Bayesian Tensor Regression with Stochastic Volatility

Qing Wang[‡]

[‡] **Ca' Foscari University of Venice**

29th Oct 2025

Model in nutshell

Tensor regression

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + \sigma \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$t = 1, \ldots, T, y_t \in \mathbb{R}, \mathcal{B}, \mathcal{X}_t \in \mathbb{R}^{p_1 \times, \ldots, \times p_N}$ are tensor coefficients and tensor covariates.

Model in nutshell

Tensor regression

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + \sigma \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$t = 1, \dots, T$, $y_t \in \mathbb{R}$, $\mathcal{B}, \mathcal{X}_t \in \mathbb{R}^{p_1 \times, \dots, \times p_N}$ are tensor coefficients and tensor covariates.

Tensor regression with stochastic volatility

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + e^{h_t/2} \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$$h_t = \alpha + \beta(h_{t-1} - \alpha) + \eta_t \qquad \eta_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

Model in nutshell

Tensor regression

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + \sigma \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$t = 1, \ldots, T$, $y_t \in \mathbb{R}$, $\mathcal{B}, \mathcal{X}_t \in \mathbb{R}^{p_1 \times, \ldots, \times p_N}$ are tensor coefficients and tensor covariates.

Tensor regression with stochastic volatility

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + e^{h_t/2} \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$$h_t = \alpha + \beta(h_{t-1} - \alpha) + \eta_t \qquad \eta_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

(i) SV-1, (ii) SV-2, (iii) SVRV-1, (iv) SVX-1

Motivation

Volatility clustering

In many applications the response variable exhibits stochastic volatility, where volatility evolves over time in a persistent yet unpredictable manner, that is the volatility clustering, a well known stylized effect in financial time series (Taylor, 1982, 1986).

Motivation

Volatility clustering

In many applications the response variable exhibits stochastic volatility, where volatility evolves over time in a persistent yet unpredictable manner, that is the volatility clustering, a well known stylized effect in financial time series (Taylor, 1982, 1986).

High dimensional time series data

Covariates naturally arise as multi-way arrays (tensors) capturing information across multiple modes (e.g., time, sector, region, and variable type) in applications, such as asset pricing, sectoral risk assessment, and macro–financial linkages.

Motivation

Volatility clustering

In many applications the response variable exhibits stochastic volatility, where volatility evolves over time in a persistent yet unpredictable manner, that is the volatility clustering, a well known stylized effect in financial time series (Taylor, 1982, 1986).

High dimensional time series data

Covariates naturally arise as multi-way arrays (tensors) capturing information across multiple modes (e.g., time, sector, region, and variable type) in applications, such as asset pricing, sectoral risk assessment, and macro–financial linkages.

Bayesian tensor regression

Bayesian tensor regression has emerged as a powerful approach for handling tensor-valued predictors while avoiding the curse of dimensionality through structured low-rank priors (Guhaniyogi et al., 2017; Billio et al., 2022, 2023; Casarin et al., 2025).

Contributions

Tensor regression + stochastic volatility

Simultaneously exploit the multi-way structure of covariates and account for the dynamic nature of volatility.

Contributions

Tensor regression + stochastic volatility

Simultaneously exploit the multi-way structure of covariates and account for the dynamic nature of volatility.

Augmented stochastic volatility

- Incorporating RV at different horizons in a HAR-RV fashion (Corsi, 2009) to improve estimation efficiency, enhance predictive performance (Bormetti et al., 2020; Bekierman and Gribisch, 2016; Koopman et al., 2005).
- Incorporating tensor-valued exogenous covariates (Harvey et al., 1994; Koopman et al., 2016) to respond to relevant external signals and add more interpretability.

Contributions

Tensor regression + stochastic volatility

Simultaneously exploit the multi-way structure of covariates and account for the dynamic nature of volatility.

Augmented stochastic volatility

- Incorporating RV at different horizons in a HAR-RV fashion (Corsi, 2009) to improve estimation efficiency, enhance predictive performance (Bormetti et al., 2020; Bekierman and Gribisch, 2016; Koopman et al., 2005).
- Incorporating tensor-valued exogenous covariates (Harvey et al., 1994; Koopman et al., 2016) to respond to relevant external signals and add more interpretability.

