A New Generative AI Approach to Forecast Solar Power in the Presence of Wildfire Smoke

Nusrat Subah Binte Shakhawat
Centre for New Energy Transition Research, Federation University Australia, Mt Helen 3353, Australia
nusrat.subah@ieee.org

Razieh Rastgoo
Electrical Engineering Department, Qatar University, Doha, Qatar
razieh.rastgoo@qu.edu.qa

Rakibuzzaman Shah
Centre for New Energy Transition Research, Federation University Australia, Mt Helen 3353, Australia
m.shah@federation.edu.au

Nima Amjady
Centre for New Energy Transition Research, Federation University Australia, Mt Helen 3353, Australia
n.amjady@federation.edu.au

S. M. Muyeen
Electrical Engineering Department, Qatar University, Doha, Qatar
sm.muyeen@qu.edu.qa

Syed Islam
Centre for New Energy Transition Research, Federation University Australia, Mt Helen 3353, Australia
s.islam@federation.edu.au

***Abstract*—Accurate forecasting of solar photovoltaic (PV) power generation is essential for maintaining stability and reliability of power systems, particularly under wildfire smoke conditions that induce rapid and nonlinear fluctuations in solar irradiance. To address this challenge, this paper introduces an energy-guided dynamic latent boundary adaptation (E-DLBA) framework, a deep generative artificial intelligence (GenAI) model designed for PV power forecasting in the presence of wildfires. The proposed E-DLBA framework ensures temporal stability and robust performance under high-volatility wildfire conditions. Additionally, a volatility-weighted variational loss (VWVL) is introduced to adapt the learning dynamics based on uncertainty levels, improving the model's sensitivity to abrupt smoke-induced variations. In addition, the proposed model leverages an enriched set of meteorological and environmental features, including wildfire-related aerosol indicators, to enhance contextual awareness. Comprehensive numerical experiments and ablation studies demonstrate that the integration of the proposed E-DLBA, VWVL, and enriched features significantly improves PV power forecasting accuracy in the presence of wildfires.**

Keywords—Deep generative AI, enriched feature set, solar power forecasting, wildfire.

I. INTRODUCTION

Wildfires have a wide-ranging impact on power grid operation and planning [1]. The frequency and intensity of wildfires have been escalating globally as a consequence of climate change [2]. Accurate estimation of the available capacity of distributed energy resources (DERs) is essential to ensure resilient grid operation during wildfire events. Since aerosol concentrations and ambient temperature, both influenced by wildfire smoke and heat transfer, directly affect solar photovoltaic (PV) generation, PV output is notably reduced under wildfire conditions [3]. Hence, developing an approach capable of reliably projecting PV generation in the presence of wildfires remains a pressing need.

Considerable progress has been made in PV power forecasting research. References [4]-[6] provide comprehensive reviews of state-of-the-art techniques, datasets, and tools used for PV power prediction. However, limited attention has been devoted to forecast PV power specifically during wildfire events. One of the earliest studies in this area [3] employed quantile regression to quantify PV capacity loss caused by wildfire smoke, enabling operators to undertake proactive measures to minimize load shedding. While [3] established a correlation between aerosol optical density (AOD) and PV output, the work in [7] incorporates the concentration of particulate matter with 2.5 μm (PM2.5) or smaller diameter —the primary smoke pollutant—together with additional meteorological inputs into a random forest machine learning (ML) model to predict PV power during wildfires. Building upon the methodology of [7], the work in [8] utilizes data for longer time periods and more geographical locations and presents an assessment of financial losses associated with PV power reduction during wildfires. In [9], six meteorological variables—absorption radiation index (ARI), temperature, humidity, visibility, wind speed, and dew point temperature—are considered alongside different PV cell technologies in a multi-linear regression model for PV power prediction during wildfires. A recent study further demonstrates that incorporating wildfire smoke forecast data can enhance PV capacity prediction when coupled with ML algorithms such as random forest and histogram-based gradient boosting [10].

These studies have substantially improved the understanding of wildfire impacts on PV systems. Recent research has also started to explore the use of GenAI-based methods to forecast PV power [4], [11]. Nevertheless, to the best of the authors' knowledge, no prior work has examined the potential of GenAI models to forecast available PV resources under wildfire conditions. Furthermore, while concentrations of particulate matter with 10 μm (PM10) or smaller diameters [12], SO_2 [13], O_3 [13], NO [13], NO_2 [13], and NH_3 [13] rise significantly during wildfires, existing PV power forecasting works have not considered these features.

