TMJM THE MINDANAWAN JOURNAL OF MATHEMATICS
OFFICIAL JOURNAL OF THE MINDANAO STATE UNIVERSITY - ILIGAN INSTITUTE OF TECHNOLOGY

Vol. 7 (2025), no. 2, pp. 149–171 ISSN: 2094-7380 (Print) | 2783-0136 (Online)
http://dx.doi.org/10.62071/tmjm.v7i2.776

A Bayesian Dynamic Spatiotemporal Model with Advection-Diffusion Motivation for Short-Term Probabilistic Forecasting of Shortwave Radiation Fields

AJ Jhones S. Lacubtan[1,2,*], Aljo Clair P. Pingal [1,2], & Arnulfo P. Supe [1,2]

[1]Department of Mathematics and Statistics
MSU-Iligan Institute of Technology, 9200 Iligan City, Philippines
ajjhones.lacubtan@g.msuiit.edu.ph

[2]Center for Computational Analytics and Modeling, PRISM
MSU-Iligan Institute of Technology, 9200 Iligan City, Philippines
aljoclair.pingal@g.msuiit.edu.ph, arnulfo.supe@g.msuiit.edu.ph

Received: 15th April 2025 Revised: 10th June 2025

Abstract

This study applies a Bayesian dynamic spatio-temporal modeling (DSTM) framework, motivated by advection–diffusion processes, to generate short-term probabilistic forecasts of shortwave radiation (SWR) fields over Mindanao, Philippines. Using the Himawari-9 Level 2 Short Wave Radiation product at 5 km spatial and 10-minute temporal resolution, we model the latent irradiance field as evolving under a stochastic partial differential equation (SPDE) with diffusion and time-varying advection components. The continuous formulation is discretized via a finite-difference scheme, resulting in sparse, linear-Gaussian dynamics suitable for state-space modeling. Inference is carried out using a Gibbs sampling algorithm that integrates a fixed-lag ensemble Kalman smoother (EnKS) to efficiently approximate posterior distributions of the high-dimensional latent states. The model is applied to satellite observations collected, and posterior summaries provide insight into the temporal evolution of advection parameters, spatial uncertainty, and irradiance exceedance probabilities. Posterior predictive simulations yield short-term forecasts that remain coherent up to 30 minutes ahead, after which forecast uncertainty increases substantially. Results demonstrate the potential of the DSTM framework for assimilating satellite data and delivering calibrated probabilistic nowcasts of irradiance in tropical environments, supporting decision-making in solar energy operations and other weather-sensitive sectors.

1 Introduction

Short-wave radiation (SWR) plays a key role in connecting Mindanao's two most climate-sensitive sectors: electricity generation and agriculture. The Philippine Energy Plan 2023–2050 projects more than 1.5 GW of new utility-scale photovoltaics on the island, raising renewable penetration on the Mindanao grid above 50 % by 2030 [8]. Operating a power system with such weather-dependent generation requires forecasts that resolve the convective cloud

*Corresponding author
2020 Mathematics Subject Classification: 62F15, 62M30, 65M06
Keywords and Phrases: Bayesian inference, spatio-temporal modeling, advection–diffusion, ensemble Kalman smoother, short-wave radiation, probabilistic forecasting
This research is supported by the DOST-ASTHRDP Grant

systems driving minute-scale irradiance ramps and, crucially, that quantify their uncertainty. Ensemble extensions of the Weather Research and Forecasting model with solar augmentations (WRF-Solar) already demonstrate that 5 km, 10-min probabilistic irradiance forecasts lower reserve procurement costs and improve reliability relative to deterministic baselines [16, 29]. Yet the governing transport physics of tropical clouds—advection by low-level winds and diffusion by turbulent mixing—are only crudely represented in existing numerical-weather-prediction (NWP) post-processing, motivating a Bayesian dynamic spatio-temporal model explicitly rooted in advection–diffusion dynamics.

Recent literature highlights two key developments relevant to short-term irradiance forecasting. First, ensemble extensions of numerical weather prediction (NWP) models like WRF-Solar have demonstrated improvements in probabilistic forecasting of solar irradiance at resolutions suitable for grid operation [16, 29]. However, such approaches typically lack explicit representation of the underlying physical processes, particularly the advection–diffusion dynamics governing cloud-driven irradiance variability. Second, high-resolution satellite data products such as the Himawari-9 Short Wave Radiation (SWR) product now offer continuous observations of irradiance fields at 10-minute and 5-kilometer resolution, but these observational data are rarely integrated within a physically interpretable, probabilistic forecasting framework. This study addresses this gap by developing a Bayesian dynamic spatiotemporal model explicitly driven by advection–diffusion physics, which assimilates Himawari-9 SWR data to generate short-term probabilistic forecasts of irradiance fields, providing a physically coherent and statistically rigorous approach that complements and extends beyond existing deterministic or purely data-driven methodologies.

Table 1 summarises representative probabilistic solar-irradiance models drawn from recent literature to clarify how existing work aligns with three complementary developments we integrate: (i) physics-driven advection–diffusion dynamics, (ii) fully dynamic spatio-temporal state-space structure, and (iii) ensemble Kalman smoothing within a Bayesian hierarchy. Calibrated machine-learning or NWP-ensemble post-processing methods deliver good short-horizon skill but generally omit an explicit advection–diffusion state-space layer at 5 km/10-min resolution—precisely the gap our framework targets.

The subsequent discussion details a physically motivated spatio-temporal statistical model that sits within the hierarchical Dynamic Spatio-Temporal Model (DSTM) framework [3, 27, 7]. In a DSTM the joint space–time dependence is induced through a pair of conditional stages—a process-evolution equation and an observation equation—so that the latent field follows a Gaussian vector autoregressive (VAR) dynamic. This conditional construction bypasses the notorious difficulty of specifying non-separable space-time covariance functions directly while retaining full probabilistic coherence. Consequently, DSTMs have become a work-horse for weather- and climate-scale problems, including multivariate extremes [18, 24] and non-stationary rainfall fields [21, 14]. The resulting Kalman-filter (or ensemble) algorithms deliver scalable inference, seamless data assimilation, and automatic propagation of parameter uncertainty—capabilities that are essential for the 5 km, 10-min short-wave-radiation forecasts targeted in this study.

We include a stochastic advection–diffusion partial differential equation (SPDE) in the process layer since variability in cloud-driven irradiance is primarily influenced by wind movement (advection) and local turbulence (diffusion). This approach produces realistic space–time correlations, clearly reflects physical processes, and—after discretization using finite differences—results in sparse, linear-Gaussian dynamics compatible with the dynamic spatiotemporal modeling (DSTM) framework [4, 26]. Recent studies indicate that such advection–diffusion models improve the accuracy of high-resolution precipitation forecasts [14] and form the basis of advanced solar irradiance prediction systems such as CIADCast [1]. Additionally, Clarotto et al. [6] demonstrated that accurate approximations to advection–diffusion SPDEs can be ob-

Table 1: Representative probabilistic solar-irradiance forecasting models

Study	Framework	Data / resolution / lead	Key finding
Zelikman et al. [30]	NGBoost + calibration	SURFRAD GHI; 5–60 min	Lowest intra-hour CRPS among tested baselines
Bai et al. [2]	Deep ensemble ConvLSTM	Multi-site GHI; 10–60 min	Better CRPS than deterministic counterparts
Feng et al. [10]	PSolarNet (CNN+BMA)	All-sky images; 10 min	5.62% nRMSE; 2.77% nCRPS at 10 min
Opoku et al. [22]	Bayesian copula regression	FAWN GHI; 15–60 min	CRPS cut from 166.2 to 27.7 vs. baseline
Sigauke et al. [25]	Spatiotemporal Gaussian-process (GP) model	AWS network; 1 h–1 d	≈12% MAE reduction over benchmarks
Kim et al. [17]	WRF-Solar ensemble + AnEn	NWP ensemble; ∼3–9 km; 1–36 h	Post-processed ensembles show higher reliability

tained through a combination of spatial finite element methods and implicit Euler methods for time discretization.

The high dimensionality of both the observation vector and the latent irradiance field renders exact Bayesian filtering impractical. We therefore pursue a simulation-based strategy that integrates data augmentation with a Gibbs–ensemble Kalman smoother (GEnKS). Missing observations are handled naturally through the data-augmentation principle of [28]. Although joint state draws could in theory be obtained by the forwards-filtering backwards-sampling (FFBS) algorithm of [5], its $O(Tn^3)$ cost is prohibitive for our $n \approx 2 \times 10^4$ grid. We instead replace the exact draw with an approximate sample generated by the ensemble Kalman smoother of [9], reducing the computational burden to nearly linear in n. All remaining model parameters have closed-form full conditional distributions, so the resulting sampler coincides with the GEnKS scheme of [15], delivering accurate posterior summaries at a fraction of the cost of traditional FFBS.

The remainder of this paper is organised as follows. Section 2 introduces the advection–diffusion vector-autoregressive process model and explains its physical rationale. The corresponding observation equation is described in Section 3. Section 4 completes the hierarchical specification by assigning prior distributions, while the simulation-based inference strategy is detailed in Section 5. The methodology is then applied to Himawari-9 short-wave-radiation data in Section 6. Finally, Section 7 summarises the main findings and outlines directions for future work. For ease of reference, a summary of the main symbols used throughout the paper is provided in Table 2 in Appendix A.

2 Process Model

We represent the latent irradiance field over Mindanao as the solution to an advection–diffusion stochastic partial differential equation (SPDE). Let $\mathcal{S} \subset \mathbb{R}^2$ denote the set of centroids corresponding to an $n_x \times n_y$ regular grid covering the island, where each centroid is given by $\{(x_i, y_j)\}$. Furthermore, let $\mathcal{T} = \{1, \ldots, T\}$ index the ten-minute time steps.

For any $(s,t) \in \mathcal{S} \times \mathcal{T}$, the continuous-time irradiance $Y(s,t)$ is governed by the linear SPDE

$$\partial_t Y = b\nabla^2 Y - \mathbf{v}(s,t) \cdot \nabla Y + \varepsilon(s,t), \tag{1}$$

where $b > 0$ is the diffusion coefficient, $\mathbf{v} = (v_x, v_y)^\top$ is a divergence-free wind field that facilitates cloud advection, and ε denotes a zero-mean Gaussian process noise.

http://dx.doi.org/10.62071/tmjm.v7i2.776

Appendix B provides a detailed account of the finite-difference discretisation of (1). Here, we present the resulting discrete dynamics.

Assuming a spatial step Δ_d and a unit time step ($\Delta_t = 1$), we define

$$\beta = \frac{b}{\Delta_d^2}, \qquad \nu_x = \frac{v_x}{2\Delta_d}, \qquad \nu_y = \frac{v_y}{2\Delta_d}.$$

Then, the irradiance at the location (x, y) and time t evolves according to

$$\begin{aligned} Y_t(x,y) = (1-4\beta)\, Y_{t-1}(x,y) + (\beta-\nu_x)\, Y_{t-1}(x+\Delta_d,y) + (\beta+\nu_x)\, Y_{t-1}(x-\Delta_d,y) \\ + (\beta-\nu_y)\, Y_{t-1}(x,y+\Delta_d) + (\beta+\nu_y)\, Y_{t-1}(x,y-\Delta_d) + \varepsilon_t(x,y), \end{aligned} \tag{2}$$

where $\varepsilon_t(x,y) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\sigma_\eta^2)$. Periodic boundary conditions are applied at the domain edges to complete the stencil.

2.1 Spatial Dependence Structure

Equation (2) implies a five-point stencil arranged around the central grid cell (x, y) as follows:

$$\begin{bmatrix} - & \beta-\nu_y & - \\ \beta+\nu_x & 1-4\beta & \beta-\nu_x \\ - & \beta+\nu_y & - \end{bmatrix}. \tag{3}$$

This stencil couples each current cell with its four cardinal neighbours only, yielding a sparse precision structure that is exploited in Section 5.