Scalable Bayesian estimation

We design an efficient Metropolis-Hastings within Gibbs sampling algorithm that alternates between the tensor regression block and the SV block, leveraging the acceptance-rejection Metroplis-Hastings of Chan and Grant (2014).

Model: A Bayesian tensor regression with stochastic volatility

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + e^{h_t/2} \varepsilon_t, \qquad \varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, 1) \tag{1}$$

1. BTRSV-1

$$h_t = \alpha + \beta(h_{t-1} - \alpha) + \eta_t, \qquad \eta_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2) \tag{2}$$

Model: A Bayesian tensor regression with stochastic volatility

$$y_t = \langle \mathcal{B}, \mathcal{X}_t \rangle + e^{h_t/2}\varepsilon_t, \qquad \varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0,1) \tag{1}$$

1. BTRSV-1

$$h_t = \alpha + \beta(h_{t-1} - \alpha) + \eta_t, \qquad \eta_t \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma^2) \tag{2}$$

2. BTRSV-2

$$h_t = \alpha + \beta(h_{t-1} - \alpha) + \gamma(h_{t-2} - \alpha) + \eta_t, \qquad \eta_t \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma^2) \tag{3}$$

Model: A Bayesian tensor regression with stochastic volatility

3. BTRSVRV-1

$$h_t = \alpha + \delta_1 \mathrm{RV}_{t-1} + \delta_2 \mathrm{RV}^{(5)}_{t-1} + \delta_3 \mathrm{RV}^{(22)}_{t-1} + \beta h_{t-1} + \eta_t, \qquad \eta_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2) \tag{4}$$

$t = 1, \ldots, T$, where RV_{t-1}, $\mathrm{RV}^{(5)}_{t-1}$ and $\mathrm{RV}^{(22)}_{t-1}$ are the daily, weekly and monthly averaged log realized volatility starting from day $t-1$.

Model: A Bayesian tensor regression with stochastic volatility

3. BTRSVRV-1

$$h_t = \alpha + \delta_1 \mathrm{RV}_{t-1} + \delta_2 \mathrm{RV}^{(5)}_{t-1} + \delta_3 \mathrm{RV}^{(22)}_{t-1} + \beta h_{t-1} + \eta_t, \qquad \eta_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2) \tag{4}$$

$t = 1, \ldots, T$, where RV_{t-1}, $\mathrm{RV}^{(5)}_{t-1}$ and $\mathrm{RV}^{(22)}_{t-1}$ are the daily, weekly and monthly averaged log realized volatility starting from day $t-1$.

4. BTRSVX-1

$$h_t = \alpha + \langle \Gamma, \mathcal{X}_t \rangle + \beta \left(h_{t-1} - \alpha - \langle \Gamma, \mathcal{X}_{t-1} \rangle \right) + \eta_t \qquad \eta_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2) \tag{5}$$

$t = 1, \ldots, T$, where $\mathcal{X}_t$ is the same tensor-valued covariates appeared in measurement equation (1), Γ is the tensor-valued coefficients for the latent log volatility.

Bayesian Inference: Tensor regression

Dimensionality reduction

We assume a Parallel Factor (PARAFAC) representation of $\mathcal{B}$ and Γ for dimensionality reduction on tensor coefficients:

$$\mathcal{B} = \sum_{d=1}^{D} \gamma_1^{(d)} \circ \cdots \circ \gamma_M^{(d)},$$

where $\circ$ denotes the *external product* of vectors, $\gamma_m^{(d)}$ are the margins from PARAFAC decomposition of tensor coefficient $\mathcal{B}$.

Bayesian Inference: Tensor regression

Dimensionality reduction

We assume a Parallel Factor (PARAFAC) representation of $\mathcal{B}$ and Γ for dimensionality reduction on tensor coefficients:

$$\mathcal{B} = \sum_{d=1}^{D} \gamma_1^{(d)} \circ \cdots \circ \gamma_M^{(d)},$$

where $\circ$ denotes the *external product* of vectors, $\gamma_m^{(d)}$ are the margins from PARAFAC decomposition of tensor coefficient $\mathcal{B}$.

- Number of parameters: $\prod_{n=1}^{N} p_n \to D(\sum_{n=1}^{N} p_n)$.

