mathematics

MDPI

Article

Full Bayesian Analysis of Seasonal Vector Autoregressive Models: An Application to Solar Radiation in Najran, Saudi Arabia

Ayman A. Amin [1,2,*] and Fatimah E. Almuhayfith [3]

[1] Department of Mathematics, College of Arts and Sciences, Najran University, Najran 66462, Saudi Arabia
[2] Science and Engineering Research Center, Najran University, Najran 66462, Saudi Arabia
[3] Department of Mathematics and Statistics, College of Science, King Faisal University, Alahsa 31982, Saudi Arabia; falmuhaifeez@kfu.edu.sa
* Correspondence: aaabdellah@nu.edu.sa

Abstract

Modelling multiple seasonal time series is a pervasive challenge across diverse domains, including energy, transportation, and financial markets. Existing Bayesian methods for seasonal vector autoregressive (SeVAR) models treat lag-order identification, coefficient estimation, and multi-step forecasting as separate tasks, lacking a unified framework. To address this limitation, we develop a comprehensive Gibbs sampling scheme capable of executing all three tasks simultaneously for SeVAR processes. We embed the SeVAR model in a hierarchical Bayes mixture framework in the spirit of group (block) stochastic search variable selection (Block-SSVS). Rather than assigning inclusion indicators to individual coefficients, a single Bernoulli-distributed binary indicator is attached to each lag matrix, ensuring that all entries within a block are either jointly retained or jointly shrunk toward zero. Following the free-parameter approximation standard in the literature, the cross (regular$\times$seasonal) lag blocks are treated as unrestricted free parameters rather than as the exact products implied by the multiplicative operator, so the fitted model is a seasonally structured subset VAR over the candidate lag set generated by that operator. A block-structured spike-and-slab matrix normal prior governs the stacked coefficient matrix conditional on these block indicators and the innovation precision matrix, and a Wishart prior completes the precision-matrix specification. Within this hierarchical structure, we establish the closed-form full conditional posteriors for all model parameters. Specifically, we derive that the coefficient matrix follows a matrix normal distribution, the precision matrix yields a Wishart posterior, and each block indicator follows a Bernoulli posterior governed by the joint, matrix-variate evidence of its block. Furthermore, the one-step and multi-step conditional predictive distributions are established as a multivariate normal structure. Leveraging these results, we construct a stochastic simulation-based recursion to produce multi-step forecasts. The resulting posterior block-inclusion probabilities furnish an automatic best-subset SeVAR model at no additional computational cost beyond the Gibbs sampler itself. A controlled Monte Carlo study confirms the accuracy of the procedure for identification, estimation, and prediction. In addition, an application to hourly solar radiation in Najran, Saudi Arabia, demonstrates the framework's identification capacity on genuinely bivariate stochastic data, with forecast accuracy reported at the data's own predictability floor.

Check for updates

Academic Editor: Mengyu Xu

Received: 21 July 2026
Revised: 31 October 2026
Accepted: 3 November 2026
Published: 5 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: SeVAR; Gibbs sampler; Block-SSVS; spike-and-slab prior; matrix normal distribution; Wishart distribution; solar radiation forecasting

MSC: 62M10; 62F15

1. Introduction

Many multivariate time series in economics, engineering, and the natural sciences exhibit recurrent patterns at fixed periodic intervals. For instance, electricity and gas demand often return to comparable levels at the same hour of the day or the same month of the year; retail sales and tourism arrivals expand and contract in step with the calendar; and hourly solar radiation series display pronounced diurnal cycles driven by the alternation of day and night. When a periodic component is superimposed on short-range autoregressive (AR) dependence, a non-seasonal vector autoregressive (VAR) specification becomes misspecified [1,2]. Such a model either disregards the periodic signal—forcing the residuals to absorb structure that is in fact predictable—or approximates seasonality with rigid deterministic dummies that fail to adapt across regimes. By contrast, introducing a stochastic seasonal AR term allows the periodic dependence itself to carry sampling uncertainty and to interact coherently with the non-seasonal dynamics [3,4]. Proper seasonal specification becomes increasingly important in a multivariate setting, where capturing seasonal structure jointly is critical. In such systems, cross-series dependence at seasonal lags frequently conveys as much information as short-run cross-correlations. Consequently, a seasonal VAR (SeVAR) specification—characterised by a multiplicative regular-seasonal lag polynomial and contemporaneously correlated innovations—provides the natural framework to accommodate these joint dynamics.

Several distinct strands of the literature relate to the Bayesian analysis of SeVAR processes, though each addresses only a subset of the problem. Within the Bayesian VAR estimation literature, Kadiyala and Karlsson [5] established an exact normal-Wishart conjugate prior structure, building upon the foundational Minnesota prior of Litterman [6] and Doan et al. [7]. This framework was subsequently scaled to large dimensions by Bańbura et al. [8] and enhanced with data-driven hyperparameter selection by Giannone et al. [9]. Extensions incorporating time-varying parameters and adaptive shrinkage were later proposed by Koop and Korobilis [10], Huber and Feldkircher [11], and Chan [12]. However, these contributions lack an explicit seasonal lag operator and do not perform individual lag selection. The lag order remains predetermined, and while some estimated coefficients appear to be shrunk, they may not be strictly excluded from the model.

The second strand focuses on the Bayesian identification of SeVAR processes. Shaarawy and Ali [13] derived closed-form posterior mass functions over non-seasonal VAR orders, and Shaarawy and Ali [14] extended this work to SeVAR processes, building on the non-seasonal VAR framework of Sims [15]. Recently, Shaarawy [16] completed the picture for non-seasonal VARMA estimation and prediction. This body of work identifies the model order by an exhaustive grid search over a discrete posterior mass function, and estimation, identification, and prediction are typically presented as separate exercises rather than one pipeline. Most recently, Amin and Almuhayfith [17] proposed the Bayesian identification of double-SeVAR processes by deriving the closed-form posterior mass function of the model order.

Stochastic search variable selection (SSVS), introduced by George and McCulloch [18] and refined in George and McCulloch [19], Chipman et al. [20], attaches a latent binary indicator to each candidate coefficient and lets its posterior inclusion probability reveal whether the coefficient is empirically active. SSVS procedure has been widely employed

in univariate seasonal AR (SAR) models [21,22] and double-SAR frameworks [23]. SSVS-type schemes have been used to restrict large non-seasonal VAR models [24,25]; to select variables in state-space models [26]; and alongside continuous shrinkage alternatives such as the Bayesian lasso [27], the normal-gamma prior [28], spike-and-slab reviews [29], and message-passing forecasting algorithms [30,31]. A closely related idea, group or block variable selection, jointly includes or excludes a pre-specified set of coefficients that share a common structural role rather than screening each coefficient in isolation [32]. Korobilis [25] applies exactly this logic to restrict entire equations or predictor blocks of a Bayesian VAR at once. None of this literature, to our knowledge, has coupled block-level selection with a seasonal AR lag polynomial, nor combined it with a forecasting recursion inside the same sampler.

In most of the Bayesian literature surveyed above, identification (choosing lag order), estimation (computing the posterior of the coefficients given the chosen order), and prediction (deriving the forecast distribution given the estimated model) are presented and implemented as separate stages, each conditioning on a point estimate handed down from the previous one. This staged approach silently discards two sources of uncertainty: uncertainty about which lags belong in the model is not propagated into the coefficient posterior, and uncertainty about the coefficients is not propagated into the forecast distribution beyond the plug-in mean. In addition, this fragmentation has a practical cost. An analyst must first search over orders, then re-estimate the chosen model by a separate procedure, and finally build a forecasting routine around the point-selected model, discarding the model uncertainty accumulated in the first two steps.

In this paper, we fill this gap by proposing a self-consistent Gibbs sampling framework that jointly executes all these tasks. This framework performs lag-level subset selection, coefficient estimation, and multi-step probabilistic forecasting for an SeVAR process within a single algorithm. Specifically, a single Gibbs sampler that draws the indicators, coefficients, precision matrix, and future values in the same sweep automatically averages over all of this uncertainty. Every retained posterior draw corresponds to one plausible subset model, one plausible coefficient matrix, and one plausible future trajectory, so that summarising the draws propagates selection, coefficient, and precision uncertainty into the credible and predictive intervals without any extra analytical work. This is the central methodological motivation for the present paper. Some of the most relevant antecedents to this paper are our earlier univariate analyses of SSVS-based seasonal autoregressive selection [21–23]. Relative to that body of work, the present contribution is new along two specific dimensions: (a) the object being selected is no longer a scalar coefficient of a univariate SAR or DSAR model but an entire $k \times k$ coefficient matrix of a multivariate SeVAR system, so the indicator, its Bernoulli conditional, and the evidence it is weighed against (a matrix-variate sum of squares rather than a scalar sum of squares) are genuinely multivariate objects with no univariate analogue; and (b) selection is performed at the level of the whole lag block rather than the individual entry, which is only a meaningful distinction once $k \geq 2$. These points make the current incremental contribution relative to our prior work explicit. Accordingly, this paper makes the following specific contributions.

(i) **Hierarchical group (block) SSVS prior.**
Rather than assigning a Bernoulli latent indicator to each scalar coefficient or regressor column, we attach a single shared indicator to every lag matrix—following the group/block selection logic of George and McCulloch [18], Korobilis [25], Yuan and Lin [32]. Conditional on these block indicators and the precision matrix, we then impose a block-structured spike-and-slab matrix-normal prior on the stacked coefficient matrix, together with a conjugate Wishart prior on the precision matrix.

(ii) **Closed-form conditional distributions.**

We derive that the full conditional posterior of the coefficient matrix is matrix-normal, that of the precision matrix is Wishart, and that each block indicator follows a Bernoulli distribution—governed by the joint matrix-variate evidence of its entire block. Moreover, the one-step and multi-step conditional predictive distributions are derived to follow a multivariate normal form.

(iii) **Unified algorithm.**
We specify a single Gibbs sampling algorithm that cycles through these conditionals once per iteration, so that subset selection, estimation, and forecasting are carried out by exactly one Markov chain rather than three disconnected procedures.

(iv) **Empirical validation.**
A Monte Carlo experiment evaluates subset-selection accuracy, estimation precision, and multi-step forecast accuracy across several parameter configurations, and an application to Najran hourly solar radiation under an intraday seasonal period demonstrates the practical value of the approach.

The remainder of the manuscript is organised as follows. Section 2 defines the SeVAR process and its matrix-regression representation. Section 3 introduces the hierarchical SSVS prior. Section 4 derives the conditional posterior distributions, while Section 5 establishes the conditional predictive distributions. Section 6 presents the Gibbs sampling algorithm for the full Bayesian analysis of SeVAR models. Section 7 reports the Monte Carlo study, and Section 8 illustrates the Najran hourly solar radiation application. Finally, Section 9 concludes the manuscript.

2. Seasonal Vector Autoregressive (SeVAR) Model

2.1. Model Definition

Let $\{\boldsymbol{y}(t)\}_{t=1}^{n}$ be a sequence of k-dimensional observable random vectors, $k \geq 2$. The backshift operator B acts by $B^r \boldsymbol{y}(t) = \boldsymbol{y}(t-r)$ for any non-negative integer r. A single seasonal period $s > 1$ is fixed and known, with non-seasonal order $p \geq 0$ and seasonal order $P \geq 0$. Two matrix-valued lag polynomials are defined as

$$\Theta_p(B) = \boldsymbol{I}_k - \sum_{i=1}^{p} \Theta_i B^i, \tag{1}$$

$$\Omega_P(B^s) = \boldsymbol{I}_k - \sum_{j=1}^{P} \Omega_j B^{js}, \tag{2}$$

where $\Theta_i, \Omega_j \in \mathbb{R}^{k\times k}$ are unknown coefficient matrices and $\boldsymbol{I}_k$ is the $k \times k$ identity matrix. We use Θ for the non-seasonal block and Ω for the seasonal block throughout, reserving $\boldsymbol{\Lambda}$ for the innovation precision matrix and $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$ for the corresponding covariance matrix; this notation is unified across the entire paper.

Definition 1 (SeVAR Process)**.** *The seasonal vector autoregressive process of non-seasonal order p, seasonal order P, period s, and dimension k, denoted* $\mathrm{SeVAR}_k(p, P)_s$*, is the k-dimensional sequence obeying*

$$\Theta_p(B)\, \Omega_P(B^s)\, \boldsymbol{y}(t) \;=\; \boldsymbol{\varepsilon}(t), \tag{3}$$

where the innovations $\{\boldsymbol{\varepsilon}(t)\}$ are independent and identically distributed as $\mathcal{N}_k(\mathbf{0}, \boldsymbol{\Lambda}^{-1})$, with $\boldsymbol{\Lambda}$ an unknown $k \times k$ positive-definite precision matrix.

Remark 1 (Special cases)**.** *Setting $P = 0$ collapses (3) to the ordinary $\mathrm{VAR}_k(p)$ model of Lütkepohl [2], Sims [15]. Setting $k = 1$ recovers the univariate seasonal autoregressive (SAR)*

model that underlies the seasonal ARMA literature. The present $\mathrm{SeVAR}_k(p,P)_s$ *specification is precisely the seasonal VAR family studied by Shaarawy and Ali [14].*

The process (3) is covariance-stationary if and only if all roots of the two determinantal equations

$$\left|\Theta_p(z)\right| = 0, \qquad \left|\Omega_P(z^s)\right| = 0 \tag{4}$$

lie strictly outside the complex unit disk, $|z| > 1$. Condition (4) characterises the multiplicative process of Definition 1; it is not imposed on—or enforced for—the seasonally structured subset VAR that is actually fitted (Remark 2).

2.2. Expanded Form and Matrix Regression Representation

Expanding the product $\Theta_p(B)\,\Omega_P(B^s)$,

$$\Theta_p(B)\,\Omega_P(B^s) = \boldsymbol{I}_k - \sum_{i=1}^{p}\Theta_i B^i - \sum_{j=1}^{P}\Omega_j B^{js} + \sum_{i=1}^{p}\sum_{j=1}^{P}\Theta_i\Omega_j\, B^{i+js}, \tag{5}$$

so that (3) may be rewritten in explicit autoregressive form as

$$\boldsymbol{y}(t) = \underbrace{\sum_{i=1}^{p}\Theta_i\,\boldsymbol{y}(t-i)}_{\text{Group 1: regular}} + \underbrace{\sum_{j=1}^{P}\Omega_j\,\boldsymbol{y}(t-js)}_{\text{Group 2: seasonal}} - \underbrace{\sum_{i=1}^{p}\sum_{j=1}^{P}\Theta_i\Omega_j\,\boldsymbol{y}(t-i-js)}_{\text{Group 3: cross-product}} + \boldsymbol{\varepsilon}(t). \tag{6}$$

Equation (6) reveals the three regressor groups generated by the multiplicative operator; their lag sets, signs, and sizes are summarised in Table 1. The total number of lag groups per equation variable is $h = p + P + pP = (1+p)(1+P) - 1$. No intercept or other deterministic component is included in (3), so the specification represents a purely multiplicative autoregression with zero mean. A non-zero mean can be accommodated, if required, by de-meaning $\boldsymbol{y}(t)$ prior to estimation; this transformation leaves all subsequent results unchanged.

Table 1. Three regressor groups in the expanded $\mathrm{SeVAR}_k(p,P)_s$ model (6).

Group	Description	Lag Indices	Sign	No. of Lags
1	Regular (non-seasonal)	$i,\ i = 1,\ldots,p$	$+$	p
2	Seasonal	$js,\ j = 1,\ldots,P$	$+$	P
3	Regular $\times$ seasonal cross	$i+js,\ i = 1,\ldots,p;\ j = 1,\ldots,P$	$-$	pP
	Total regressor groups per variable:			h

Conditioning on the first $p + Ps$ observations, the SeVAR model (6) takes the multivariate linear regression form

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\mathcal{K}} + \boldsymbol{U}, \tag{7}$$

where the response matrix $\boldsymbol{Y} \in \mathbb{R}^{n^*\times k}$ stacks the row vectors $\boldsymbol{y}^\top(t_0),\ldots,\boldsymbol{y}^\top(n)$, with $t_0 = p + Ps + 1$, effective sample size $n^* = n - p - Ps$, and the regressor matrix $\boldsymbol{X} \in \mathbb{R}^{n^*\times q}$, $q = kh$, and has a t-th row equal to the $1\times q$ vector

$$\begin{aligned}\boldsymbol{x}^\top(t-1) = \Big[&\boldsymbol{y}^\top(t-1),\ldots,\boldsymbol{y}^\top(t-p),\ \boldsymbol{y}^\top(t-s),\ldots,\boldsymbol{y}^\top(t-Ps),\\ &-\boldsymbol{y}^\top(t-s-1),\ldots,-\boldsymbol{y}^\top(t-Ps-p)\Big].\end{aligned} \tag{8}$$

The stacked coefficient matrix $\mathcal{K} \in \mathbb{R}^{q\times k}$ partitions into three blocks conforming to Table 1,

$$\mathcal{K} = \begin{bmatrix} \boldsymbol{\kappa}_1(p) \\ \boldsymbol{\kappa}_2(P) \\ \boldsymbol{\kappa}_3(p,P) \end{bmatrix}_{q\times k}, \qquad \boldsymbol{\kappa}_1(p) = [\Theta_1^\top, \ldots, \Theta_p^\top]^\top, \qquad \boldsymbol{\kappa}_2(P) = [\Omega_1^\top, \ldots, \Omega_P^\top]^\top, \quad (9)$$

with $\boldsymbol{\kappa}_3(p,P) \in \mathbb{R}^{pPk\times k}$ treated as a free cross-product block, following the free-parameter approximation of Shaarawy and Ali [14]. It is therefore regarded as an unrestricted component, independent of $\{\Theta_i\}$ and $\{\Omega_j\}$. Under this treatment, the representation in (7) becomes an exact multivariate linear regression, fully amenable to conjugate Bayesian analysis. The noise matrix $\boldsymbol{U} \in \mathbb{R}^{n^*\times k}$ has rows $\boldsymbol{\varepsilon}^\top(t) \overset{\text{i.i.d.}}{\sim} \mathcal{N}_k(\mathbf{0}, \boldsymbol{\Lambda}^{-1})$.

Remark 2 (The fitted object is a seasonally structured subset VAR)**.** *We state plainly the consequence of the free-parameter approximation just introduced. Under the exact multiplicative operator of Definition 1, the cross-block is not free. It equals $\Theta_i\Omega_j$ (group 3 of Table 1), so that switching off Θ_i or Ω_j mechanically forces the corresponding cross-term to zero as well. Once $\boldsymbol{\kappa}_3(p,P)$ is treated as an unrestricted matrix, this constraint is relaxed, and $\mathcal{K}$ in (9) indexes a* seasonally structured subset VAR *over the candidate lag set $\{1,\ldots,p\} \cup \{s,\ldots,Ps\} \cup \{i+js : i \le p, j \le P\}$ generated by the multiplicative operator, rather than the multiplicative model itself.*

3. Hierarchical Bayesian Model with Group (Block) Stochastic Search Variable Selection

3.1. From Element-Wise to Group (Block) Latent Indicators

The total number of free scalar parameters in $\mathcal{K}$ is $qk = k^2h$, which increases rapidly with the model orders p, P, and the dimension k. For example, in a bivariate system ($k = 2$) with $p = P = 2$ one obtains $q = kh = 16$ regressors per equation and hence $qk = 32$ free parameters. In practice the true orders p and P are unknown, so conservative upper bounds $p_{\max}$ and $P_{\max}$ are typically assumed, further exacerbating the dimensionality challenge. Stochastic search variable selection (SSVS) [18,19] is designed to mitigate this problem by attaching a latent binary indicator to each coefficient and letting the posterior distribution reveal which coefficients are empirically supported.

A naïve, fully disaggregated (*element-wise*) implementation of SSVS in the present setting would assign an independent indicator $\gamma_{ij} \in \{0,1\}$ to *every individual entry* of every lag matrix, so that the latent indicator space has dimension hk^2 (the total number of scalar coefficients in $\mathcal{K}$). We do not adopt this scheme. Instead, following the group (or block) variable-selection principle [25,32] and its use in Bayesian VAR restriction [24], we adopt *group (block) SSVS*. Specifically, Block-SSVS assigns a single, shared indicator $\delta_m \in \{0,1\}$ to the entire $k \times k$ coefficient matrix associated with lag m (Table 1)—that is, to a full Θ_i, a full Ω_j, or a full regular×seasonal cross-product matrix. Under this scheme the joint, matrix-variate density of the entire lag block is evaluated, so that all k^2 parameters within the block are jointly retained under the slab prior or jointly shrunk toward zero under the spike prior. Employing Block-SSVS for SeVAR models offers three distinct advantages.

1. **Preservation of systemic structural integrity within a block.**
   The SeVAR model is fundamentally defined by the multiplicative operator $\Theta_p(B)\,\Omega_P(B^s)$ of Definition 1. Block-SSVS strictly honours this definition at the level of a single lag block: a regular lag, a seasonal cycle, or a cross-lag represents a single systemic state of the joint process. So, if this lag is active it is theoretically sound that it influences the interconnected k-variate system as a whole, rather than operating independently

equation by equation. This coherence is enforced *within* each block and does not extend *across* blocks under the free-parameter approximation.

2. **Prevention of artificial matrix truncation.**
   Element-wise selection often yields "half-dead" matrices, where a lag is deemed significant for one series but shrunk to zero for another within the same $k \times k$ block—fragmenting the very cross-equation feedback that the SeVAR representation exists to capture. By evaluating the aggregate, $\mathbf{\Lambda}$-weighted sum of squares of the entire block rather than one entry at a time, Block-SSVS captures the block's overall signal-to-noise ratio and preserves genuine cross-variable interactions within active lags intact.
3. **Enhanced parsimony and MCMC convergence.**
   Synchronising the inclusion probability across the k equations and the k source variables of a lag reduces the dimension of the latent indicator space from hk^2 (element-wise) to simply h (Block-SSVS). This drastically restricts the combinatorial space of candidate models the sampler must explore, improving the mixing, efficiency, and convergence speed of the Gibbs sampler. This drastically restricts the combinatorial space of candidate models the sampler must explore. This is a statement about the model space rather than about chain behaviour. The expected mixing benefit is consistent with our indirect evidence, but we do not claim a direct chain-level demonstration here.

Robustness to intra-block heterogeneity

Block-SSVS's whole-matrix evidence, $\mathrm{tr}[\boldsymbol{\mathcal{K}}_{(m)}^{\top}\boldsymbol{\mathcal{K}}_{(m)}\mathbf{\Lambda}]$ in (38), is an *aggregate* signal-to-noise statistic. It sums the (weighted) squared magnitude of all k^2 entries of the block before comparing to the spike/slab penalty. This is the source of the parsimony and convergence gains listed above. However, it also means Block-SSVS implicitly assumes that a genuinely active lag block contributes broadly diffuse signal across its entries, rather than concentrating all of its explanatory power in a single, or a few, entries while the remaining entries are essentially noise. If a lag block is only partially active—e.g., one row of Θ_i carries genuine signal and the other row(s) are pure noise—the aggregate statistic still reflects the active row's contribution and Block-SSVS will tend to retain the whole block, correctly identifying the lag as relevant. This is at the cost of not shrinking the genuinely inactive entries within that block toward zero as aggressively as an element-wise scheme would. This is the intra-block heterogeneity trade-off, and it is the price paid for the coherence and mixing gains above. One considerations bears on its practical severity in the present setting. For SeVAR systems specifically, a lag block corresponds to a single physical mechanism (a regular AR effect, or a full seasonal cycle) acting jointly on a k-vector of physically related series (e.g., jointly driven meteorological quantities, such as the solar irradiance and air temperature). We consider it more plausible a priori that such a mechanism, if active at all, has some non-negligible loading on every series in the system than in a generic multivariate regression with unrelated predictors, which mitigates (without eliminating) the heterogeneity concern. A numerical comparison against element-wise SSVS, including a data-generating process with partial within-block signal by design, is introduced in the simulation study in Section 7, which confirms the "half-dead matrix" failure mode described above. Table 2 below summarises the conceptual trade-off to make the comparison concrete, ahead of the numerical evidence in Section 7.3.

Table 2. Conceptual comparison of coefficient-wise (element-wise) SSVS and block (group) SSVS for a $\text{SeVAR}_k(p, P)_s$ model with h candidate lags.

