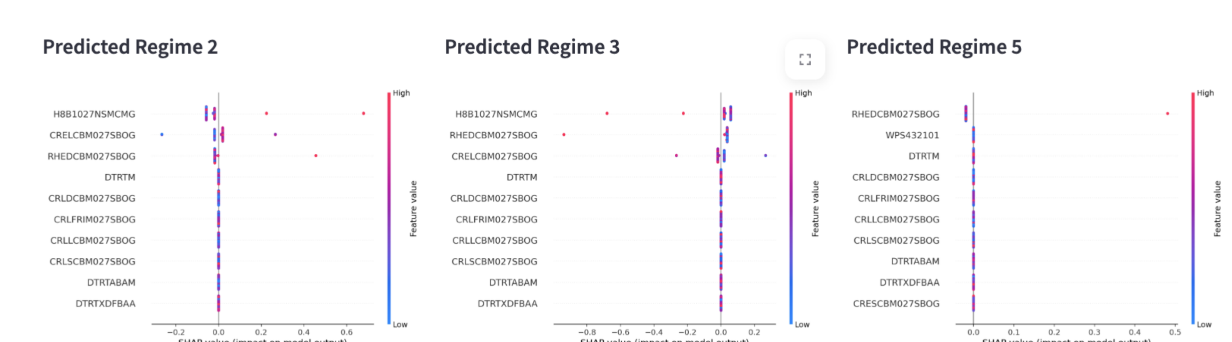


Figure 4. *

SHAP summary plot: variable contributions to regime prediction.

Decision Tree with Soft Boundary: A Multi-Perspective Interpretation

To further interpret regime-switching patterns, we analyze the decision trees along three dimensions, focusing on boundary sharpness, feature robustness, and soft generalization.

1. Feature Value vs. Split Threshold: We plot individual samples along the feature axis, marking tree thresholds to reveal whether decisions are near the boundary or well-separated.

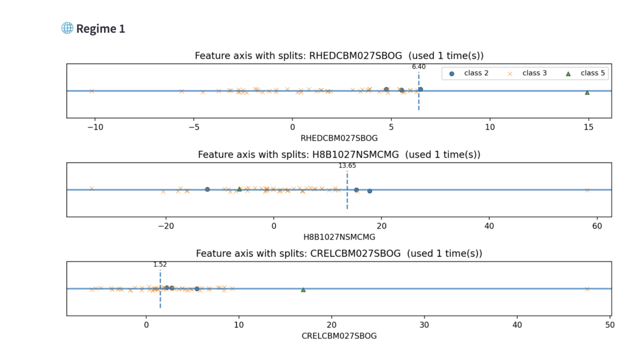


Figure 5. *

Samples in Regime 1 vs. splitting thresholds.

2. Feature Perturbation Sensitivity: For each important variable, we introduce local noise and measure prediction drop. A steeper accuracy drop indicates higher dependence on that variable.

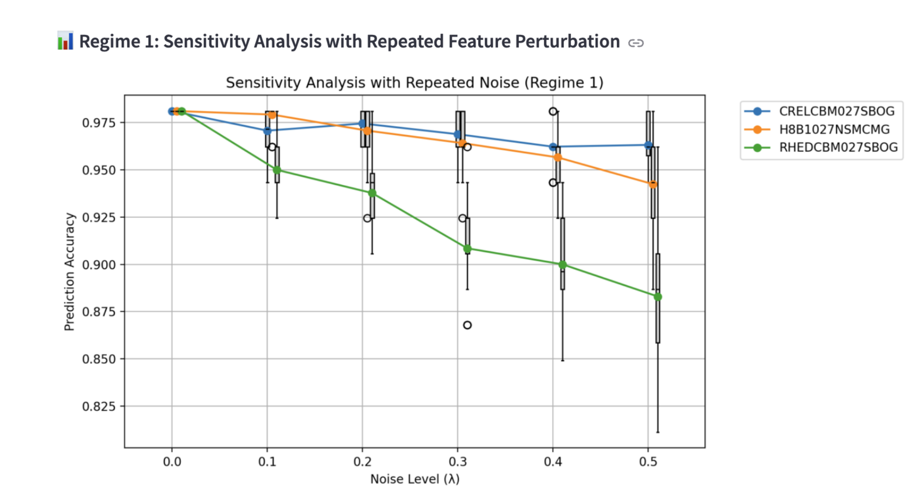
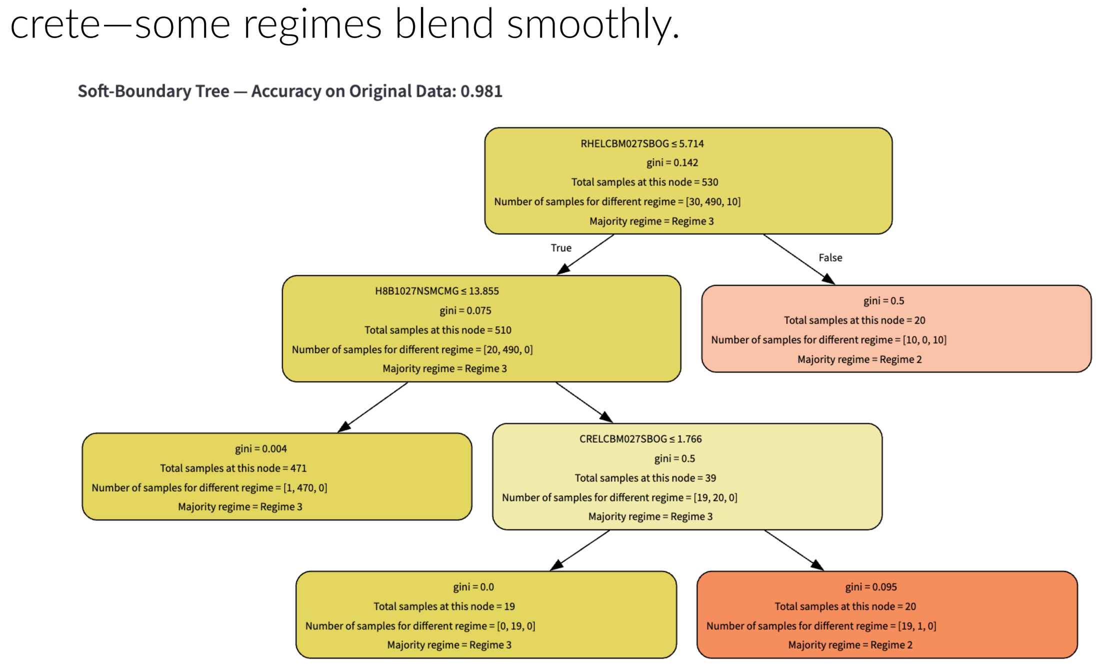


Figure 6. *

Prediction drop after injecting noise into key features.

3. Soft Tree with Gaussian Boundary: We fit a soft decision model with Gaussian-weighted splits to capture fuzzier transitions. This reveals that boundaries are not strictly discrete—some regimes blend smoothly.



yiting.liu@bfh.ch