2.2 State–Space Representation

Let $N = n_x n_y$ and stack the grid in row-major order to obtain the state vector $\mathbf{Y}_t \in \mathbb{R}^N$, $\mathbf{Y}_t = \big(Y_t(x_1,y_1),\ldots,Y_t(x_{n_x},y_{n_y})\big)^\top$. Write $\boldsymbol{\nu}_t = (\nu_x(t),\nu_y(t))^\top$ for the time-varying advection parameters and collect them in the state equation

$$\mathbf{Y}_t = \mathbf{M}(\boldsymbol{\nu}_t)\,\mathbf{Y}_{t-1} + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0},\sigma_\eta^2\mathbf{I}_N), \tag{4}$$

where $\mathbf{M}(\boldsymbol{\nu}_t)$ encodes the stencil (3) under periodic boundaries. The advection parameters themselves follow a first-order autoregression

$$\boldsymbol{\nu}_t = \alpha\,\boldsymbol{\nu}_{t-1} + \boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \sim \mathcal{N}(\mathbf{0},\sigma_\nu^2\mathbf{I}_2), \tag{5}$$

with $|\alpha| < 1$ to ensure stationarity. Equations (4)–(5) constitute the Gaussian state-space model that underpins the Bayesian inference strategy developed in Section 5.

3 Observation Model

This section specifies how measured Short Wave Radiation (SWR) data are linked to the latent irradiance field introduced in Section 2. In practice, some measurements are missing or discarded due to data-quality issues; hence, the model incorporates missing entries via a data-augmentation scheme. At each grid cell $s=(x,y) \in \mathcal{S}$ and time $t \in \mathcal{T}$, the observed SWR $Z_t(s)$ is modeled as

$$Z_t(s) \;=\; Y_t(s) \;+\; \varepsilon_t(s), \tag{6}$$

where $Y_t(s)$ is the latent irradiance from the process model (4) and $\varepsilon_t(s) \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is the measurement error. To allow for missing or invalid readings, we define:

$$Z_t^{\text{obs}}(s) \;=\; \begin{cases} Z_t(s), & \text{if observed,} \\ \text{NA}, & \text{if missing,} \end{cases}$$

and treat any missing $Z_t(s)$ as an unobserved variable $Z_t^{\text{mis}}(s)$, which we impute via the conditional distribution

$$Z_t^{\text{mis}}(s) \;\big|\; Y_t(s), \sigma_\varepsilon^2 \;\sim\; \mathcal{N}\big(Y_t(s),\, \sigma_\varepsilon^2\big).$$

This data-augmentation procedure preserves the joint Gaussianity of the observation vector when some components are missing.

Collecting all spatial locations into $\mathbf{Z}_t = (Z_t(s_1), \ldots, Z_t(s_n))^\top$ and similarly $\mathbf{Y}_t$ from (4), the vectorized observation model is

$$\mathbf{Z}_t \;=\; \mathbf{Y}_t \;+\; \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \;\sim\; \mathcal{N}_n\big(\mathbf{0},\, \sigma_\varepsilon^2\, \mathbf{I}_n\big), \tag{7}$$

where the missing entries of $\mathbf{Z}_t$ are treated as latent variables. We assume a constant error variance σ_ε^2, reflecting the fairly uniform methodology used to derive SWR from Himawari-9 data; note that this variance is distinct from the process innovation variance σ_η^2 and the advection innovation variance σ_ν^2 appearing in the state equations. Although errors are taken independent across space and time, this can be relaxed if future analyses reveal systematic correlation patterns.

In summary, the data model closes the hierarchical specification by linking the advection–diffusion state process of Section 2 to real-world SWR measurements. In combination with the stochastic process model, this treatment of missing observations enables a fully Bayesian inference scheme that assimilates irregularly sampled data without sacrificing the Gaussian structure.

4 Bayesian Hierarchical Specification

With the process model in (4)–(5) and the observation model in (7) in hand, we now collect them into a single Bayesian hierarchy. Let $\mathbf{Z}_{1:T} = \{\mathbf{Z}_t\}_{t=1}^T$ denote the observed (short-wave-radiation) data, $\mathbf{Y}_{0:T} = \{\mathbf{Y}_t\}_{t=0}^T$ the latent irradiance states, and $\boldsymbol{\nu}_{0:T} = \{\boldsymbol{\nu}_t\}_{t=0}^T$ the time-varying advection parameters. So we have,

$$\begin{aligned} \mathbf{Z}_t \mid \mathbf{Y}_t, \sigma_\varepsilon^2 &\overset{\text{ind}}{\sim} \mathcal{N}_N\big(\mathbf{Y}_t,\, \sigma_\varepsilon^2 \mathbf{I}_n\big), \\ \mathbf{Y}_t \mid \mathbf{Y}_{t-1}, \boldsymbol{\nu}_t, \sigma_\eta^2 &\overset{\text{ind}}{\sim} \mathcal{N}_N\big(\mathbf{M}(\boldsymbol{\nu}_t)\, \mathbf{Y}_{t-1},\, \sigma_\eta^2 \mathbf{I}_N\big), \\ \boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1}, \alpha, \sigma_\nu^2 &\overset{\text{ind}}{\sim} \mathcal{N}_2\big(\alpha\, \boldsymbol{\nu}_{t-1},\, \sigma_\nu^2 \mathbf{I}_2\big), \end{aligned} \tag{8}$$

for $t = 1, \ldots, T$. The process innovations $\boldsymbol{\eta}_t = \mathbf{Y}_t - \mathbf{M}(\boldsymbol{\nu}_t)\mathbf{Y}_{t-1}$ and advection innovations $\boldsymbol{\xi}_t = \boldsymbol{\nu}_t - \alpha\boldsymbol{\nu}_{t-1}$ are assumed independent, zero-mean Gaussian, and independent across time:

$$\boldsymbol{\eta}_t \sim \mathcal{N}_N(\mathbf{0}, \sigma_\eta^2 \mathbf{I}_N), \qquad \boldsymbol{\xi}_t \sim \mathcal{N}_2(\mathbf{0}, \sigma_\nu^2 \mathbf{I}_2).$$

Here σ_η^2 controls the magnitude of process innovations in the latent irradiance field, σ_ν^2 governs the variability of the advection parameters, and σ_ε^2 (from Section 3) represents measurement-error variance in the satellite-retrieved SWR. These three variance components play distinct roles in the hierarchy and are not used interchangeably. For the initial states we specify diffuse Gaussian priors

$$\mathbf{Y}_0 \sim \mathcal{N}_N(\mathbf{m}_0, \mathbf{C}_0), \qquad \boldsymbol{\nu}_0 \sim \mathcal{N}_2(\mathbf{m}_{\nu 0}, \mathbf{C}_{\nu 0}),$$

http://dx.doi.org/10.62071/tmjm.v7i2.776

where $\mathbf{m}_0, \mathbf{m}_{\nu 0}$ and $\mathbf{C}_0, \mathbf{C}_{\nu 0}$ are chosen to encode vague prior information about irradiance levels and prevailing wind velocities, respectively. The hierarchy is completed by assigning prior distributions to $\boldsymbol{\Theta}$; these and the resulting joint posterior density are introduced in the next subsection.

4.1 Prior Distribution

In keeping with the Gaussian state–space structure, we use weakly informative priors that respect basic physical constraints. The initial advection vector is assigned

$$\boldsymbol{\nu}_0 \sim \mathcal{N}_2(\mathbf{0}, 0.001\,\mathbf{I}_2),$$

and the initial irradiance field

$$\mathbf{Y}_0 \sim \mathcal{N}_N(\mathbf{m}_0, \mathbf{C}_0),$$

where $(\mathbf{m}_0, \mathbf{C}_0)$ are obtained from a short pre-training window of Himawari-9 frames and provide a vague, data-driven prior for the illumination field (see Section 6). The diffusion coefficient is given a stability-aware prior

$$\beta \sim \mathrm{Unif}(0, 0.25),$$

restricting it to the Courant–Friedrichs–Lewy stable range implied by the finite-difference stencil, while the advection autoregressive parameter follows

$$\alpha \sim \mathrm{TN}(0.5, 0.01; 0, 1),$$

favouring moderately persistent yet stationary dynamics. For the variance components we fix σ_ε^2 using external Himawari-8/9 validation studies and choose σ_ν^2 so that the implied variability in $\boldsymbol{\nu}_t$ corresponds to realistic low-level winds; attempts to estimate all variance terms freely led to weak identifiability and poor mixing in pilot runs. The state-innovation variance σ_η^2 is treated as a tuning parameter and selected from a small grid of candidates using DIC and WAIC in the empirical application, but the framework can accommodate conjugate Inverse-Gamma priors if full hierarchical learning is desired.

4.2 Joint Posterior Density

Combining the observation equation (7) with the state dynamics (4) and (5) and the priors specified above yields the joint posterior distribution of the unknown quantities $\{\mathbf{Y}_{0:T}, \boldsymbol{\nu}_{0:T}, \beta, \alpha\}$ given the data $\mathbf{Z}_{1:T}$:

$$\begin{aligned}
\pi\big(\mathbf{Y}_{0:T}, \boldsymbol{\nu}_{0:T}, \beta, \alpha \mid \mathbf{Z}_{1:T}\big) \;\propto\; & \underbrace{\prod_{t=1}^{T} \mathcal{N}_n\big(\mathbf{Z}_t \mid \mathbf{Y}_t, \sigma_\varepsilon^2 \mathbf{I}_n\big)}_{\text{likelihood}} \\
& \times \underbrace{\prod_{t=1}^{T} \mathcal{N}_N\big(\mathbf{Y}_t \mid \mathbf{M}(\boldsymbol{\nu}_t)\mathbf{Y}_{t-1}, \sigma_\eta^2 \mathbf{I}_N\big)}_{\text{process}} \\
& \times \underbrace{\prod_{t=1}^{T} \mathcal{N}_2\big(\boldsymbol{\nu}_t \mid \alpha\, \boldsymbol{\nu}_{t-1}, \sigma_\nu^2 \mathbf{I}_2\big)}_{\text{advection}} \\
& \times \mathcal{N}_N(\mathbf{Y}_0 \mid \mathbf{m}_0, \mathbf{C}_0)\, \mathcal{N}_2(\boldsymbol{\nu}_0 \mid \mathbf{m}_{\nu 0}, \mathbf{C}_{\nu 0}) \\
& \times \mathcal{N}(\beta \mid m_\beta, v_\beta)\, \mathrm{TN}(\alpha \mid m_\alpha, v_\alpha; 0, 1). \qquad (9)
\end{aligned}$$

Define the augmented state vector $\mathbf{x}_t = \left(\mathbf{Y}_t^\top, \boldsymbol{\nu}_t^\top\right)^\top \in \mathbb{R}^{N+2}$ and the block matrices

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_N & \mathbf{0}_{N\times 2} \end{bmatrix}, \qquad \mathbf{G}_t(\beta, \alpha) = \begin{bmatrix} \mathbf{M}(\boldsymbol{\nu}_t) & \mathbf{0}_{N\times 2} \\ \mathbf{0}_{2\times N} & \alpha \mathbf{I}_2 \end{bmatrix}.$$

Equations (7), (4) and (5) can then be rewritten in the familiar dynamic linear model (DLM) form

$$\mathbf{Z}_t = \mathbf{H}\,\mathbf{x}_t + \boldsymbol{\varepsilon}_t, \qquad \mathbf{x}_t = \mathbf{G}_t\,\mathbf{x}_{t-1} + \mathbf{w}_t,$$

with $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_n)$ and $\mathbf{w}_t = \left(\boldsymbol{\eta}_t^\top, \boldsymbol{\xi}_t^\top\right)^\top$ Gaussian and block-diagonal in covariance. Because the joint distribution of $\left\{\mathbf{x}_{0:T}, \mathbf{Z}_{1:T}\right\}$ is Gaussian, the full conditional density $\pi(\mathbf{Y}_{0:T} \mid \cdot)$ (and analogously $\pi(\boldsymbol{\nu}_{0:T} \mid \cdot)$) is available in closed form and, in moderate dimensions, a draw can be obtained via the forwards-filtering backwards-sampling (FFBS) algorithm of [12]. In our application the latent field dimension is $N = 124 \times 176 = 21{,}824$, so FFBS would require manipulating $21{,}824 \times 21{,}824$ covariance matrices at every iteration—an intractable burden. Instead we replace the exact FFBS step with an approximate draw produced by the ensemble Kalman smoother (EnKS), yielding the Gibbs–EnKS sampler introduced in Section 5. The remaining full conditional distributions for β and α follow immediately from (9) and admit direct sampling because both priors are conjugate (Gaussian or truncated Gaussian).