Property	Element-Wise SSVS	Block (Group) SSVS
Indicator space size	2^{hk^2}	2^h
Bernoulli draws per sweep	hk^2	h
Evidence statistic per indicator	One squared entry	Whole-block $\text{tr}[\boldsymbol{\mathcal{K}}_{(m)}^\top \boldsymbol{\mathcal{K}}_{(m)} \boldsymbol{\Lambda}]$
Can select a "half-dead" matrix	Yes	No, by construction
Robust to diffuse within-block signal	Yes	Yes
Robust to sparse within-block signal	Yes	Only partially
Respects multiplicative-operator coherence	Only within a scalar, not a block	Within a block (Remark 2)

Let $m = 1, \ldots, h$ index the h lags of Table 1, and let $\mathcal{I}_m \subset \{1, \ldots, q\}$ denote the set of k consecutive row-indices of $\boldsymbol{\mathcal{K}}$ (equivalently, the k columns of $\boldsymbol{X}$ spanned by the k source variables) associated with lag m, so that

$$\boldsymbol{\mathcal{K}}_{(m)} \;:=\; \boldsymbol{\mathcal{K}}[\mathcal{I}_m, :] \;\in\; \mathbb{R}^{k \times k} \tag{10}$$

is exactly the coefficient matrix of lag m—that is, $\boldsymbol{\mathcal{K}}_{(m)} = \Theta_i$ for a group 1 lag; $\boldsymbol{\mathcal{K}}_{(m)} = \Omega_j$ for a group 2 lag; and $\boldsymbol{\mathcal{K}}_{(m)}$ equal to the free cross-product matrix of group 3, associated with the pair (i, j). For each $m = 1, \ldots, h$ introduce a single binary indicator $\delta_m \in \{0, 1\}$ and collect them in $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_h)^\top \in \{0, 1\}^h$, so that the latent indicator space has dimension h rather than $q = kh$ or $qk = hk^2$.

For each lag m, let $\eta_m > 0$ denote the *spike scale* and $\zeta_m \gg 1$ a unitless *slab inflation factor*, and define the block-constant scale $g_m(\delta_m)\eta_m$ as

$$g_m(\delta_m)\eta_m = \begin{cases} \eta_m, & \delta_m = 0 \quad \text{(spike)}, \\ \zeta_m\, \eta_m, & \delta_m = 1 \quad \text{(slab)} \end{cases} \tag{11}$$

For the spike case, each entry of $\boldsymbol{\mathcal{K}}_{(m)}$ has prior SD η_m, producing strong shrinkage toward zero; in contrast, under the slab each entry has prior SD $\zeta_m \eta_m \gg \eta_m$, allowing the block to vary essentially without restriction. Let the $q \times q$ diagonal spike-and-slab scale matrix be

$$\boldsymbol{H}_{\boldsymbol{\delta}} = \text{diag}\big(\underbrace{g_1\eta_1, \ldots, g_1\eta_1}_{k \text{ rows in } \mathcal{I}_1},\ \underbrace{g_2\eta_2, \ldots, g_2\eta_2}_{k \text{ rows in } \mathcal{I}_2},\ \cdots,\ \underbrace{g_h\eta_h, \ldots, g_h\eta_h}_{k \text{ rows in } \mathcal{I}_h}\big), \tag{12}$$

i.e., all k rows of $\boldsymbol{\mathcal{K}}$ within block m share the identical scale $g_m(\delta_m)\eta_m$. When $\delta_m = 0$, the entire block $\boldsymbol{\mathcal{K}}_{(m)}$ is tightly concentrated near zero; when $\delta_m = 1$, all k^2 entries of $\boldsymbol{\mathcal{K}}_{(m)}$ are jointly free to take any appreciable value. Following George and McCulloch [18], Chipman et al. [20], typical fixed choices are $\eta_m = 0.01$ and $\zeta_m = 100$ for all m (so, the slab SD is $\zeta_m\eta_m = 1.0$). We do not standardise the regressors before applying these scales; consequently, $\boldsymbol{H}_{\boldsymbol{\delta}}$ is not scale-free, and the same numerical (η_m, ζ_m) pair implies a tighter relative penalty for a series measured on a large scale than for one on a small scale. We flag this explicitly as a modelling choice rather than a neutral default.

3.2. Prior Specification, Likelihood Function, and Joint Posterior

Prior on $\mathcal{K}$ given $(\boldsymbol{\Lambda}, \boldsymbol{\delta})$

Conditional on the precision matrix $\boldsymbol{\Lambda}$ and the block indicators $\boldsymbol{\delta} \in \{0,1\}^h$, the stacked coefficient matrix follows a matrix normal distribution,

$$\mathcal{K} \mid \boldsymbol{\Lambda}, \boldsymbol{\delta} \;\sim\; \mathcal{MN}_{q\times k}\big(\mathcal{D}_{\boldsymbol{\delta}},\; \mathcal{Q}_{\boldsymbol{\delta}}^{-1},\; \boldsymbol{\Lambda}^{-1}\big), \tag{13}$$

with density

$$f(\mathcal{K} \mid \boldsymbol{\Lambda}, \boldsymbol{\delta}) \propto |\boldsymbol{\Lambda}|^{q/2}\, |\mathcal{Q}_{\boldsymbol{\delta}}|^{k/2} \exp\Big(-\tfrac{1}{2}\operatorname{tr}\big[(\mathcal{K}-\mathcal{D}_{\boldsymbol{\delta}})^{\top}\mathcal{Q}_{\boldsymbol{\delta}}(\mathcal{K}-\mathcal{D}_{\boldsymbol{\delta}})\,\boldsymbol{\Lambda}\big]\Big), \tag{14}$$

where $\mathcal{D}_{\boldsymbol{\delta}} \in \mathbb{R}^{q\times k}$ is set to $\mathbf{0}$ (reflecting no prior preference for the sign or magnitude of any coefficient) and $\mathcal{Q}_{\boldsymbol{\delta}} = \boldsymbol{H}_{\boldsymbol{\delta}}^{-\top}\boldsymbol{V}_0^{-1}\boldsymbol{H}_{\boldsymbol{\delta}}^{-1}$ is the $q \times q$ prior precision, with $\boldsymbol{V}_0 = \boldsymbol{I}_q$ in the absence of prior information about inter-lag correlation. Because $\boldsymbol{H}_{\boldsymbol{\delta}}$ is constant within each block $\mathcal{I}_m$ (Equation (12)), so too is $\mathcal{Q}_{\boldsymbol{\delta}}$: writing $\tau_m(\delta_m) = [g_m(\delta_m)\eta_m]^{-2}$ for the common scalar precision of block m,

$$\mathcal{Q}_{\boldsymbol{\delta}} = \operatorname{diag}\big(\tau_1(\delta_1)\boldsymbol{I}_k,\; \tau_2(\delta_2)\boldsymbol{I}_k,\; \ldots,\; \tau_h(\delta_h)\boldsymbol{I}_k\big), \tag{15}$$

so that $\mathcal{Q}_{\boldsymbol{\delta}}$ is itself block-diagonal, with one $k \times k$ scaled identity block per lag m. This block structure is what allows the matrix normal density (14) to factorise across lags, and it is the key device exploited in Section 4 to derive the block-level conditional posterior of $\boldsymbol{\delta}$.

Prior on $\boldsymbol{\Lambda}$

The precision matrix is assigned a conjugate Wishart prior,

$$\boldsymbol{\Lambda} \;\sim\; \mathcal{W}_k\big(\kappa_0,\, \Xi_0\big), \qquad f(\boldsymbol{\Lambda}) \propto |\boldsymbol{\Lambda}|^{(\kappa_0-k-1)/2} \exp\big(-\tfrac{1}{2}\operatorname{tr}[\Xi_0\,\boldsymbol{\Lambda}]\big), \tag{16}$$

where $\kappa_0 > k-1$ is the prior degrees of freedom and Ξ_0 is a $k \times k$ positive-definite scale matrix. A weakly informative default is $\kappa_0 = k+1$ and $\Xi_0 = \boldsymbol{I}_k$. We note, however, that "weakly informative" describes only the marginal Wishart prior (16) on $\boldsymbol{\Lambda}$ in isolation. Because $\mathcal{K}$ is given a prior conditional on $\boldsymbol{\Lambda}$ (Equation (13)), the coefficient prior contributes its own $q = kh$ degrees of freedom to the conditional posterior of $\boldsymbol{\Lambda}$ (Theorem 2: $\boldsymbol{\Lambda} \mid \mathcal{K}, \boldsymbol{\delta}, \mathcal{S}_n \sim \mathcal{W}_k(n^* + q + \kappa_0, \cdot)$). When q is not small relative to n^*, this contribution is not negligible, and describing the overall precision-matrix prior as weakly informative is accordingly optimistic. It is the marginal prior on $\boldsymbol{\Lambda}$ that is weak, not its net effect once $\mathcal{K}$'s own prior degrees of freedom are folded in through the hierarchy.

Prior on $\boldsymbol{\delta}$

Each block indicator is independently Bernoulli,

$$p(\delta_m = 1) = \rho_m, \qquad p(\delta_m = 0) = 1-\rho_m, \qquad m = 1, \ldots, h, \tag{17}$$

with joint mass $p(\boldsymbol{\delta}) = \prod_{m=1}^{h} \rho_m^{\delta_m}(1-\rho_m)^{1-\delta_m}$. Setting $\rho_m = 1/2$ for all m assigns equal prior odds to inclusion and exclusion of each lag block, letting the data alone determine each posterior block-inclusion probability [18].

Joint Prior

Combining the three components, the joint prior density of $(\mathcal{K}, \boldsymbol{\Lambda}, \boldsymbol{\delta})$ is

$$f(\mathcal{K}, \boldsymbol{\Lambda}, \boldsymbol{\delta}) \propto p(\boldsymbol{\delta})\, |\mathcal{Q}_{\boldsymbol{\delta}}|^{k/2}\, |\boldsymbol{\Lambda}|^{(q+\kappa_0-k-1)/2} \exp\Big(-\tfrac{1}{2}\operatorname{tr}\big[\big((\mathcal{K}-\mathcal{D}_{\boldsymbol{\delta}})^{\top}\mathcal{Q}_{\boldsymbol{\delta}}(\mathcal{K}-\mathcal{D}_{\boldsymbol{\delta}}) + \Xi_0\big)\boldsymbol{\Lambda}\big]\Big), \tag{18}$$

with $\delta \in \{0,1\}^h$ and $\mathcal{Q}_\delta$ as in (15). This differs from a conventional element-wise SSVS prior only through the coarser, block-constant parametrisation of $\mathcal{Q}_\delta$; the matrix normal/Wishart conjugate form, and hence every downstream posterior derivation, is otherwise unaffected.

Likelihood Function and Its Structural Role in the Block-SSVS Hierarchy

Given the regressor matrix (8), the conditional likelihood of $(\mathcal{K}, \Lambda)$ given the observed sample $\mathcal{S}_n = [\boldsymbol{y}(1), \ldots, \boldsymbol{y}(n)]$ is the standard matrix-normal regression kernel

$$\mathcal{L}(\mathcal{K}, \Lambda \mid \mathcal{S}_n) \propto |\Lambda|^{n^*/2} \exp\Big(-\tfrac{1}{2} \operatorname{tr}\big[(\boldsymbol{Y} - \boldsymbol{X}\mathcal{K})^\top (\boldsymbol{Y} - \boldsymbol{X}\mathcal{K})\, \Lambda\big]\Big). \tag{19}$$

We deliberately retain this conventional form and place all model-specific innovation in the prior, because three structural properties of (19) are precisely what the Block-SSVS hierarchy exploits downstream.

(i) *Gram-matrix sufficiency.*
The likelihood depends on $(\mathcal{K}, \Lambda)$ only through the three Gram matrices $\boldsymbol{X}^\top \boldsymbol{X}$, $\boldsymbol{X}^\top \boldsymbol{Y}$, and $\boldsymbol{Y}^\top \boldsymbol{Y}$. The first two combine additively with their prior counterparts in $\mathcal{A}_\delta$ and $\mathcal{M}_\delta$ (Equations (22)–(23)), so the matrix-normal update of Theorem 1 holds for *every* configuration $\delta \in \{0,1\}^h$, with the block-constant precisions $\tau_m(\delta_m)$ of (15) controlling the prior–data weighting block by block. No conditional requires re-derivation as the indicators switch.

(ii) *A single shared quadratic form.*
Combined with the prior terms of (18), the exponent of (19) produces the single quadratic form $\boldsymbol{G}(\mathcal{K}, \delta)$ in (21). Completing the square in $\mathcal{K}$ yields the matrix-normal conditional of Theorem 1, while the same matrix, evaluated at the current draw of $\mathcal{K}$, is the scale of the Wishart conditional of Theorem 2. The indicators enter $\boldsymbol{G}$ only through the closed-form objects $\mathcal{Q}_\delta$ and $\mathcal{D}_\delta$, so both conditionals are available analytically for any δ.

(iii) *Independence of the likelihood from δ.*
The likelihood (19) does not contain the block indicators; data information reaches δ only through the current draw of $\mathcal{K}$ in the chain. Consequently, in the Bernoulli update of Theorem 3, the likelihood cancels exactly between numerator and denominator, and the posterior odds of each block collapse to the prior odds times the ratio of the two matrix-normal block densities (36)—the closed-form ratio (38) evaluated in $\mathcal{O}(k^3)$ per block in the proposed algorithm. This is the property that keeps the indicator sweep free of any matrix inversion or quadratic form involving $\boldsymbol{X}$.

The associated least-squares quantities,

$$\hat{\mathcal{K}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}, \qquad \boldsymbol{S}_e = (\boldsymbol{Y} - \boldsymbol{X}\hat{\mathcal{K}})^\top (\boldsymbol{Y} - \boldsymbol{X}\hat{\mathcal{K}}), \tag{20}$$

retain an operational rather than decorative role: $\hat{\mathcal{K}}$ and $\boldsymbol{S}_e/n^*$ supply the OLS initialisation of the Gibbs chain in the proposed algorithm, placing the starting values in the data-supported region of the parameter space and accelerating burn-in.

Joint Posterior

The joint posterior of $(\mathcal{K}, \Lambda, \delta)$ given $\mathcal{S}_n$ is proportional to the product of the likelihood (19) and the joint prior (18):

$$\pi(\mathcal{K}, \Lambda, \delta \mid \mathcal{S}_n) \propto p(\delta)\, |\mathcal{Q}_\delta|^{k/2}\, |\Lambda|^{(n^*+q+\kappa_0-k-1)/2} \exp\Big(-\tfrac{1}{2} \operatorname{tr}\big[\boldsymbol{G}(\mathcal{K}, \delta)\, \Lambda\big]\Big), \tag{21}$$

where $\boldsymbol{G}(\mathcal{K}, \delta) = (\boldsymbol{Y} - \boldsymbol{X}\mathcal{K})^\top (\boldsymbol{Y} - \boldsymbol{X}\mathcal{K}) + (\mathcal{K} - \mathcal{D}_\delta)^\top \mathcal{Q}_\delta (\mathcal{K} - \mathcal{D}_\delta) + \Xi_0$. This joint posterior (21) admits no closed form, as it integrates continuously over $(\mathcal{K}, \Lambda)$ and discretely over δ. Gibbs sampling [33,34] offers a practical resolution by iteratively updating from

the full conditional posteriors, which are derived in the next section. We define here three posterior summary matrices, each depending on δ through $\mathcal{Q}_\delta$:

$$\mathcal{A}_\delta = \mathcal{Q}_\delta + X^\top X, \tag{22}$$

$$\mathcal{M}_\delta = \mathcal{Q}_\delta \mathcal{D}_\delta + X^\top Y, \tag{23}$$

$$\mathcal{C}_\delta = \mathcal{D}_\delta^\top \mathcal{Q}_\delta \mathcal{D}_\delta + \Xi_0 + Y^\top Y - \mathcal{M}_\delta^\top \mathcal{A}_\delta^{-1} \mathcal{M}_\delta, \tag{24}$$

$\mathcal{A}_\delta$ is the precision-weighted combination of prior and sample information, $\mathcal{M}_\delta$ the corresponding cross-product, and $\mathcal{C}_\delta$ the posterior scale matrix.

4. Conditional Posterior Distributions

4.1. Conditional Posteriors of the Coefficient and Precision Matrices

Theorem 1 (Conditional posterior of $\mathcal{K}$)**.** *Under the likelihood* (19) *and the hierarchical prior* (18)*, the full conditional posterior of* $\mathcal{K}$ *given* $(\Lambda, \delta, \mathcal{S}_n)$ *is*

$$\mathcal{K} \mid \Lambda, \delta, \mathcal{S}_n \;\sim\; \mathcal{MN}_{q\times k}\big(\widetilde{\mathcal{K}}_\delta,\; \mathcal{A}_\delta^{-1},\; \Lambda^{-1}\big), \tag{25}$$

with conditional mean $\mathbb{E}[\mathcal{K} \mid \Lambda, \delta, \mathcal{S}_n] = \widetilde{\mathcal{K}}_\delta$ *and conditional covariance* $\mathrm{Var}[\mathrm{vec}(\mathcal{K}) \mid \Lambda, \delta, \mathcal{S}_n] = \Sigma \otimes \mathcal{A}_\delta^{-1}$, $\Sigma = \Lambda^{-1}$.

Proof. Collecting the terms of (21) that involve $\mathcal{K}$,

$$\pi(\mathcal{K} \mid \Lambda, \delta, \mathcal{S}_n) \propto \exp\Big(-\tfrac{1}{2}\,\mathrm{tr}\big[G(\mathcal{K}, \delta)\,\Lambda\big]\Big). \tag{26}$$

Expand $G(\mathcal{K}, \delta)$ in powers of $\mathcal{K}$:

$$G(\mathcal{K}, \delta) = \mathcal{K}^\top \mathcal{A}_\delta \mathcal{K} - 2\mathcal{K}^\top \mathcal{M}_\delta + Y^\top Y + \mathcal{D}_\delta^\top \mathcal{Q}_\delta \mathcal{D}_\delta + \Xi_0. \tag{27}$$

Since $\mathcal{A}_\delta \succ 0$, completing the square in $\mathcal{K}$ gives

$$G(\mathcal{K}, \delta) = (\mathcal{K} - \widetilde{\mathcal{K}}_\delta)^\top \mathcal{A}_\delta (\mathcal{K} - \widetilde{\mathcal{K}}_\delta) + \mathcal{C}_\delta, \tag{28}$$

with $\widetilde{\mathcal{K}}_\delta = \mathcal{A}_\delta^{-1} \mathcal{M}_\delta$ and $\mathcal{C}_\delta$ as in (24). Substituting (28) back,

$$\pi(\mathcal{K} \mid \Lambda, \delta, \mathcal{S}_n) \propto \exp\Big(-\tfrac{1}{2}\,\mathrm{tr}\big[(\mathcal{K} - \widetilde{\mathcal{K}}_\delta)^\top \mathcal{A}_\delta (\mathcal{K} - \widetilde{\mathcal{K}}_\delta)\,\Lambda\big]\Big), \tag{29}$$

which is the kernel of $\mathcal{MN}_{q\times k}(\widetilde{\mathcal{K}}_\delta, \mathcal{A}_\delta^{-1}, \Lambda^{-1})$. The stated moments follow from the standard matrix normal mean and covariance identities. □

Theorem 2 (Conditional posterior of Λ)**.** *The full conditional posterior of* Λ *given* $(\mathcal{K}, \delta, \mathcal{S}_n)$ *is*

$$\Lambda \mid \mathcal{K}, \delta, \mathcal{S}_n \;\sim\; \mathcal{W}_k\big(n^* + q + \kappa_0,\; G(\mathcal{K}, \delta)\big), \tag{30}$$

with conditional mean $\mathbb{E}[\Lambda \mid \mathcal{K}, \delta, \mathcal{S}_n] = (n^* + q + \kappa_0)\,[G(\mathcal{K}, \delta)]^{-1}$.

Proof. Isolating the Λ-dependent terms of (21),

$$\pi(\Lambda \mid \mathcal{K}, \delta, \mathcal{S}_n) \propto |\Lambda|^{(n^*+q+\kappa_0-k-1)/2} \exp\big(-\tfrac{1}{2}\,\mathrm{tr}[G(\mathcal{K}, \delta)\,\Lambda]\big), \tag{31}$$

which is the $\mathcal{W}_k(n^* + q + \kappa_0, G(\mathcal{K}, \delta))$ kernel. The stated mean follows from the standard Wishart expectation. □

4.2. Conditional Posterior of Each Block Indicator

Recall from (10) that $\mathcal{K}_{(m)} = \mathcal{K}[\mathcal{I}_m, :] \in \mathbb{R}^{k\times k}$ denotes the entire coefficient matrix of lag m, and that $\mathcal{Q}_\delta$ is block-diagonal with common scalar precision $\tau_m(\delta_m)$ on block m (Equation (15)). Because $\mathcal{D}_\delta = \mathbf{0}$, the exponent of the matrix normal density (14) decomposes additively across the h lags:

$$\operatorname{tr}\big[(\mathcal{K}-\mathcal{D}_\delta)^\top \mathcal{Q}_\delta(\mathcal{K}-\mathcal{D}_\delta)\,\Lambda\big] = \sum_{\ell=1}^{h} \tau_\ell(\delta_\ell)\,\operatorname{tr}\big[\mathcal{K}_{(\ell)}^\top \mathcal{K}_{(\ell)}\,\Lambda\big], \tag{32}$$

and likewise $|\mathcal{Q}_\delta|^{k/2} = \prod_{\ell=1}^{h} \tau_\ell(\delta_\ell)^{k^2/2}$, since block ℓ contributes k diagonal entries, each equal to $\tau_\ell(\delta_\ell)$. This factorisation is the formal expression of the "jointly in or jointly out" logic of Block-SSVS: only the single summand $\ell = m$ in (32) and the single factor $\tau_m(\delta_m)^{k^2/2}$ depend on δ_m.

Theorem 3 (Conditional posterior of δ_m)**.** *For each $m = 1, \ldots, h$, the full conditional posterior of δ_m given all remaining unknowns and $\mathcal{S}_n$ is Bernoulli,*

$$\delta_m \mid \mathcal{K}, \Lambda, \delta_{(-m)}, \mathcal{S}_n \;\sim\; \text{Bernoulli}\!\left(\frac{\alpha_m}{\alpha_m+\beta_m}\right), \tag{33}$$

where $\delta_{(-m)}$ collects all block indicators except δ_m,

$$\alpha_m = \rho_m\, f_m(\mathcal{K}_{(m)} \mid \Lambda, \delta_m = 1), \tag{34}$$

$$\beta_m = (1-\rho_m)\, f_m(\mathcal{K}_{(m)} \mid \Lambda, \delta_m = 0), \tag{35}$$

and

$$f_m(\mathcal{K}_{(m)} \mid \Lambda, \delta_m = d) \propto |\Lambda|^{k/2}\, \tau_m(d)^{k^2/2} \exp\!\Big(-\tfrac{1}{2}\,\tau_m(d)\,\operatorname{tr}\big[\mathcal{K}_{(m)}^\top \mathcal{K}_{(m)}\,\Lambda\big]\Big) \tag{36}$$

is the matrix normal density of the entire $k \times k$ block $\mathcal{K}_{(m)}$, evaluated at its current draw, under row-precision $\tau_m(d)\mathbf{I}_k$ and column covariance Λ^{-1}.