Dimensionality reduction

We assume a Parallel Factor (PARAFAC) representation of $\mathcal{B}$ and Γ for dimensionality reduction on tensor coefficients:

$$\mathcal{B} = \sum_{d=1}^{D} \gamma_1^{(d)} \circ \cdots \circ \gamma_M^{(d)},$$

where $\circ$ denotes the *external product* of vectors, $\gamma_m^{(d)}$ are the margins from PARAFAC decomposition of tensor coefficient $\mathcal{B}$.

- Number of parameters: $\prod_{n=1}^{N} p_n \to D(\sum_{n=1}^{N} p_n)$.
- The hierarchical prior distribution includes two stages.

Bayesian Inference: Tensor regression

Priors - first stage

We assume that the margins from the PARAFAC decomposition are independent and follow multivariate normal distributions

$$\gamma_m^{(d)} \sim \mathcal{N}_{q_m}(0, \tau\zeta^{(d)} W_m^{(d)}), \qquad m = 1, \ldots, M, \quad d = 1, \ldots, D \tag{6}$$

Bayesian Inference: Tensor regression

Priors - first stage

We assume that the margins from the PARAFAC decomposition are independent and follow multivariate normal distributions

$$\gamma_m^{(d)} \sim \mathcal{N}_{q_m}(0, \tau\zeta^{(d)} W_m^{(d)}), \qquad m = 1, \ldots, M, \quad d = 1, \ldots, D \tag{6}$$

- τ: global shrinkage parameters.
- $\zeta^{(d)}$: component specific shrinkage parameter, allow a subset of the D factors to contribute more while leaving the values of other components close to zero.
- $W_m^{(d)} = \mathrm{diag}(w_{m,1}^{(d)}, \ldots, w_{m,j_m}^{(d)}, \ldots, w_{m,q_m}^{(d)})$: element specific shrinkage parameter.

Bayesian Inference: Tensor regression

Priors - second stage

We modify the priors from Guhaniyogi and Dunson (2015) and further assume the following prior distributions for the scales:

$$\tau \sim \mathcal{IG}(a_\tau, b_\tau), \quad w_{m,j_m}^{(d)} \sim \mathcal{E}xp((\lambda_m^{(d)})^2/2) \tag{7}$$

$$\lambda_m^{(d)} \sim \mathcal{G}a(a_\lambda, b_\lambda), \quad (\zeta^{(1)}, \ldots, \zeta^{(D)}) \sim \mathcal{D}ir(\alpha, \ldots, \alpha) \tag{8}$$

Bayesian LASSO: the priors on $w_{m,j_m}^{(d)}$ and $\lambda_m^{(d)}$ lead to Bayesian LASSO type penalty on $\gamma_m^{(d)}$: $\gamma_{m,j_m}^{(d)} \sim \mathcal{DE}\left(0, \sqrt{\tau\zeta^{(d)}}/\lambda_m^{(d)}\right)$ (Park and Casella, 2008).

Priors for h_t

If we stack all the latent equations by t, we obtain the matrix form of log-volatility process:

$$\boldsymbol{h} = H^{-1}(\boldsymbol{b} + \boldsymbol{\eta}), \qquad \boldsymbol{\eta} \sim \mathcal{N}(0, \Omega)$$

where $\boldsymbol{h} = (h_1, \ldots, h_T)^\top$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_T)^\top$, $\boldsymbol{b} = (\alpha, \alpha(1-\beta), \ldots, \alpha(1-\beta))^\top$ is a $T \times 1$ vector, $\Omega = \mathrm{diag}(\sigma^2/(1-\beta^2), \sigma^2, \ldots, \sigma^2)$ is a $T \times T$ covariance matrix. H is a $T \times T$ banded matrix. Thus,

$$\boldsymbol{h} \sim \mathcal{N}(H^{-1}\boldsymbol{b}, (H^\top \Omega^{-1} H)^{-1}), \tag{9}$$

Bayesian Inference: Stochastic volatility (BTRSV-1)

Priors for h_t

If we stack all the latent equations by t, we obtain the matrix form of log-volatility process:

$$\boldsymbol{h} = H^{-1}(\boldsymbol{b} + \boldsymbol{\eta}), \qquad \boldsymbol{\eta} \sim \mathcal{N}(0, \Omega)$$

where $\boldsymbol{h} = (h_1, \ldots, h_T)^\top$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_T)^\top$, $\boldsymbol{b} = (\alpha, \alpha(1-\beta), \ldots, \alpha(1-\beta))^\top$ is a $T \times 1$ vector, $\Omega = \text{diag}(\sigma^2/(1-\beta^2), \sigma^2, \ldots, \sigma^2)$ is a $T \times T$ covariance matrix. H is a $T \times T$ banded matrix. Thus,

$$\boldsymbol{h} \sim \mathcal{N}(H^{-1}\boldsymbol{b}, (H^\top \Omega^{-1} H)^{-1}), \tag{9}$$