5 Posterior Computation

Inference proceeds by drawing samples from the joint posterior density in (9). Although closed-form full conditional distributions (FCDs) are available for all unknown quantities, the high dimensionality of the latent state $\mathbf{Y}_{0:T}$ (with $N = 21\,824$ per time point) makes exact joint sampling via forwards-filtering backwards-sampling (FFBS) computationally prohibitive. Element-wise updates, while avoiding expensive covariance computations, would lead to slow mixing. Similarly, designing a global Metropolis–Hastings proposal for the high-dimensional state would be challenging and likely yield low acceptance rates.

To address these issues, we propose a hybrid strategy: the latent state $\mathbf{Y}_{0:T}$ is updated using a fixed-lag ensemble Kalman smoother (EnKS), which provides an approximate draw from $\pi(\mathbf{Y}_{0:T} \mid \cdot)$, while the advection parameters $\boldsymbol{\nu}_{0:T}$ and the static parameters β and α are updated via their closed-form full conditional distributions. The resulting Gibbs sampler alternates between the EnKS step for the latent state and direct Gibbs updates for the remaining parameters, yielding a computationally tractable and efficient scheme. The next subsection details the EnKS step and its integration into the overall sampler.

5.1 State Sampling using Fixed–lag Ensemble Kalman Smoother

Our state sampling algorithm is based on the ensemble Kalman smoother (EnKS) of [9], which has proven effective in nonlinear and high-dimensional settings. Key to the successful implementation of the EnKS is the accurate estimation of the (cross-)covariance matrices needed to compute the Kalman gain, particularly when the ensemble size N_e is small relative to the state or observation dimensions. In our approach, these covariance matrices are approximated from the so-called deterministic ensemble, thereby ensuring robust performance in practice.

http://dx.doi.org/10.62071/tmjm.v7i2.776

Algorithm 1 Fixed–lag Ensemble Kalman Smoother (EnKS)

Require: Observations $\mathbf{Z}_{1:T}$, ensemble size N_e, lag L, model parameters $\{\boldsymbol{\nu}_t\}$, noise variances $\sigma_\eta^2, \sigma_\varepsilon^2$, initial ensemble $\{\mathbf{Y}_0^{(j)}\}_{j=1}^{N_e}$.

for $t = 1, \ldots, T$ **do**

Forecast. For each $j = 1, \ldots, N_e$, draw process noise $\boldsymbol{\eta}_t^{(j)} \sim \mathcal{N}(\mathbf{0}, \sigma_\eta^2 \mathbf{I})$ and propagate

$$\mathbf{Y}_{t|t-1}^{(j)} = \mathbf{M}(\boldsymbol{\nu}_t)\, \mathbf{Y}_{t-1}^{(j)} + \boldsymbol{\eta}_t^{(j)}.$$

Generate pseudo-observations

$$\mathbf{Z}_{t|t-1}^{(j)} = \mathbf{Y}_{t|t-1}^{(j)} + \boldsymbol{\varepsilon}_t^{(j)}, \quad \boldsymbol{\varepsilon}_t^{(j)} \sim \mathcal{N}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}).$$

Form ensemble matrices $\mathbf{Y}_f = \left[\mathbf{Y}_{t|t-1}^{(1)}, \ldots, \mathbf{Y}_{t|t-1}^{(N_e)}\right]$ and $\mathbf{Z}_f = \left[\mathbf{Z}_{t|t-1}^{(1)}, \ldots, \mathbf{Z}_{t|t-1}^{(N_e)}\right]$.

Compute ensemble means $\bar{\mathbf{y}}_t$ and $\bar{\mathbf{z}}_t$, and anomaly matrices

$$\mathbf{Y}' = \mathbf{Y}_f - \bar{\mathbf{y}}_t \mathbf{1}^\top, \qquad \mathbf{Z}' = \mathbf{Z}_f - \bar{\mathbf{z}}_t \mathbf{1}^\top.$$

Compute the Kalman gain

$$\mathbf{K}_t = \frac{1}{N_e - 1} \mathbf{Y}' \mathbf{Z}'^\top \left(\frac{1}{N_e - 1} \mathbf{Z}' \mathbf{Z}'^\top + \sigma_\varepsilon^2 \mathbf{I} \right)^{-1}.$$

Analysis update (smoothing). For each lag $l = \max(1, t - L + 1), \ldots, t$:

Compute the state anomaly matrix $\mathbf{Y}_l'$ from the ensemble $\{\mathbf{Y}_l^{(j)}\}_{j=1}^{N_e}$.

Compute the smoother gain

$$\mathbf{K}_{l,t} = \frac{1}{N_e - 1} \mathbf{Y}_l' \mathbf{Z}'^\top \left(\frac{1}{N_e - 1} \mathbf{Z}' \mathbf{Z}'^\top + \sigma_\varepsilon^2 \mathbf{I} \right)^{-1}.$$

Update the state ensemble at time l:

$$\mathbf{Y}_l^{(j)} \leftarrow \mathbf{Y}_l^{(j)} + \mathbf{K}_{l,t} \left(\mathbf{Z}_t^{\text{obs}} - \mathbf{Z}_{t|t-1}^{(j)} \right), \quad j = 1, \ldots, N_e.$$

end for

Output: Smoothed ensemble trajectories $\{\mathbf{Y}_{0:T}^{(j)}\}_{j=1}^{N_e}$.

Algorithm 1 outlines the fixed–lag EnKS procedure. In each assimilation step, an ensemble of latent states is first propagated forward according to the advection–diffusion process model (4). Pseudo-observations are then generated using the Gaussian observation model (7), and ensemble anomalies are computed to estimate the covariance matrices. The Kalman gain is calculated using a regularised inversion via a Sherman–Morrison–Woodbury scheme, ensuring computational efficiency. Finally, for a fixed lag L, the analysis update retroactively adjusts the state estimates over previous time steps to incorporate the new information, yielding a smoothed ensemble approximation of $\pi(\mathbf{Y}_{0:T} \mid \cdot)$. When $L = 0$ the procedure reduces to the standard ensemble Kalman filter (EnKF); choosing $L > 0$ distributes each innovation backward in time, producing a smoothed ensemble that approximates the joint draw from $\pi(\mathbf{Y}_{0:T} \mid \cdot)$ required by the Gibbs sampler.

5.2 MCMC Sampling Algorithm

Our MCMC sampling algorithm targets the joint posterior distribution

$$\pi\Big(\mathbf{Y}_{0:T}, \boldsymbol{\nu}_{0:T}, \beta, \alpha \mid \mathbf{Z}_{1:T}\Big)$$

as given in (9). In our Gibbs sampler, the latent state $\mathbf{Y}_{0:T}$ is updated using the fixed–lag ensemble Kalman smoother (EnKS), following the approach in [15]—in which a single state trajectory is selected uniformly from the ensemble produced by the EnKS. The advection parameters $\boldsymbol{\nu}_{0:T}$ and the static parameters β and α are then updated from their closed–form full conditional distributions (see Appendix C).

Algorithm 2 Gibbs Sampling for the Joint Posterior

Require: Data $\mathbf{Z}_{1:T}$, initial values $\mathbf{Y}_{0:T}^{(0)}$, $\boldsymbol{\nu}_{0:T}^{(0)}$, $\beta^{(0)}$, $\alpha^{(0)}$, number of iterations K.

1: **for** $k = 1, \ldots, K$ **do**

2: **State update:** Sample $\mathbf{Y}_{0:T}^{(k,1:N_e)}$ from $\mathbf{Y}_{0:T}^{(k)} \sim \pi\Big(\mathbf{Y}_{0:T} \mid \mathbf{Z}_{1:T}, \boldsymbol{\nu}_{0:T}^{(k-1)}, \beta^{(k-1)}, \alpha^{(k-1)}\Big)$ using the Algorithm 1, and set $\mathbf{Y}_{0:T}^{(k)} = Y_{0:T}^{(k,l)}$ for a j sampled uniformly at random from $\{1, \ldots, N_e\}$

3: **Advection update:** For each $t = 0, \ldots, T$, sample

$$\boldsymbol{\nu}_t^{(k)} \sim \pi\Big(\boldsymbol{\nu}_t \mid \mathbf{Y}_{0:T}^{(k)}, \beta^{(k-1)}, \alpha^{(k-1)}, \mathbf{Z}_{1:T}\Big).$$

4: **Static parameters update:** Sample

$$\beta^{(k)} \sim \pi\Big(\beta \mid \mathbf{Y}_{0:T}^{(k)}, \boldsymbol{\nu}_{0:T}^{(k)}, \mathbf{Z}_{1:T}\Big)$$

and

$$\alpha^{(k)} \sim \pi\Big(\alpha \mid \mathbf{Y}_{0:T}^{(k)}, \boldsymbol{\nu}_{0:T}^{(k)}, \mathbf{Z}_{1:T}\Big).$$

5: **end for**

6: **Output:** Posterior samples $\{\mathbf{Y}_{0:T}^{(k)}, \boldsymbol{\nu}_{0:T}^{(k)}, \beta^{(k)}, \alpha^{(k)}\}_{k=1}^{K}$.

Despite the serial nature of Gibbs samplers, we achieved substantial runtime reductions by parallelising the EnKS computations within each iteration. The EnKS core is implemented in Julia and called from a Python driver that handles the remaining Gibbs updates. All experiments were run on a Paperspace Linux server with a single Intel® Xeon Gold 6226R CPU (12 cores at 2.9 GHz) and 16 GB RAM. For the Mindanao experiment ($N = 21{,}824$, $T = 29$, ensemble size $N_{\text{ens}} = 100$), 2,000 GEnKS iterations (with 1,000 burn-in) required approximately 3–4 hours of wall-clock time and comfortably fit within the available memory; storing 500 thinned posterior draws of the latent fields and advection parameters occupied about 2.5 GB of compressed disk space. Since the dominant EnKS operations scale approximately linearly in N and T under the sparse stencil, the framework can be extended to moderately larger or more frequent grids, or to near-real-time operation, by modestly increasing computing resources or reducing the ensemble size.

http://dx.doi.org/10.62071/tmjm.v7i2.776

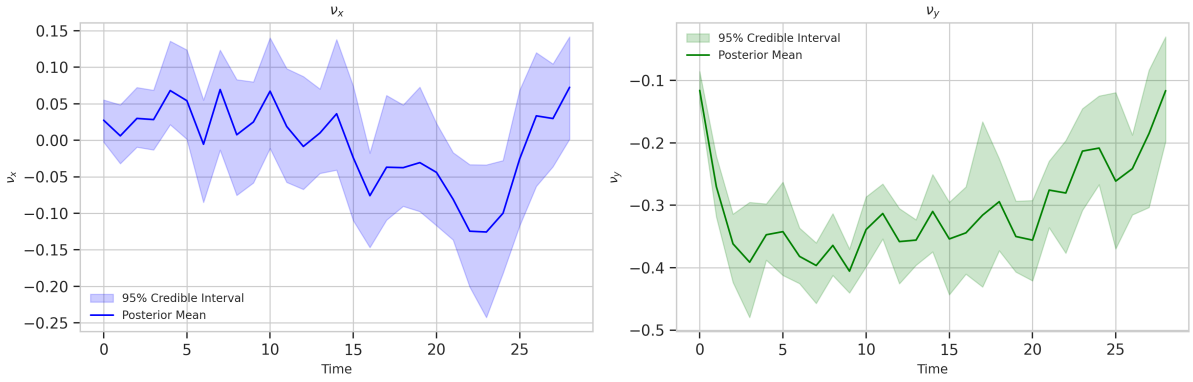

Figure 1: Posterior distributions of the advection parameters ν_x (left panel) and ν_y (right panel). The shaded regions represent the 95% credible intervals computed from the posterior samples (excluding the last time point), and the solid lines indicate the posterior means. This visualization summarizes the evolution of the advection parameters over time.