Proof. By Bayes' theorem applied to the conditional distribution of δ_m,

$$p(\delta_m = 1 \mid \mathcal{K}, \Lambda, \delta_{(-m)}, \mathcal{S}_n) = \frac{p(\delta_m = 1)\, f(\mathcal{K} \mid \Lambda, \delta_{(-m)}, \delta_m = 1)}{\sum_{d\in\{0,1\}} p(\delta_m = d)\, f(\mathcal{K} \mid \Lambda, \delta_{(-m)}, \delta_m = d)}, \tag{37}$$

where $f(\mathcal{K} \mid \Lambda, \cdot)$ is the full matrix normal density (14). The likelihood $\mathcal{L}(\mathcal{K}, \Lambda \mid \mathcal{S}_n)$ and every factor of the joint posterior that does not depend on δ_m are identical in numerator and denominator and cancel. By the block factorisation (32), all summands $\ell \neq m$ (and the corresponding factors of $|\mathcal{Q}_\delta|^{k/2}$) are also common to both terms, since $\delta_{(-m)}$ is held fixed, and cancel as well. What remains in each term is exactly the block density $f_m(\mathcal{K}_{(m)} \mid \Lambda, \delta_m = d)$ of (36), evaluated at $d = 1$ in the numerator and $d = 0$ in the denominator, weighted, respectively, by ρ_m and $1-\rho_m$. Identifying these two weighted terms with α_m and β_m and normalising yields the stated Bernoulli probability. □

Remark 3 (Closed-form block indicator ratio)**.** *Only the ratio α_m/β_m is needed in implementation. From (36),*

$$\frac{\alpha_m}{\beta_m} = \frac{\rho_m}{1-\rho_m} \cdot \left(\frac{\tau_m(1)}{\tau_m(0)}\right)^{k^2/2} \cdot \exp\!\Big(-\tfrac{1}{2}\big[\tau_m(1)-\tau_m(0)\big]\,\operatorname{tr}\big[\mathcal{K}_{(m)}^\top \mathcal{K}_{(m)}\,\Lambda\big]\Big). \tag{38}$$

The quantity $\mathrm{tr}[\boldsymbol{\mathcal{K}}_{(m)}^{\top}\boldsymbol{\mathcal{K}}_{(m)}\boldsymbol{\Lambda}]$ *is precisely the* $\boldsymbol{\Lambda}$*-weighted (generalised) sum of squares—the aggregate "signal"—of the entire* $k \times k$ *block* $\boldsymbol{\mathcal{K}}_{(m)}$*, so a lag is retained exactly when its whole-block evidence outweighs the slab-to-spike penalty* $[\tau_m(1) - \tau_m(0)]$*. Evaluating (38) costs* $\mathcal{O}(k^3)$ *per block (dominated by the matrix product forming the trace), so a full sweep over all* h *block indicators costs* $\mathcal{O}(hk^3)$ *per Gibbs iteration—the same floating-point order as an element-wise sweep over* $q = kh$ *rows, but requiring only* h *(rather than* $q = kh$*, or* hk^2 *under full element-wise SSVS) Bernoulli draws, which is what shrinks the combinatorial indicator space from* 2^{hk^2} *or* 2^{kh} *down to* 2^h *and improves mixing.*

Remark 4 (Posterior block-inclusion probability and best-subset selection)**.** *The quantity* $\mathrm{PIP}_m = \alpha_m/(\alpha_m + \beta_m)$ *is the posterior inclusion probability of the entire lag block* m*—that is, of the whole matrix* Θ_i*,* Ω_j*, or cross-product block, rather than of an individual coefficient. Following the standard SSVS convention [18,20], a value exceeding one-half is interpreted as evidence favouring inclusion of that block, and the set of blocks with* $\mathrm{PIP}_m > 0.50$ *defines the identified best-subset* SeVAR *model: each selected block enters with all* k^2 *of its coefficients, and each excluded block is dropped in its entirety, so the identified model is by construction free of the "half-dead" matrices that can arise under element-wise selection. Because* $\widehat{\mathrm{PIP}}_m$ *is estimated as the sampled proportion of Gibbs iterations with* $\delta_m = 1$*, model selection is obtained as a by-product of the sampler with no additional computation.*

Remark 5 (Analytic sensitivity of block selection to the hyperparameters)**.** *Equation (38) makes the dependence of block selection on the fixed hyperparameters* $(\eta_m, \zeta_m, \rho_m)$ *fully explicit, so its qualitative sensitivity can be read off directly rather than only explored numerically. Substituting* $\tau_m(1) = (\zeta_m\eta_m)^{-2}$ *and* $\tau_m(0) = \eta_m^{-2}$ *gives* $\tau_m(1)/\tau_m(0) = \zeta_m^{-2}$*, so*

$$\frac{\alpha_m}{\beta_m} = \underbrace{\frac{\rho_m}{1-\rho_m}}_{\text{prior odds}} \cdot \underbrace{\zeta_m^{-k^2}}_{\text{Occam penalty}} \cdot \exp\Big(\underbrace{\tfrac{1}{2}\eta_m^{-2}(1-\zeta_m^{-2})\ \mathrm{tr}[\boldsymbol{\mathcal{K}}_{(m)}^{\top}\boldsymbol{\mathcal{K}}_{(m)}\boldsymbol{\Lambda}]}_{\text{evidence}}\Big).$$

Three consequences follow directly, without requiring a numerical grid search. First, the prior inclusion probability ρ_m *enters only through the multiplicative prior-odds factor* $\rho_m/(1-\rho_m)$*; at the default* $\rho_m = 0.5$ *this factor is exactly* 1 *and is therefore neutral, while* $\rho_m \to 1$ *or* $\rho_m \to 0$ *shifts every block's posterior odds up or down by the same constant factor regardless of the data. So, an extreme choice of* ρ_m *can retain spurious lags or discard genuine ones irrespective of their evidence; this is why we adopt the neutral* $\rho_m = 0.5$ *throughout rather than an informative alternative. Second, the Occam penalty* $\zeta_m^{-k^2}$ *decays geometrically in the* squared *block dimension* k^2 *(not* k*), so the slab-to-spike inflation factor* ζ_m *required to retain a block of a given evidence strength grows sharply with* k*. Third, the evidence term scales with* $\eta_m^{-2}(1-\zeta_m^{-2}) \approx \eta_m^{-2}$ *for* $\zeta_m \gg 1$*, so* increasing *the spike scale* η_m *by an order of magnitude reduces the evidence exponent by two orders of magnitude, meaning inclusion decisions are far more sensitive to* η_m *than the linear-in-*ρ_m *dependence noted above. Under the values* $\eta_m = 0.01$ *and* $\zeta_m = 100$ *used throughout this paper, the implied slab standard deviation is* $\zeta_m\eta_m = 1.0$*, which is comfortably diffuse relative to the coefficient magnitudes generated in the simulation study in Section 7 without being so wide as to erase the spike–slab contrast entirely, while the spike standard deviation* $\eta_m = 0.01$ *remains tight enough to concentrate inactive blocks sharply at zero. Section 7.3(d) reports a direct, fully numerical sensitivity check against this analytic argument.*

https://doi.org/10.3390/math14214067

5. Conditional Predictive Distributions

5.1. One-Step-Ahead Conditional Predictive Distribution

Because the SeVAR model is purely autoregressive, the regressor vector at the forecast origin n, namely, $\boldsymbol{x}(n)$ formed from (8) with $t = n+1$, is fully observed. Conditional on $(\boldsymbol{\mathcal{K}}, \boldsymbol{\Lambda})$,

$$\boldsymbol{y}(n+1) \mid \boldsymbol{\mathcal{K}}, \boldsymbol{\Lambda}, \mathcal{S}_n \;\sim\; \mathcal{N}_k\big(\boldsymbol{\mathcal{K}}^\top \boldsymbol{x}(n),\; \boldsymbol{\Lambda}^{-1}\big). \tag{39}$$

Theorem 4 (One-step conditional predictive distribution)**.** *Given $(\boldsymbol{\Lambda}, \delta, \mathcal{S}_n)$, the predictive distribution obtained by integrating (39) over the conditional posterior of $\boldsymbol{\mathcal{K}}$ (Theorem 1) is*

$$\boldsymbol{y}(n+1) \mid \boldsymbol{\Lambda}, \delta, \mathcal{S}_n \;\sim\; \mathcal{N}_k\big(\widetilde{\boldsymbol{\mathcal{K}}}_\delta^\top \boldsymbol{x}(n),\; F_\delta^{-1}\boldsymbol{\Lambda}^{-1}\big), \tag{40}$$

where $F_\delta = \big(1 + \boldsymbol{x}(n)^\top \boldsymbol{\mathcal{A}}_\delta^{-1} \boldsymbol{x}(n)\big)^{-1} \in (0,1)$ is a scalar inflation factor accounting for parameter uncertainty.

Proof. Since $\boldsymbol{y}(n+1) \mid \boldsymbol{\mathcal{K}}, \boldsymbol{\Lambda} \sim \mathcal{N}_k(\boldsymbol{\mathcal{K}}^\top \boldsymbol{x}(n), \boldsymbol{\Lambda}^{-1})$ and $\boldsymbol{\mathcal{K}} \mid \boldsymbol{\Lambda}, \delta, \mathcal{S}_n \sim \mathcal{MN}(\widetilde{\boldsymbol{\mathcal{K}}}_\delta, \boldsymbol{\mathcal{A}}_\delta^{-1}, \boldsymbol{\Lambda}^{-1})$, the tower property gives

$$\mathbb{E}[\boldsymbol{y}(n+1) \mid \boldsymbol{\Lambda}, \delta, \mathcal{S}_n] = \mathbb{E}_{\boldsymbol{\mathcal{K}}}\big[\boldsymbol{\mathcal{K}}^\top \boldsymbol{x}(n)\big] = \widetilde{\boldsymbol{\mathcal{K}}}_\delta^\top \boldsymbol{x}(n), \tag{41}$$

$$\begin{aligned} \mathrm{Var}[\boldsymbol{y}(n+1) \mid \boldsymbol{\Lambda}, \delta, \mathcal{S}_n] &= \boldsymbol{\Lambda}^{-1} + \boldsymbol{x}(n)^\top \boldsymbol{\mathcal{A}}_\delta^{-1} \boldsymbol{x}(n)\, \boldsymbol{\Lambda}^{-1} \\ &= \big(1 + \boldsymbol{x}(n)^\top \boldsymbol{\mathcal{A}}_\delta^{-1} \boldsymbol{x}(n)\big)\boldsymbol{\Lambda}^{-1} = F_\delta^{-1}\boldsymbol{\Lambda}^{-1}. \end{aligned} \tag{42}$$

Since $\boldsymbol{y}(n+1)$ is a linear-Gaussian mixture over a matrix normal $\boldsymbol{\mathcal{K}}$, the resulting predictive distribution is itself normal with these moments, completing the proof. □

5.2. Multi-Step Prediction

A purely analytic r-step formula built by substituting the plug-in predictive means $\hat{\boldsymbol{y}}(n+j) = \widetilde{\boldsymbol{\mathcal{K}}}_\delta^\top \boldsymbol{x}^{(j)}(n)$ into $\boldsymbol{x}^{(r)}(n)$—the classical device of Shaarawy [16] and Broemeling and Shaarawy [35]—has an important limitation. The resulting variance inflation factor $F_\delta^{(r),-1}$ reflects only *parameter* uncertainty, entering through $\boldsymbol{\mathcal{A}}_\delta^{-1}$. It does not incorporate the innovation uncertainty that accumulates over intermediate periods $n+1, \ldots, n+r-1$, each of which contributes its own unobserved shock $\varepsilon(n+j)$. In contrast, the standard r-step VAR forecast-error variance is a growing sum of moving-average-weighted innovation covariances [2]. We therefore do not employ mean substitution in the implemented algorithm; instead, we rely on genuine stochastic forward simulation. Within each retained Gibbs sweep t, the regressor $\boldsymbol{x}^{(r)}(n)$ for horizon r is constructed from the previously drawn values $\boldsymbol{y}^{(t)}(n+1), \ldots, \boldsymbol{y}^{(t)}(n+r-1)$ rather than from their posterior means. Consequently, $\boldsymbol{y}^{(t)}(n+r)$ is drawn conditionally on one simulated trajectory, using the current parameter draw $(\boldsymbol{\mathcal{K}}^{(t)}, \boldsymbol{\Lambda}^{(t)})$. This is precisely the standard mechanism by which Bayesian VAR forecast densities are constructed [6]. No closed-form correction is required: each sweep injects a fresh $\mathcal{N}_k(\mathbf{0}, [\boldsymbol{\Lambda}^{(t)}]^{-1})$ shock at every intermediate step, and the parameters $(\boldsymbol{\mathcal{K}}^{(t)}, \boldsymbol{\Lambda}^{(t)}, \delta^{(t)})$ vary across sweeps according to their joint posterior. The empirical dispersion of $\{\boldsymbol{y}^{(t)}(n+r)\}_{t=1}^{T}$ therefore accumulates both innovation uncertainty and parameter uncertainty as r increases, without requiring a closed-form derivation. The corollary below thus states the one-step conditional law used at each recursive simulation step, conditional on the trajectory realised so far. It does not claim a closed-form marginal law for $\boldsymbol{y}(n+r)$; the true marginal predictive distribution is defined in Remark 6 and approximated by the implemented algorithm through simulation.

Corollary 1 (Conditional one-step law used at each step of the r-step recursion)**.** *For any $r \geq 1$ and any Gibbs sweep t, conditional on $(\mathcal{K}^{(t)}, \boldsymbol{\Lambda}^{(t)}, \boldsymbol{\delta}^{(t)}, \mathcal{S}_n)$ and on the already-simulated trajectory $\boldsymbol{y}^{(t)}(n+1), \ldots, \boldsymbol{y}^{(t)}(n+r-1)$ used to build $\boldsymbol{x}^{(t),(r)}(n)$ via (8),*

$$\boldsymbol{y}^{(t)}(n+r) \mid \mathcal{K}^{(t)}, \boldsymbol{\Lambda}^{(t)}, \boldsymbol{\delta}^{(t)}, \boldsymbol{x}^{(t),(r)}(n) \;\sim\; \mathcal{N}_k\big([\mathcal{K}^{(t)}]^\top \boldsymbol{x}^{(t),(r)}(n),\; [\boldsymbol{\Lambda}^{(t)}]^{-1}\big), \tag{43}$$

the direct analogue of (39) at horizon r. Marginalising (43) over the conditional posterior of $\mathcal{K}^{(t)}$ alone (holding $\boldsymbol{\Lambda}^{(t)}, \boldsymbol{\delta}^{(t)}$ and the trajectory fixed) recovers the parameter-uncertainty-inflated form

$$\boldsymbol{y}(n+r) \mid \boldsymbol{\Lambda}, \boldsymbol{\delta}, \boldsymbol{x}^{(r)}(n) \;\sim\; \mathcal{N}_k\big(\widetilde{\mathcal{K}}_{\boldsymbol{\delta}}^\top \boldsymbol{x}^{(r)}(n),\; F_{\boldsymbol{\delta}}^{(r),-1} \boldsymbol{\Lambda}^{-1}\big), \qquad F_{\boldsymbol{\delta}}^{(r)} = \big(1 + \boldsymbol{x}^{(r)}(n)^\top \boldsymbol{\mathcal{A}}_{\boldsymbol{\delta}}^{-1} \boldsymbol{x}^{(r)}(n)\big)^{-1}, \tag{44}$$

by the same argument as Theorem 4, and it is this conditional-on-trajectory form, evaluated at the drawn (not averaged) $\boldsymbol{x}^{(t),(r)}(n)$, that the proposed algorithm samples from the fully conditional law (43) at every sweep and every horizon—conditional on the current draw $\mathcal{K}^{(t)}$, not marginalised over it.

Proof. Equation (43) restates (39) with $\boldsymbol{x}(n)$ replaced by $\boldsymbol{x}^{(t),(r)}(n)$, which is a fixed, known vector once the trajectory $\boldsymbol{y}^{(t)}(n+1), \ldots, \boldsymbol{y}^{(t)}(n+r-1)$ and the current draw are held fixed; no approximation is involved at this conditional level, since $\boldsymbol{x}^{(t),(r)}(n)$ plays exactly the role of $\boldsymbol{x}(n)$ in (39). For the marginal-over-$\mathcal{K}$ form, substitute the plug-in regressor $\boldsymbol{x}^{(r)}(n)$ for the true (partially unobserved) regressor and repeat the argument of Theorem 4 verbatim. □

Within the Gibbs sampler, every retained iteration draws a complete stochastic trajectory $\{\boldsymbol{y}^{(t)}(n+r)\}_{r=1}^{r_{\max}}$ by repeated application of the conditional law (43), using the drawn (not averaged) intermediate values as described above. Averaging across iterations marginalises out $(\mathcal{K}, \boldsymbol{\Lambda}, \boldsymbol{\delta})$ jointly and, because the trajectory itself is re-simulated at every sweep, also marginalises out the sequence of intermediate innovations. So, the resulting predictive summary automatically reflects both parameter and subset-selection uncertainty and the accumulating innovation uncertainty of a genuine multi-step forecast.

Remark 6 (Fully marginal predictive distribution)**.** *The fully marginal predictive distribution $p(\boldsymbol{y}(n+r) \mid \mathcal{S}_n)$, obtained by averaging over all unknowns including the discrete indicators, is approximated by the empirical distribution of the retained Gibbs draws $\{\boldsymbol{y}^{(t)}(n+r)\}_{t=1}^{T}$. The resulting credible intervals widen automatically to reflect model-selection uncertainty, which a procedure that first selects a single best subset and only then forecasts from it cannot deliver.*

Remark 7 (Reported predictive SDs are empirical, not a plug-in formula)**.** *Theorem 4 and Corollary 1 both condition on $\boldsymbol{\Lambda}$ (or, in the stochastic recursion, use the current draw $\boldsymbol{\Lambda}^{(t)}$); the predictive SDs $\hat{\mathbf{s}}$ reported in the simulation study are the empirical standard deviations of the retained Gibbs draws $\{\boldsymbol{y}^{(t)}(n+r)\}_{t=1}^{T}$ (Remark 6), which fully marginalise over $\boldsymbol{\Lambda}$ (and $\mathcal{K}, \boldsymbol{\delta}$) rather than evaluate $F_{\boldsymbol{\delta}}^{-1}\boldsymbol{\Lambda}^{-1}$ at a single plug-in value such as the posterior mean $\hat{\boldsymbol{\Lambda}}$. Because matrix inversion is a convex map on the positive-definite cone, Jensen's inequality gives $\mathbb{E}[\boldsymbol{\Lambda}^{-1}] \succeq (\mathbb{E}[\boldsymbol{\Lambda}])^{-1}$ in the Loewner order, so a naive comparison of the reported $\hat{\mathbf{s}}$ against $\sqrt{[\hat{\boldsymbol{\Lambda}}^{-1}]_{ii}}$ computed from the* posterior mean $\hat{\boldsymbol{\Lambda}}$ *alone need not match, even under a correctly implemented sampler, and should not be read as a diagnostic of an implementation error.*

6. Gibbs Sampling Algorithm for Full Bayesian Analysis of SeVAR Models

As noted earlier, the joint posterior (21) has no closed form because it mixes continuously over $(\mathcal{K}, \boldsymbol{\Lambda})$ and discretely over $\boldsymbol{\delta}$. Gibbs sampling [33,34] resolves this by cycling through the full conditionals established in Theorems 1–3, each of which is a standard

distribution that can be drawn directly. Under mild regularity conditions the resulting chain converges to the target joint posterior [33,36]. Embedding the predictive recursion of Corollary 1 in the same sweep, as Algorithm 1 below does, means that a single chain delivers subset selection, coefficient estimation, and multi-step forecasting simultaneously—the unifying design goal of this paper.

Algorithm 1 Gibbs sampler for full Bayesian analysis of $\mathrm{SeVAR}_k(p, P)_s$

Inputs: data $\mathcal{S}_n$; maximum orders $p_{\max}, P_{\max}$; seasonal period s; forecast horizon $r_{\max}$; block hyperparameters $\kappa_0, \Xi_0, \{\eta_m\}_{m=1}^h, \{\zeta_m\}_{m=1}^h, \{\rho_m\}_{m=1}^h$; sampler design $T_{\mathrm{sim}}, T_{\mathrm{burn}}, T_{\mathrm{thin}}$.
Initialise: construct $\boldsymbol{X}, \boldsymbol{Y}$; identify the h lag row-blocks $\mathcal{I}_1, \ldots, \mathcal{I}_h$ of $\boldsymbol{X}$ (Table 1); compute OLS starting values $\boldsymbol{\mathcal{K}}^{(0)}, \boldsymbol{\Lambda}^{(0)}$; set $\boldsymbol{\delta}^{(0)} = \mathbf{1}_h$.
for $t = 1 : T_{\mathrm{sim}}$ **do**
    Update coefficient matrix: draw $\boldsymbol{\mathcal{K}}^{(t)} \sim \mathcal{MN}_{q\times k}\big(\widetilde{\boldsymbol{\mathcal{K}}}_{\boldsymbol{\delta}^{(t-1)}}, \boldsymbol{\mathcal{A}}^{-1}_{\boldsymbol{\delta}^{(t-1)}}, [\boldsymbol{\Lambda}^{(t-1)}]^{-1}\big)$ (Theorem 1).
    Update precision matrix: draw $\boldsymbol{\Lambda}^{(t)} \sim \mathcal{W}_k\big(n^* + q + \kappa_0, \boldsymbol{G}(\boldsymbol{\mathcal{K}}^{(t)}, \boldsymbol{\delta}^{(t-1)})\big)$ (Theorem 2).
    for $m = 1 : h$ **do**
        Update m-th block indicator: extract the current block $\boldsymbol{\mathcal{K}}^{(t)}_{(m)} = \boldsymbol{\mathcal{K}}^{(t)}[\mathcal{I}_m, :]$; compute α_m, β_m via (34)–(35) (or the closed-form ratio (38)); draw $\delta_m^{(t)} \sim \mathrm{Bernoulli}\big(\alpha_m/(\alpha_m + \beta_m)\big)$, applying the outcome jointly to all k^2 entries of block m (Theorem 3).
    end for
    for $r = 1 : r_{\max}$ **do**
        Update r-step prediction: build $\boldsymbol{x}^{(t),(r)}(n)$ using the previously *drawn* values $\boldsymbol{y}^{(t)}(n+j)$, $j < r$ (not their means); draw $\boldsymbol{y}^{(t)}(n+r) \sim \mathcal{N}_k\big([\boldsymbol{\mathcal{K}}^{(t)}]^\top \boldsymbol{x}^{(t),(r)}(n), [\boldsymbol{\Lambda}^{(t)}]^{-1}\big)$ (Corollary 1, conditional form (43))
    end for
end for
Post-processing: discard the first T_{burn} draws; retain every T_{thin}-th remaining iteration.
Convergence assessment: compute autocorrelations, Raftery–Lewis diagnostics, and Geweke Z-statistics for all chains.
Point estimates and credible intervals (Parameters): compute the sample mean, standard deviation, and 2.5th/97.5th percentiles over retained draws of $\boldsymbol{\mathcal{K}}^{(t)}$ and $\boldsymbol{\Lambda}^{(t)}$.
Point forecasts and predictive intervals (Outputs): for each horizon $r = 1, \ldots, r_{\max}$, compute the sample mean of the retained draws in $\boldsymbol{Y}^{(r)}_{\mathrm{pred}}$ to obtain the r-step point forecast $\hat{\boldsymbol{y}}(n+r)$, and extract the 2.5th/97.5th percentiles to obtain the 95% predictive intervals.
Subset selection: set $\widehat{\mathrm{PIP}}_m = \bar{\delta}_m$ (retained sample mean of δ_m); flag lag block m (and all k^2 of its coefficients jointly) as selected whenever $\widehat{\mathrm{PIP}}_m > 0.50$.

Algorithm Summary

Algorithm 1 synthesises the theoretical results of the previous section into a practical computational engine for joint selection, estimation, and prediction. By exploiting the conjugate structure of the prior, every full conditional is a standard distribution, entirely bypassing the need for Metropolis–Hastings tuning. Crucially, the sampler integrates subset selection, parameter estimation, and forecasting into a single iterative sweep. Rather than estimating a single selected model and subsequently forecasting from it, the r-step predictions are generated in tandem with the parameter draws. This design ensures that the uncertainty from the block inclusion indicators $\boldsymbol{\delta} \in \{0,1\}^h$, the coefficient matrix $\boldsymbol{\mathcal{K}}$, and the precision matrix $\boldsymbol{\Lambda}$ is natively propagated into the predictive distributions. To initialise the chain, ordinary least squares supplies starting values close to the data-supported region of the parameter space, specifically $\boldsymbol{\mathcal{K}}^{(0)} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y}$ and $[\boldsymbol{\Lambda}^{(0)}]^{-1} = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\mathcal{K}}^{(0)})^\top(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\mathcal{K}}^{(0)})/n^*$, while all h block indicators are set to $\boldsymbol{\delta}^{(0)} = \mathbf{1}_h$ (i.e., the full model with every lag active) [18,20]. Post-processing the retained draws then yields

exact posterior block-inclusion probabilities for lag selection, posterior means for coefficient estimation, and empirical predictive intervals for the multi-step forecasts, all derived from one unified Markov chain. We emphasise that Algorithm 1 already performs Bayesian model averaging over the 2^h candidate subset models indexed by δ. Estimation and forecasting are never conditioned on a single selected order, and the best-subset model of Remark 4 is reported solely as an interpretable summary of the sampler's output, not as the basis for the coefficient or predictive posteriors.

Convergence Diagnostics

We assess convergence with three complementary tools. Autocorrelation estimates measure serial dependence within each chain; rapid decay to zero indicates good mixing. Raftery–Lewis diagnostics [37] give a data-driven estimate of the required burn-in and total draw count, together with the inefficiency factor I-stat, with values below 5 considered satisfactory. Geweke's Z-statistic [38] compares the means of the first 10% and last 50% of each chain; $|Z| < 1.96$ indicates that stationarity has been reached. These three tools are applied throughout this paper to the continuous parameters, as is standard practice. However, we note explicitly that they are not the natural diagnostic for the *discrete* block indicators $\delta \in \{0,1\}^h$, which govern the paper's central claim of improved model-space mixing relative to element-wise selection. Discrete Bernoulli indicators can exhibit poor mixing even when every continuous parameter mixes well, since a chain can become "stuck" at one indicator configuration for long stretches while the continuous parameters conditional on that configuration continue to move freely. Two main diagnostics specific to δ are especially informative: (i) the number of distinct configurations $\delta^{(t)} \in \{0,1\}^h$ visited over the retained draws, and how this count depends on the starting configuration; and (ii) stability of the estimated $\widehat{\mathrm{PIP}}_m$ across independent chains launched from different starting points, in particular from the empty model ($\delta^{(0)} = \mathbf{0}_h$) as well as from the full model ($\delta^{(0)} = \mathbf{1}_h$), since starting exclusively from the full model could in principle bias the chain toward over-inclusion in short runs.

7. Monte Carlo Simulation Study

7.1. Experimental Design

We evaluate Algorithm 1 along three dimensions, namely, subset selection, coefficient estimation, and multi-step prediction, using a controlled Monte Carlo experiment. Data are generated from the SeVAR model under three parameter configurations, Models I–III, whose true coefficient and precision matrices are given in Table 3.

Table 3. True parameter matrices in the simulation study.