Priors for α, β, σ

We assume the following priors for α, β, σ^2:

$$\alpha \sim \mathcal{N}(\alpha_0, \sigma_\alpha^2), \qquad \beta \sim \mathcal{N}(\beta_0, \sigma_\beta^2)\mathbb{I}(|\beta| < 1), \qquad \sigma^2 \sim \mathcal{IG}(a_\sigma, b_\sigma), \tag{10}$$

We impose the stationarity condition $|\beta| < 1$ through the prior on β.

Posterior approximation - First block

We sample the tensor coefficients $\mathcal{B}$ from $f(\mathcal{B} \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{h})$ and the hyperparameters of its hierarchical prior by using a similar strategy as in the Bayesian tensor regression proposed in Casarin et al. (2025); Papadogeorgou et al. (2021); Guhaniyogi et al. (2017).

Posterior approximation - First block

We sample the tensor coefficients $\mathcal{B}$ from $f(\mathcal{B} \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{h})$ and the hyperparameters of its hierarchical prior by using a similar strategy as in the Bayesian tensor regression proposed in Casarin et al. (2025); Papadogeorgou et al. (2021); Guhaniyogi et al. (2017).

1. Draw $\gamma_m^{(d)}$ from a multivariate normal distribution $f(\boldsymbol{\gamma}_m^{(d)} \mid \boldsymbol{y}, \mathcal{X}, \boldsymbol{\gamma}_{-m}, \tau, \boldsymbol{\zeta}, \boldsymbol{w}, \boldsymbol{h})$ for $d \in \{1, \ldots, D\}$ and $m \in \{1, \ldots, M\}$.
2. Draw $\zeta^{(d)}$ from the GIG distribution $f(\zeta^{(d)} \mid \boldsymbol{\gamma}^{(d)}, \tau, \boldsymbol{w}^{(d)})$.
3. Draw τ from the IG distribution $f(\tau \mid \boldsymbol{\gamma}, \boldsymbol{\zeta}, \boldsymbol{w})$.
4. Draw $\lambda_m^{(d)}$ from a Gamma distribution $f(\lambda_m^{(d)} \mid \boldsymbol{\gamma}_m^{(d)}, \tau, \zeta^{(d)})$.
5. Draw $w_{m,j_m}^{(d)}$ from the GIG distribution $f(w_{m,j_m}^{(d)} \mid \gamma_{m,j_m}^{(d)}, \lambda_m^{(d)}, \tau, \zeta^{(d)})$.

Posterior approximation - Second block

We sample latent log-volatility $\{h_t\}_{t=1}^T$ using Metroplis-Hastings and sample other parameters α, β and σ^2 from their full conditionals using MH within Gibbs procedure.

Posterior approximation - Second block

We sample latent log-volatility $\{h_t\}_{t=1}^T$ using Metroplis-Hastings and sample other parameters α, β and σ^2 from their full conditionals using MH within Gibbs procedure.

6. Draw $\boldsymbol{h}$ from $f(\boldsymbol{h} \mid \boldsymbol{y}, \mathcal{X}, \mathcal{B}, \alpha, \beta, \sigma^2)$.
7. Draw α from a normal distribution $f(\alpha \mid \boldsymbol{h}, \beta, \sigma^2)$.
8. Draw β from a normal distribution $f(\beta \mid \boldsymbol{h}, \alpha, \sigma^2)$.
9. Draw σ^2 from $\mathcal{IG}$ distribution $f(\sigma^2 \mid \boldsymbol{h}, \alpha, \beta)$.

In step 6, the joint conditional density $f(\boldsymbol{h} \mid \boldsymbol{y}, \mathcal{X}, \mathcal{B}, \alpha, \beta, \sigma^2)$ is high dimensional and non-standard. We follow Chan and Grant (2014) to simulate from this density using MH algorithm.