6 Empirical Application to Himawari-9 Shortwave Radiation over Mindanao

This study employs the Short Wave Radiation (SWR) Level 2 product Version 2.1 derived from Himawari-9 data provided by the P-Tree System, Japan Aerospace Exploration Agency (JAXA). The SWR data, expressed in watts per square meter ($\mathrm{W\,m^{-2}}$), cover the full-disk observation area with a spatial resolution of 5 km and temporal resolution of 10 minutes. Specifically, data corresponding to April 4, 2025, spanning from 00:00 to 23:50 UTC, were downloaded and subsequently subsetted to the geographic extent covering Mindanao, Philippines, defined by the bounding coordinates 4.47° N to 10.67° N latitude and 117.86° E to 126.66° E longitude. The Himawari-9 SWR product relies on algorithms detailed in Frouin and Murakami [11] and utilizes ancillary ozone data sourced from the Japan Meteorological Agency's global chemical transport model MRI-CCM2 [13]. The product employed, originally developed for Himawari-8, incorporates vicarious calibration coefficients updated in Version 2.1 to enhance the temporal accuracy of irradiance estimates.

For model training and evaluation, the dataset was partitioned temporally into two segments following Philippine Standard Time (PST). The training set spans from 9:00 AM to 2:00 PM PST, and the test set covers from 2:00 PM to 3:00 PM PST, both on April 4, 2025. Consequently, the resulting length of the training time series is $T = 29$ time steps, with each time step comprising $N = 124 \times 176 = 21,824$ spatial observations.

Our primary objective is to infer the latent irradiance field at ground level. Since direct ground truth observations are unavailable, we account for the measurement uncertainty of the satellite-derived irradiance by fixing the observation error variance to $\sigma_\varepsilon^2 = 1200$, corresponding to a standard deviation of about $35\,\mathrm{W\,m^{-2}}$. This choice is informed by Himawari-8/9 validation studies, which report typical midday uncertainties in the range 30–100 $\mathrm{W/m^2}$ [19, 20]. The remaining priors follow Section 4.1: $\boldsymbol{\nu}_0 \sim \mathcal{N}(\mathbf{0}, 0.001\,\mathbf{I}_2)$, $\beta \sim \mathrm{Unif}(0, 0.25)$, and $\alpha \sim \mathrm{TN}(0.5, 0.01; 0, 1)$. For the initial latent field $\mathbf{Y}_0$, we adopt a data-driven strategy analogous to Primiceri [23], using a short pre-training period to form an empirical mean image and adding Gaussian perturbations from $\mathcal{N}(\mathbf{0}, 5000\,\mathbf{I})$ to encode vague but finite uncertainty around this mean.

To determine an appropriate value for the innovation variance parameter σ_η^2, we considered

three candidate values: $\sigma_\eta^2 \in \{500, 800, 1200\}$. Posterior draws were obtained using the MCMC algorithm described in Section 5, consisting of 2000 iterations, with the first 1000 iterations discarded as burn-in. Convergence was assessed through visual inspection of the trace plots of the complete data log-likelihood, the static parameters, and a randomly selected subset of latent states.

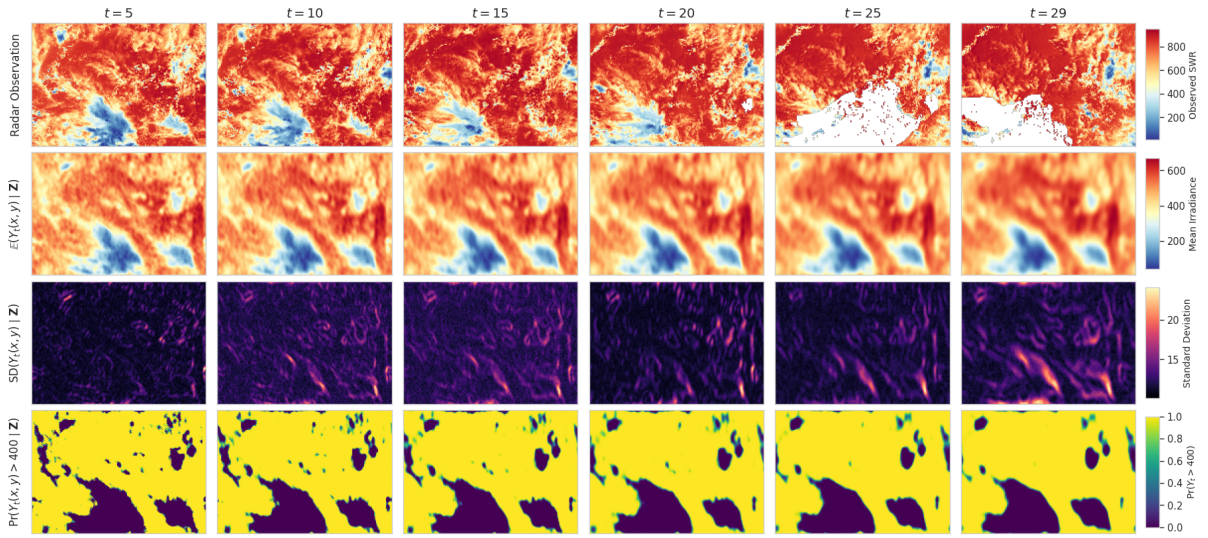

Figure 2: Spatiotemporal summary of the posterior latent irradiance field at six selected time points ($t = 5, 10, 15, 20, 25, 29$). Rows show: (top) radar observations, (second) posterior mean $\mathbb{E}(Y_t(x, y) \mid \mathbf{Z})$, (third) posterior standard deviation $\mathrm{SD}(Y_t(x, y) \mid \mathbf{Z})$, and (bottom) exceedance probability $\Pr(Y_t(x, y) > 400 \mid \mathbf{Z})$.

Model fit was evaluated using DIC and WAIC. Both criteria strongly favored the model with $\sigma_\eta^2 = 500$, which yielded the lowest values for DIC and WAIC among the candidates (DIC $= 1.87 \times 10^9$; WAIC $= 1.87 \times 10^9$). These results indicate that a lower innovation variance provides a better balance between model fit and complexity in capturing the underlying spatiotemporal dynamics of the irradiance field. The posterior mean and standard deviation for the autoregressive parameter α are $\mathbb{E}(\alpha \mid \mathbf{Z}) = 0.84$ and $\mathrm{SD}(\alpha \mid \mathbf{Z}) = 0.05$, with a 95% highest density interval (HDI) of $[0.74,\ 0.94]$, indicating strong temporal persistence in the evolution of the advection parameters. For the diffusion coefficient β, we obtain a posterior mean of $\mathbb{E}(\beta \mid \mathbf{Z}) = 0.23$ and standard deviation $\mathrm{SD}(\beta \mid \mathbf{Z}) = 0.0022$, with a narrow 95% HDI of $[0.23,\ 0.23]$, suggesting high certainty in the estimation of the spatial smoothing effect induced by the diffusion mechanism.

The time-varying posterior estimates of the advection parameters $\boldsymbol{\nu}_t = (\nu_{x,t}, \nu_{y,t})^\top$ are shown in Figure 1, where shaded regions represent the 95% pointwise credible intervals and the solid lines indicate the posterior means. The ν_x component (left panel) fluctuates around zero for the first half of the window, exhibiting both positive and negative excursions, before gradually increasing and attaining more positive values toward the latter time steps. This behavior suggests minor east–west transport effects in the early hours, transitioning to stronger eastward tendencies closer to noon. In contrast, ν_y (right panel) displays a more stable pattern, with predominantly negative values throughout the window, indicating persistent southward advection. However, a gradual upward trend is observed after $t = 15$, with ν_y approaching closer to zero by the final time steps. The dynamic evolution captured here reflects the spatiotemporal variability of the advective effects governing the latent irradiance field over Mindanao during

http://dx.doi.org/10.62071/tmjm.v7i2.776

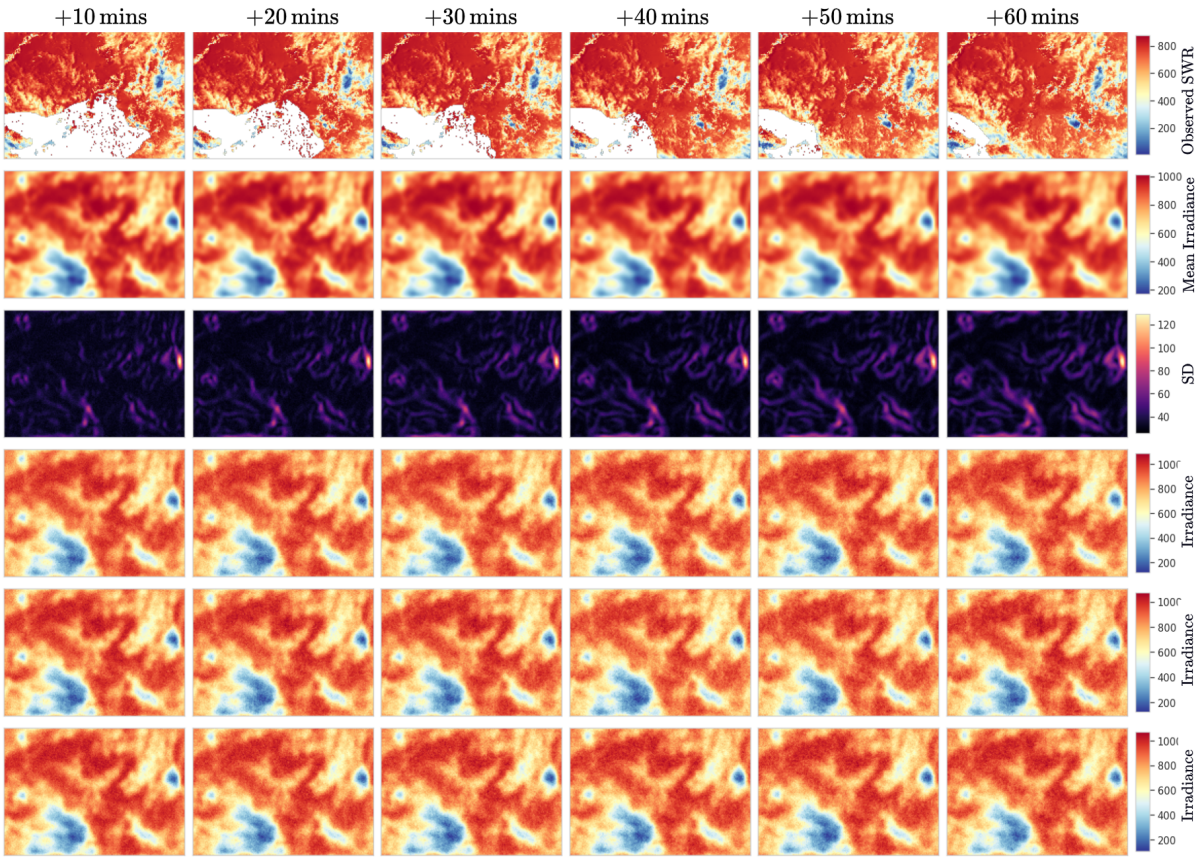

Figure 3: Posterior predictive distribution of shortwave radiation (SWR) forecasts over Mindanao. The top row shows actual radar-based SWR observations at forecast horizons of +10 to +60 minutes. The second and third rows display the posterior predictive mean and standard deviation, respectively. The last three rows present three representative predictive draws from the posterior predictive distribution, highlighting spatial variability and forecast uncertainty.

the modeled period.