Model	Θ_1	Θ_2	Ω_1	Λ
I	$\begin{bmatrix} 0.70 & 0.50 \\ -0.30 & -0.60 \end{bmatrix}$	—	$\begin{bmatrix} 0.40 & 0.30 \\ 0.30 & 0.40 \end{bmatrix}$	$\begin{bmatrix} 1.00 & 0.30 \\ 0.30 & 1.00 \end{bmatrix}$
II	$\begin{bmatrix} 0.50 & 0.30 \\ -0.30 & -0.40 \end{bmatrix}$	—	$\begin{bmatrix} 0.40 & -0.40 \\ -0.30 & -0.80 \end{bmatrix}$	$\begin{bmatrix} 1.00 & 0.25 \\ 0.25 & 1.00 \end{bmatrix}$
III	$\begin{bmatrix} 0.70 & 0.50 \\ -0.30 & -0.60 \end{bmatrix}$	$\begin{bmatrix} -0.35 & 0.30 \\ 0.25 & -0.40 \end{bmatrix}$	$\begin{bmatrix} 0.40 & -0.30 \\ 0.40 & -0.30 \end{bmatrix}$	$\begin{bmatrix} 1.00 & -0.20 \\ -0.20 & 1.00 \end{bmatrix}$

The three configurations differ in the strength of contemporaneous cross-series dependence. Specifically, the off-diagonal element of the precision matrix is 0.3 in Model I (corresponding to a correlation of -0.3), 0.25 in Model II (correlation $= -0.25$), and -0.2 in Model III (correlation $= 0.2$). This variation allows us to assess whether performance depends on the degree of cross-series linkage. In addition, the configurations differ in

their seasonal periods, with $s = 5$, 12, and 6 in Models I–III, respectively. For each model, $R = 500$ independent bivariate series of length $n = 200$ are generated, and the maximum SeVAR orders are set to $p_{\max} = 4$ and $P_{\max} = 4$. Because every candidate model is fitted with the same maximum orders regardless of the true generating orders, the effective (post-presample) sample size available for estimation, $n^* = n - p_{\max} - P_{\max}s$, differs across designs purely through the seasonal period s. $n^* = 176$ for Model I ($s = 5$), $n^* = 148$ for Model II ($s = 12$), and $n^* = 172$ for Model III ($s = 6$). Each replicate is analysed with Algorithm 1 using $T_{\text{sim}} = 12{,}000$, $T_{\text{burn}} = 2000$, and $T_{\text{thin}} = 10$, retaining 1000 draws. Hyperparameters are $\eta_m = 0.01$ and $\zeta_m = 100$ for all m, $\rho_m = 0.5$ (equal prior odds), $\kappa_0 = k+1$, and $\Xi_0 = \boldsymbol{I}_k$. The data-generating processes for Models I–III follow the exact multiplicative operator of Definition 1: $\boldsymbol{y}(t) = \Theta_1\boldsymbol{y}(t-1) + \Omega_1\boldsymbol{y}(t-s) - \Theta_1\Omega_1\boldsymbol{y}(t-1-s) + \varepsilon(t)$ for Models I–II, and the corresponding three-lag expansion with Θ_2 added for Model III. Consequently, the cross-product blocks are genuinely active in every design, and their true values are the literal matrix products of their two active parent blocks.

Performance is evaluated as follows. **(a) Subset selection**: For each of the lag blocks and each replicate PIP is recorded and averaged over the 500 replicates, and the proportion of replicates for which the estimated orders match the true data-generating orders is calculated. **(b) Estimation**: Posterior means, standard deviations, and 95% credible interval limits are averaged across replicates, with element-wise MAE and RMSE of the posterior mean relative to the truth. **(c) Prediction**: MAE and RMSE of the multi-step predictive means at horizons $r = 1, \ldots, 5$ are computed and averaged across replicates, along with reporting the empirical 95% coverage probability (CP).

7.2. Illustration on a Single Representative Replicate from Model I

Subset selection

Table 4 reports the posterior block-inclusion probabilities for lags in one representative dataset drawn from Model I. Under Block-SSVS each PIP now summarises the joint evidence for an entire 2×2 coefficient matrix rather than for a single column. The regular block Θ_1, the seasonal block Ω_1, and their cross-product block $\Theta_1\Omega_1$—all three truly active under the generating $\text{SeVAR}_2(1,1)_5$—are retained in every retained draw ($\widehat{\text{PIP}} = 1.000$, bounded below by 0.997 at $T_{\text{ret}} = 1000$). The remaining inactive blocks are excluded in every retained draw ($\widehat{\text{PIP}} = 0.000$, bounded above by 0.003), except for the inactive cross-product block $\Theta_1\Omega_3$, which is excluded with $\widehat{\text{PIP}} = 0.031$ (entries of 1.000 (0.000) indicate that no retained draw visited the spike (slab) state. With $T_{\text{ret}} = (T_{\text{sim}} - T_{\text{burn}})/T_{\text{thin}}$ retained draws, the rule of three bounds the posterior inclusion probability at $\geq 1 - 3/T_{\text{ret}}$ ($\leq 3/T_{\text{ret}}$) at the 95% level [39]). The posterior therefore identifies exactly the true structure in this replicate—regular, seasonal, and multiplicative cross-product components alike—with no ambiguity at the 0.50 threshold. This identification is also insensitive to the initialisation of the chain. Re-running the sampler from the opposite starting point $\delta^{(0)} = \mathbf{0}_h$ reproduces the same best subset and the same $\widehat{\text{PIP}}$ values for all h blocks. This confirms that the sampler explores the indicator space without stagnation, so the extreme PIPs reflect genuine posterior concentration rather than a convergence problem in the discrete model space.

Table 4. Posterior block-inclusion probabilities (PIPs) for lag blocks in one representative replicate from Model I.

Θ_i	PIP	Ω_j	PIP	$\Theta_i\Omega_1$	PIP	$\Theta_i\Omega_2$	PIP	$\Theta_i\Omega_3$	PIP	$\Theta_i\Omega_4$	PIP
Θ_1	1.000	Ω_1	1.000	$\Theta_1\Omega_1$	1.000	$\Theta_1\Omega_2$	0.000	$\Theta_1\Omega_3$	0.031	$\Theta_1\Omega_4$	0.000
Θ_2	0.000	Ω_2	0.000	$\Theta_2\Omega_1$	0.000	$\Theta_2\Omega_2$	0.000	$\Theta_2\Omega_3$	0.000	$\Theta_2\Omega_4$	0.000
Θ_3	0.000	Ω_3	0.000	$\Theta_3\Omega_1$	0.000	$\Theta_3\Omega_2$	0.000	$\Theta_3\Omega_3$	0.000	$\Theta_3\Omega_4$	0.000
Θ_4	0.000	Ω_4	0.000	$\Theta_4\Omega_1$	0.000	$\Theta_4\Omega_2$	0.000	$\Theta_4\Omega_3$	0.000	$\Theta_4\Omega_4$	0.000

Estimation

Table 5 presents posterior summaries of the coefficient and precision matrices. The posterior means closely track the true values, with the largest element-wise deviation being 0.066, and every true entry lying within its 95% credible interval. Moreover, the 95% credible intervals for the inactive lags consistently include zero. The diagnostics in Table 6 indicate satisfactory MCMC convergence. The Geweke z-scores fall inside $(-1.96, 1.96)$, the Raftery–Lewis test recommends about 1000 iterations, and the thinned chain exhibits negligible autocorrelation across parameters. This assessment is additionally confirmed by the trace and marginal posterior plots in Figures 1 and 2.

Table 5. Posterior summaries of the coefficient and precision matrices for one representative dataset from Model I. $\hat{\mu}$: posterior mean; $\hat{s}$: posterior SD; $(\hat{L}, \hat{U})$: 95% credible interval.

Param.	True	$\hat{\mu}$	$\hat{s}$	$\hat{L}$	$\hat{U}$
Θ_1	$\begin{bmatrix} 0.70 & 0.50 \\ -0.30 & -0.60 \end{bmatrix}$	$\begin{bmatrix} 0.747 & 0.566 \\ -0.306 & -0.635 \end{bmatrix}$	$\begin{bmatrix} 0.060 & 0.069 \\ 0.061 & 0.069 \end{bmatrix}$	$\begin{bmatrix} 0.636 & 0.438 \\ -0.425 & -0.773 \end{bmatrix}$	$\begin{bmatrix} 0.862 & 0.710 \\ -0.190 & -0.499 \end{bmatrix}$
Θ_2	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} 0.002 & -0.002 \\ -0.002 & 0.002 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.011 \\ 0.010 & 0.011 \end{bmatrix}$	$\begin{bmatrix} -0.018 & -0.024 \\ -0.023 & -0.018 \end{bmatrix}$	$\begin{bmatrix} 0.021 & 0.019 \\ 0.019 & 0.023 \end{bmatrix}$
Θ_3	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} 0.001 & 0.000 \\ 0.003 & -0.003 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.010 \\ 0.011 & 0.011 \end{bmatrix}$	$\begin{bmatrix} -0.020 & -0.020 \\ -0.018 & -0.024 \end{bmatrix}$	$\begin{bmatrix} 0.020 & 0.020 \\ 0.026 & 0.018 \end{bmatrix}$
Θ_4	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} -0.001 & 0.000 \\ 0.004 & -0.001 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.010 \\ 0.011 & 0.011 \end{bmatrix}$	$\begin{bmatrix} -0.021 & -0.020 \\ -0.017 & -0.021 \end{bmatrix}$	$\begin{bmatrix} 0.019 & 0.019 \\ 0.024 & 0.020 \end{bmatrix}$
Ω_1	$\begin{bmatrix} 0.40 & 0.30 \\ 0.30 & 0.40 \end{bmatrix}$	$\begin{bmatrix} 0.450 & 0.269 \\ 0.256 & 0.330 \end{bmatrix}$	$\begin{bmatrix} 0.061 & 0.062 \\ 0.063 & 0.062 \end{bmatrix}$	$\begin{bmatrix} 0.330 & 0.154 \\ 0.132 & 0.209 \end{bmatrix}$	$\begin{bmatrix} 0.569 & 0.391 \\ 0.375 & 0.453 \end{bmatrix}$
Ω_2	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} -0.001 & 0.001 \\ 0.001 & -0.000 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.011 \\ 0.011 & 0.010 \end{bmatrix}$	$\begin{bmatrix} -0.021 & -0.019 \\ -0.020 & -0.020 \end{bmatrix}$	$\begin{bmatrix} 0.021 & 0.022 \\ 0.022 & 0.022 \end{bmatrix}$
Ω_3	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} 0.004 & -0.003 \\ 0.000 & 0.000 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.010 \\ 0.010 & 0.011 \end{bmatrix}$	$\begin{bmatrix} -0.015 & -0.023 \\ -0.021 & -0.020 \end{bmatrix}$	$\begin{bmatrix} 0.024 & 0.017 \\ 0.021 & 0.022 \end{bmatrix}$
Ω_4	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} -0.002 & 0.002 \\ 0.002 & -0.003 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.010 \\ 0.011 & 0.010 \end{bmatrix}$	$\begin{bmatrix} -0.023 & -0.018 \\ -0.019 & -0.024 \end{bmatrix}$	$\begin{bmatrix} 0.018 & 0.024 \\ 0.023 & 0.017 \end{bmatrix}$
$\Theta_1\Omega_1$	$\begin{bmatrix} 0.43 & 0.41 \\ -0.30 & -0.33 \end{bmatrix}$	$\begin{bmatrix} 0.456 & 0.424 \\ -0.298 & -0.312 \end{bmatrix}$	$\begin{bmatrix} 0.070 & 0.073 \\ 0.071 & 0.074 \end{bmatrix}$	$\begin{bmatrix} 0.317 & 0.279 \\ -0.441 & -0.466 \end{bmatrix}$	$\begin{bmatrix} 0.589 & 0.559 \\ -0.163 & -0.170 \end{bmatrix}$
$\Theta_2\Omega_1$	$\begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \end{bmatrix}$	$\begin{bmatrix} -0.002 & 0.000 \\ 0.002 & -0.003 \end{bmatrix}$	$\begin{bmatrix} 0.010 & 0.010 \\ 0.011 & 0.010 \end{bmatrix}$	$\begin{bmatrix} -0.022 & -0.020 \\ -0.019 & -0.024 \end{bmatrix}$	$\begin{bmatrix} 0.017 & 0.019 \\ 0.021 & 0.017 \end{bmatrix}$
Λ	$\begin{bmatrix} 1.00 & 0.30 \\ 0.30 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.993 & 0.262 \\ 0.262 & 0.963 \end{bmatrix}$	$\begin{bmatrix} 0.107 & 0.075 \\ 0.075 & 0.103 \end{bmatrix}$	$\begin{bmatrix} 0.802 & 0.122 \\ 0.122 & 0.777 \end{bmatrix}$	$\begin{bmatrix} 1.211 & 0.415 \\ 0.415 & 1.181 \end{bmatrix}$

Table 6. Convergence diagnostics of the MCMC chain for one representative dataset from Model I.

	Raftery–Lewis Diagnostic			Geweke Z-Statistic		Autocorrelations			
Param	*Ntotal*	*Nmin*	I-Stat	Z-Score	p-Value	Lag 1	Lag 5	Lag 10	Lag 50
$\Theta_{1,(1,1)}$	893	937	0.953	−0.664	0.507	0.014	−0.008	0.011	0.048
$\Theta_{1,(2,1)}$	969	937	1.034	−1.405	0.160	0.015	−0.004	−0.013	0.043
$\Theta_{1,(1,2)}$	1053	937	1.124	−0.643	0.520	0.050	0.011	−0.022	−0.008
$\Theta_{1,(2,2)}$	1143	937	1.220	0.570	0.569	0.029	−0.070	−0.004	0.043
$\Omega_{1,(1,1)}$	893	937	0.953	0.579	0.562	0.010	−0.005	0.087	0.001
$\Omega_{1,(2,1)}$	893	937	0.953	−1.879	0.060	0.006	0.013	−0.030	−0.025
$\Omega_{1,(1,2)}$	893	937	0.953	1.316	0.188	0.044	−0.008	−0.018	−0.036
$\Omega_{1,(2,2)}$	893	937	0.953	−0.670	0.503	−0.010	0.013	0.013	−0.008
$\Theta_1\Omega_{1,(1,1)}$	893	937	0.953	0.464	0.643	0.069	0.005	0.061	0.015
$\Theta_2\Omega_{1,(2,1)}$	893	937	0.953	1.787	0.074	0.022	0.012	0.007	−0.057
$\Theta_3\Omega_{1,(1,2)}$	893	937	0.953	−0.159	0.874	0.042	0.022	−0.028	−0.053

Table 6. *Cont.*

	Raftery–Lewis Diagnostic			Geweke Z-Statistic		Autocorrelations			
Param	***Ntotal***	***Nmin***	**I-Stat**	**Z-Score**	***p*-Value**	**Lag 1**	**Lag 5**	**Lag 10**	**Lag 50**
$\Theta_4\Omega_{1,(2,2)}$	893	937	0.953	1.000	0.317	0.050	0.009	0.020	0.004
$\Theta_1\Omega_{2,(1,1)}$	969	937	1.034	−0.850	0.395	−0.045	0.018	−0.050	−0.045
$\Theta_2\Omega_{2,(2,1)}$	969	937	1.034	0.697	0.486	0.004	−0.025	−0.020	−0.014
$\Theta_3\Omega_{2,(1,2)}$	1143	937	1.220	0.477	0.633	0.025	−0.004	0.018	−0.074
$\Theta_4\Omega_{2,(2,2)}$	893	937	0.953	0.343	0.732	−0.027	−0.017	0.004	−0.001
$\mathbf{\Lambda}_{(1,1)}$	969	937	1.034	1.588	0.112	−0.025	−0.061	0.006	−0.010
$\mathbf{\Lambda}_{(2,1)}$	893	937	0.953	0.346	0.729	−0.011	−0.022	−0.024	−0.004
$\mathbf{\Lambda}_{(1,2)}$	893	937	0.953	0.346	0.729	−0.011	−0.022	−0.024	−0.004
$\mathbf{\Lambda}_{(2,2)}$	893	937	0.953	0.526	0.599	0.001	0.035	0.012	0.000

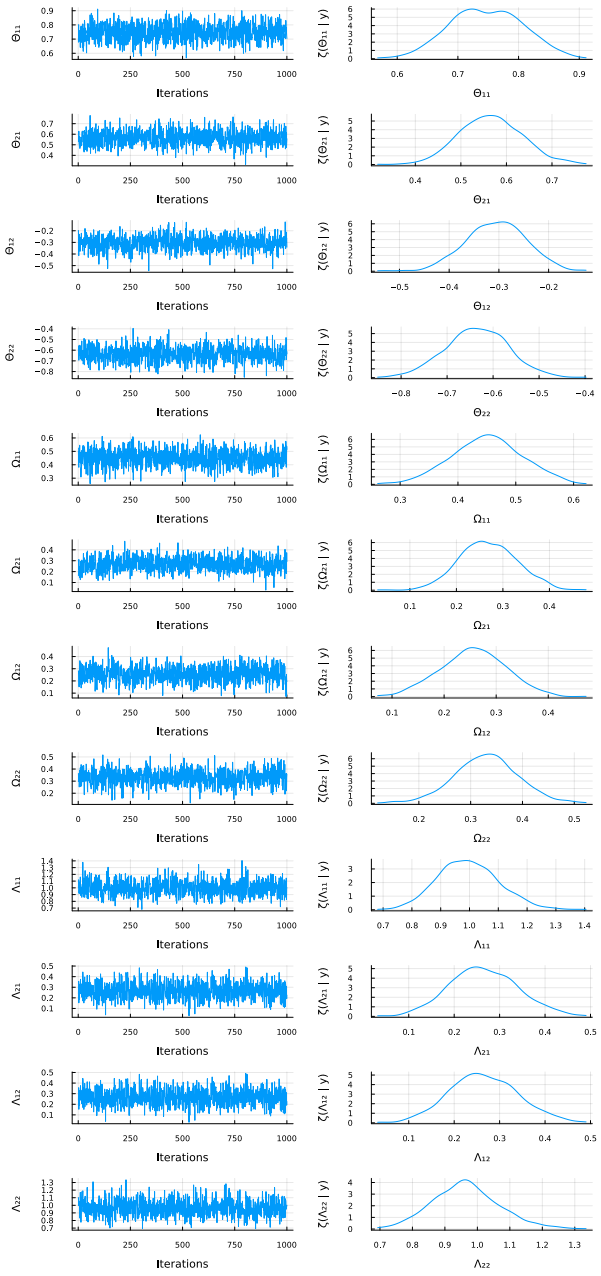


Figure 1. Trace plots and marginal posterior distributions of active primitive lags and precision for one representative dataset from Model I.

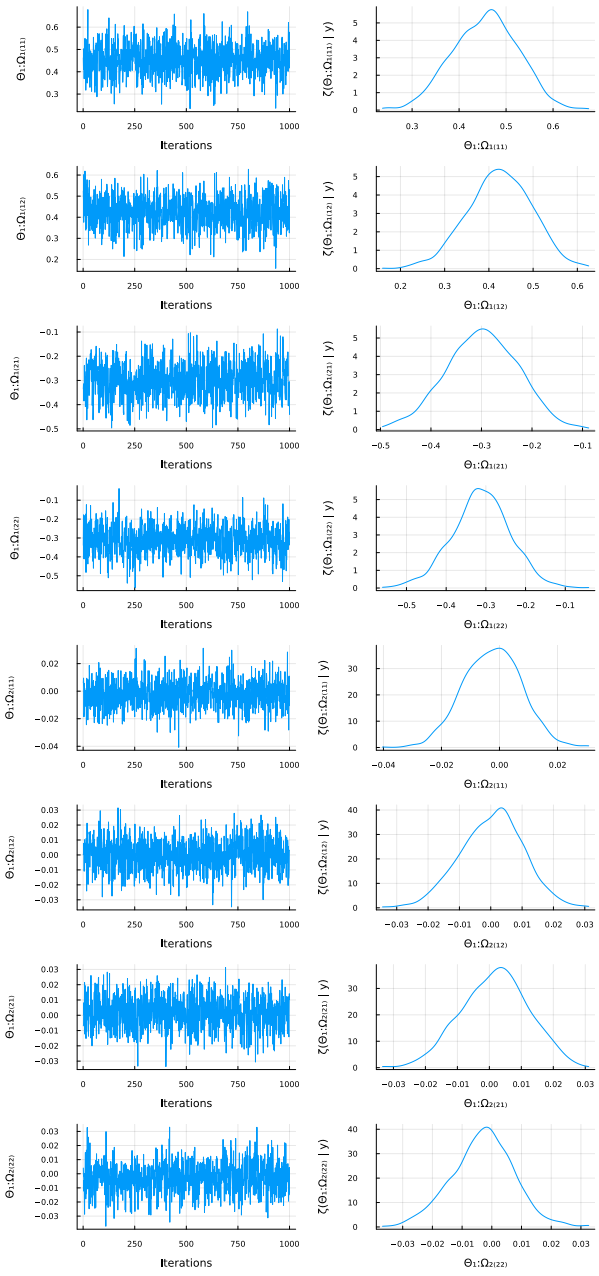


Figure 2. Trace plots and marginal posterior distributions of some cross-product lags for one representative dataset from Model I.

Prediction

Table 7 presents the multi-step predictive summaries. The realised value falls inside its reported 95% interval at every one of the ten (horizon $\times$ series) combinations in this replicate. The point forecasts $\hat{\boldsymbol{\mu}}$ track the pattern of the realised trajectory without reproducing it exactly. For instance, at $r = 1$ the realised value for series 1 (-2.575) is considerably more negative than the forecast mean (-1.189), a standardised deviation of about 1.28 posterior predictive standard deviations, the largest such deviation across all ten combinations and still well inside the interval $[-3.218, 0.886]$. This is expected as a single realisation reflects one draw of the innovation $\varepsilon(n+1)$, which the model cannot anticipate, and a deviation of this size is unremarkable under a correctly calibrated 95% interval. The posterior predictive standard deviations $\hat{\boldsymbol{s}}$ widen only moderately across the horizon, from about 1.08–1.09 at $r = 1$ to about 1.31–1.46 at $r = 5$, consistent with the stochastic forward-simulation recursion of Section 5.2 correctly accumulating both parameter and innovation uncertainty as

the horizon grows. This single replicate is, of course, only illustrative; Section 7.3(c) below reports the empirical coverage probability aggregated over all $R = 500$ replicates, which is the appropriate basis for assessing calibration formally rather than any one realisation.

Table 7. Multi-step predictive summaries for one representative dataset from Model I.

$y(n+r)$	True	$\hat{\mu}$	$\hat{s}$	$\hat{L}$	$\hat{U}$
$\begin{bmatrix} y(n+1) \\ y(n+2) \\ y(n+3) \\ y(n+4) \\ y(n+5) \end{bmatrix}$	$\begin{bmatrix} -2.575 & 1.116 \\ -0.100 & -0.993 \\ -0.966 & 1.734 \\ -0.983 & 1.127 \\ 0.164 & -1.201 \end{bmatrix}$	$\begin{bmatrix} -1.189 & 0.055 \\ -0.127 & 0.200 \\ -0.311 & 0.256 \\ -0.027 & 0.353 \\ -0.854 & -0.416 \end{bmatrix}$	$\begin{bmatrix} 1.081 & 1.094 \\ 1.296 & 1.321 \\ 1.439 & 1.335 \\ 1.421 & 1.337 \\ 1.463 & 1.313 \end{bmatrix}$	$\begin{bmatrix} -3.218 & -1.933 \\ -2.760 & -2.368 \\ -3.084 & -2.446 \\ -2.848 & -2.185 \\ -3.736 & -3.143 \end{bmatrix}$	$\begin{bmatrix} 0.886 & 2.275 \\ 2.153 & 2.831 \\ 2.641 & 2.735 \\ 2.994 & 3.030 \\ 1.985 & 2.033 \end{bmatrix}$

7.3. Aggregated Monte Carlo Results

(a) Subset-selection performance

Table 8 reports aggregated Block-SSVS subset-selection results in terms of mean posterior block-inclusion probabilities.

Table 8. Aggregated Block-SSVS subset-selection results in the simulated SeVAR models: mean posterior block-inclusion probability ($\overline{\text{PIP}}$) for each lag block.