To address the existing literature gaps, this paper presents the following novel contributions:

1) Enriched input feature set: An enriched input feature set that includes PM10, SO_2, O_3, NO, NO_2, and NH_3 with other relevant features is designed and utilized to represent the influence of wildfire heat transfer and smoke on PV power generation more effectively.

2) Energy-guided dynamic latent boundary adaptation (E-DLBA): A new deep GenAI model with a novel latent regulation mechanism is proposed for PV power forecasting during wildfire conditions. The proposed E-DLBA mechanism integrates energy-aware latent modulation with adaptive boundary evolution within the conditional variational auto-encoder (CVAE) backbone. The proposed E-DLBA mechanism dynamically adjusts the latent manifold boundaries according to temporal energy variations and volatility, ensuring stable yet flexible latent trajectories under wildfire-induced uncertainty.

3) Volatility-weighted variational loss (VWVL): A volatility-adaptive optimization strategy is proposed for the deep GenAI model to balance reconstruction fidelity and latent regularization according to local temporal volatility. The proposed VWVL reinforces latent robustness during wildfire events while maintaining accuracy in stable conditions, resul-

979-8-3315-8138-1

ting in a resilient and physically consistent PV power forecasting framework.

The next section presents the proposed methodology, followed by results and discussion in Section III. Section IV concludes the paper and outlines future directions.

II. METHODOLOGY

Details of the proposed framework are presented in this Section in two main parts: proposed data enrichment and a deep GenAI model.

A. Data Enrichment

An enriched set of input features, x for the deep GenAI model is proposed in this study, expressed in (1) and (2).

$$x = [x_r, x_s] \tag{1}$$

$$x_s = [x_{s1}, x_{s2}] \tag{2}$$

where x_r=[wind speed, wind direction, humidity, clear sky global horizontal irradiance (GHI), clear sky direct horizontal irradiance (DHI), clear sky direct normal irradiance (DNI), GHI with cloud factor, DHI with cloud factor, DNI with cloud factor, dew point, solar zenith angle, surface albedo, atmospheric pressure, precipitable water], x_{s1}=[temperature, AOD, PM2.5], and x_{s2}=[PM10, SO_2, O_3, NO, NO_2, NH_3]. In (1), x_r represents "Regular features", which are commonly employed in PV power forecasting regardless of wildfire presence [4]-[6]; furthermore, x_s is "Special features", which are not only notably influenced by wildfire smoke [12]-[13], but also affect PV output. x_s has two subgroups: x_{s1} and x_{s2}. Moreover, x_{s1} has been previously examined to establish the relationship between PV generation and wildfires [3], [7], [8]. In contrast, x_{s2} (despite their substantial impacts on PV performance [14]-[16]) has not yet been explored in wildfire context. While PM10, SO_2, O_3, NO and NO_2 attenuate solar radiation through scattering and absorption, thereby reducing the irradiance incident on PV panels [14-15], NH_3 contributes to increased soiling and haze formation [16], both of which diminish PV output.

B. Deep GenAI Model

The foundation of the proposed deep GenAI model is a CVAE model. It aims to learn a probabilistic mapping function between the input features of PV power forecasting process and their latent representations, conditioned on exogenous contextual factors such as meteorological variables and wildfire indicators. Unlike a standard variational autoencoder (VAE), which models an unconditional latent prior, the proposed CVAE model introduces conditioning on the context (c_t), enabling the latent variables to capture context-dependent variability in solar generation under environmental perturbations. An overview of the model is illustrated in Fig. 1. Further details are presented in the following.