Figure 2 provides a spatiotemporal summary of the posterior latent irradiance field across selected time points, offering insight into both the central tendency and uncertainty of the inferred surface-level shortwave radiation. The posterior mean $\mathbb{E}(Y_t(x, y) \mid \mathbf{Z})$ (second row) closely mirrors the spatial structure of the observed data (top row), capturing key irradiance gradients and cloud-shadowed regions while exhibiting smoother patterns due to model regularization. The posterior standard deviations (third row) remain relatively low (mostly below 20 W/m^2), indicating high certainty in the state estimates across most locations and time points, although uncertainty increases slightly in areas with substantial missingness or abrupt spatial transitions. The bottom row displays the posterior probability $\Pr(Y_t(x, y) > 400 \mid \mathbf{Z})$, revealing well-defined high-irradiance regions—shown in yellow—that persist throughout the day and align well with expected clear-sky conditions. These summaries illustrate the model's ability to yield spatially coherent irradiance fields while quantifying uncertainty and enabling probabilistic interpretation of threshold exceedance.

Figure 3 presents spatial posterior predictive forecasts for radar-based shortwave radiation (SWR) over Mindanao at six forecast horizons: +10, +20, +30, +40, +50, and +60 minutes. The top row displays the actual radar observations, serving as a reference. The second and

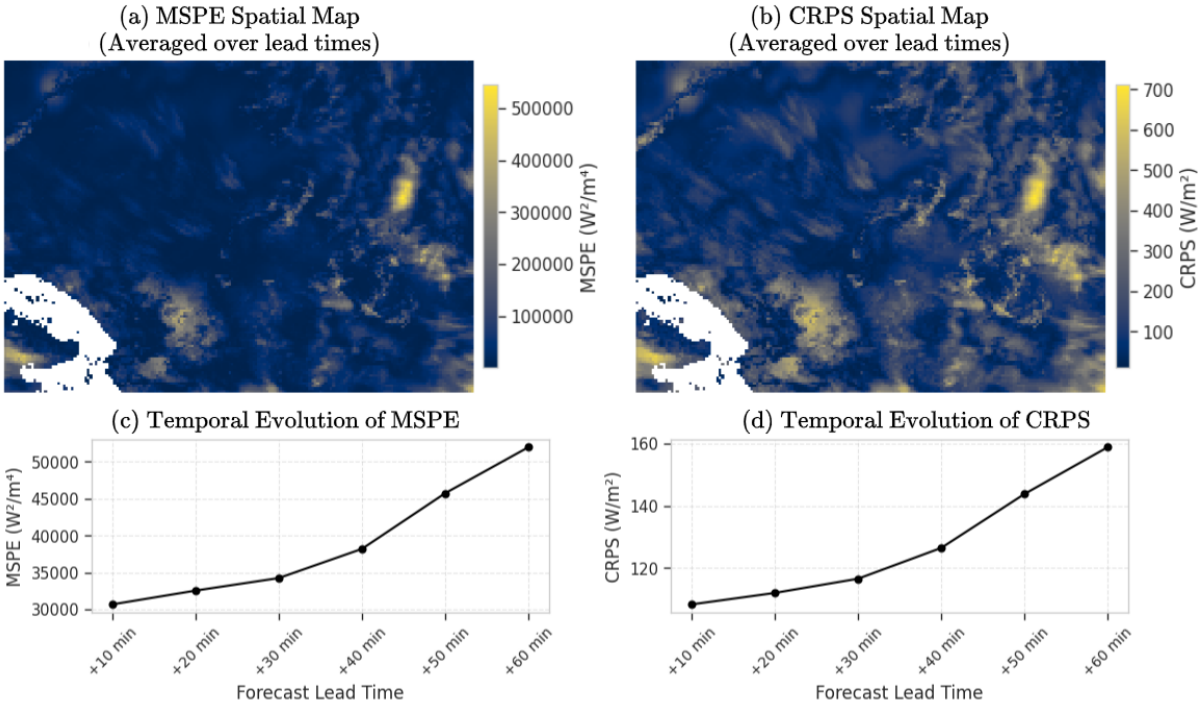

Figure 4: Spatial and temporal evaluation of forecast performance. Panels (a)–(b) show the spatial distribution of MSPE and CRPS averaged over lead times; panels (c)–(d) display the evolution of domain-averaged MSPE and CRPS across forecast horizons from +10 to +60 minutes.

third rows summarize the forecast distribution via the posterior predictive mean and standard deviation, respectively, while the last three rows illustrate three representative predictive draws. These forecasts are generated by forward simulation from the fitted hierarchical model (8). Aggregated over the one-hour hold-out period, the domain-averaged mean squared prediction error (MSPE) increases from about 3.1×10^4 $(\mathrm{W\,m^{-2}})^2$ at +10 minutes to roughly 4.0×10^4 $(\mathrm{W\,m^{-2}})^2$ at +30 minutes, with a sharper rise thereafter; the corresponding Continuous Ranked Probability Score (CRPS) grows from approximately 100 to 160 $\mathrm{W\,m^{-2}}$ over the same range and continues to worsen beyond +30 minutes. Figure 4 summarises these patterns spatially and by lead time. Together, these diagnostics confirm that the model delivers reliable probabilistic nowcasts up to about +30 minutes, with forecast uncertainty and loss of spatial coherence becoming pronounced at longer lead times.

7 Conclusions

In this study, we developed a Bayesian dynamic spatio-temporal model for short-term probabilistic forecasting of shortwave radiation (SWR) fields over Mindanao. Our approach integrates an advection–diffusion process, informed by the physical dynamics of cloud movement and turbulent mixing, with a hierarchical model that assimilates satellite-derived observations through a Gaussian state-space formulation. A computationally efficient Gibbs sampler—incorporating a fixed-lag ensemble Kalman smoother—was used to perform posterior inference on high-dimensional latent states and time-varying advection parameters. Posterior summaries revealed strong temporal persistence in advection dynamics and spatial coherence in the recovered latent

http://dx.doi.org/10.62071/tmjm.v7i2.776

irradiance fields, with reliable nowcasts up to 30 minutes.

For future work, several methodological extensions can be considered. First, alternative MCMC sampling algorithms may be explored to improve posterior sample mixing and convergence, especially for static parameters where autocorrelation remains high. Second, different classes of stochastic partial differential equations (SPDEs) may be investigated to better reflect physical processes beyond advection–diffusion, including those with non-stationary or anisotropic components. Lastly, the model could be expanded to incorporate additional covariates such as cloud optical depth, relative humidity, or topography, either as predictors in the observation equation (through additive regression terms) or as forcings in the process layer that modulate the local drift and diffusion coefficients. This would retain the existing hierarchical structure while enabling more flexible and accurate forecasts under varying meteorological conditions. These extensions will further enhance the applicability of the proposed framework for operational solar energy forecasting and broader environmental modeling tasks.

A Notation Summary

This section summarises the main symbols and notation used throughout the paper for quick reference.

Table 2: Summary of main notation used in the model

Symbol	Description
$Y_t(s)$	Latent irradiance at location s and time t
$Z_t(s)$	Observed SWR at location s and time t
$\mathbf{Y}_t$	Vector of latent irradiance over all locations at time t
$\mathbf{Z}_t$	Vector of observations over all locations at time t
$\boldsymbol{\nu}_t = (\nu_{x,t}, \nu_{y,t})^\top$	Advection parameters (scaled wind components) at time t
β	Scaled diffusion coefficient in the advection–diffusion stencil
$\mathbf{M}(\boldsymbol{\nu}_t)$	Sparse transition matrix implied by the five-point stencil and advection parameters
$\boldsymbol{\eta}_t$	Process innovations in the latent field, variance σ_η^2
$\boldsymbol{\xi}_t$	Innovations in the advection process, variance σ_ν^2
$\varepsilon_t(s)$, $\boldsymbol{\varepsilon}_t$	Measurement error at location s (or vector form), variance σ_ε^2
α	Autoregressive coefficient in the advection dynamics
$N = n_x n_y$	Total number of spatial grid cells
N_{ens}	Ensemble size used in the EnKS

B Finite–Difference Discretization of the Advection–Diffusion SPDE

To derive the discrete evolution equation from the continuous advection–diffusion equation (2), we approximate equation (1) using the forward-time centered-space (FTCS) method. Consider a regular lattice with spatial step size $\Delta_d \equiv \Delta_x = \Delta_y$ and a unit temporal step $\Delta_t = 1$.

First, we express the advection term in two dimensions as

$$-v \cdot \nabla Y = -v_x \frac{\partial Y}{\partial x} - v_y \frac{\partial Y}{\partial y}. \tag{10}$$

Next, we approximate the required partial derivatives using finite differences. For the temporal derivative, we use a forward difference:

$$\frac{\partial Y}{\partial t} \approx \frac{Y_{t+1}(x,y) - Y_t(x,y)}{\Delta_t}. \tag{11}$$

For the second-order spatial (diffusion) terms, we use the centered difference approximation:

$$\frac{\partial^2 Y}{\partial x^2} \approx \frac{Y_t(x+\Delta_d, y) - 2Y_t(x,y) + Y_t(x-\Delta_d, y)}{(\Delta_d)^2}, \tag{12}$$

$$\frac{\partial^2 Y}{\partial y^2} \approx \frac{Y_t(x, y+\Delta_d) - 2Y_t(x,y) + Y_t(x, y-\Delta_d)}{(\Delta_d)^2}. \tag{13}$$

Similarly, the first-order spatial (advection) derivatives are approximated by

$$\frac{\partial Y}{\partial x} \approx \frac{Y_t(x+\Delta_d, y) - Y_t(x-\Delta_d, y)}{2\Delta_d}, \tag{14}$$

$$\frac{\partial Y}{\partial y} \approx \frac{Y_t(x, y+\Delta_d) - Y_t(x, y-\Delta_d)}{2\Delta_d}. \tag{15}$$

Substituting these finite-difference approximations into the continuous advection–diffusion equation (1), we obtain

$$\frac{Y_{t+1}(x,y) - Y_t(x,y)}{\Delta_t} = b\left(\frac{\partial^2 Y}{\partial x^2} + \frac{\partial^2 Y}{\partial y^2}\right) - v_x \frac{\partial Y}{\partial x} - v_y \frac{\partial Y}{\partial y} + \epsilon_t(x,y). \tag{16}$$

Assuming $\Delta_t = 1$ for simplicity, substituting equations (12)–(15) into (16) yields

$$\begin{aligned} Y_{t+1}(x,y) = Y_t(x,y) + b\Bigg(&\frac{Y_t(x+\Delta_d, y) - 2Y_t(x,y) + Y_t(x-\Delta_d, y)}{(\Delta_d)^2} \\ &+ \frac{Y_t(x, y+\Delta_d) - 2Y_t(x,y) + Y_t(x, y-\Delta_d)}{(\Delta_d)^2}\Bigg) \\ &- v_x \frac{Y_t(x+\Delta_d, y) - Y_t(x-\Delta_d, y)}{2\Delta_d} - v_y \frac{Y_t(x, y+\Delta_d) - Y_t(x, y-\Delta_d)}{2\Delta_d} + \epsilon_t(x,y). \end{aligned} \tag{17}$$

To simplify the notation, we define

$$\beta = \frac{b}{(\Delta_d)^2}, \quad \nu_x = \frac{v_x}{2\Delta_d}, \quad \nu_y = \frac{v_y}{2\Delta_d}. \tag{18}$$

Substituting these definitions into (17), the evolution equation becomes

$$\begin{aligned} Y_{t+1}(x,y) = Y_t(x,y) + \beta\Big(&Y_t(x+\Delta_d, y) - 2Y_t(x,y) + Y_t(x-\Delta_d, y) \\ &+ Y_t(x, y+\Delta_d) - 2Y_t(x,y) + Y_t(x, y-\Delta_d)\Big) \\ &- \nu_x\Big(Y_t(x+\Delta_d, y) - Y_t(x-\Delta_d, y)\Big) - \nu_y\Big(Y_t(x, y+\Delta_d) - Y_t(x, y-\Delta_d)\Big) + \epsilon_t(x,y). \end{aligned} \tag{19}$$

Grouping like terms, we finally obtain the discrete evolution equation in its compact form

$$\begin{aligned} Y_{t+1}(x,y) = (1-4\beta)\, Y_t(x,y) + (\beta - \nu_x)\, Y_t(x+\Delta_d, y) + (\beta + \nu_x)\, Y_t(x-\Delta_d, y) \\ + (\beta - \nu_y)\, Y_t(x, y+\Delta_d) + (\beta + \nu_y)\, Y_t(x, y-\Delta_d) + \epsilon_t(x,y). \end{aligned} \tag{20}$$

http://dx.doi.org/10.62071/tmjm.v7i2.776

C Full Conditional Distributions (FCD)

This appendix provides a complete derivation of the full conditional distributions required in the posterior sampling procedure.