Θ_i	$\overline{\text{PIP}}$	Ω_j	$\overline{\text{PIP}}$	$\Theta_i\Omega_1$	$\overline{\text{PIP}}$	$\Theta_i\Omega_2$	$\overline{\text{PIP}}$	$\Theta_i\Omega_3$	$\overline{\text{PIP}}$	$\Theta_i\Omega_4$	$\overline{\text{PIP}}$
Model I											
Θ_1	0.999	Ω_1	0.999	$\Theta_1\Omega_1$	0.988	$\Theta_1\Omega_2$	0.023	$\Theta_1\Omega_3$	0.014	$\Theta_1\Omega_4$	0.012
Θ_2	0.016	Ω_2	0.034	$\Theta_2\Omega_1$	0.022	$\Theta_2\Omega_2$	0.019	$\Theta_2\Omega_3$	0.007	$\Theta_2\Omega_4$	0.008
Θ_3	0.007	Ω_3	0.007	$\Theta_3\Omega_1$	0.008	$\Theta_3\Omega_2$	0.012	$\Theta_3\Omega_3$	0.007	$\Theta_3\Omega_4$	0.009
Θ_4	0.008	Ω_4	0.006	$\Theta_4\Omega_1$	0.004	$\Theta_4\Omega_2$	0.007	$\Theta_4\Omega_3$	0.005	$\Theta_4\Omega_4$	0.006
Model II											
Θ_1	0.997	Ω_1	0.999	$\Theta_1\Omega_1$	0.994	$\Theta_1\Omega_2$	0.052	$\Theta_1\Omega_3$	0.036	$\Theta_1\Omega_4$	0.024
Θ_2	0.023	Ω_2	0.065	$\Theta_2\Omega_1$	0.035	$\Theta_2\Omega_2$	0.030	$\Theta_2\Omega_3$	0.029	$\Theta_2\Omega_4$	0.017
Θ_3	0.017	Ω_3	0.031	$\Theta_3\Omega_1$	0.024	$\Theta_3\Omega_2$	0.020	$\Theta_3\Omega_3$	0.020	$\Theta_3\Omega_4$	0.015
Θ_4	0.012	Ω_4	0.031	$\Theta_4\Omega_1$	0.021	$\Theta_4\Omega_2$	0.020	$\Theta_4\Omega_3$	0.025	$\Theta_4\Omega_4$	0.012
Model III											
Θ_1	0.999	Ω_1	0.999	$\Theta_1\Omega_1$	0.999	$\Theta_1\Omega_2$	0.018	$\Theta_1\Omega_3$	0.008	$\Theta_1\Omega_4$	0.003
Θ_2	0.998	Ω_2	0.012	$\Theta_2\Omega_1$	0.046	$\Theta_2\Omega_2$	0.014	$\Theta_2\Omega_3$	0.006	$\Theta_2\Omega_4$	0.003
Θ_3	0.021	Ω_3	0.005	$\Theta_3\Omega_1$	0.011	$\Theta_3\Omega_2$	0.002	$\Theta_3\Omega_3$	0.005	$\Theta_3\Omega_4$	0.008
Θ_4	0.013	Ω_4	0.004	$\Theta_4\Omega_1$	0.012	$\Theta_4\Omega_2$	0.007	$\Theta_4\Omega_3$	0.006	$\Theta_4\Omega_4$	0.001

From Table 8, for the regular lags, Θ_1 (lag 1) is consistently selected across all three models, with $\overline{\text{PIP}}$ values exceeding 0.99, confirming its strong structural relevance. By contrast, higher-order regular lags (Θ_2–Θ_4) show negligible inclusion probabilities in Models I and II, whereas Θ_2 is retained with near certainty ($\overline{\text{PIP}} = 0.998$) under Model III, exactly as it should be, since Θ_2 is genuinely active in that design's generating process. For the seasonal lags, Ω_1 (lag s) is strongly supported across all models, while longer seasonal lags (Ω_2–Ω_4) remain essentially excluded, with probabilities close to zero. The cross-product blocks tell a correspondingly clean story once their generating status is taken into account (Section 7.1). The genuinely active block $\Theta_1\Omega_1$ is retained with $\overline{\text{PIP}} \geq 0.988$ in every one of the three designs, on a par with Θ_1 and Ω_1 themselves, while every cross-product block that is either zero or negligible in the true model—including $\Theta_2\Omega_1$ in Model III, whose true entries are an order of magnitude smaller than $\Theta_1\Omega_1$'s—remains below 0.065. The one partial exception is $\Theta_2\Omega_1$ itself ($\overline{\text{PIP}} = 0.046$): correctly excluded on average, but a genuine (if practically negligible) non-zero true value. So, this is a case where the procedure

conservatively treats a true signal as noise once its magnitude falls near the sampling floor at $n = 200$, rather than a false exclusion of a substantively important block. Overall, the results highlight a parsimonious lag structure dominated by the first regular and first seasonal components, together with their multiplicative cross-product, with only limited evidence for additional lag dependence under Model III. Importantly, these selection patterns align closely with the true generating SeVAR orders, confirming that the Block-SSVS procedure successfully recovers the underlying SeVAR model structure, including the genuinely active cross-product block that the exact multiplicative data-generating process implies. Moreover, across all cases the percentage of correctly identified SeVAR models is not less than 82%, underscoring the robustness and reliability of the selection mechanism.

(b) Estimation performance

Tables 9 and 10 report the aggregated posterior summaries and evaluation metrics of coefficient and precision matrices, restricted to the active lags in the simulated SeVAR Models I–III.

Table 9. Aggregated posterior summaries of coefficient and precision matrices for the active lags in the simulated SeVAR models.

Param.	True	$\hat{\mu}$	$\hat{s}$	$\hat{L}$	$\hat{U}$
			Model I		
Θ_1	$\begin{bmatrix} 0.70 & 0.50 \\ -0.30 & -0.60 \end{bmatrix}$	$\begin{bmatrix} 0.678 & 0.489 \\ -0.300 & -0.592 \end{bmatrix}$	$\begin{bmatrix} 0.064 & 0.071 \\ 0.065 & 0.070 \end{bmatrix}$	$\begin{bmatrix} 0.556 & 0.355 \\ -0.428 & -0.730 \end{bmatrix}$	$\begin{bmatrix} 0.802 & 0.633 \\ -0.175 & -0.456 \end{bmatrix}$
Ω_1	$\begin{bmatrix} 0.40 & 0.30 \\ 0.30 & 0.40 \end{bmatrix}$	$\begin{bmatrix} 0.377 & 0.287 \\ 0.294 & 0.381 \end{bmatrix}$	$\begin{bmatrix} 0.063 & 0.064 \\ 0.063 & 0.063 \end{bmatrix}$	$\begin{bmatrix} 0.253 & 0.167 \\ 0.171 & 0.253 \end{bmatrix}$	$\begin{bmatrix} 0.500 & 0.414 \\ 0.415 & 0.505 \end{bmatrix}$
$\Theta_1\Omega_1$	$\begin{bmatrix} 0.43 & 0.41 \\ -0.30 & -0.33 \end{bmatrix}$	$\begin{bmatrix} 0.418 & 0.399 \\ -0.289 & -0.312 \end{bmatrix}$	$\begin{bmatrix} 0.073 & 0.075 \\ 0.073 & 0.074 \end{bmatrix}$	$\begin{bmatrix} 0.271 & 0.250 \\ -0.436 & -0.466 \end{bmatrix}$	$\begin{bmatrix} 0.555 & 0.538 \\ -0.152 & -0.169 \end{bmatrix}$
$\mathbf{\Lambda}$	$\begin{bmatrix} 1.00 & 0.30 \\ 0.30 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 1.063 & 0.312 \\ 0.312 & 1.071 \end{bmatrix}$	$\begin{bmatrix} 0.115 & 0.083 \\ 0.083 & 0.116 \end{bmatrix}$	$\begin{bmatrix} 0.858 & 0.160 \\ 0.160 & 0.862 \end{bmatrix}$	$\begin{bmatrix} 1.304 & 0.485 \\ 0.485 & 1.321 \end{bmatrix}$
			Model II		
Θ_1	$\begin{bmatrix} 0.50 & 0.30 \\ -0.30 & -0.40 \end{bmatrix}$	$\begin{bmatrix} 0.472 & 0.289 \\ -0.290 & -0.388 \end{bmatrix}$	$\begin{bmatrix} 0.077 & 0.079 \\ 0.076 & 0.080 \end{bmatrix}$	$\begin{bmatrix} 0.324 & 0.136 \\ -0.444 & -0.545 \end{bmatrix}$	$\begin{bmatrix} 0.623 & 0.445 \\ -0.142 & -0.235 \end{bmatrix}$
Ω_1	$\begin{bmatrix} 0.40 & -0.40 \\ -0.30 & -0.80 \end{bmatrix}$	$\begin{bmatrix} 0.393 & -0.397 \\ -0.292 & -0.788 \end{bmatrix}$	$\begin{bmatrix} 0.063 & 0.050 \\ 0.063 & 0.048 \end{bmatrix}$	$\begin{bmatrix} 0.272 & -0.495 \\ -0.413 & -0.883 \end{bmatrix}$	$\begin{bmatrix} 0.515 & -0.300 \\ -0.167 & -0.693 \end{bmatrix}$
$\Theta_1\Omega_1$	$\begin{bmatrix} 0.11 & -0.44 \\ 0.00 & 0.44 \end{bmatrix}$	$\begin{bmatrix} 0.108 & -0.419 \\ 0.004 & 0.423 \end{bmatrix}$	$\begin{bmatrix} 0.070 & 0.090 \\ 0.072 & 0.093 \end{bmatrix}$	$\begin{bmatrix} -0.029 & -0.599 \\ -0.137 & 0.239 \end{bmatrix}$	$\begin{bmatrix} 0.244 & -0.245 \\ 0.140 & 0.600 \end{bmatrix}$
$\mathbf{\Lambda}$	$\begin{bmatrix} 1.00 & 0.25 \\ 0.25 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 1.098 & 0.262 \\ 0.262 & 1.092 \end{bmatrix}$	$\begin{bmatrix} 0.133 & 0.092 \\ 0.092 & 0.130 \end{bmatrix}$	$\begin{bmatrix} 0.863 & 0.093 \\ 0.093 & 0.860 \end{bmatrix}$	$\begin{bmatrix} 1.381 & 0.453 \\ 0.453 & 1.372 \end{bmatrix}$
			Model III		
Θ_1	$\begin{bmatrix} 0.70 & 0.50 \\ -0.30 & -0.60 \end{bmatrix}$	$\begin{bmatrix} 0.689 & 0.505 \\ -0.310 & -0.579 \end{bmatrix}$	$\begin{bmatrix} 0.061 & 0.064 \\ 0.060 & 0.063 \end{bmatrix}$	$\begin{bmatrix} 0.573 & 0.378 \\ -0.431 & -0.704 \end{bmatrix}$	$\begin{bmatrix} 0.811 & 0.631 \\ -0.195 & -0.456 \end{bmatrix}$
Θ_2	$\begin{bmatrix} -0.35 & 0.30 \\ 0.25 & -0.40 \end{bmatrix}$	$\begin{bmatrix} -0.339 & 0.311 \\ 0.279 & -0.360 \end{bmatrix}$	$\begin{bmatrix} 0.056 & 0.060 \\ 0.055 & 0.059 \end{bmatrix}$	$\begin{bmatrix} -0.447 & 0.194 \\ 0.168 & -0.475 \end{bmatrix}$	$\begin{bmatrix} -0.232 & 0.429 \\ 0.389 & -0.244 \end{bmatrix}$
Ω_1	$\begin{bmatrix} 0.40 & -0.30 \\ 0.40 & -0.30 \end{bmatrix}$	$\begin{bmatrix} 0.393 & -0.301 \\ 0.386 & -0.284 \end{bmatrix}$	$\begin{bmatrix} 0.055 & 0.054 \\ 0.056 & 0.053 \end{bmatrix}$	$\begin{bmatrix} 0.283 & -0.412 \\ 0.275 & -0.387 \end{bmatrix}$	$\begin{bmatrix} 0.498 & -0.198 \\ 0.490 & -0.179 \end{bmatrix}$
$\Theta_1\Omega_1$	$\begin{bmatrix} 0.48 & -0.36 \\ -0.36 & 0.27 \end{bmatrix}$	$\begin{bmatrix} 0.476 & -0.360 \\ -0.359 & 0.255 \end{bmatrix}$	$\begin{bmatrix} 0.067 & 0.060 \\ 0.068 & 0.061 \end{bmatrix}$	$\begin{bmatrix} 0.342 & -0.481 \\ -0.498 & 0.131 \end{bmatrix}$	$\begin{bmatrix} 0.606 & -0.242 \\ -0.226 & 0.368 \end{bmatrix}$
$\Theta_2\Omega_1$	$\begin{bmatrix} -0.02 & 0.015 \\ -0.06 & 0.045 \end{bmatrix}$	$\begin{bmatrix} -0.003 & 0.001 \\ -0.005 & 0.006 \end{bmatrix}$	$\begin{bmatrix} 0.016 & 0.015 \\ 0.017 & 0.016 \end{bmatrix}$	$\begin{bmatrix} -0.033 & -0.028 \\ -0.040 & -0.022 \end{bmatrix}$	$\begin{bmatrix} 0.026 & 0.031 \\ 0.023 & 0.040 \end{bmatrix}$
$\mathbf{\Lambda}$	$\begin{bmatrix} 1.00 & -0.20 \\ -0.20 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 1.077 & -0.207 \\ -0.207 & 1.086 \end{bmatrix}$	$\begin{bmatrix} 0.118 & 0.082 \\ 0.082 & 0.118 \end{bmatrix}$	$\begin{bmatrix} 0.867 & -0.369 \\ -0.369 & 0.865 \end{bmatrix}$	$\begin{bmatrix} 1.324 & -0.051 \\ -0.051 & 1.324 \end{bmatrix}$

From Tables 9 and 10, we observe that posterior means for Θ_1, Θ_2, Ω_1, and $\mathbf{\Lambda}$ are close to unbiased across all SeVAR models, with the largest absolute deviation from the true value not exceeding roughly 1.2 posterior standard deviations (element $(2,2)$ of Θ_2 in Model III). Every true entry falls comfortably within its reported 95% credible interval.

Table 10. Evaluation metrics of coefficient and precision matrices for the active lags in the simulated SeVAR models.

Metric	Θ_1	Θ_2	Ω_1	$\Theta_1\Omega_1$	$\Theta_2\Omega_1$	Λ
			Model I			
MAE	$\begin{bmatrix} 0.055 & 0.058 \\ 0.055 & 0.055 \end{bmatrix}$	—	$\begin{bmatrix} 0.055 & 0.055 \\ 0.054 & 0.056 \end{bmatrix}$	$\begin{bmatrix} 0.060 & 0.062 \\ 0.066 & 0.061 \end{bmatrix}$	—	$\begin{bmatrix} 0.106 & 0.066 \\ 0.066 & 0.105 \end{bmatrix}$
RMSE	$\begin{bmatrix} 0.071 & 0.074 \\ 0.070 & 0.069 \end{bmatrix}$	—	$\begin{bmatrix} 0.071 & 0.069 \\ 0.069 & 0.071 \end{bmatrix}$	$\begin{bmatrix} 0.083 & 0.086 \\ 0.086 & 0.080 \end{bmatrix}$	—	$\begin{bmatrix} 0.136 & 0.083 \\ 0.083 & 0.136 \end{bmatrix}$
			Model II			
MAE	$\begin{bmatrix} 0.069 & 0.063 \\ 0.065 & 0.065 \end{bmatrix}$	—	$\begin{bmatrix} 0.054 & 0.040 \\ 0.050 & 0.039 \end{bmatrix}$	$\begin{bmatrix} 0.060 & 0.075 \\ 0.058 & 0.078 \end{bmatrix}$	—	$\begin{bmatrix} 0.132 & 0.079 \\ 0.079 & 0.125 \end{bmatrix}$
RMSE	$\begin{bmatrix} 0.091 & 0.080 \\ 0.081 & 0.083 \end{bmatrix}$	—	$\begin{bmatrix} 0.069 & 0.053 \\ 0.063 & 0.053 \end{bmatrix}$	$\begin{bmatrix} 0.075 & 0.096 \\ 0.071 & 0.100 \end{bmatrix}$	—	$\begin{bmatrix} 0.172 & 0.099 \\ 0.099 & 0.161 \end{bmatrix}$
			Model III			
MAE	$\begin{bmatrix} 0.049 & 0.053 \\ 0.051 & 0.052 \end{bmatrix}$	$\begin{bmatrix} 0.044 & 0.048 \\ 0.052 & 0.057 \end{bmatrix}$	$\begin{bmatrix} 0.046 & 0.045 \\ 0.046 & 0.047 \end{bmatrix}$	$\begin{bmatrix} 0.054 & 0.046 \\ 0.052 & 0.050 \end{bmatrix}$	$\begin{bmatrix} 0.022 & 0.018 \\ 0.059 & 0.045 \end{bmatrix}$	$\begin{bmatrix} 0.107 & 0.069 \\ 0.069 & 0.112 \end{bmatrix}$
RMSE	$\begin{bmatrix} 0.061 & 0.066 \\ 0.065 & 0.067 \end{bmatrix}$	$\begin{bmatrix} 0.055 & 0.061 \\ 0.066 & 0.071 \end{bmatrix}$	$\begin{bmatrix} 0.058 & 0.056 \\ 0.058 & 0.058 \end{bmatrix}$	$\begin{bmatrix} 0.067 & 0.059 \\ 0.065 & 0.065 \end{bmatrix}$	$\begin{bmatrix} 0.031 & 0.027 \\ 0.061 & 0.047 \end{bmatrix}$	$\begin{bmatrix} 0.140 & 0.087 \\ 0.087 & 0.146 \end{bmatrix}$

Element-wise MAE and RMSE remain close to one another throughout, with MAE/RMSE ratios consistently near the Gaussian benchmark $\sqrt{2/\pi} \approx 0.798$ for Θ_1, Θ_2, Ω_1, and $\mathbf{\Lambda}$ alike. This indicates a symmetric, near-Gaussian sampling distribution of the posterior mean, consistent with the matrix-normal conditional posterior of Theorem 1.

Model III, which activates a second regular lag Θ_2 alongside Θ_1 and Ω_1, is estimated with accuracy comparable to the two-block specifications of Models I and II. The RMSEs for Θ_2 range from 0.055 to 0.071, comparable to (not systematically larger than) the 0.056–0.146 range spanning Θ_1, Ω_1, and $\mathbf{\Lambda}$ within the same model. So, we find no evidence that adding a third active block degrades estimation of the other blocks at this sample size. The genuinely (if only marginally) active cross-product block $\Theta_2\Omega_1$—whose true entries are an order of magnitude smaller than $\Theta_1\Omega_1$'s—is recovered with RMSE 0.027–0.061, appropriately small given that its posterior mean correctly sits close to zero. This is the expected behaviour of an estimator that treats a near-negligible true coefficient as belonging to the spike component on average. Overall, the results confirm that the proposed Gibbs sampler recovers all active coefficient blocks with accurate, well-calibrated uncertainty quantification, even as the number of active lags increases from two to three, and appropriately shrinks a near-negligible active block toward zero without materially degrading the accuracy of its substantively active neighbours.

Turning to the precision matrix, $\hat{\mathbf{\Lambda}}$ shows a small bias in its diagonal entries across all three models (e.g., 1.00 true versus 1.063–1.098 posterior mean for the leading diagonal), with all true entries remaining well within their respective 95% credible intervals throughout [40,41]. Across all replicates and all SeVAR simulation models, MCMC chains pass the convergence diagnostics, verifying that the reported Bayesian results are based on converged chains.

(c) Prediction performance

Table 11 reports the evaluation metrics of multi-step-ahead predictions for the simulated SeVAR Models I–III.

Table 11. Evaluation metrics of multi-step-ahead predictions for the simulated SeVAR models.

$y(n+r)$	$y(n+1)$	$y(n+2)$	$y(n+3)$	$y(n+4)$	$y(n+5)$
			Model I		
MAE	$\begin{bmatrix}0.866\\0.867\end{bmatrix}$	$\begin{bmatrix}1.076\\0.974\end{bmatrix}$	$\begin{bmatrix}1.141\\0.936\end{bmatrix}$	$\begin{bmatrix}1.143\\0.990\end{bmatrix}$	$\begin{bmatrix}1.157\\1.007\end{bmatrix}$
RMSE	$\begin{bmatrix}1.068\\1.063\end{bmatrix}$	$\begin{bmatrix}1.331\\1.219\end{bmatrix}$	$\begin{bmatrix}1.442\\1.192\end{bmatrix}$	$\begin{bmatrix}1.441\\1.250\end{bmatrix}$	$\begin{bmatrix}1.436\\1.255\end{bmatrix}$
CP (%)	$\begin{bmatrix}94.8\\94.0\end{bmatrix}$	$\begin{bmatrix}92.4\\95.4\end{bmatrix}$	$\begin{bmatrix}93.0\\96.0\end{bmatrix}$	$\begin{bmatrix}93.0\\95.8\end{bmatrix}$	$\begin{bmatrix}93.8\\94.2\end{bmatrix}$
			Model II		
MAE	$\begin{bmatrix}0.830\\0.878\end{bmatrix}$	$\begin{bmatrix}0.912\\0.924\end{bmatrix}$	$\begin{bmatrix}0.980\\0.856\end{bmatrix}$	$\begin{bmatrix}0.909\\0.928\end{bmatrix}$	$\begin{bmatrix}0.940\\0.912\end{bmatrix}$
RMSE	$\begin{bmatrix}1.035\\1.080\end{bmatrix}$	$\begin{bmatrix}1.146\\1.149\end{bmatrix}$	$\begin{bmatrix}1.210\\1.104\end{bmatrix}$	$\begin{bmatrix}1.155\\1.168\end{bmatrix}$	$\begin{bmatrix}1.179\\1.139\end{bmatrix}$
CP (%)	$\begin{bmatrix}94.8\\93.6\end{bmatrix}$	$\begin{bmatrix}94.6\\94.4\end{bmatrix}$	$\begin{bmatrix}94.8\\94.4\end{bmatrix}$	$\begin{bmatrix}94.2\\94.8\end{bmatrix}$	$\begin{bmatrix}94.8\\94.8\end{bmatrix}$
			Model III		
MAE	$\begin{bmatrix}0.822\\0.890\end{bmatrix}$	$\begin{bmatrix}1.168\\1.069\end{bmatrix}$	$\begin{bmatrix}1.130\\1.051\end{bmatrix}$	$\begin{bmatrix}1.175\\1.191\end{bmatrix}$	$\begin{bmatrix}1.213\\1.162\end{bmatrix}$
RMSE	$\begin{bmatrix}1.049\\1.105\end{bmatrix}$	$\begin{bmatrix}1.469\\1.332\end{bmatrix}$	$\begin{bmatrix}1.453\\1.325\end{bmatrix}$	$\begin{bmatrix}1.465\\1.533\end{bmatrix}$	$\begin{bmatrix}1.512\\1.464\end{bmatrix}$
CP (%)	$\begin{bmatrix}92.6\\92.8\end{bmatrix}$	$\begin{bmatrix}92.0\\93.2\end{bmatrix}$	$\begin{bmatrix}94.0\\94.0\end{bmatrix}$	$\begin{bmatrix}95.4\\89.6\end{bmatrix}$	$\begin{bmatrix}95.4\\92.6\end{bmatrix}$

Across all three models and all five forecast horizons, the ratio MAE/RMSE lies uniformly between 0.775 and 0.816, matching the Gaussian benchmark $\sqrt{2/\pi} \approx 0.798$ essentially everywhere. This confirms that the multi-step conditional predictive distribution of Corollary 1 remains well calibrated as the number of active blocks grows from two (Models I and II) to three (Model III), and is not degraded by the additional parameter and subset-selection uncertainty introduced by the extra block. The RMSE grows with horizon in almost every case, as expected of a correctly calibrated Bayesian multi-step forecasting procedure [42–44]. A small number of mild, isolated dips appear, e.g., series 2 of Model I at $r = 3$ (RMSE falling from 1.219 at $r = 2$ to 1.192 at $r = 3$ before resuming its rise), and series 2 of Model II at $r = 3$. None of these horizons is a seasonal multiple, so we interpret these fluctuations as ordinary Monte Carlo variability across the $R = 500$ replicates. Separately, comparing across models, Model III exhibits larger long-horizon RMSEs than the two-lag Models I and II (1.46–1.51 at $r = 5$ versus 1.14–1.44), reflecting the additional parameter and propagation uncertainty that a richer active lag structure introduces into multi-step forecasts.

Empirical coverage probability

Table 11 reports, in its CP row, the empirical coverage probability at each horizon—the proportion of the $R = 500$ replicates for which the realised $\boldsymbol{y}_i(n+r)$ falls inside its reported 95% predictive interval, computed component-wise as $\widehat{\mathrm{CP}}(r) = R^{-1}\sum_{i=1}^{R} \mathbb{1}\{y_i(n+r) \in [\hat{L}_i(n+r), \hat{U}_i(n+r)]\}$. Across all 30 model×horizon×series combinations, CP ranges from 89.6% to 96.0% with a mean of 94.0%, close to the nominal 95% level and directly confirming, rather than merely arguing indirectly for, the calibration of the stochastic forecasting recursion of Section 5.2. This also resolves the single-replicate pattern discussed in Section 7.2: nominal 95% intervals are calibrated on average across replicates, and the empirical rates here show that the average is close to nominal at every horizon, not merely at

the one replicate illustrated there. The single lowest value, 89.6% (Model III, series 2, $r = 4$), is the only one of the thirty cells more than about three percentage points from nominal. Given that the five horizons within a model reuse the same 500 simulated trajectories and are therefore correlated draws rather than independent replications, an isolated cell of this size is consistent with ordinary Monte Carlo variability rather than a systematic under-coverage problem, particularly as neither $r = 3$ nor $r = 5$ for the same series shows a comparable dip (94.0% and 92.6%, respectively). We do not observe a systematic tendency for CP to drift away from nominal as the horizon increases in any of the three models, which is the pattern that would indicate the multi-step recursion progressively under- or over-states its own uncertainty.