Simulations

We carry out simulation studies to demonstrate the validity of the Bayesian inference procedures for SV-1 and SV-2. Let $\mu = \langle \mathcal{B}, \mathcal{X}_t \rangle$:

	μ	α	σ	β	γ
SV-1	0	-1	0.2	0.95	
SV-2	0	-1	0.2	0.5	0.4

Simulations

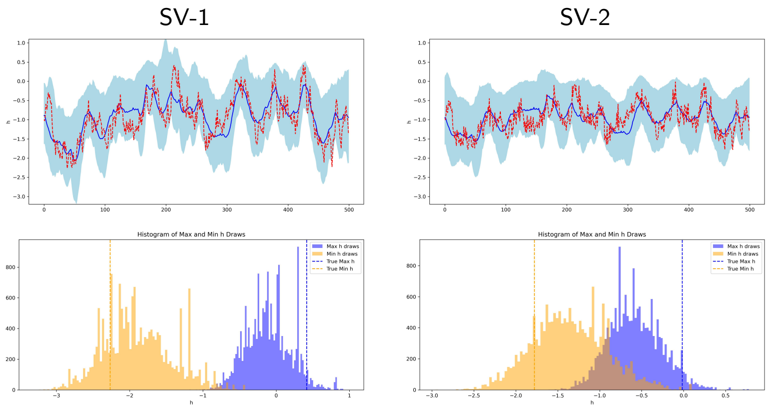

Figure: Estimated log-volatility and approximated posterior distributions of minimum and maximum log volatility for the two stochastic volatility models: SV-1 and SV-2. First row: estimated log-volatility (blue solid line) together with the true log-volatility (red dashed line) and the 95% credible interval (blue shaded area). Second row: histograms of the MCMC draws of maximum (purple) and minimum (yellow) log-volatility and their true values (vertical dashed lines).

Empirical experiment: forecasting market returns

Objective

We apply the Bayesian tensor regression with stochastic volatility to model the log-return of S&P 500, and compare the performances of different SV specifications.

Empirical experiment: forecasting market returns

Objective

We apply the Bayesian tensor regression with stochastic volatility to model the log-return of S&P 500, and compare the performances of different SV specifications.

Variables

- Oil price volatility: Good oil volatility (GV), Bad oil volatility (BV).
- Other covariates: USD index (ER), TED spread (IR), VIX index (VI), T-bill rate (TB) and bond spread (BD). In case of BTRSVX-1, an extra variable of daily realized volatility (RV) is also included in the tensor covariates.

Empirical experiment: forecasting market returns

Specification

- We model the log-return following a Mixed Data Sampling strategy (Rodriguez and Puggioni, 2010).
- y_t is monthly log-return of S&P 500, covariates are sampled daily at 1-day to 22-day before month t: $t - 1/22, t - 2/22, \dots, t - 1$.
- $\mathcal{X}_t \in \mathbb{R}^{7\times 22\times 4}$ is a mode-3 tensor. First mode: variables. Second mode: daily data points. Third mode: monthly lag.

Empirical experiment: forecasting market returns

Specification

- We model the log-return following a Mixed Data Sampling strategy (Rodriguez and Puggioni, 2010).
- y_t is monthly log-return of S&P 500, covariates are sampled daily at 1-day to 22-day before month t: $t - 1/22, t - 2/22, \ldots, t - 1$.
- $\mathcal{X}_t \in \mathbb{R}^{7\times 22\times 4}$ is a mode-3 tensor. First mode: variables. Second mode: daily data points. Third mode: monthly lag.