C.1 Advection Parameter FCD

In this subsection, we derive the full conditional distribution for the advection parameter vector $\boldsymbol{\nu}_t = (\nu_{x,t}, \nu_{y,t})^\top$, for each time step $t = 0, \ldots, T$. We initially provide a detailed derivation for the general case $t = 1, \ldots, T-1$, and subsequently outline the special cases for the initial time ($t = 0$) and final time ($t = T$).

C.1.1 General Case

For the general case $t = 1, \ldots, T-1$, the full conditional distribution for $\boldsymbol{\nu}_t$ is proportional to the product of the following three densities:

- the process model for $\boldsymbol{\nu}_t$ given by $p(\boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1}, \alpha, \sigma_\nu^2)$;
- the process model for $\boldsymbol{\nu}_{t+1}$ given by $p(\boldsymbol{\nu}_{t+1} \mid \boldsymbol{\nu}_t, \alpha, \sigma_\nu^2)$; and
- the latent field process model for $\mathbf{Y}_{t+1}$ given by $p(\mathbf{Y}_{t+1} \mid \mathbf{Y}_t, \beta, \boldsymbol{\nu}_t, \sigma_\eta^2)$.

Explicitly, we have

$$\begin{aligned} p(\boldsymbol{\nu}_t \mid \cdot) &\propto p(\boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1}, \alpha, \sigma_\nu^2)\, p(\boldsymbol{\nu}_{t+1} \mid \boldsymbol{\nu}_t, \alpha, \sigma_\nu^2)\, p(\mathbf{Y}_{t+1} \mid \mathbf{Y}_t, \beta, \boldsymbol{\nu}_t, \sigma_\eta^2) \\ &\propto \exp\left(-\frac{1}{2\sigma_\nu^2}\|\boldsymbol{\nu}_t - \alpha\boldsymbol{\nu}_{t-1}\|^2\right) \exp\left(-\frac{1}{2\sigma_\nu^2}\|\boldsymbol{\nu}_{t+1} - \alpha\boldsymbol{\nu}_t\|^2\right) \\ &\quad \times \exp\left(-\frac{1}{2\sigma_\eta^2}\sum_{i=1}^{N}(Y_{t+1,i} - \mu_{t,i})^2\right), \end{aligned} \tag{21}$$

where the conditional mean $\mu_{t,i}$ is given by the discretized advection–diffusion model:

$$\mu_{t,i} = (1 - 4\beta)Y_{t,i} + (\beta - \nu_{x,t})Y_{t,j_1} + (\beta + \nu_{x,t})Y_{t,j_2} + (\beta - \nu_{y,t})Y_{t,j_3} + (\beta + \nu_{y,t})Y_{t,j_4}.$$

Here, j_1, j_2, j_3, j_4 denote indices of the right, left, up, and down neighbors of location i.

Expanding the quadratic terms:

$$\begin{aligned} p(\boldsymbol{\nu}_t|\cdot) \propto \exp\Bigg(&-\frac{1}{2\sigma_\nu^2}(\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\boldsymbol{\nu}_{t-1}^\top\boldsymbol{\nu}_t) - \frac{1}{2\sigma_\nu^2}(\alpha^2\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\boldsymbol{\nu}_{t+1}^\top\boldsymbol{\nu}_t) \\ &- \frac{1}{2\sigma_\eta^2}\sum_{i=1}^{N}(Y_{t+1,i} - \mu_{t,i})^2\Bigg) \end{aligned} \tag{22}$$

Now, we focus on the last term of (22). We can expand it as

$$\begin{aligned} \sum_{i=1}^{N}(Y_{t+1,i} - \mu_{t,i})^2 &= \sum_{i=1}^{N}(Y_{t+1,i} - (1-4\beta)Y_{t,i} - (\beta - \nu_{x,t})Y_{t,j_1} - (\beta + \nu_{x,t})Y_{t,j_2} \\ &\quad - (\beta - \nu_{y,t})Y_{t,j_3} - (\beta + \nu_{y,t})Y_{t,j_4})^2 \\ &= \sum_{i=1}^{N}(A_i - \nu_{x,t}B_i - \nu_{y,t}C_i)^2 \end{aligned} \tag{23}$$

where

$$A_i = Y_{t+1,i} - (1-4\beta)Y_{t,i} - \beta(Y_{t,j_1} + Y_{t,j_2} + Y_{t,j_3} + Y_{t,j_4})$$
$$B_i = Y_{t,j_2} - Y_{t,j_1}$$
$$C_i = Y_{t,j_4} - Y_{t,j_3}$$

Expanding this quadratic term (23), we have

$$\begin{aligned}
\sum_{i=1}^{N}(A_i - \nu_{x,t}B_i - \nu_{y,t}C_i)^2 &= \sum_{i=1}^{N}(A_i^2 + \nu_{x,t}^2B_i^2 + \nu_{y,t}^2C_i^2 - 2A_i\nu_{x,t}B_i - 2A_i\nu_{y,t}C_i + 2\nu_{x,t}\nu_{y,t}B_iC_i) \\
&= \sum_{i=1}^{N}A_i^2 + \nu_{x,t}^2\sum_{i=1}^{N}B_i^2 + \nu_{y,t}^2\sum_{i=1}^{N}C_i^2 - 2\nu_{x,t}\sum_{i=1}^{N}A_iB_i - 2\nu_{y,t}\sum_{i=1}^{N}A_iC_i \\
&\quad + 2\nu_{x,t}\nu_{y,t}\sum_{i=1}^{N}B_iC_i \\
&= \mathbf{A}^\top\mathbf{A} + \nu_{x,t}^2\mathbf{B}^\top\mathbf{B} + \nu_{y,t}^2\mathbf{C}^\top\mathbf{C} - 2\nu_{x,t}\mathbf{A}^\top\mathbf{B} - 2\nu_{y,t}\mathbf{A}^\top\mathbf{C} + 2\nu_{x,t}\nu_{y,t}\mathbf{B}^\top\mathbf{C} \\
&= \mathbf{A}^\top\mathbf{A} + [\nu_{x,t}, \nu_{y,t}]\begin{bmatrix}\mathbf{B}^\top\mathbf{B} & \mathbf{B}^\top\mathbf{C} \\ \mathbf{C}^\top\mathbf{B} & \mathbf{C}^\top\mathbf{C}\end{bmatrix}\begin{bmatrix}\nu_{x,t} \\ \nu_{y,t}\end{bmatrix} - 2[\nu_{x,t}, \nu_{y,t}]\begin{bmatrix}\mathbf{B}^\top\mathbf{A} \\ \mathbf{C}^\top\mathbf{A}\end{bmatrix} \\
&= \mathbf{A}^\top\mathbf{A} + \boldsymbol{\nu}_t^\top\mathbf{R}^\top\mathbf{R}\boldsymbol{\nu}_t - 2\boldsymbol{\nu}_t^\top\mathbf{R}^\top\mathbf{A} \\
&= \boldsymbol{\nu}_t^\top\mathbf{R}^\top\mathbf{R}\boldsymbol{\nu}_t - 2\mathbf{A}^\top\mathbf{R}\boldsymbol{\nu}_t + \mathbf{A}^\top\mathbf{A} \qquad (24)
\end{aligned}$$

where $\mathbf{R} = [\mathbf{B}, \mathbf{C}]$ is an $N \times 2$ matrix with columns $\mathbf{B}$ and $\mathbf{C}$, and $\mathbf{A}$ is the vector of A_i values. Incorporating (24) to (22), we have

$$\begin{aligned}
p(\boldsymbol{\nu}_t|\cdot) &\propto \exp\Bigg(-\frac{1}{2\sigma_\nu^2}(\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\boldsymbol{\nu}_{t-1}^\top\boldsymbol{\nu}_t) - \frac{1}{2\sigma_\nu^2}(\alpha^2\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\boldsymbol{\nu}_{t+1}^\top\boldsymbol{\nu}_t) \\
&\quad - \frac{1}{2\sigma_\eta^2}(\boldsymbol{\nu}_t^\top\mathbf{R}^\top\mathbf{R}\boldsymbol{\nu}_t - 2\mathbf{A}^\top\mathbf{R}\boldsymbol{\nu}_t + \mathbf{A}^\top\mathbf{A})\Bigg) \\
&\propto \exp\Bigg(-\frac{1}{2\sigma_\nu^2}(\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\boldsymbol{\nu}_{t-1}^\top\boldsymbol{\nu}_t) - \frac{1}{2\sigma_\nu^2}(\alpha^2\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\boldsymbol{\nu}_{t+1}^\top\boldsymbol{\nu}_t) \\
&\quad - \frac{1}{2\sigma_\eta^2}(\boldsymbol{\nu}_t^\top\mathbf{R}^\top\mathbf{R}\boldsymbol{\nu}_t - 2\mathbf{A}^\top\mathbf{R}\boldsymbol{\nu}_t)\Bigg) \\
&= \exp\Bigg(-\frac{1}{2\sigma_\nu^2}\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t + \frac{1}{\sigma_\nu^2}\alpha\boldsymbol{\nu}_{t-1}^\top\boldsymbol{\nu}_t - \frac{1}{2\sigma_\nu^2}\alpha^2\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t + \frac{1}{\sigma_\nu^2}\alpha\boldsymbol{\nu}_{t+1}^\top\boldsymbol{\nu}_t \\
&\quad - \frac{1}{2\sigma_\eta^2}\boldsymbol{\nu}_t^\top\mathbf{R}^\top\mathbf{R}\boldsymbol{\nu}_t - \frac{1}{\sigma_\eta^2}\mathbf{A}^\top\mathbf{R}\boldsymbol{\nu}_t)\Bigg) \\
&= \exp\Bigg(-\frac{1}{2}\boldsymbol{\nu}_t^\top\left(\frac{\boldsymbol{\nu}_t}{\sigma_\nu^2} + \frac{\alpha^2\boldsymbol{\nu}_t}{\sigma_\nu^2} + \frac{\mathbf{R}^\top\mathbf{R}\boldsymbol{\nu}_t}{\sigma_\eta^2}\right) + \left(\frac{\alpha\boldsymbol{\nu}_{t-1}^\top}{\sigma_\nu^2} + \frac{\alpha\boldsymbol{\nu}_{t+1}^\top}{\sigma_\nu^2} + \frac{\mathbf{A}^\top\mathbf{R}}{\sigma_\eta^2}\right)\boldsymbol{\nu}_t\Bigg) \\
&= \exp\Bigg(-\frac{1}{2}\boldsymbol{\nu}_t^\top\left(\frac{1+\alpha^2}{\sigma_\nu^2}\mathbf{I} + \frac{\mathbf{R}^\top\mathbf{R}}{\sigma_\eta^2}\right)\boldsymbol{\nu}_t + \boldsymbol{\nu}_t^\top\left(\frac{\alpha\boldsymbol{\nu}_{t-1}}{\sigma_\nu^2} + \frac{\alpha\boldsymbol{\nu}_{t+1}}{\sigma_\nu^2} + \frac{\mathbf{R}^\top\mathbf{A}}{\sigma_\eta^2}\right)\Bigg) \\
&= \exp\Bigg(-\frac{1}{2}\boldsymbol{\nu}_t^\top\left(\frac{1+\alpha^2}{\sigma_\nu^2}\mathbf{I} + \frac{\mathbf{R}^\top\mathbf{R}}{\sigma_\eta^2}\right)\boldsymbol{\nu}_t + \boldsymbol{\nu}_t^\top\left(\frac{\alpha}{\sigma_\nu^2}(\boldsymbol{\nu}_{t-1} + \boldsymbol{\nu}_{t+1}) + \frac{\mathbf{R}^\top\mathbf{A}}{\sigma_\eta^2}\right)\Bigg)
\end{aligned}$$