(d) Hyperparameter sensitivity

To quantitatively assess the robustness of posterior inference and forecasting to the choice of hyperparameters, the Model I simulation experiment ($n = 200$, $R = 500$ replicates) was repeated under two alternative hyperparameter specifications: (i) $\eta_m = 0.01$ and $\zeta_m = 10$ (a ten-fold weaker Occam penalty than the paper's default), and (ii) $\eta_m = 0.10$ and $\zeta_m = 100$ (a ten-fold wider spike than the paper's default). Table 12 reports aggregated Block-SSVS subset selection results in terms of mean posterior block-inclusion probabilities under each specification. In addition, Table 13 reports the coefficient and precision matrix RMSE and MAE, while Table 14 reports the multi-step forecast RMSE, MAE, and CP under each specification. These results give a direct numerical counterpart to the analytic argument of Remark 5, and confirm both of its qualitative predictions sharply.

Table 12. Aggregated Block-SSVS subset selection results in the simulated SeVAR Model I under two hyperparameter specifications: mean posterior block-inclusion probability ($\overline{\text{PIP}}$) for each lag block.

Θ_i	$\overline{\text{PIP}}$	Ω_j	$\overline{\text{PIP}}$	$\Theta_i\Omega_1$	$\overline{\text{PIP}}$	$\Theta_i\Omega_2$	$\overline{\text{PIP}}$	$\Theta_i\Omega_3$	$\overline{\text{PIP}}$	$\Theta_i\Omega_4$	$\overline{\text{PIP}}$
Model I with $\eta_m = 0.01$ and $\zeta_m = 10$											
Θ_1	0.999	Ω_1	0.999	$\Theta_1\Omega_1$	0.724	$\Theta_1\Omega_2$	0.386	$\Theta_1\Omega_3$	0.240	$\Theta_1\Omega_4$	0.283
Θ_2	0.442	Ω_2	0.647	$\Theta_2\Omega_1$	0.255	$\Theta_2\Omega_2$	0.240	$\Theta_2\Omega_3$	0.250	$\Theta_2\Omega_4$	0.258
Θ_3	0.315	Ω_3	0.356	$\Theta_3\Omega_1$	0.190	$\Theta_3\Omega_2$	0.207	$\Theta_3\Omega_3$	0.207	$\Theta_3\Omega_4$	0.245
Θ_4	0.568	Ω_4	0.231	$\Theta_4\Omega_1$	0.208	$\Theta_4\Omega_2$	0.301	$\Theta_4\Omega_3$	0.217	$\Theta_4\Omega_4$	0.235
Model I with $\eta_m = 0.1$ and $\zeta_m = 100$											
Θ_1	0.776	Ω_1	0.397	$\Theta_1\Omega_1$	0.017	$\Theta_1\Omega_2$	0.000	$\Theta_1\Omega_3$	0.000	$\Theta_1\Omega_4$	0.000
Θ_2	0.000	Ω_2	0.000	$\Theta_2\Omega_1$	0.000	$\Theta_2\Omega_2$	0.000	$\Theta_2\Omega_3$	0.000	$\Theta_2\Omega_4$	0.000
Θ_3	0.000	Ω_3	0.000	$\Theta_3\Omega_1$	0.000	$\Theta_3\Omega_2$	0.000	$\Theta_3\Omega_3$	0.000	$\Theta_3\Omega_4$	0.000
Θ_4	0.000	Ω_4	0.000	$\Theta_4\Omega_1$	0.000	$\Theta_4\Omega_2$	0.000	$\Theta_4\Omega_3$	0.000	$\Theta_4\Omega_4$	0.000

Under specification (i) ($\zeta_m = 10$), the Occam penalty $\zeta_m^{-k^2} = 10^{-4}$ is four orders of magnitude weaker than the paper's default $100^{-4} = 10^{-8}$. Table 12 shows the consequence directly: the three genuinely active blocks are still retained on average ($\overline{\text{PIP}} = 0.999, 0.999, 0.724$ for Θ_1, Ω_1, $\Theta_1\Omega_1$), but every one of the 21 genuinely inactive blocks now has a materially inflated mean PIP (average 0.299, maximum 0.647, versus average 0.011 and maximum 0.034 under the paper's default $\zeta_m = 100$ in Table 8). Two of the twenty-one—Θ_4 ($\overline{\text{PIP}} = 0.568$) and Ω_2 ($\overline{\text{PIP}} = 0.647$)—would be spuriously retained on average at the conventional 0.50 cut-off. This is precisely the false-positive inflation the $\zeta_m^{-k^2}$ Occam-penalty term of Remark 5 predicts under a weaker slab-to-spike contrast, and it is the reason we adopt $\zeta_m = 100$ rather than the smaller $\zeta_m = 10$ that is a more conventional default elsewhere in the SSVS literature [18,20]. Tables 13 and 14 show the further consequence: because several inactive blocks are, on average, half-included, the RMSEs of Θ_1, Ω_1, and $\Theta_1\Omega_1$ under $\zeta_m = 10$ are uniformly larger than under the paper's

default specification (e.g., Θ_1 RMSE: 0.186–0.321 versus 0.069–0.074; full comparison in Table 13).

Table 13. Evaluation metrics of coefficient and precision matrices for the active lags in the simulated SeVAR Model I under two hyperparameter specifications.

Metric	$\boldsymbol{\Theta_1}$	$\boldsymbol{\Omega_1}$	$\boldsymbol{\Theta_1\Omega_1}$	$\boldsymbol{\Lambda}$
	Model I with $\eta_m = 0.01$ and $\zeta_m = 10$			
MAE	$\begin{bmatrix} 0.317 & 0.282 \\ 0.180 & 0.268 \end{bmatrix}$	$\begin{bmatrix} 0.210 & 0.112 \\ 0.106 & 0.198 \end{bmatrix}$	$\begin{bmatrix} 0.344 & 0.308 \\ 0.213 & 0.270 \end{bmatrix}$	$\begin{bmatrix} 0.163 & 0.052 \\ 0.052 & 0.110 \end{bmatrix}$
RMSE	$\begin{bmatrix} 0.321 & 0.286 \\ 0.186 & 0.272 \end{bmatrix}$	$\begin{bmatrix} 0.216 & 0.123 \\ 0.117 & 0.205 \end{bmatrix}$	$\begin{bmatrix} 0.347 & 0.313 \\ 0.219 & 0.273 \end{bmatrix}$	$\begin{bmatrix} 0.181 & 0.065 \\ 0.065 & 0.130 \end{bmatrix}$
	Model I with $\eta_m = 0.1$ and $\zeta_m = 100$			
MAE	$\begin{bmatrix} 0.151 & 0.141 \\ 0.100 & 0.121 \end{bmatrix}$	$\begin{bmatrix} 0.195 & 0.110 \\ 0.106 & 0.176 \end{bmatrix}$	$\begin{bmatrix} 0.278 & 0.252 \\ 0.172 & 0.219 \end{bmatrix}$	$\begin{bmatrix} 0.120 & 0.084 \\ 0.084 & 0.144 \end{bmatrix}$
RMSE	$\begin{bmatrix} 0.190 & 0.170 \\ 0.122 & 0.155 \end{bmatrix}$	$\begin{bmatrix} 0.216 & 0.125 \\ 0.121 & 0.199 \end{bmatrix}$	$\begin{bmatrix} 0.283 & 0.259 \\ 0.181 & 0.225 \end{bmatrix}$	$\begin{bmatrix} 0.155 & 0.107 \\ 0.107 & 0.184 \end{bmatrix}$

Table 14. Evaluation metrics of multi-step-ahead predictions for the simulated SeVAR Model I under two hyperparameter specifications.

$\boldsymbol{y(n+r)}$	$\boldsymbol{y(n+1)}$	$\boldsymbol{y(n+2)}$	$\boldsymbol{y(n+3)}$	$\boldsymbol{y(n+4)}$	$\boldsymbol{y(n+5)}$
	Model I with $\eta_m = 0.01$ and $\zeta_m = 10$				
MAE	$\begin{bmatrix} 0.957 \\ 0.934 \end{bmatrix}$	$\begin{bmatrix} 1.097 \\ 1.021 \end{bmatrix}$	$\begin{bmatrix} 1.173 \\ 0.950 \end{bmatrix}$	$\begin{bmatrix} 1.166 \\ 0.987 \end{bmatrix}$	$\begin{bmatrix} 1.196 \\ 1.030 \end{bmatrix}$
RMSE	$\begin{bmatrix} 1.177 \\ 1.150 \end{bmatrix}$	$\begin{bmatrix} 1.365 \\ 1.263 \end{bmatrix}$	$\begin{bmatrix} 1.477 \\ 1.212 \end{bmatrix}$	$\begin{bmatrix} 1.476 \\ 1.257 \end{bmatrix}$	$\begin{bmatrix} 1.487 \\ 1.291 \end{bmatrix}$
CP (%)	$\begin{bmatrix} 96.200 \\ 95.000 \end{bmatrix}$	$\begin{bmatrix} 92.400 \\ 94.400 \end{bmatrix}$	$\begin{bmatrix} 93.000 \\ 95.200 \end{bmatrix}$	$\begin{bmatrix} 91.400 \\ 94.600 \end{bmatrix}$	$\begin{bmatrix} 91.800 \\ 92.600 \end{bmatrix}$
	Model I with $\eta_m = 0.1$ and $\zeta_m = 100$				
MAE	$\begin{bmatrix} 0.933 \\ 0.932 \end{bmatrix}$	$\begin{bmatrix} 1.118 \\ 1.048 \end{bmatrix}$	$\begin{bmatrix} 1.201 \\ 0.980 \end{bmatrix}$	$\begin{bmatrix} 1.221 \\ 1.049 \end{bmatrix}$	$\begin{bmatrix} 1.221 \\ 1.102 \end{bmatrix}$
RMSE	$\begin{bmatrix} 1.161 \\ 1.157 \end{bmatrix}$	$\begin{bmatrix} 1.399 \\ 1.282 \end{bmatrix}$	$\begin{bmatrix} 1.515 \\ 1.253 \end{bmatrix}$	$\begin{bmatrix} 1.529 \\ 1.331 \end{bmatrix}$	$\begin{bmatrix} 1.526 \\ 1.370 \end{bmatrix}$
CP (%)	$\begin{bmatrix} 94.2 \\ 93.6 \end{bmatrix}$	$\begin{bmatrix} 91.6 \\ 94.2 \end{bmatrix}$	$\begin{bmatrix} 91.0 \\ 94.2 \end{bmatrix}$	$\begin{bmatrix} 90.4 \\ 93.4 \end{bmatrix}$	$\begin{bmatrix} 91.6 \\ 90.8 \end{bmatrix}$

In addition, the one-step forecast RMSE is correspondingly larger (1.18, 1.15 at $r = 1$ versus the paper's default 1.07, 1.06 in Table 11). Although CP remains close to nominal throughout (91.4–96.2%), the wider slab inflates the posterior uncertainty of both genuine and spurious coefficients. As a result, the reported intervals widen enough to preserve coverage, but at the cost of a less parsimonious and less precise fitted model.

Under specification (ii) ($\eta_m = 0.10$), the spike itself is ten times wider than the default, so the evidence exponent $\eta_m^{-2}(1 - \zeta_m^{-2})$ shrinks by two orders of magnitude relative to the default $\eta_m = 0.01$. Table 12 shows the predicted evidence collapse directly. Even the two most strongly active blocks are severely under-included on average, with $\overline{\text{PIP}}$ falling to 0.776 for Θ_1 and, more strikingly, to just 0.397—below the 0.50 inclusion threshold—for Ω_1. Meanwhile, the genuinely active cross-block $\Theta_1\Omega_1$ is essentially never retained ($\overline{\text{PIP}} = 0.017$). All twenty-one inactive blocks are correctly excluded (every mean PIP is exactly 0.000), so this specification makes no false-positive errors; but it makes severe false-negative errors on genuine signal, exactly as the η_m^{-2} evidence term of Remark 5 predicts once the spike is widened enough to blur the spike/slab contrast for coefficients

of the magnitude generated here (0.3–0.8). Again, the RMSEs of Θ_1, Ω_1, and $\Theta_1\Omega_1$ under $\eta_m = 0.10$ are uniformly larger than under the paper's default specification (e.g., Θ_1 RMSE 0.155–0.190 versus 0.069–0.074). Forecast CP again remains close to nominal (90.4–94.2%; Table 14), underscoring that predictive coverage alone, absent the selection diagnostics, would not have revealed this specification's failure to recover the true active-block structure.

Taken together, these results show that the paper's default $(\eta_m, \zeta_m) = (0.01, 100)$ lies in a regime that avoids both failure modes identified analytically in Remark 5. It prevents the false-positive inflation associated with an insufficiently separated slab, and it avoids the false-negative collapse induced by an over-widened spike. This holds for coefficient magnitudes of the order generated in this study and, by extension, of the order estimated in the application of Section 8. We would not expect this specific numerical pair to be optimal for a system with substantially larger or smaller true coefficient magnitudes, precisely because the evidence term in Remark 5 is scaled by the coefficients themselves. A practitioner applying this method to data on a different scale should therefore re-examine the sensitivity using the same two-specification protocol employed here, rather than adopt $(0.01, 100)$ by default.

(e) Comparison with element-wise SSVS subset selection

To compare with the element-wise SSVS subset selection, the Model II simulation experiment ($n = 200$, $R = 500$ replicates) was repeated using the element-wise SSVS procedure. Under element-wise selection, each of the $qk = hk^2$ individual coefficients carries its own Bernoulli indicator, rather than the h block-level indicators used in Block-SSVS. Accordingly, Table 15 reports a separate mean PIP for every entry of every candidate matrix. In addition, Table 16 reports the coefficient and precision matrix RMSE and MAE, while Table 17 reports the multi-step forecast RMSE, MAE, and CP obtained under the element-wise SSVS subset selection.

Table 15. Aggregated element-wise SSVS subset selection results in the simulated SeVAR Model II: mean posterior block-inclusion probability ($\overline{\text{PIP}}$) for each lag block.

$\boldsymbol{\Theta_i}$	$\overline{\text{PIP}}$	$\boldsymbol{\Omega_j}$	$\overline{\text{PIP}}$	$\boldsymbol{\Theta_i\Omega_1}$	$\overline{\text{PIP}}$	$\boldsymbol{\Theta_i\Omega_2}$	$\overline{\text{PIP}}$	$\boldsymbol{\Theta_i\Omega_3}$	$\overline{\text{PIP}}$	$\boldsymbol{\Theta_i\Omega_4}$	$\overline{\text{PIP}}$
Θ_1	0.951 0.745 0.703 0.829	Ω_1	0.982 0.929 0.990 1.000	$\Theta_1\Omega_1$	0.297 0.776 0.138 0.773	$\Theta_1\Omega_2$	0.139 0.115 0.161 0.150	$\Theta_1\Omega_3$	0.094 0.113 0.121 0.108	$\Theta_1\Omega_4$	0.115 0.111 0.097 0.101
Θ_2	0.118 0.114 0.090 0.088	Ω_2	0.114 0.126 0.169 0.180	$\Theta_2\Omega_1$	0.116 0.109 0.097 0.092	$\Theta_2\Omega_2$	0.113 0.103 0.087 0.093	$\Theta_2\Omega_3$	0.115 0.109 0.088 0.095	$\Theta_2\Omega_4$	0.101 0.097 0.078 0.082
Θ_3	0.117 0.100 0.086 0.080	Ω_3	0.102 0.108 0.131 0.106	$\Theta_3\Omega_1$	0.114 0.108 0.089 0.088	$\Theta_3\Omega_2$	0.095 0.100 0.085 0.092	$\Theta_3\Omega_3$	0.098 0.106 0.092 0.086	$\Theta_3\Omega_4$	0.114 0.100 0.086 0.087
Θ_4	0.088 0.107 0.088 0.088	Ω_4	0.122 0.106 0.100 0.112	$\Theta_4\Omega_1$	0.105 0.113 0.083 0.090	$\Theta_4\Omega_2$	0.097 0.101 0.088 0.098	$\Theta_4\Omega_3$	0.104 0.108 0.090 0.088	$\Theta_4\Omega_4$	0.094 0.114 0.080 0.082

Table 16. Evaluation metrics of coefficient and precision matrices for the active lags in the simulated SeVAR Model II with element-wise SSVS.

Metric	$\boldsymbol{\Theta_1}$	$\boldsymbol{\Omega_1}$	$\boldsymbol{\Theta_1\Omega_1}$	$\boldsymbol{\Lambda}$
MAE	0.122 0.128 0.116 0.129	0.062 0.048 0.062 0.050	0.081 0.165 0.015 0.161	0.117 0.087 0.087 0.107
RMSE	0.163 0.158 0.150 0.174	0.084 0.067 0.084 0.066	0.088 0.212 0.030 0.213	0.154 0.110 0.110 0.138

Table 17. Evaluation metrics of multi-step-ahead predictions for the simulated SeVAR Model II with element-wise SSVS.

$y(n+r)$	$y(n+1)$	$y(n+2)$	$y(n+3)$	$y(n+4)$	$y(n+5)$
MAE	$\begin{bmatrix}0.838\\0.878\end{bmatrix}$	$\begin{bmatrix}0.917\\0.933\end{bmatrix}$	$\begin{bmatrix}0.970\\0.867\end{bmatrix}$	$\begin{bmatrix}0.936\\0.915\end{bmatrix}$	$\begin{bmatrix}0.954\\0.914\end{bmatrix}$
RMSE	$\begin{bmatrix}1.040\\1.080\end{bmatrix}$	$\begin{bmatrix}1.152\\1.164\end{bmatrix}$	$\begin{bmatrix}1.207\\1.104\end{bmatrix}$	$\begin{bmatrix}1.178\\1.161\end{bmatrix}$	$\begin{bmatrix}1.189\\1.155\end{bmatrix}$
CP (%)	$\begin{bmatrix}96.4\\94.4\end{bmatrix}$	$\begin{bmatrix}96.0\\94.6\end{bmatrix}$	$\begin{bmatrix}94.4\\95.2\end{bmatrix}$	$\begin{bmatrix}94.2\\95.8\end{bmatrix}$	$\begin{bmatrix}94.8\\94.6\end{bmatrix}$

The central finding is the "half-dead matrix" pattern anticipated in Section 3. For the genuinely active cross-product block $\Theta_1\Omega_1$, the four entry-wise mean PIPs are 0.297, 0.776, 0.138, and 0.773. Two of these four entries fall below the 0.50 inclusion threshold even though the block as a whole is genuinely active. Consequently, an element-wise best-subset rule would retain this block only partially, breaking the cross-equation coherence that motivates the block construction of Section 3. The same block under Block-SSVS is retained as a coherent whole with $\overline{\text{PIP}} = 0.994$ (Table 8). The regular and seasonal blocks Θ_1 and Ω_1 fare better under element-wise selection in this design (entry-wise PIPs: 0.70–0.95 and 0.93–1.00, respectively, all above threshold). The half-dead-matrix failure observed here is therefore specific to the smaller-magnitude cross-product block rather than universal across all active blocks. This is consistent with Remark 5's point that inclusion evidence scales with the coefficient's own magnitude, and the entries of $\Theta_1\Omega_1$ (up to 0.44) are smaller than those of Ω_1 (up to 0.80) in this design. On the inactive side, element-wise selection performs comparably well to Block-SSVS. Every entry of every genuinely inactive block has mean PIP below 0.19, so element-wise selection is not making false-positive errors here, only the false-negative, block-fragmenting error on part of $\Theta_1\Omega_1$ described above.

Coefficient and precision-matrix accuracy, however, tell a sharper and more informative story (Table 16, to be read against the block-wise figures of Table 10, Model II). Two distinct effects are visible. First, even for the two blocks with no internal heterogeneity—Θ_1 and Ω_1, whose true entries are all of comparable, non-negligible magnitude and whose element-wise PIPs are all comfortably above the 0.50 threshold (Table 15)—element-wise RMSE is uniformly larger than Block-SSVS RMSE: by a factor of 1.8–2.1 for every entry of Θ_1, and by a smaller but still consistent factor of 1.2–1.3 for every entry of Ω_1. Since neither block has a genuinely inactive entry to fragment, this gap is attributable specifically to the loss of evidence-pooling. Block-SSVS decides inclusion from the joint, matrix-variate evidence of all k^2 entries at once (mean PIP: 0.997 for Θ_1 and 0.999 for Ω_1; Table 8). In contrast, element-wise selection decides each entry from its own evidence alone, so individually weaker per-entry signals (0.70–0.95 for the four entries of Θ_1; Table 15) leave more posterior mass on the spike component in any given sweep. This inflates the coefficient's posterior variance even when the modal decision remains correct inclusion. Second, the cross-product block $\Theta_1\Omega_1$—which genuinely is heterogeneous, with one true entry at 0.00 and three at $0.11, -0.44, 0.44$—shows the predicted bifurcation directly. The near-zero entry is estimated more accurately under element-wise selection (RMSE $= 0.030$ versus Block-SSVS's 0.071), since the element-wise procedure can shrink that single entry without being tied to its more active neighbours. In contrast, the two largest-magnitude entries (-0.44 and 0.44, both with element-wise PIP ≈ 0.77–0.78, i.e., correctly but only inconsistently included) are estimated substantially worse (RMSE $= 0.212, 0.213$ versus Block-SSVS's $0.096, 0.100$, a factor of about 2.1–2.2). An entry that is included in roughly three sweeps out of four, and shrunk to the tight spike in the remaining quarter, has a distinctly bimodal, high-variance posterior—something a decisively and uniformly in-

cluded coefficient (Block-SSVS's $\overline{\text{PIP}} = 0.994$ for the whole block) does not exhibit. The precision matrix $\mathbf{\Lambda}$, which is not subject to block selection under either scheme, is estimated comparably well by both (RMSE ratios: 0.86–1.11, no systematic direction), confirming that the differences above are attributable to the selection mechanism specifically and not to some general advantage of one Gibbs implementation over the other.

For forecast accuracy, Table 17 shows element-wise selection performs essentially on par with Block-SSVS at every horizon (e.g., one-step RMSE of 1.040, 1.080 under element-wise versus 1.035, 1.080 under Block-SSVS; CP within one to two percentage points throughout), which is unsurprising given that predictive accuracy is a coarser, sample-averaged criterion that a partially retained cross-block can still largely satisfy even when the underlying selection is structurally incoherent.

Taken together, this is a complete and internally consistent picture. Element-wise selection is never more than mildly advantageous (and only for the one genuinely near-zero entry it can shrink in isolation), while it is substantially disadvantageous—by a factor of roughly two—for every entry of a fully active block and for the majority-included entries of a partially active one. Block-SSVS's joint evidence-pooling is therefore not merely a matter of computational convenience or matrix-level interpretability; it materially improves coefficient-estimation accuracy for the fully and majority-active blocks that dominate this design, at the modest cost of a slightly worse fit to the one entry that is genuinely negligible.

Overall, this comparison provides direct empirical support for the block-coherence argument of Section 3. Block-SSVS and element-wise selection are comparably accurate for out-of-sample prediction and comparably reliable at excluding genuinely inactive coefficients, but element-wise selection (i) can fragment a genuinely active, moderate-magnitude cross-product block into a partially included, partially excluded matrix, and (ii) even absent such fragmentation, estimates every entry of a fully active block with materially larger RMSE than Block-SSVS, because deciding inclusion entry-by-entry forgoes the sharper, pooled evidence that deciding it block-by-block provides. Block-SSVS retains or excludes a block coherently as a whole, by construction, and this comparison shows that that coherence translates into a real gain in coefficient-estimation accuracy, not only into cleaner matrix-level interpretability.

8. Application to Hourly Solar Radiation in Najran, Saudi Arabia

We illustrate the proposed methodology on a meteorological dataset from Najran, a city in the arid southwestern interior of Saudi Arabia. The region ranks among the sunniest in the Arabian Peninsula, a consequence of its low latitude, persistently clear skies, and dry atmosphere, which makes it a valuable testbed both for high-frequency radiometric time series analysis and for solar-resource assessment in practice. Najran's southern districts, near 17.5°N, receive sustained direct-normal irradiance and sit at the centre of several utility-scale solar developments supporting Saudi Arabia's Vision 2030 renewable energy targets [45].

8.1. Data and Construction

The data comprise two co-measured atmospheric quantities, recorded hourly over 90 consecutive days and restricted to a fixed daylight twelve-hour clock window (06:00–17:00 local standard time; Saudi Arabia does not observe daylight saving), identical for every day. This yields $n = 1080$ hourly observations per series, organised as $n_d = 90$ day-windows of $s = 12$ consecutive hours, so that the seasonal lag maps every observation to the same clock-hour of the previous day. The restriction to daytime hours is the standard construction in solar applications. Nighttime irradiance is identically zero, and the daytime regime is the operationally relevant one for photovoltaic production. It is a modelling

choice for temperature, whose daytime values are well defined. The resulting window seams are absorbed by the deterministic profile of Section 8.2 and examined out-of-sample in Section 8.4. The bivariate vector

$$\boldsymbol{y}(t) = \big(y_1(t),\, y_2(t)\big)^{\top} \tag{45}$$

collects the following:

- **Global horizontal irradiance (GHI)**
  $[y_1(t)]$: The combined direct and diffuse shortwave energy striking a horizontal surface, in $\mathrm{Wh/m^2}$, recorded by a ground-based pyranometer at the Najran station [46]. GHI is the primary determinant of photovoltaic energy yield and the standard forecasting target in solar-energy applications.
- **Air temperature**
  $[y_2(t)]$, in °C: A genuinely stochastic meteorological quantity governed by the surface energy balance, thermal inertia, and synoptic advection. The pair is chosen deliberately so that both components are stochastic and jointly driven, providing the cross-equation feedback structure that the Block-SSVS framework is designed to identify.