1) Conditional Generative Process: Given a time-series input $x_{1:T} = \{x_t\}_{t=1}^T$ representing PV power output and corresponding contextual features $c_{1:T} = \{c\}_{t=1}^T$, the conditional generative process is formulated as [17]:

$$p_\theta(x_{1:T}|c_{1:T}) = \int p_\theta(x_{1:T}|z_{1:T}, c_{1:T}) p_\theta(z_{1:T}|c_{1:T})\, dz_{1:T} \tag{3}$$

where $z_{1:T}$ denotes the latent variables; $p_\theta(x_{1:T}|z_{1:T}, c_{1:T})$ indicates the likelihood decoder; $p_\theta(z_{1:T}|c_{1:T})$ represents a context prior conditioned on the latent dynamics: and θ indicates the learnable parameters of the generative model. This probabilistic formulation allows the model to represent not only the expected PV trajectory but also its variability under unc-

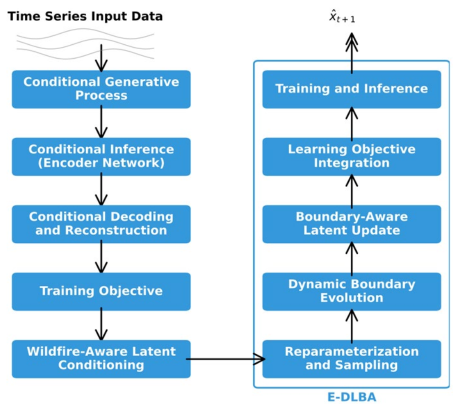

Fig. 1. Overview of the proposed CVAE model.

ertain wildfire-affected conditions.

2) Conditional Inference (Encoder Network): The encoder $q_\phi(z_{1:T}|x_{1:T}, c_{1:T})$ maps observed PV sequences and their conditioning context to a latent posterior distribution. It is typically parameterized as a diagonal Gaussian [17]:

$$q_\phi(z_t|x_{1:t}, c_{1:t}) = \mathcal{N}\left(\mu_\phi(x_{1:t}, c_{1:t}), diag\left(\sigma_\phi^2(x_{1:t}, c_{1:t})\right)\right) \tag{4}$$

where $\mu_\phi(.,.)$ represents the mean of the latent Gaussian distribution $\mathcal{N}(.,.)$; σ_ϕ^2 indicates a vector containing the variance for each latent dimension independently; and $diag(.,.)$ converts the variance vector into a diagonal matrix. To ensure temporal coherence, a temporal latent propagation mechanism is proposed as follows:

$$z_t = f_{enc}(x_t, c_t, h_{t-1}) \tag{5}$$

$$h_t = TCN(z_t, h_{t-1}) \tag{6}$$

where f_{enc} denotes the encoder transformation that maps the input–context pair and previous hidden state (h_{t-1}) into the latent representation z_t. It is implemented as a temporal convolutional network (TCN), which captures long-term dependencies. This preserves the temporal continuity of latent representations and ensures that latent variability aligns with observed volatility patterns.

3) Conditional Decoding and Reconstruction: The proposed likelihood decoder $p_\theta(x_t|z_t, c_t)$ reconstructs the predicted PV output at each time step, incorporating both z_t and c_t as below:

$$\hat{x}_t = f_\theta(z_t, c_t) \tag{7}$$

where $\hat{x}_t$ represents the reconstructed PV power output at time step t. This likelihood decoder $f_\theta(.,.)$ is implemented as a TCN as well, which is then modeled as follows to use in the reconstruction loss:

$$p_\theta(x_t|z_t, c_t) = \mathcal{N}(\hat{x}_t, \sigma_\theta^2) \tag{8}$$

where σ_θ^2 indicates the variance of the Gaussian distribution. By conditioning on c_t, the proposed likelihood decoder explicitly learns nonlinear responses to wildfire-related contextual variations (e.g., AOD, PM10).

4) Training Objective: The proposed CVAE maximizes the conditional evidence lower bound (ELBO) as follows [17]:

$$\mathcal{L}_{VAE} = \mathbb{E}_{q_\phi(z_{1:T}|x_{1:T}, c_{1:T})}[\log p_\theta(x_{1:T}|z_{1:T}, c_{1:T})] - D_{KL}\left(q_\phi(z_{1:T}|x_{1:T}, c_{1:T})||p_\theta(z_{1:T}|c_{1:T})\right) \tag{9}$$

where $\mathbb{E}[.]$ indicates the expectation value function and $D_{KL}(.)$ is the Kullback-Leibler divergence to penalize unlikely latent configurations and promote a well-structured, physically plausible latent manifold; $\mathbb{E}[.]$ promotes accurate reconstruction of the PV signal, while $D_{KL}(.)$ regularizes the learned posterior to remain close to the conditional prior, encouraging generalization under unseen wildfire-impact scenarios.