http://dx.doi.org/10.62071/tmjm.v7i2.776

Collecting all terms involving $\boldsymbol{\nu}_t$

$$p(\boldsymbol{\nu}_t|\cdot) \propto \exp\Bigg(-\frac{1}{2}\boldsymbol{\nu}_t^\top \left(\frac{1+\alpha^2}{\sigma_\nu^2}\mathbf{I} + \frac{1}{\sigma_\eta^2}\mathbf{R}^\top\mathbf{R}\right)\boldsymbol{\nu}_t + \boldsymbol{\nu}_t^\top\left(\frac{\alpha}{\sigma_\nu^2}(\boldsymbol{\nu}_{t-1}+\boldsymbol{\nu}_{t+1}) + \frac{1}{\sigma_\eta^2}\mathbf{R}^\top\mathbf{A}\right)\Bigg)$$

This is a quadratic form in $\boldsymbol{\nu}_t$, indicating a Gaussian distribution. We define

$$\mathbf{Q}_{\nu,t} = \frac{1+\alpha^2}{\sigma_\nu^2}\mathbf{I} + \frac{1}{\sigma_\eta^2}\mathbf{R}^\top\mathbf{R} \tag{25}$$

$$\mathbf{b}_{\nu,t} = \frac{\alpha}{\sigma_\nu^2}(\boldsymbol{\nu}_{t-1}+\boldsymbol{\nu}_{t+1}) + \frac{1}{\sigma_\eta^2}\mathbf{R}^\top\mathbf{A} \tag{26}$$

Therefore, the full conditional distribution is:

$$p(\boldsymbol{\nu}_t|\cdot) \sim \mathcal{N}(\mathbf{Q}_{\nu,t}^{-1}\mathbf{b}_{\nu,t}, \mathbf{Q}_{\nu,t}^{-1})$$

C.1.2 Initial Condition

At $t = 0$ only the forward transition is present and the prior is given by

$$\boldsymbol{\nu}_0 \sim \mathcal{N}_2\Big(\mathbf{m}_{\nu 0}, \mathbf{C}_{\nu 0}\Big).$$

Thus, from the forward transition we have

$$p\Big(\boldsymbol{\nu}_1 \mid \boldsymbol{\nu}_0\Big) \propto \exp\left\{-\frac{1}{2\sigma_\nu^2}\Big\|\boldsymbol{\nu}_1 - \alpha\,\boldsymbol{\nu}_0\Big\|^2\right\}.$$

Define $\mathbf{A}_0$ and $\mathbf{R}_0$ as the counterparts of $\mathbf{A}$ and $\mathbf{R}$ using data at $t = 0$; that is, for each spatial index i

$$A_{0,i} = Y_{1,i} - (1-4\beta)Y_{0,i} - \beta\Big(Y_{0,j_1} + Y_{0,j_2} + Y_{0,j_3} + Y_{0,j_4}\Big),$$

and $\mathbf{R}_0$ is the $N \times 2$ matrix with columns defined by

$$B_{0,i} = Y_{0,j_2} - Y_{0,j_1}, \quad \text{and} \quad C_{0,i} = Y_{0,j_4} - Y_{0,j_3}.$$

Then, by similar derivations as for the general case (see Section C.1.1), the kernel of the full conditional distribution for $\boldsymbol{\nu}_0$ becomes

$$\boldsymbol{\nu}_0 \mid \cdot \propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\nu}_0^\top\Big(\frac{1}{\sigma_\eta^2}\mathbf{R}_0^\top\mathbf{R}_0 + \frac{\alpha^2}{\sigma_\nu^2}\mathbf{I}_2 + \mathbf{C}_{\nu 0}^{-1}\Big)\boldsymbol{\nu}_0 - 2\,\boldsymbol{\nu}_0^\top\Big(\frac{1}{\sigma_\eta^2}\mathbf{R}_0^\top\mathbf{A}_0 + \frac{\alpha}{\sigma_\nu^2}\boldsymbol{\nu}_1 + \mathbf{C}_{\nu 0}^{-1}\mathbf{m}_{\nu 0}\Big)\right]\right\}.$$

Thus, the full conditional distribution for $\boldsymbol{\nu}_0$ is

$$\boldsymbol{\nu}_0 \mid \cdot \sim \mathcal{N}_2\Big(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0\Big),$$

with

$$\boldsymbol{\Sigma}_0 = \Big(\frac{1}{\sigma_\eta^2}\mathbf{R}_0^\top\mathbf{R}_0 + \frac{\alpha^2}{\sigma_\nu^2}\mathbf{I}_2 + \mathbf{C}_{\nu 0}^{-1}\Big)^{-1}, \quad \boldsymbol{\mu}_0 = \boldsymbol{\Sigma}_0\Big(\frac{1}{\sigma_\eta^2}\mathbf{R}_0^\top\mathbf{A}_0 + \frac{\alpha}{\sigma_\nu^2}\boldsymbol{\nu}_1 + \mathbf{C}_{\nu 0}^{-1}\mathbf{m}_{\nu 0}\Big).$$

C.1.3 Terminal Condition

At $t = T$ only the backward evolution is present,

$$p(\boldsymbol{\nu}_T \mid \cdot) \propto \exp\left\{ -\frac{1}{2\sigma_\nu^2} \|\boldsymbol{\nu}_T - a\,\boldsymbol{\nu}_{T-1}\|^2 \right\},$$

which directly yields

$$\boldsymbol{\nu}_T \mid \cdot \sim \mathcal{N}_2\Big(a\,\boldsymbol{\nu}_{T-1},\, \sigma_\nu^2\,\mathbf{I}_2\Big).$$

C.2 Diffusion Parameter FCD

The full conditional distribution for the diffusion parameter β is proportional to the product of the likelihood from the latent field process model and the prior for β. That is,

$$p(\beta \mid \cdot) \propto \left[\prod_{t=1}^{T} p(\mathbf{Y}_t \mid \mathbf{Y}_{t-1}, \beta, \boldsymbol{\nu}_{t-1}, \sigma_\eta^2) \right] \times p(\beta). \tag{27}$$

Expanding the right-hand side, we obtain

$$p(\beta \mid \cdot) \propto \exp\left\{ -\frac{1}{2\sigma_\eta^2} \sum_{t=1}^{T} \sum_{i=1}^{N} \Big(Y_t(i) - \mu_t(i) \Big)^2 \right\} \times \exp\left\{ -\frac{1}{2v_\beta} \Big(\beta - m_\beta \Big)^2 \right\}. \tag{28}$$

Here, the conditional mean $\mu_t(i)$ for the latent state at location i is given by the discretized evolution equation:

$$\begin{aligned} \mu_t(i) = (1 - 4\beta) Y_{t-1}(i) + (\beta - \nu_{x,t-1}) Y_{t-1}(j_1) + (\beta + \nu_{x,t-1}) Y_{t-1}(j_2) \\ + (\beta - \nu_{y,t-1}) Y_{t-1}(j_3) + (\beta + \nu_{y,t-1}) Y_{t-1}(j_4). \end{aligned} \tag{29}$$

To simplify the exposition, we rewrite the residual $(Y_t(i) - \mu_t(i))$ in the compact form:

$$Y_t(i) - \mu_t(i) = A_t(i) - \beta\, B_t(i), \tag{30}$$

where

$$\begin{aligned} A_t(i) &= Y_t(i) - Y_{t-1}(i) + \nu_{x,t-1} \Big[Y_{t-1}(j_2) - Y_{t-1}(j_1) \Big] \\ &\quad + \nu_{y,t-1} \Big[Y_{t-1}(j_4) - Y_{t-1}(j_3) \Big], \end{aligned} \tag{31}$$

$$B_t(i) = -4 Y_{t-1}(i) + Y_{t-1}(j_1) + Y_{t-1}(j_2) + Y_{t-1}(j_3) + Y_{t-1}(j_4). \tag{32}$$

Substituting (30) into (28) and expanding the square yields

$$p(\beta \mid \cdot) \propto \exp\left\{ -\frac{1}{2\sigma_\eta^2} \sum_{t=1}^{T} \sum_{i=1}^{N} \Big[A_t(i)^2 - 2 A_t(i) B_t(i) \beta + B_t(i)^2 \beta^2 \Big] - \frac{1}{2v_\beta} \Big(\beta^2 - 2 m_\beta \beta + m_\beta^2 \Big) \right\}. \tag{33}$$

http://dx.doi.org/10.62071/tmjm.v7i2.776

Collecting the terms that are quadratic and linear in β, we define the precision and the associated weighted sum as follows:

$$Q_\beta = \frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}\sum_{i=1}^{N} B_t(i)^2 + \frac{1}{v_\beta},$$

$$b_\beta = \frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}\sum_{i=1}^{N} A_t(i)B_t(i) + \frac{m_\beta}{v_\beta}.$$

Thus, the kernel in (33) can be written as

$$p(\beta \mid \cdot) \propto \exp\left\{-\frac{1}{2}\left[Q_\beta\,\beta^2 - 2b_\beta\,\beta\right]\right\}. \tag{34}$$

Recognizing the kernel of a univariate Gaussian distribution, we conclude that the full conditional distribution for β is

$$\beta \mid \cdot \sim \mathcal{N}\left(\frac{b_\beta}{Q_\beta}, \frac{1}{Q_\beta}\right).$$

C.3 Autoregressive Parameter FCD

The full conditional distribution for α is proportional to the product of the transition densities for the advection parameters and the prior for α. That is,

$$p(\alpha \mid \cdot) \propto \left[\prod_{t=1}^{T} p\big(\boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1}, \alpha, \sigma_\nu^2\big)\right] \times p(\alpha), \tag{35}$$

where for each $t = 1, \ldots, T$ the transition density is given by

$$p\big(\boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1}, \alpha, \sigma_\nu^2\big) \propto \exp\left\{-\frac{1}{2\sigma_\nu^2}\left\|\boldsymbol{\nu}_t - \alpha\,\boldsymbol{\nu}_{t-1}\right\|^2\right\},$$

and the prior for α is assumed to be Gaussian (truncated on the interval $(0, 1)$),

$$p(\alpha) \propto \exp\left\{-\frac{1}{2v_\alpha}\Big(\alpha - m_\alpha\Big)^2\right\} I_{(0,1)}(\alpha).$$