$$y_t = \sum_{i_3=1}^{4} \left\langle B_{\tilde{I}(i_3)}, \begin{pmatrix} \mathrm{GV}_{t-\frac{1}{22}-i_3+1} & \mathrm{GV}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{GV}_{t-\frac{21}{22}-i_3+1} & \mathrm{GV}_{t-i_3} \\ \mathrm{BV}_{t-\frac{1}{22}-i_3+1} & \mathrm{BV}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{BV}_{t-\frac{21}{22}-i_3+1} & \mathrm{BV}_{t-i_3} \\ \mathrm{ER}_{t-\frac{1}{22}-i_3+1} & \mathrm{ER}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{ER}_{t-\frac{21}{22}-i_3+1} & \mathrm{ER}_{t-i_3} \\ \mathrm{IR}_{t-\frac{1}{22}-i_3+1} & \mathrm{IR}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{IR}_{t-\frac{21}{22}-i_3+1} & \mathrm{IR}_{t-i_3} \\ \mathrm{VI}_{t-\frac{1}{22}-i_3+1} & \mathrm{VI}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{VI}_{t-\frac{21}{22}-i_3+1} & \mathrm{VI}_{t-i_3} \\ \mathrm{TB}_{t-\frac{1}{22}-i_3+1} & \mathrm{TB}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{TB}_{t-\frac{21}{22}-i_3+1} & \mathrm{TB}_{t-i_3} \\ \mathrm{BD}_{t-\frac{1}{22}-i_3+1} & \mathrm{BD}_{t-\frac{2}{22}-i_3+1} & \cdots & \mathrm{BD}_{t-\frac{21}{22}-i_3+1} & \mathrm{BD}_{t-i_3} \end{pmatrix} \right\rangle + e^{h_t/2}\varepsilon_t, \quad \varepsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0,1) \tag{11}$$

where $\tilde{I}(i_3) = \{(i_1, i_2, i_3), i_h \in \{1, \ldots, p_h\}, \forall h \neq 3\}$ and $B_{\tilde{I}(i_3)}$ denotes the i_3th slice of tensor coefficients B along the third mode.

Empirical experiment: forecasting market returns

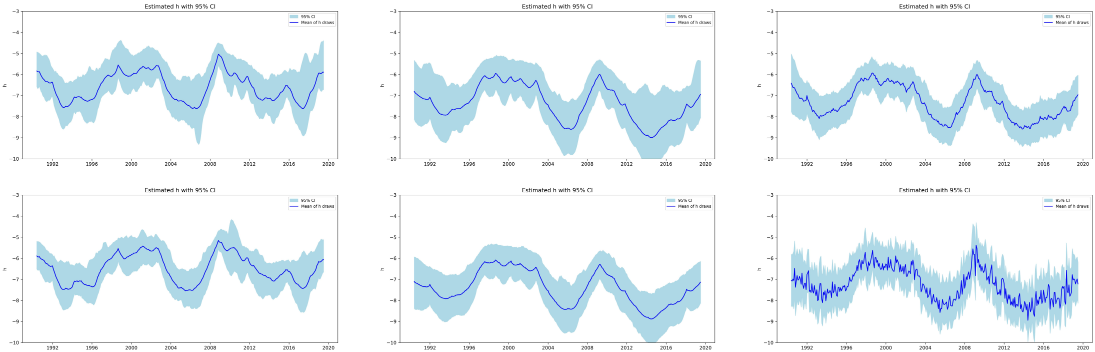

Figure: Estimated log-volatility and their 95% quantiles. Posterior mean of the log-volatility draws (blue line) and 95% credible interval (blue shaded area) for SV-1 and SV-2 (first column), BTRSV-1 and BTRSV-2 (second column), BTRSVRV-1 and BTRSVX-1 (third column).

Empirical experiment: forecasting market returns

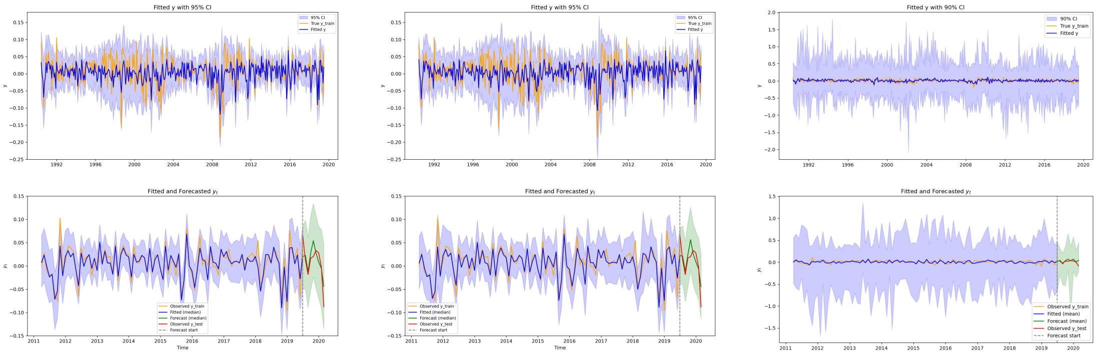

Figure: In-sample and out-of-sample performance. First row: in-sample fitting for BTRSVRV-1 (1st column), BTRSVX-1 (2nd column) and BTR (3rd column). Second row: out-of-sample forecasting for the same models. The observed training response y_t is shown in orange solid line, the posterior medium of estimated response is shown in blue solid line and the 95% credible interval is shown in blue shaded area. The observed test responses are shown in red solid line and the posterior medium of the estimated responses are shown in green solid line with the 95% shown in green shaded area.