5) Wildfire-Aware Latent Conditioning: To capture wildfire influence without requiring external physical aerosol models, a volatility-conditioned latent modulation c_t' is introduced. The volatility v_t (measured from short-term reconstruction residuals) serves as an adaptive conditioning signal as below:

$$c_t' = [c_t; v_t] \tag{10}$$

$$v_t = \sqrt{\frac{1}{K}\sum_{i=t-K+1}^{t}\left(|x_i - \hat{x}_i| - \frac{1}{K}\sum_{j=t-k+1}^{t}|x_j - \hat{x}_j|\right)^2} \tag{11}$$

where $[.;.]$ denotes vector concatenation; and K represents the length of the short look-back window used to compute local volatility (v_t). This augmented conditioning vector c_t' is fed to both the encoder and decoder parts of the proposed CVAE, allowing the latent variable z_t to encode context-dependent volatility responses. This mechanism is crucial because wildfire events create abrupt, short-lived deviations that cannot be captured by simple contextual embeddings.

6) Reparameterization and Sampling: Sampling from the approximate posterior uses the reparameterization technique as below:

$$z_t = \mu_\phi(x_t, c_t) + \sigma_\phi(x_t, c_t) \odot \epsilon, \;\; \epsilon \sim \mathcal{N}(0, I) \tag{12}$$

where $\odot$ denotes the elementwise (Hadamard) product and ϵ is a random noise vector drawn from a standard multivariate normal distribution $\mathcal{N}(0, I)$, where I is the identity matrix. This noise term introduces stochasticity into the latent variable generation process, allowing the model to learn a smooth, continuous latent space. The decoder reconstructs the future PV sequence as below:

$$\hat{x}_{t+1} = f_\theta(z_t, c_t) \tag{13}$$

where f_θ is implemented via a TCN. The likelihood decoder is modeled as a Gaussian distribution as below [17]:

$$p_\theta(x_t | z_t, c_t) = \mathcal{N}(\hat{x}_t, \sigma_\theta^2 I) \tag{14}$$

This yields a reconstruction loss function $\mathcal{L}_{rec}$ as follows:

$$\mathcal{L}_{rec} = \|x_t - \hat{x}_t\|_2^2 \tag{15}$$

7) E-DLBA: To ensure stable temporal generalization and resilience to wildfire-induced volatility, the proposed E-DLBA mechanism is incorporated within the latent manifold learned by the proposed CVAE framework. This component is designed to adaptively reshape the latent boundary field so that latent trajectories neither diverge under abrupt perturbations nor collapse into overly smooth representations during stable intervals. The proposed E-DLBA thus acts as a temporal boundary regulator that continuously re-parameterizes the latent distribution in response to volatility and energy variations.

To have a latent boundary state formulation, let the latent representation sequence inferred from the proposed CVAE encoder be $Z = \{z_t\}_{t=1}^{T}$ and let $\mathcal{B}_t$ denotes the functional latent boundary state associated with z_t. Each boundary state $\mathcal{B}_t$ is defined as a learnable deformation of the nominal latent coordinate as below:

$$\mathcal{B}_t = z_t + \Delta_t \tag{16}$$

where Δ_t represents an adaptive boundary offset that is dynamically computed, intended to capture transient displacements in the latent trajectory caused by high-volatility phenomena such as wildfire haze or aerosol loading. Furthermore, Δ_t is modeled as a gradient-guided modulation derived from the functional energy field as follows:

$$\Delta_t = -\gamma_t \nabla_{z_t} E_\theta(z_t, \hat{z}_t) \tag{17}$$

$$\hat{z}_t = z_t - z_{t-1} \tag{18}$$

$$E_\theta(z_t, \hat{z}_t) = \frac{1}{2}\|z_t - z_{mean}\|^2 + \frac{\lambda}{2}\|\hat{z}_t\|^2 \tag{19}$$

where $E_\theta(.,.)$ denotes the functional energy-manifold generation, which constructs a parametric energy field over the latent space; z_{mean} and λ are the latent mean and a hyperparameter controlling the regularization term, respectively; $\hat{z}_t$ is the latent variation; and γ_t is an adaptive scaling factor governed by local volatility as below:

$$\gamma_t = \sigma_\gamma(\alpha_1 v_t + \alpha_2 \|\hat{z}_t\|) \tag{20}$$

with $\sigma_\gamma(\cdot)$ being a soft-plus activation function [18]. The associated loss function $\mathcal{L}_E$ ensures consistency between reconstructed energy flow and latent modulation, defined as below:

$$\mathcal{L}_E = \frac{1}{T}\sum_{t=1}^{T}\left\|\nabla_{z_t} E_\theta(z_t, \hat{z}_t)\right\|_2^2 \tag{21}$$

In this way, the magnitude of boundary deformation is automatically increased in regions of large reconstruction volatility or rapid latent movement. More details of the proposed E-DLBA are explained in three main parts: dynamic boundary evolution, boundary-aware latent update, and learning objective integration.