Table 18 lists summary statistics and Figure 3 displays the two raw series with their autocorrelation functions (ACFs). The two series differ in scale by more than an order of magnitude (GHI standard deviation: 320 $\mathrm{Wh/m^2}$ versus 6.5 °C) and are moderately, not mechanically, coupled: the contemporaneous correlation is 0.646, reflecting genuinely joint but non-tautological dynamics.

Table 18. Summary statistics of the hourly Najran GHI and air-temperature series.

Series	Mean	Median	Std	Min	Max
GHI ($\mathrm{Wh/m^2}$)	518.03	564.5	320.15	6.0	1022.0
Temperature (°C)	34.78	37.21	6.49	18.72	44.22

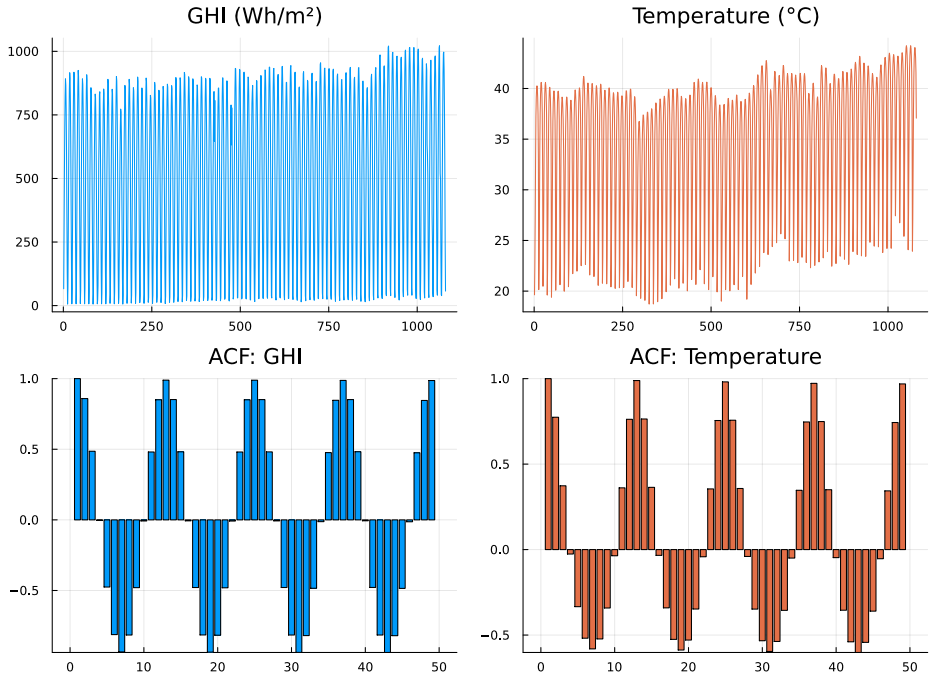


Figure 3. Hourly GHI and air temperature at Najran: raw series (**top**) and sample autocorrelation functions to lag 48 (**bottom**).

The daylight-hour ACFs of both series in Figure 3 exhibit pronounced recurring peaks at multiples of $s = 12$, motivating the seasonal operator of Definition 1 with $s = 12$. The series are partitioned into a training set comprising the first 70 days (840 hourly observations) and an evaluation set of the final 20 days (240 observations), used for the rolling-origin assessment of Section 8.4.

8.2. Transformation: Deterministic Seasonality and Its Evolution

Because the dominant feature of both series is a deterministic daily cycle, the appropriate transformation question is whether any *stochastic* seasonal structure remains. The levels of both series are therefore examined with the HEGY seasonal unit-root test [47], which separates the unit-root hypothesis frequency by frequency—zero, π, and each seasonal harmonic—and with the Canova–Hansen test [48] of stable seasonality, which examines whether the seasonal pattern is constant over the sample. The lag order is selected by AIC (12 for irradiance, 4 for temperature; the BIC rule yields 11 and 1), and inference is by block bootstrap. The HEGY null is rejected for small bootstrap p-values at the zero frequency and at the seasonal frequency pairs (joint Wald statistics), and for large negative t-statistics at π, while the Canova–Hansen null of stable seasonality is rejected for small p. The outcomes are decisive, as presented in Table 19.

Table 19. HEGY seasonal unit-root and Canova–Hansen stable-seasonality tests on levels; bootstrap p-values; lag order by AIC.

Series	t_0	p (0)	t_π	p (π)	p (12 h)	p (6 h)	p (4 h)	p (3 h)	p (2.4 h)	CH p
GHI	−2.82	0.323	6.38	0.999	0.140	0.019	0.001	0.001	0.001	0.001
Temperature	−2.54	0.539	5.64	0.999	0.008	0.001	0.001	0.001	0.001	0.001

HEGY tests at $s = 12$ reject unit roots at all five seasonal harmonic frequencies for temperature (bootstrap $p \leq 0.008$) and at four of five for irradiance ($p \leq 0.019$; the fundamental seasonal frequency is the sole non-rejecting cell, $p = 0.14$). The zero frequency null is not rejected for either series ($p \geq 0.323$), and the π-frequency null is likewise not rejected ($p = 0.999$). The case against seasonal differencing rests on the seasonal-frequency rejections alone, not on the zero frequency. The fitted model operates on profile-removed levels, for which the zero frequency question is immaterial. Moreover, the Canova–Hansen test rejects stable seasonality for both series ($p = 0.001$), formally confirming that the daily profile is deterministic but time-evolving: across sample thirds, the per-hour profile of GHI shifts by up to 134 Wh/m^2 and that of temperature by up to 3.9 °C. This drift is the evidence on which the time-evolving transformation below rests.

Accordingly, we remove the deterministic component by a per-hour profile with a polynomial trend in day index (degree 1 for GHI, degree 2 for temperature), fitted on training days only, rather than by differencing. The variance consequences of each candidate transformation quantify the choice. Removing the static hourly profile reduces the standard deviation of GHI from 320.2 to 51.6 Wh/m^2 and that of temperature from 6.5 to 1.9 °C. The deterministic daily cycle accounts for roughly 97% (GHI) and 91% (temperature) of the total variance. Allowing the profile to evolve—through the per-hour polynomial trend in day index—removes a further layer of deterministic structure, lowering these to 40.1 Wh/m^2 and 1.24 °C, respectively. This further reduction is the variance-space counterpart of the profile drift that the Canova–Hansen rejections detect. By contrast, seasonal differencing leaves the GHI standard deviation at 41.2 Wh/m^2—no better than profile removal—while distorting the dynamics. Two further diagnostics support the choice. First, the differenced specifications are empirically self-defeating: seasonally differenced series yield coefficient estimates at or beyond the stationarity boundary, and double differencing whitens the

series until every fitted coefficient is indistinguishable from zero. Second, after profile removal the day-over-day persistence of the anomalies is 0.09 (95% CI $[-0.14, 0.32]$) for GHI—a near-white day-level—and 0.69 ($[0.53, 0.86]$) for temperature, indicating a genuine stochastic day-level effect that the seasonal AR block of the SeVAR is designed to capture. This contrast is the empirical basis for expecting equation-asymmetric seasonal selection in Section 8.3.

8.3. Primary SeVAR Fit on Profile-Removed Series

We apply the proposed Algorithm 1 with $p_{\max} = P_{\max} = 4$ to the drift-profile-removed training series, giving a candidate set of $h = 24$ lag blocks (four regular, four seasonal, sixteen cross-product), with the Block-SSVS hyperparameters of Section 7.1. The application chain is lengthened to $T_{\text{sim}} = 32{,}000$ iterations ($T_{\text{burn}} = 2000$, $T_{\text{thin}} = 30$), retaining 1000 draws, and it is initialised at the full model, $\boldsymbol{\delta}^{(0)} = \mathbf{1}_{24}$. Table 20 reports the posterior block-inclusion probabilities, and Table 21 presents the posterior summaries of the coefficient and precision matrices.

Table 20. Posterior block-inclusion probabilities (PIPs) for SeVAR on the drift-profile-removed series.

Θ_i	PIP	Ω_j	PIP	$\Theta_i\Omega_1$	PIP	$\Theta_i\Omega_2$	PIP	$\Theta_i\Omega_3$	PIP	$\Theta_i\Omega_4$	PIP
Θ_1	1.000	Ω_1	1.000	$\Theta_1\Omega_1$	0.018	$\Theta_1\Omega_2$	0.001	$\Theta_1\Omega_3$	0.004	$\Theta_1\Omega_4$	0.009
Θ_2	0.997	Ω_2	0.007	$\Theta_2\Omega_1$	0.804	$\Theta_2\Omega_2$	0.005	$\Theta_2\Omega_3$	0.002	$\Theta_2\Omega_4$	0.011
Θ_3	0.013	Ω_3	0.013	$\Theta_3\Omega_1$	0.122	$\Theta_3\Omega_2$	0.004	$\Theta_3\Omega_3$	0.002	$\Theta_3\Omega_4$	0.003
Θ_4	0.002	Ω_4	0.005	$\Theta_4\Omega_1$	0.005	$\Theta_4\Omega_2$	0.011	$\Theta_4\Omega_3$	0.001	$\Theta_4\Omega_4$	0.001

Table 21. Posterior summaries of the coefficient and precision matrices for SeVAR on the drift-profile-removed series. $\hat{\boldsymbol{\mu}}$: posterior mean; $\hat{\boldsymbol{s}}$: posterior SD; $(\hat{L}, \hat{U})$: 95% credible interval.

Param.	$\hat{\mu}$	$\hat{s}$	$\hat{L}$	$\hat{U}$
Θ_1	$\begin{bmatrix} 0.836 & 0.734 \\ 0.002 & 0.674 \end{bmatrix}$	$\begin{bmatrix} 0.036 & 1.077 \\ 0.001 & 0.029 \end{bmatrix}$	$\begin{bmatrix} 0.771 & -1.417 \\ 0.000 & 0.618 \end{bmatrix}$	$\begin{bmatrix} 0.905 & 2.746 \\ 0.004 & 0.728 \end{bmatrix}$
Θ_2	$\begin{bmatrix} -0.201 & -0.282 \\ -0.000 & 0.142 \end{bmatrix}$	$\begin{bmatrix} 0.045 & 1.094 \\ 0.001 & 0.030 \end{bmatrix}$	$\begin{bmatrix} -0.287 & -2.375 \\ -0.003 & 0.085 \end{bmatrix}$	$\begin{bmatrix} -0.118 & 1.945 \\ 0.002 & 0.200 \end{bmatrix}$
Θ_3	$\begin{bmatrix} 0.258 & -0.040 \\ -0.002 & -0.000 \end{bmatrix}$	$\begin{bmatrix} 0.044 & 0.250 \\ 0.001 & 0.005 \end{bmatrix}$	$\begin{bmatrix} 0.167 & -0.452 \\ -0.004 & -0.010 \end{bmatrix}$	$\begin{bmatrix} 0.335 & 0.355 \\ 0.001 & 0.009 \end{bmatrix}$
Θ_4	$\begin{bmatrix} -0.111 & -0.018 \\ -0.000 & 0.001 \end{bmatrix}$	$\begin{bmatrix} 0.034 & 0.195 \\ 0.001 & 0.005 \end{bmatrix}$	$\begin{bmatrix} -0.179 & -0.374 \\ -0.002 & -0.009 \end{bmatrix}$	$\begin{bmatrix} -0.044 & 0.366 \\ 0.002 & 0.010 \end{bmatrix}$
Ω_1	$\begin{bmatrix} -0.009 & -0.301 \\ -0.001 & 0.199 \end{bmatrix}$	$\begin{bmatrix} 0.036 & 0.917 \\ 0.001 & 0.028 \end{bmatrix}$	$\begin{bmatrix} -0.083 & -2.114 \\ -0.002 & 0.140 \end{bmatrix}$	$\begin{bmatrix} 0.062 & 1.413 \\ 0.001 & 0.252 \end{bmatrix}$
Ω_2	$\begin{bmatrix} 0.009 & -0.006 \\ -0.001 & 0.002 \end{bmatrix}$	$\begin{bmatrix} 0.034 & 0.194 \\ 0.001 & 0.006 \end{bmatrix}$	$\begin{bmatrix} -0.060 & -0.369 \\ -0.003 & -0.008 \end{bmatrix}$	$\begin{bmatrix} 0.073 & 0.392 \\ 0.001 & 0.012 \end{bmatrix}$
Ω_3	$\begin{bmatrix} -0.038 & -0.027 \\ -0.001 & -0.003 \end{bmatrix}$	$\begin{bmatrix} 0.035 & 0.205 \\ 0.001 & 0.006 \end{bmatrix}$	$\begin{bmatrix} -0.105 & -0.414 \\ -0.002 & -0.012 \end{bmatrix}$	$\begin{bmatrix} 0.034 & 0.335 \\ 0.001 & 0.007 \end{bmatrix}$
Ω_4	$\begin{bmatrix} 0.150 & -0.028 \\ 0.001 & -0.001 \end{bmatrix}$	$\begin{bmatrix} 0.035 & 0.205 \\ 0.001 & 0.007 \end{bmatrix}$	$\begin{bmatrix} 0.081 & -0.386 \\ -0.001 & -0.011 \end{bmatrix}$	$\begin{bmatrix} 0.223 & 0.367 \\ 0.002 & 0.009 \end{bmatrix}$
$\Theta_1\Omega_1$	$\begin{bmatrix} 0.004 & 0.024 \\ 0.000 & 0.001 \end{bmatrix}$	$\begin{bmatrix} 0.046 & 0.291 \\ 0.001 & 0.010 \end{bmatrix}$	$\begin{bmatrix} -0.087 & -0.435 \\ -0.002 & -0.010 \end{bmatrix}$	$\begin{bmatrix} 0.091 & 0.417 \\ 0.002 & 0.012 \end{bmatrix}$
$\Theta_2\Omega_1$	$\begin{bmatrix} 0.016 & 0.073 \\ 0.001 & 0.084 \end{bmatrix}$	$\begin{bmatrix} 0.043 & 0.878 \\ 0.001 & 0.047 \end{bmatrix}$	$\begin{bmatrix} -0.069 & -1.702 \\ -0.002 & -0.003 \end{bmatrix}$	$\begin{bmatrix} 0.099 & 1.889 \\ 0.003 & 0.150 \end{bmatrix}$
Λ	$\begin{bmatrix} 0.003 & -0.008 \\ -0.008 & 3.984 \end{bmatrix}$	$\begin{bmatrix} 0.000 & 0.004 \\ 0.004 & 0.209 \end{bmatrix}$	$\begin{bmatrix} 0.002 & -0.015 \\ -0.015 & 3.596 \end{bmatrix}$	$\begin{bmatrix} 0.003 & -0.000 \\ -0.000 & 4.404 \end{bmatrix}$

Table 20 identifies a sharp and structurally coherent subset model. Three primitive blocks are retained with near certainty—Θ_1 and Ω_1 at PIP $= 1.000$ and Θ_2 at 0.997—together with the cross-product block $\Theta_2\Omega_1$ at PIP $= 0.804$ (Monte Carlo standard error 0.06; the block's inclusion probability is interior). Every remaining block is excluded decisively: no regular or seasonal lag beyond the second exceeds 0.02, and of the sixteen cross-product candidates only the immediate neighbour $\Theta_3\Omega_1$ attracts non-negligible support (0.122; all others are at most 0.018). The selection is also coherent in a sense worth emphasising. A cross-product block survives only when both of its parents survive, and the sole surviving interaction is precisely the one coupling the two retained regular lags with the retained seasonal lag. The resulting model is $\mathrm{SeVAR}_2(2,1)_{12}$ with the $(2,1)$ cross-block—a genuinely multiplicative seasonal structure, selected by the data rather than imposed.

Table 21 separates clearly into identified and unidentified entries, and the identified subset carries the application's central finding. In the seasonal block Ω_1, the equation asymmetry is stark. The coefficient of the GHI equation on lag-12 GHI is -0.009 $[-0.083, 0.062]$—statistically zero—while the coefficient of the temperature equation on lag-12 temperature is 0.199 $[0.140, 0.252]$. The stochastic day-over-day effect is therefore temperature-specific, in exact agreement with the profile-removed daily persistence measured in Section 8.2. After the evolving deterministic profile is removed, irradiance carries no day-level memory for a seasonal block to capture, whereas temperature retains a genuine stochastic day effect, and the sampler allocates the seasonal structure to that equation alone. The retained regular blocks tell the same short-memory story. GHI dynamics are carried by $\Theta_1(1,1) = 0.836$ $[0.771, 0.905]$ and $\Theta_2(1,1) = -0.201$ $[-0.287, -0.118]$, and temperature dynamics by 0.674 $[0.618, 0.728]$ and 0.142 $[0.085, 0.200]$, with the GHI→temperature entries pinned at essentially zero (0.002 $[0.000, 0.004]$ in Θ_1; °C per Wh/m^2). At the anomaly level, the cross-series coupling is negligible, because the strong physical coupling between the two series (heating leads temperature by one to two hours) is carried by the shared deterministic profile, not by anomaly-to-anomaly feedback.

Two further features of Table 21 deserve explicit comment. First, the retained cross-block illustrates the free-parameter parametrisation of Section 2.2. Its temperature-side entry, 0.084 $[-0.003, 0.150]$, exceeds the product $\hat{\Theta}_2\,\hat{\Omega}_1 \approx 0.028$ implied by the primitive estimates, confirming that the free block captures interaction content beyond the constrained product form. Its selection completes the structural inventory. The fitted model employs all three mechanisms that the operator $\Theta_p(B)\,\Omega_P(B^s)$ affords—regular persistence, seasonal memory, and their multiplicative interaction. Second, the four temperature→GHI entries (first row, second column of each block: the temperature-source regressors in the irradiance equation) have credible intervals of width 3.5–4.3, spanning the prior-supported range rather than likelihood-supported estimates. The anomaly variance of temperature is minuscule relative to the innovation variance of irradiance, so the likelihood is nearly flat in those directions. Consequently, the prior correctly reports uninformativeness rather than spurious precision. We therefore report these entries as unidentified and do not interpret them.

A related phenomenon, in the opposite direction, appears among the excluded blocks and merits one sentence of interpretation. Individual entries of excluded blocks can remain individually estimable—most notably the GHI–GHI entry of Θ_3, 0.258 $[0.167, 0.335]$—even though the block is excluded at PIP $= 0.013$. The two criteria measure different things. Block inclusion is decided by the innovation-standardised, likelihood-level evidence of the entire $k \times k$ block (38), in which source-column contributions are weighted by the corresponding innovation precision. Consequently, the irradiance-source entries, discounted by GHI's large innovation variance ($\Lambda_{11} = 0.003$), contribute almost nothing to the block evidence,

whereas an entry-level credible interval conditions on the drawn state mixture and can exclude zero on raw-scale estimability alone. Selection here is a likelihood-based statement of predictive relevance, not an entry-wise significance screen—precisely the block-level logic that distinguishes the proposed procedure from element-wise SSVS.

Finally, the fitted model passes its internal diagnostics. The residuals of the GHI equation are white at the lags of interest (autocorrelation 0.04 at lag 1 and 0.03 at lag 12), confirming that the selected structure captures the within-window dynamics and the near-absent seasonal structure of that series. The temperature residuals show no remaining seasonal autocorrelation (lag-12: 0.13, down from 0.64 before fitting). The contemporaneous innovation correlation is 0.076—the strong observed coupling between the two series ($r = 0.65$ in levels) is dynamic and deterministic in origin. The fitted effective lag polynomial is covariance-stationary in 94.6% of posterior draws (mean maximal companion eigenvalue: 0.968, 95% interval $[0.929, 1.006]$). The proximity to the boundary reflects the strength of the temperature day effect rather than a unit root (profile-removed daily AR(1) CI excludes one). Chain convergence of the continuous parameters is supported by the Geweke and Raftery–Lewis diagnostics, and the thinned chain exhibits negligible empirical autocorrelation across all parameters. Discrete-chain diagnostics for the 24 block indicators over the 30,000 unthinned post-burn-in sweeps are reported in Table 22.

Table 22. Discrete-chain diagnostics for the 24 block indicators of SeVAR on the drift-profile-removed series. PIP_A: chain initialised at the full model ($\delta^{(0)} = \mathbf{1}_{24}$); PIP_B: chain initialised at the empty model ($\delta^{(0)} = \mathbf{0}_{24}$); $|\Delta\mathrm{PIP}|$: absolute cross-start difference; switches: number of state changes of the indicator over the unthinned post-burn-in sweeps.

Block	PIP_A	PIP_B	$\lvert\Delta\mathrm{PIP}\rvert$	Switches
Θ_1	**1.000**	**1.000**	0.000	0
Θ_2	**0.983**	**1.000**	0.017	4
Θ_3	0.010	0.009	0.001	114
Θ_4	0.003	0.003	0.000	54
Ω_1	**1.000**	**1.000**	0.000	0
Ω_2	0.003	0.008	0.005	35
Ω_3	0.008	0.012	0.003	72
Ω_4	0.001	0.004	0.003	36
$\Theta_1\Omega_1$	0.010	0.018	0.008	50
$\Theta_1\Omega_2$	0.003	0.004	0.001	48
$\Theta_1\Omega_3$	0.002	0.002	0.000	44
$\Theta_1\Omega_4$	0.007	0.011	0.004	100
$\Theta_2\Omega_1$	**0.768**	**0.662**	0.105	23
$\Theta_2\Omega_2$	0.005	0.003	0.002	40
$\Theta_2\Omega_3$	0.003	0.002	0.001	60
$\Theta_2\Omega_4$	0.009	0.010	0.001	68
$\Theta_3\Omega_1$	0.111	0.151	0.040	49
$\Theta_3\Omega_2$	0.006	0.013	0.008	44
$\Theta_3\Omega_3$	0.001	0.001	0.000	40
$\Theta_3\Omega_4$	0.004	0.003	0.001	48
$\Theta_4\Omega_1$	0.005	0.008	0.003	62
$\Theta_4\Omega_2$	0.009	0.014	0.005	68
$\Theta_4\Omega_3$	0.004	0.002	0.000	48
$\Theta_4\Omega_4$	0.001	0.001	0.000	46

Table 22 points to three regimes. The blocks Θ_1 and Ω_1 are active in every post-burn-in sweep of both chains (PIP = 1.000, no switches) and Θ_2 in 98.3% and 100% of sweeps under the full-model and empty-model starts, so these indicators are effectively constant, while 19 blocks are decided in favour of exclusion in both chains, with PIPs of at most 0.018 and only brief excursions from the excluded state (at most 114 switches, fewer than four per 1000 sweeps). All remaining posterior mass concentrates on the two interior blocks $\Theta_2\Omega_1$

(PIP = 0.768 and 0.662) and $\Theta_3\Omega_1$ (0.111 and 0.151), which carry the model uncertainty of the application. Their indicators leave and re-enter the included state 23 and 49 times over the 30,000 sweeps of the full-start chain (0.8 and 1.6 switches per 1000 sweeps), so the sampler revisits both models many times instead of committing to a single configuration. At the configuration level, the full-start chain visits 113 distinct indicator configurations and the empty-start chain 121—vanishing fractions of the $2^{24} \approx 1.7 \times 10^7$ candidate models—and the modal configuration of both chains coincides with the selected subset $\{\Theta_1, \Theta_2, \Omega_1, \Theta_2\Omega_1\}$, visited in 70.3% and 59.7% of sweeps. Most importantly for start-point independence, the two chains are initialised at opposite corners of the model space and driven by independent seeds, so they are genuinely independent rather than two segments of one random stream. Every include/exclude decision at the 0.50 threshold agrees across the chains; block-level PIPs differ by at most 0.008 for 21 of the 24 blocks; and the three larger discrepancies arise precisely for the indicators that move most ($|\Delta\text{PIP}| = 0.017$ for the near-saturated Θ_2, 0.040 for $\Theta_3\Omega_1$ and 0.105 for $\Theta_2\Omega_1$). The empty-start chain, which must first discover the active blocks, devotes relatively more mass to simpler configurations: its second-most-frequent configuration drops $\Theta_2\Omega_1$ altogether ($\{\Theta_1, \Theta_2, \Omega_1\}$; 15.5%), whereas the full-start chain's runner-up exchanges $\Theta_2\Omega_1$ for $\Theta_3\Omega_1$ ($\{\Theta_1, \Theta_2, \Omega_1, \Theta_3\Omega_1\}$; 9.5%). The identified model and the posterior ranking of the candidate blocks therefore do not depend on the initialisation or the seed of the indicator chain, completing the validation anticipated in Section 6. It is worth noting that Table 20 averages the thinned retained draws (1000 values), the paper-wide convention, whereas PIP_A in Table 22 averages all 30,000 unthinned post-burn-in sweeps of one and the same chain. The small differences between the two summaries (e.g., Θ_2: 0.997 versus 0.983; $\Theta_2\Omega_1$: 0.804 versus 0.768) are the Monte Carlo noise of the coarser thinned estimator.