Expanding the likelihood term in (35), we have

$$p(\alpha \mid \cdot) \propto \exp\left\{-\frac{1}{2\sigma_\nu^2}\sum_{t=1}^{T}\Big(\boldsymbol{\nu}_t - \alpha\,\boldsymbol{\nu}_{t-1}\Big)^\top\Big(\boldsymbol{\nu}_t - \alpha\,\boldsymbol{\nu}_{t-1}\Big)\right\}$$

$$\times \exp\left\{-\frac{1}{2v_\alpha}\Big(\alpha - m_\alpha\Big)^2\right\} I_{(0,1)}(\alpha). \tag{36}$$

Expanding the quadratic form in the exponent, we obtain

$$\sum_{t=1}^{T}\Big(\boldsymbol{\nu}_t - \alpha\,\boldsymbol{\nu}_{t-1}\Big)^\top\Big(\boldsymbol{\nu}_t - \alpha\,\boldsymbol{\nu}_{t-1}\Big) = \sum_{t=1}^{T}\Big[\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_t - 2\alpha\,\boldsymbol{\nu}_t^\top\boldsymbol{\nu}_{t-1} + \alpha^2\,\boldsymbol{\nu}_{t-1}^\top\boldsymbol{\nu}_{t-1}\Big].$$

Thus, equation (36) becomes

$$p(\alpha \mid \cdot) \propto \exp\Bigg\{ -\frac{1}{2\sigma_\nu^2} \sum_{t=1}^{T} \Big[\boldsymbol{\nu}_t^\top \boldsymbol{\nu}_t - 2\alpha\, \boldsymbol{\nu}_t^\top \boldsymbol{\nu}_{t-1} + \alpha^2\, \boldsymbol{\nu}_{t-1}^\top \boldsymbol{\nu}_{t-1} \Big] - \frac{1}{2v_\alpha} \Big(\alpha^2 - 2\alpha\, m_\alpha + m_\alpha^2 \Big) \Bigg\} I_{(0,1)}(\alpha).$$

Collecting the terms involving α, we can write

$$p(\alpha \mid \cdot) \propto \exp\Bigg\{ -\frac{1}{2} \Bigg[\left(\frac{1}{\sigma_\nu^2} \sum_{t=1}^{T} \boldsymbol{\nu}_{t-1}^\top \boldsymbol{\nu}_{t-1} + \frac{1}{v_\alpha} \right) \alpha^2 - 2 \left(\frac{1}{\sigma_\nu^2} \sum_{t=1}^{T} \boldsymbol{\nu}_t^\top \boldsymbol{\nu}_{t-1} + \frac{m_\alpha}{v_\alpha} \right) \alpha \Bigg] \Bigg\} I_{(0,1)}(\alpha). \tag{37}$$

Defining the precision and weighted sum as

$$Q_\alpha = \frac{1}{\sigma_\nu^2} \sum_{t=1}^{T} \boldsymbol{\nu}_{t-1}^\top \boldsymbol{\nu}_{t-1} + \frac{1}{v_\alpha},$$

$$b_\alpha = \frac{1}{\sigma_\nu^2} \sum_{t=1}^{T} \boldsymbol{\nu}_t^\top \boldsymbol{\nu}_{t-1} + \frac{m_\alpha}{v_\alpha},$$

the kernel in equation (37) becomes

$$p(\alpha \mid \cdot) \propto \exp\Bigg\{ -\frac{1}{2} \Big[Q_\alpha\, \alpha^2 - 2 b_\alpha\, \alpha \Big] \Bigg\} I_{(0,1)}(\alpha). \tag{38}$$

Equation (38) corresponds to the kernel of a normal distribution, except that the support is truncated to the interval $(0, 1)$. Therefore, the full conditional distribution for α is a truncated normal distribution:

$$\alpha \mid \cdot \sim \mathrm{TN}\left(\frac{b_\alpha}{Q_\alpha}, \frac{1}{Q_\alpha}; 0, 1 \right),$$

which completes the derivation.

Acknowledgements

The authors are grateful to the Japan Aerospace Exploration Agency (JAXA) and the P-Tree System for providing the Himawari-9 Short Wave Radiation (SWR) product used in this research. We also acknowledge the Department of Science and Technology (DOST), Republic of the Philippines, for their financial support through the Graduate Scholarship Program. Lastly, we thank the anonymous referees for their valuable comments and suggestions, which greatly improved the clarity and rigor of this manuscript.

References

[1] C. Arbizu-Barrena, J. A. Ruiz-Arias, F. J. Rodriguez-Benitez, D. Pozo-Vazquez, and J. Tovar-Pescador, *Short-Term Solar Radiation Forecasting by Advecting and Diffusing* MSG *Cloud Index*, Solar Energy **155** (2017), 1092–1103, doi:10.1016/j.solener.2017.07.045.

[2] M. Bai, X. Zhao, Z. Long, J. Liu, and D. Yu, *Short-term probabilistic photovoltaic power forecast based on deep convolutional long short-term memory network and kernel density estimation*, arXiv preprint arXiv:2107.01343 (2021), https://arxiv.org/abs/2107.01343.

[3] L. M. Berliner, *Physical-statistical modeling in geophysics*, J. Geophys. Res. Atmos. **108** (2003), no. D24, doi:10.1029/2002JD002865.

[4] P. E. Brown, G. O. Roberts, K. F. Karesen, and S. Tonellato, *Blur-generated non-separable space-time models*, J. Roy. Stat. Soc. Ser. B **62** (2000), no. 4, 847–860.

[5] C. K. Carter and R. Kohn, *On Gibbs sampling for state space models*, Biometrika **81** (1994), no. 3, 541–553.

[6] L. Clarotto, D. Allard, T. Romary, and N. Desassis, *The SPDE approach for spatio-temporal datasets with advection and diffusion*, Spatial Statistics (2024), 100847.

[7] N. Cressie and C. K. Wikle, *Statistics for spatio-temporal data*, John Wiley & Sons, 2011.

[8] Department of Energy, Republic of the Philippines, *Philippine Energy Plan 2023–2050*, Government report, 2023, https://doe.gov.ph/pep/philippine-energy-plan-2023-2050-0.

[9] G. Evensen and P. J. van Leeuwen, *An ensemble Kalman smoother for nonlinear dynamics*, Mon. Weather Rev. **128** (2000), no. 6, 1852–1867.

[10] C. Feng, W. Zhang, B.-M. Hodge, and Y. Zhang, *Occlusion-Perturbed Deep Learning for Probabilistic Solar Forecasting via Sky Images*, Proc. IEEE Power and Energy Society General Meeting (2022), 1–5, doi:10.1109/PESGM48719.2022.9917222.

[11] R. Frouin and H. Murakami, *Estimating photosynthetically available radiation at the ocean surface from ADEOS-II Global Imager data*, J. Oceanogr. **63** (2007), no. 3, 493–503, doi:10.1007/s10872-007-0044-3.

[12] S. Fruhwirth-Schnatter, *Data augmentation and dynamic linear models*, J. Time Ser. Anal. **15** (1994), no. 2, 183–202.

[13] Japan Meteorological Agency, *Outline of the Operational Numerical Weather Prediction at the Japan Meteorological Agency*, Tech. Rep., March 2019, https://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2019-nwp/index.htm.

[14] S. R. Johnson, S. E. Heaps, K. J. Wilson, and D. J. Wilkinson, *A Bayesian spatio-temporal model for short-term forecasting of precipitation fields*, Environmetrics **34** (2023), e2824, doi:10.1002/env.2824.

[15] M. Katzfuss, J. R. Stroud, and C. K. Wikle, *Ensemble Kalman Methods for High-Dimensional Hierarchical Dynamic Space-Time Models*, J. Amer. Stat. Assoc. **115** (2020), no. 530, 866–885, doi:10.1080/01621459.2019.1592753.

[16] J.-H. Kim, P. A. Jimenez, J. Dudhia, J. Yang, M. Sengupta, and Y. Xie, *Probabilistic Forecast of All-Sky Solar Radiation Using Enhanced WRF-Solar*, Proc. 37th European Photovoltaic Solar Energy Conference and Exhibition (2020).

[17] J.-H. Kim, P. A. Jimenez, M. Sengupta, J. Yang, J. Dudhia, S. Alessandrini, and Y. Xie, *The WRF-Solar Ensemble Prediction System to Provide Solar Irradiance Probabilistic Forecasts*, IEEE Journal of Photovoltaics **12** (2022), no. 1, 141–144, doi:10.1109/JPHOTOV.2021.3117904.

[18] W. Kleiber, S. Sain, L. Madaus, and P. Harr, *Stochastic tropical cyclone precipitation field generation*, Environmetrics **34** (2023), no. 1, e2766.

[19] H. Letu, T. Y. Nakajima, T. Wang, H. Shang, R. Ma, K. Yang, A. J. Baran, J. Riedi, H. Ishimoto, and M. Yoshida, *A new benchmark for surface radiation products over the East Asia–Pacific region retrieved from the Himawari-8/AHI next-generation geostationary satellite*, Bull. Amer. Meteorol. Soc. **103** (2022), no. 3, E873–E888.

[20] H. Letu, R. Ma, T. Y. Nakajima, C. Shi, M. Hashimoto, T. M. Nagao, A. J. Baran, T. Nakajima, J. Xu, and T. Wang, *Surface solar radiation compositions observed from Himawari-8/9 and Fengyun-4 series*, Bull. Amer. Meteorol. Soc. **104** (2023), no. 10, E1772–E1789.

[21] G. Mastrantonio, G. J. Lasinio, A. Pollice, L. Teodonio, and G. Capotorti, *A Dirichlet process model for change-point detection with multivariate bioclimatic data*, Environmetrics **33** (2022), no. 1, e2699.

[22] K. Opoku, S. Lucemo, Q. Z. Sun, and A. Dimitrovski, *A Bayesian Approach to Probabilistic Solar Irradiance Forecasting*, arXiv preprint arXiv:2209.00792 (2022), https://arxiv.org/abs/2209.00792.

[23] G. E. Primiceri, *Time varying structural vector autoregressions and monetary policy*, Rev. Econ. Stud. **72** (2005), no. 3, 821–852.

[24] J. Richards and J. L. Wadsworth, *Spatial deformation for nonstationary extremal dependence*, Environmetrics **32** (2021), no. 5, e2671.

[25] C. Sigauke, E. Chandiwana, and A. Bere, *Spatio-Temporal Forecasting of Global Horizontal Irradiance Using Bayesian Inference*, Applied Sciences **13** (2023), no. 1, 201, doi:10.3390/app13010201.

[26] F. Sigrist, H. R. Kunsch, and W. A. Stahel, *Stochastic partial differential equation based modelling of large space–time data sets*, J. Roy. Stat. Soc. Ser. B **77** (2015), no. 1, 3–33.

[27] J. R. Stroud, M. L. Stein, B. M. Lesht, D. J. Schwab, and D. Beletsky, *An Ensemble Kalman Filter and Smoother for Satellite Data Assimilation*, J. Amer. Stat. Assoc. **105** (2010), no. 491, 978–990, doi:10.1198/jasa.2010.ap07636.

[28] M. A. Tanner and W. H. Wong, *The calculation of posterior distributions by data augmentation*, J. Amer. Stat. Assoc. **82** (1987), no. 398, 528–540.

[29] Q. Wang, A. Tuohy, M. Ortega-Vazquez, M. Bello, E. Ela, D. Kirk-Davidoff, W. B. Hobbs, D. J. Ault, and R. Philbrick, *Quantifying the Value of Probabilistic Forecasting for Power System Operation Planning*, Applied Energy **343** (2023), 121254, doi:10.1016/j.apenergy.2023.121254.

[30] E. Zelikman, S. Zhou, J. Irvin, C. Raterink, H. Sheng, A. Avati, J. Kelly, R. Rajagopal, A. Y. Ng, and D. Gagne, *Short-Term Solar Irradiance Forecasting Using Calibrated Probabilistic Models*, arXiv preprint arXiv:2010.04715 (2020), https://arxiv.org/abs/2010.04715.