(a) Dynamic Boundary Evolution: To prevent abrupt transitions between adjacent latent boundaries, a temporal smoothing constraint is introduced:

$$\mathcal{B}_t^* = (1 - \lambda_b)\mathcal{B}_t + \frac{\lambda_b}{2}(\mathcal{B}_{t-1} + \mathcal{B}_{t+1}) \tag{22}$$

where $0 < \lambda_b < 1$ regulates the smoothness of boundary propagation and $\mathcal{B}_t^*$ indicates the optimal value of $\mathcal{B}_t$. This discrete formulation can be interpreted as a latent operator that distributes boundary corrections over neighboring time steps, ensuring temporal coherence while preserving the fine-scale response to volatility.

(b) Boundary-Aware Latent Update: The adapted latent variable ($\tilde{z}_t$) employed by the decoder is expressed as:

$$\tilde{z}_t = z_t + \mathcal{W}_b \odot (\mathcal{B}_t^* - z_t) \tag{23}$$

where $\mathcal{W}_b$ is a learnable weighting vector that controls dimension-wise sensitivity to boundary adaptation. Hence, dimensions with higher sensitivity to volatility receive stronger boundary corrections, allowing the decoder to focus on latent components most affected by wildfire events.

(c) Learning Objective Integration: The proposed E-DLBA mechanism contributes an auxiliary regularization term to the total training loss:

$$\mathcal{L}_{E-DLBA} = \eta_1 \mathbb{E}_t[\|\mathcal{B}_t - z_t\|_2^2] + \eta_2 \mathbb{E}_t[\|\mathcal{B}_t - \mathcal{B}_{t-1}\|_2^2] + \eta_3 \mathbb{E}_t[\mathbb{E}_\theta(z_t, \hat{z}_t)] \tag{24}$$

where η_1, η_2, and η_3 represent scalar coefficients, fine-tuned using the grid search algorithm [11]; and $\mathbb{E}_t[.]$ is the time-dependent expected value operator. The first term in (24) penalizes excessive boundary displacement, the second one enforces smooth temporal evolution, and the third one links the loss function of E-DLBA to the proposed energy function $E_\theta(.,.)$, ensuring consistency between latent energy dissipation and boundary modulation.

8) Total Loss Function: The total loss function $\mathcal{L}_{Total}$, which is used for the end-to-end training of the proposed deep

GenAI model, can be formulated as follows:

$$\mathcal{L}_{Total} = \mathcal{L}_{CVAE} + \omega_1 \mathcal{L}_E + \omega_2 \mathcal{L}_{E-DLBA} \quad (25)$$

where the weighting coefficients ω_1 and ω_2 control the relative contributions of $\mathcal{L}_E$ and $\mathcal{L}_{E-DLBA}$ regularization, enabling the proposed deep GenAI model to produce reconstructions that are both accurate and physically plausible. These weighting coefficients are fine-tuned using the grid search algorithm [11]. More specifically, $\mathcal{L}_{CVAE}$ drives the reconstruction of observed PV sequences while enforcing a context-conditioned latent prior, allowing latent variables to capture variability due to meteorological and wildfire factors. The functional energy-manifold loss $\mathcal{L}_E$ regularizes the latent space via a learnable energy field, promoting smooth and energy-consistent latent trajectories. Finally, the dynamic latent boundary adaptation loss $\mathcal{L}_{E-DLBA}$ adaptively reshapes latent boundaries in response to local volatility and latent energy, ensuring temporal coherence and preventing collapse or divergence of latent trajectories.