Seasonal operator versus a non-seasonal VAR

The central question the application is positioned to answer is whether the seasonal lag operator of Definition 1 is supported by the data, or whether an ordinary VAR suffices. We therefore estimate the non-seasonal special case of the same estimator ($P = 0$), i.e., a Bayesian VAR($p_{\max}$) with identical priors, identical Block-SSVS lag selection, and identical Gibbs machinery, fitted to the same drift-profile-removed series. The only structural difference between the two models is the seasonal factor in the lag polynomial. So, any divergence in fit is attributable to that factor alone. Table 23 reports the VAR's subset selection and Figure 4 the residual ACFs of both fitted models in the temperature equation, delivering the comparison that motivates the seasonal operator.

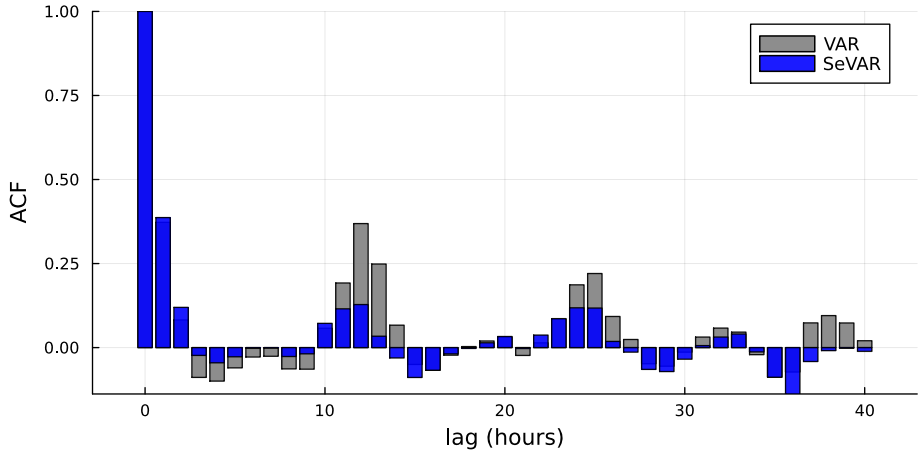


Figure 4. Residual ACFs of the temperature equation: VAR (grey bars) versus SeVAR (blue bars). The VAR leaves the seasonal signature (peak at lag 12) that the SeVAR removes.

Table 23. Posterior block-inclusion probabilities (PIPs) for non-seasonal VAR on the drift-profile-removed series.

Regular Block	PIP	Selected.
Θ_1	1.000	✓
Θ_2	0.993	✓
Θ_3	0.016	×
Θ_4	0.003	×

The residual ACFs of the temperature equation in Figure 4 show that the two models agree closely on the regular structure. From Table 23, the VAR retains Θ_1 and Θ_2 with posterior inclusion probabilities of 1.000 and 0.993, mirroring the SeVAR's own primitive selections (1.000 and 0.997), and likewise excludes Θ_3 and Θ_4 as decisively as the SeVAR does. Where the two models differ is exactly where their lag polynomials differ. The VAR's residuals retain a pronounced seasonal signature: 0.369 at lag 12, with the day effect leaking into the neighbouring lags (0.192 at lag 11 and 0.249 at lag 13). This produces a three-lag band of unresolved seasonal autocorrelation. The SeVAR's residuals show no such band: 0.115 at lag 11, 0.128 at lag 12, and 0.034 at lag 13. The non-seasonal model leaves exactly the seasonal signature that the SeVAR removes.

The reason is representability, not estimation effort. The VAR(4)'s candidate set comprises sixteen coefficients across lags 1–4. Lag 12 is simply absent from its span, so the day-over-day dependence that temperature carries is not merely unparsed but unrepresentable within the non-seasonal class. The SeVAR expresses the same dependence through the retained seasonal block Ω_1 and the interaction block $\Theta_2\Omega_1$ at lag 14, whose selection is precisely what absorbs the residual seasonal content. A VAR with sufficiently many regular lags could in principle restore representability; for example, a VAR(14) reaches lag 12, but only by admitting a long ladder of candidate lags. It would still remain silent on the multiplicative structure between regular and seasonal dynamics that the selected cross-product block expresses.

The comparison therefore resolves the question posed at the outset of this paragraph. At short lags, the two specifications are interchangeable; at the seasonal lags, they diverge, and the residual evidence localises the difference precisely at the frequencies that the operator $\Theta_p(B)\,\Omega_P(B^s)$ of Definition 1 is designed to capture.

8.4. Rolling-Origin Forecast Evaluation

Out-of-sample accuracy is assessed by an expanding-window rolling-origin design over the final 20 days (one origin per day, horizon $s = 12$). At each origin, the deterministic profile and the model are re-estimated from all data available up to that point, and all forecasts are compared on the level scale. Eight candidate specifications are compared. The proposed SeVAR estimator is evaluated under five data transformations: fitted to raw levels (SeVAR-levels); to series with the static hourly profile removed (SeVAR-static); to series with the time-evolving profile removed (SeVAR-profiled, the primary specification of Section 8.3); and to seasonally and regularly differenced series (SeVAR-sdiff, SeVAR-rdiff). The non-seasonal special case is fitted to the same profile-removed series (VAR-profiled). Two naive benchmarks are also included: persistence (Naive-P, $\hat{y}(n+r) = y(n)$) and the seasonal naive (Naive-S, $\hat{y}(n+r) = y(n+r-s)$). The twelve-step rolling forecasts use all retained draws; no stationarity filter is applied. As reported in Section 8.3, explosive draws account for only 5.4% of the posterior and lie marginally beyond unity, so their contribution to the pooled forecasts is negligible.

All accuracy figures for the race are pooled over origins and horizons. For series i and specification S, with level-scale forecast errors $e_i^{(o)}(r)$ as

$$e_i^{(o)}(r) \;=\; y_i(n_o + r) \;-\; \hat{y}_i^{(o)}(n_o + r), \tag{46}$$

the pooled RMSE is the quadratic mean of all $O \times s = 240$ error cells, given as

$$\mathrm{RMSE}_{i,S}^{\mathrm{pool}} = \sqrt{(Os)^{-1} \sum_{o=1}^{O} \sum_{r=1}^{s} \left[e_i^{(o)}(r)\right]^2}, \tag{47}$$

and the accompanying skill score is

$$\mathrm{Skill}_{i,S} = 1 - \mathrm{RMSE}_{i,S}^{\mathrm{pool}} / \mathrm{RMSE}_{i,\text{Naive-P}}^{\mathrm{pool}}, \tag{48}$$

and the per-horizon RMSEs, by contrast, average across origins only. Table 24 reports the pooled RMSE together with skill scores relative to Naive-P. Table 25 presents the Diebold–Mariano tests [49] comparing the primary specification against each rival. Figure 5 shows the per-horizon RMSE; Naive-P is omitted from the figure for scale.

Table 24. Rolling-origin evaluation (20 origins, horizon s = 12): pooled RMSE and skill vs persistence.

Specification	RMSE (GHI)	Skill (GHI)	RMSE (T)	Skill (T)
SeVAR-profiled	52.8	0.913	0.82	0.868
SeVAR-static	51.2	0.916	0.86	0.862
SeVAR-levels	51.1	0.916	0.90	0.855
SeVAR-sdiff	49.6	0.919	0.83	0.866
SeVAR-rdiff	52.5	0.914	0.89	0.857
VAR-profiled	53.4	0.912	0.86	0.862
Naive-P	609.6	0.000	6.21	0.000
Naive-S	53.9	0.912	0.80	0.871

Table 25. Diebold–Mariano tests of SeVAR-profiled against each rival (negative statistic: SeVAR-profiled more accurate).

Rival	DM Statistic	*p*-Value
SeVAR-static (GHI)	0.64	0.5240
SeVAR-static (T)	−0.78	0.4363
SeVAR-levels (GHI)	0.45	0.6537
SeVAR-levels (T)	−1.19	0.2352
SeVAR-sdiff (GHI)	1.02	0.3079
SeVAR-sdiff (T)	−0.20	0.8405
SeVAR-rdiff (GHI)	0.06	0.9500
SeVAR-rdiff (T)	−0.80	0.4258
VAR-profiled (GHI)	−0.63	0.5257
VAR-profiled (T)	−0.79	0.4314
Naive-P (GHI)	−38.59	0.0000
Naive-P (T)	−30.96	0.0000
Naive-S (GHI)	−0.20	0.8399
Naive-S (T)	0.26	0.7976

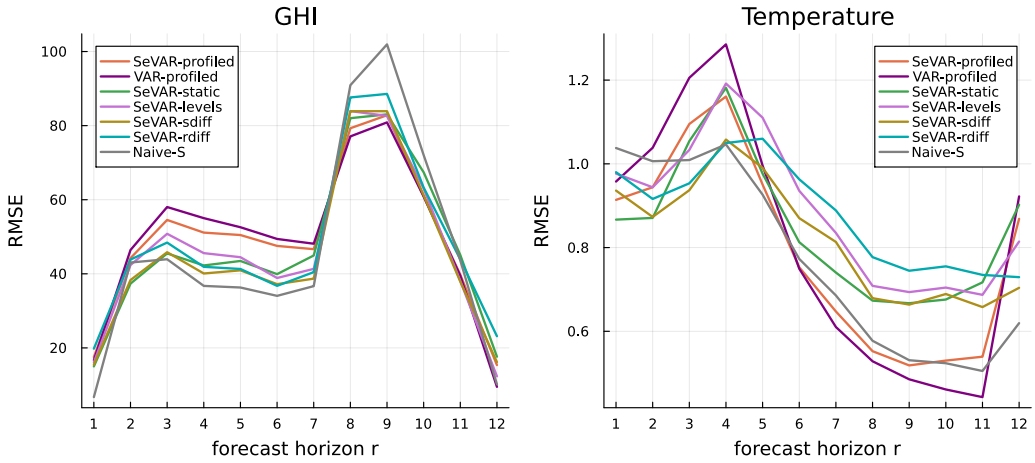


Figure 5. Per-horizon rolling-origin RMSE by specification.

Table 24 yields three findings. First, every specification dominates persistence overwhelmingly. Naive-P produces pooled RMSE of 609.6 Wh/m^2 and 6.21 °C, whereas all fitted models lie in a narrow 49.6–53.9 Wh/m^2 and 0.80–0.90 °C band. The Diebold–Mariano statistics, -38.6 and -31.0 ($p < 0.001$), confirm that all fitted specifications capture the deterministic diurnal structure that anchors persistence to the previous dusk. Second, no statistically significant difference separates the primary specification from any rival. Every Diebold–Mariano test of SeVAR-profiled returns $|t| \leq 1.19$ with $p \geq 0.24$, including the non-seasonal VAR-profiled ($p = 0.43$–0.53) and Naive-S ($p = 0.80$ and 0.84). The one additional pairwise contrast of interest, SeVAR-sdiff against Naive-S, is likewise insignificant ($p = 0.31$). The pooled RMSEs span only 8% for irradiance and 13% for temperature, and this compression is not a failure of discrimination but a property of the data that Section 8.3 has already quantified. After removal of the deterministic profile, the day-over-day level of irradiance is nearly white (AR(1)$= 0.09$). As a result, the remaining synoptic variation lies close to the innovation-noise floor, and no conditional-mean specification—seasonal or not—can improve materially on it. The forecast race therefore corroborates the dynamics identified in Section 8.3 from the outside. The model's own estimated persistence predicts that competent forecasts must converge, and they do.

Third, the horizon-resolved results in Figure 5 are consistent with the data geometry documented above. Naive-S is the most accurate specification at the daylight-window boundaries ($r = 1$: 6.7 versus 15.0–19.8 Wh/m^2 across the SeVAR specifications). The lag-s mapping tracks the solar-time slide of the fixed daylight window mechanically, whereas profile-based forecasts rely on a one-day extrapolation of the deterministic profile. The SeVAR specifications are the most accurate in the afternoon convective regime ($r = 8$–9, where Naive-S reaches 90.9 and 101.9 Wh/m^2 against 79.3 and 82.8 for SeVAR-profiled), where persistence of the level fails and dynamic structure contributes. SeVAR-sdiff delivers the lowest SeVAR RMSE at the ramp horizons $r = 4$, 5, and 7 (and is within 0.5 Wh/m^2 of the best SeVAR variant at $r = 3$ and 6), indicating that day-over-day shape persistence—discarded by construction under profile removal—carries descriptive information; the corresponding pooled contrast, however, does not attain significance ($p = 0.31$).

Taken together, the evaluation supports the framing advanced throughout this section. On data whose stochastic day-level dynamics are this weak, the seasonal structure is not a forecasting device—and the paper does not claim it as one. It is the correct description of the dynamics, as established by selection and by the residual signatures of the non-seasonal rival; and the forecast evaluation, far from contradicting that claim, closes it. A model whose estimated innovation process leaves almost nothing predictable beyond the deterministic profile must forecast at the floor that its own parameters describe.

The application therefore demonstrates the framework's three capacities on genuinely stochastic, jointly driven data. It *identifies* the structure that is present—a deterministic, time-evolving daily profile; a stochastic seasonal autoregression confined to one equation; and a multiplicative cross-product block. It also *quantifies* the deterministic–stochastic decomposition, with the transformation justified by formal tests and confirmed by the estimator's own subset selections. Finally, it *prices* the modelling choices in forecast terms—competitive accuracy at the innovation-noise floor, with the trade-offs of each transformation made explicit.

9. Conclusions

In this paper, we developed a unified Gibbs sampling framework for the full Bayesian analysis of SeVAR models, integrating automatic lag-level subset selection, coefficient estimation, and multi-step probabilistic forecasting within a single algorithm rather than using them as three disconnected procedures. We introduced a hierarchical Group-SSVS prior that attaches one Bernoulli latent indicator to each $k \times k$ lag matrix, rather than to every scalar coefficient or regressor column. Accordingly, we derived closed-form full conditionals: matrix normal for the coefficient matrix, Wishart for the precision matrix, Bernoulli for each block indicator, and multivariate normal for the one-step and r-step conditional predictive distributions. The matrix-normal and Wishart conditionals extend standard Bayesian-VAR conjugacy results to the block-indexed setting; the principal new result is the block-level Bernoulli conditional and its closed-form ratio, which is what makes joint selection, estimation, and forecasting tractable in a single conjugate Gibbs sweep.

The Monte Carlo study confirmed all three inferential dimensions. Block-level subset selection achieved a high proportion of correct identification across lag blocks and simulated SeVAR models. Posterior means of coefficient and precision matrices were nearly unbiased, with relatively small element-wise MAEs and RMSEs. Multi-step forecast errors were modest and increased monotonically with horizon, validating the forecasting accuracy of the proposed framework. Convergence diagnostics indicated rapid mixing throughout, consistent with the sharply reduced combinatorial indicator space induced by Block-SSVS. In the empirical application, applying Block-SSVS to the joint hourly dynamics of solar irradiance and air temperature—after removing the dominant time-evolving daily cycle—retained four of the candidate lag blocks, including the multiplicative cross-product block, a genuinely discovered interaction rather than one imposed by the simulation design. A direct comparison against the non-seasonal special case of the same estimator confirms that this seasonal structure is not an artefact of the selection procedure. The non-seasonal VAR leaves a pronounced, unrepresentable seasonal signature in its temperature residuals that the SeVAR removes almost entirely. An expanding-window rolling-origin forecast race further shows that the fitted SeVAR overwhelmingly outperforms a naive persistence benchmark but is not significantly more accurate than a seasonal-naive benchmark, the non-seasonal VAR, or alternative data transformations. The application's contribution is accordingly one of correct structural identification and reported forecast performance at the data's own predictability floor, not a claim of forecast superiority, and it demonstrates the practical applicability of the proposed full Bayesian SeVAR framework on genuinely bivariate stochastic data.

Future directions include the following: (i) exploring continuous global-local shrinkage alternatives to the discrete SSVS prior, such as the Bayesian lasso [27], and message-passing forecasting schemes for very large k [30], including group-wise continuous-shrinkage priors as a direct alternative to the discrete Block-SSVS scheme developed here [50]; (ii) considering stochastic-volatility or time-varying-coefficient extensions of the innovation covariance [10–12,31], in the spirit of the stochastic-volatility Bayesian VAR literature more broadly [51]; and (iii) a formal large-sample consistency result for the block-inclusion prob-

abilities, and an ergodicity argument for the Gibbs chain over the joint continuous-discrete state space, would strengthen the theoretical contribution beyond the conjugacy results derived here and are natural companions to the Monte Carlo evidence already reported.

Author Contributions: Conceptualisation, A.A.A.; methodology, A.A.A.; software, A.A.A.; validation, A.A.A. and F.E.A.; writing—original draft, A.A.A.; writing—review and editing, F.E.A.; project administration, F.E.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU266474].

Data Availability Statement: The dataset used in this study is publicly accessible through the GitHub repository at https://github.com/aymanaamin/solar_radiation_najran/ (accessed on 21 July 2026).

Acknowledgments: The authors are thankful to the Deanship of Graduate Studies and Scientific Research at Najran University for funding this work under the Najran Research Funding Program grant code (NU/NRP/SERC/14/4440-1).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ghysels, E.; Osborn, D.R. *The Econometric Analysis of Seasonal Time Series*; Cambridge University Press: Cambridge, UK, 2001.
2. Lütkepohl, H. *New Introduction to Multiple Time Series Analysis*; Springer: Berlin, Germany, 2006.
3. Canova, F.; Ghysels, E. Changes in seasonal patterns: Are they cyclical? *J. Econ. Dyn. Control* **1994**, *18*, 1143–1171.
4. Osborn, D.R. Seasonality and the order of integration for consumption. *Oxf. Bull. Econ. Stat.* **1988**, *50*, 361–377.
5. Kadiyala, K.R.; Karlsson, S. Numerical methods for estimation and inference in Bayesian VAR-models. *J. Appl. Econom.* **1997**, *12*, 99–132.
6. Litterman, R.B. Forecasting with Bayesian vector autoregressions—Five years of experience. *J. Bus. Econ. Stat.* **1986**, *4*, 25–38.
7. Doan, T.; Litterman, R.; Sims, C. Forecasting and conditional projection using realistic prior distributions. *Econom. Rev.* **1984**, *3*, 1–100.
8. Bańbura, M.; Giannone, D.; Reichlin, L. Large Bayesian vector auto regressions. *J. Appl. Econom.* **2010**, *25*, 71–92.
9. Giannone, D.; Lenza, M.; Primiceri, G.E. Prior selection for vector autoregressions. *Rev. Econ. Stat.* **2015**, *97*, 436–451.
10. Koop, G.; Korobilis, D. Large time-varying parameter VARs. *J. Econom.* **2013**, *177*, 185–198.
11. Huber, F.; Feldkircher, M. Adaptive shrinkage in Bayesian vector autoregressive models. *J. Bus. Econ. Stat.* **2019**, *37*, 27–39.
12. Chan, J.C.C. Large Bayesian VARs: A flexible Kronecker error covariance structure. *J. Bus. Econ. Stat.* **2020**, *38*, 68–79.
13. Shaarawy, S.M.; Ali, S.S. Bayesian identification of multivariate autoregressive processes. *Commun. Stat. Methods* **2008**, *37*, 791–802.
14. Shaarawy, S.M.; Ali, S.S. Bayesian identification of seasonal multivariate autoregressive processes. *Commun. Stat. Methods* **2015**, *44*, 823–836.
15. Sims, C.A. Macroeconomics and reality. *Econometrica* **1980**, *48*, 1–48.
16. Shaarawy, S.M. Bayesian modeling and forecasting of vector autoregressive moving average processes. *Commun. Stat.-Theory Methods* **2023**, *52*, 3795–3815.
17. Amin, A.A.; Almuhayfith, F.E. Bayesian Identification of Double-Seasonal Vector Autoregressive Models. *Mathematics* **2026**, *14*, 2476. https://doi.org/10.3390/math14142476.
18. George, E.I.; McCulloch, R.E. Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* **1993**, *88*, 881–889.
19. George, E.I.; McCulloch, R.E. Approaches for Bayesian variable selection. *Stat. Sin.* **1997**, *7*, 339–373.
20. Chipman, H.; George, E.I.; McCulloch, R.E. The practical implementation of Bayesian model selection. In *Model Selection*; *IMS Lecture Notes—Monograph Series*; Institute of Mathematical Statistics: Waite Hill, OH, USA, 2001; Volume 38, pp. 65–134.
21. Amin, A.A.; Emam, W.; Tashkandy, Y.; Chesneau, C. Bayesian subset selection of seasonal autoregressive models. *Mathematics* **2023**, *11*, 2878.
22. Amin, A.A. Full Bayesian analysis of seasonal autoregressive models under scale-mixtures of normal errors. *J. Stat. Comput. Simul.* **2026**, *96*, 184–204.
23. Amin, A.A. Full Bayesian analysis of double seasonal autoregressive models with real applications. *J. Appl. Stat.* **2024**, *51*, 1524–1544.
24. George, E.I.; Sun, D.; Ni, S. Bayesian stochastic search for VAR model restrictions. *J. Econom.* **2008**, *142*, 553–580.
25. Korobilis, D. VAR forecasting using Bayesian variable selection. *J. Appl. Econom.* **2013**, *28*, 204–230.

26. Frühwirth-Schnatter, S.; Wagner, H. Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *J. Econom.* **2010**, *154*, 85–100.
27. Park, T.; Casella, G. The Bayesian lasso. *J. Am. Stat. Assoc.* **2008**, *103*, 681–686.
28. Griffin, J.E.; Brown, P.J. Inference with normal-gamma prior distributions in regression problems. *Bayesian Anal.* **2010**, *5*, 171–188.
29. Ishwaran, H.; Rao, J.S. Spike and slab variable selection: Frequentist and Bayesian strategies. *Ann. Stat.* **2005**, *33*, 730–773.
30. Korobilis, D. Forecasting with many predictors using message passing algorithms. In *Essex Finance Centre Working Papers 19565*; University of Essex, Essex Business School: Colchester, UK, 2017. https://doi.org/10.2139/ssrn.2977838.
31. Carriero, A.; Clark, T.E.; Marcellino, M. Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *J. Econom.* **2019**, *212*, 137–154. Corrigendum in *J. Econom.* **2019**, *227*, 506–512.
32. Yuan, M.; Lin, Y. Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B* **2006**, *68*, 49–67.
33. Geman, S.; Geman, D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 721–741.
34. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2013.
35. Broemeling, L.D.; Shaarawy, S.M. Time series: A Bayesian analysis in the time domain. In *Bayesian Analysis of Time Series and Dynamic Models*; Spall, J.C., Ed.; Marcel Dekker: New York, NY, USA, 1988; pp. 1–22.
36. Carter, C.K.; Kohn, R. On Gibbs sampling for state space models. *Biometrika* **1994**, *81*, 541–553.
37. Raftery, A.E.; Lewis, S.M. The number of iterations, convergence diagnostics and generic Metropolis algorithms. *Pract. Markov Chain Monte Carlo* **1995**, *7*, 763–773.
38. Geweke, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian Stat.* **1992**, *4*, 169–193.
39. Hanley, J.A.; Lippman-Hand, A. If Nothing Goes Wrong, Is Everything All Right? Interpreting Zero Numerators. *JAMA* **1983**, *249*, 1743–1745. https://doi.org/10.1001/jama.1983.03330370053031.
40. Box, G.E.P.; Tiao, G.C. *Bayesian Inference in Statistical Analysis*; Addison-Wesley: Reading, MA, USA, 1973.
41. Zellner, A. Models, prior information, and Bayesian analysis. *J. Econom.* **1996**, *75*, 51–68.
42. Chatfield, C. Calculating interval forecasts. *J. Bus. Econ. Stat.* **1993**, *11*, 121–135.
43. West, M.; Harrison, J.; Migon, H.S. Dynamic generalised linear models and Bayesian forecasting. *J. Am. Stat. Assoc.* **1985**, *80*, 73–83.
44. Makridakis, S.; Hibon, M. The M3-competition: Results, conclusions and implications. *Int. J. Forecast.* **2000**, *16*, 451–476.
45. Almasoud, A.H.; Gandayh, H.M. Future of solar energy in Saudi Arabia. *J. King Saud Univ.-Eng. Sci.* **2015**, *27*, 153–157.
46. Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renew. Sustain. Energy Rev.* **2013**, *27*, 65–76.
47. Hylleberg, S.; Engle, R.F.; Granger, C.W.J.; Yoo, B.S. Seasonal integration and cointegration. *J. Econom.* **1990**, *44*, 215–238.
48. Canova, F.; Hansen, B.E. Are seasonal patterns constant over time? A test for seasonal stability. *J. Bus. Econ. Stat.* **1995**, *13*, 237–252.
49. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263.
50. Gruber, L.; Kastner, G. Forecasting macroeconomic data with Bayesian VARs: Sparse or dense? It depends! *Int. J. Forecast.* **2025**, *41*, 1589–1619. https://doi.org/10.1016/j.ijforecast.2025.02.001.
51. Clark, T.E.; Mertens, E. Stochastic Volatility in Bayesian Vector Autoregressions. In *Oxford Research Encyclopedia of Economics and Finance*; Oxford University Press: Oxford, UK, 2023. https://doi.org/10.1093/acrefore/9780190625979.013.919.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.