Empirical experiment: forecasting market returns

We report the Root Mean Square Error (RMSE) and Continuous Ranked Probability Score (CRPS) (Gneiting and Raftery, 2007) to evaluate the in-sample and out-of-sample performance of the seven different models.

Table: RMSE and CRPS of in- and out-of-sample performances for different models.

	RMSE		CRPS	
	in-sample	out-of-sample	in-sample	out-of-sample
BTRSV-1	0.0271	0.0226	0.0136	0.0134
BTRSV-2	**0.0263**	0.0235	**0.0133**	0.0137
BTRSVRV-1	0.0270	**0.0217**	0.0136	0.0131
BTRSVX-1	0.0275	0.0220	0.0139	**0.0127**
SV-1	0.0416	0.0383	0.0216	0.0213
SV-2	0.0416	0.0378	0.0217	0.0207
BTR	0.0633	0.0918	0.0397	0.0464

Conclusion

- We introduce a unified and flexible Bayesian framework that integrates tensor regression with various forms of stochastic volatility modeling.
- We introduce new SV models incorporating RV and tensor-valued exogenous variables.
- We provide a scalable and fully Bayesian estimation strategy based on a Metropolis-Hastings within Gibbs sampler.
- Empirical study using S&P 500 returns and a large panel of mixed-frequency financial indicators confirms that models that incorporate stochastic volatility exhibit enhanced responsiveness to market conditions and better uncertainty quantification.

References I

Bekierman, J. and Gribisch, B. (2016). Estimating stochastic volatility models using realized measures. *Studies in Nonlinear Dynamics & Econometrics*, 20(3):279–300.

Billio, M., Casarin, R., and Iacopini, M. (2022). Bayesian Markov-Switching Tensor Regression for Time-Varying Networks. *Journal of the American Statistical Association*, pages 1–13.

Billio, M., Casarin, R., Iacopini, M., and Kaufmann, S. (2023). Bayesian dynamic tensor regression. *Journal of Business & Economic Statistics*, 41(2):429–439.

Bormetti, G., Casarin, R., Corsi, F., and Livieri, G. (2020). A Stochastic Volatility Model With Realized Measures for Option Pricing. *Journal of Business & Economic Statistics*, 38(4):856–871.

Casarin, R., Craiu, R. V., and Wang, Q. (2025). Markov switching multiple-equation tensor regressions. *Journal of Multivariate Analysis*, 208:105427.

Chan, J. C. and Grant, A. (2014). Issues in comparing stochastic volatility models using the deviance information criterion.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of financial econometrics*, 7(2):174–196.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378.

Guhaniyogi, R. and Dunson, D. B. (2015). Bayesian Compressed Regression. *Journal of the American Statistical Association*, 110(512):1500–1514.

References II

Guhaniyogi, R., Qamar, S., and Dunson, D. B. (2017). Bayesian tensor regression. *Journal of Machine Learning Research*, 18(79):1–31.

Harvey, A., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.

Koopman, S. J., Jungbacker, B., and Hol, E. (2005). Forecasting daily variability of the s&p 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3):445–475.

Koopman, S. J., Lucas, A., and Scharth, M. (2016). Predicting time-varying parameters with parameter-driven and observation-driven models. *Review of Economics and Statistics*, 98(1):97–110.

Papadogeorgou, G., Zhang, Z., and Dunson, D. B. (2021). Soft tensor regression. *Journal of Machine Learning Research*, 22(219):1–53.

Park, T. and Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, 103(482):681–686.

Rodriguez, A. and Puggioni, G. (2010). Mixed frequency models: Bayesian approaches to estimation and prediction. *International Journal of Forecasting*, 26(2):293–311.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. *Time series analysis: theory and practice*, 1:203–226.

Taylor, S. J. (1986). *Modelling Financial Time Series*. John Wiley & Sons, Chichester, UK.