9) Training and Inference: The proposed model is trained end-to-end using $\mathcal{L}_{Total}$ (as in (25)). During training, observed PV sequences $x_{1:T}$ and their corresponding contextual features $c_{1:T}$, including wildfire indicators (i.e., the enriched set of input features), are fed into the encoder to infer the approximate posterior $q_\phi(z_{1:T}|x_{1:T}, c_{1:T})$. Latent variables $z_{1:T}$ are sampled using the reparameterization technique and passed through the decoder to reconstruct $\hat{x}_{1:T}$. The proposed E-DLBA mechanism dynamically adapts latent boundaries according to local volatility, while the proposed functional energy-manifold generation enforces energy-consistent latent trajectories. Gradients of the total loss are backpropagated to update both encoder and decoder parameters (i.e., ϕ and θ) as well as the energy-field and boundary adaptation modules, ensuring temporally coherent and physically plausible latent representations. During inference, only the contextual inputs $c_{1:T}$ and the learned conditional prior $p_\theta(z_{1:T}|c_{1:T})$ are used to sample latent trajectories, which are then decoded to generate PV forecasts $\hat{x}_{1:T}$ under unseen wildfire conditions. This enables the proposed model to produce robust, context-aware PV power point forecasts while capturing uncertainty in latent dynamics.

III. Results and Discussions

The dataset used in this study as input to the proposed deep GenAI model comprises real-world historical data from 2019 and 2020 for Parramatta North, New South Wales, Australia. Meteorological variables were obtained from [19]- [21], measured real-time, with point-level spatial resolution. Wildfire occurrence records, used to identify wildfire days, were sourced from [22]. After data collection with hourly temporal resolution, missing values were filled using linear interpolation. The proposed methodology has been tested on a PV power plant with a capacity of 2.5 MW. The specifications of this plant have been collected from a distribution network service provider in Australia [23].

All numerical experiments of this paper were implemented in Python and executed on NVIDIA Tesla K80 GPUs. For fair comparisons, both the proposed model and all benchmark counterparts were trained and evaluated under identical experimental settings, including the same training/validation/testing datasets, the same normalization procedure (with zero mean and unit variance [17]), and the same evaluation metrics. Hyperparameter optimization for all models was performed using a grid search strategy, as [11], where the optimal parameters were selected based on the lowest validation error and subsequently applied during testing. Additionally, an identical early stopping mechanism was adopted across all models to prevent overfitting [17]. This uniform setup ensured that the performance differences arose solely from model design rather than implementation bias. For one year of hourly training data, the proposed model required approximately 2.5 hours of training time on an NVIDIA Tesla K80 GPU, while inference on one year of test data was completed in less than 5 seconds, corresponding to an average per day inference cost of <0.014 s, confirming the practical efficiency of the proposed framework for operational day-ahead forecasting. Additionally, the proposed model has approximately 1.2 million trainable parameters, and a linear computational cost due to encoding–decoding based on temporal convolutional network and E-DLBA updates, resulting in an average inference cost of 0.8 Giga Floating-Point Operations per forecasting day (24 forecasting steps), which remains well within real-time operational limits.

In this study, in-sample (IS) and out-of-Sample (OoS) data refer to the training/validation and testing data, respectively. The two-year dataset (2019 and 2020) is divided into training, validation, and test sets with a ratio of 3:1:4. Thus, the testing set (year 2020) is completely unseen for the proposed and comparative forecast methods. The two years have 61 and 49 wildfire days, respectively. Due to the large number of wildfire events, the 2019-2020 wildfires in Australia are known as the "Black summer fires" [24]. Moreover, the forecasting step and forecasting horizon are set to one hour and one day (24 hours), respectively. The predictive performance of the proposed and comparative models is quantitatively assessed using two widely adopted error metrics: mean absolute error (MAE) and mean absolute percentage error (MAPE) [11].

A. Ablation Analysis Based on Methodological Components

To evaluate the proposed model, an ablation analysis is conducted to investigate the individual and joint contributions of the key components of the proposed framework: E-DLBA and VWVL. To this end, three comparative configurations are designed and evaluated as in Table I. To visualize the model's and its comparative configurations' performances under wildfire conditions, a severe wildfire day (February 14, 2019) was selected. Fig. 2 presents a time-series visualization comparing the actual PV power values, and the values forecasted using the three methods in Table I for this day.

In the first configuration, E-DLBA is retained while reverting to the original loss, which corresponds to the conventional VAE objective treating all temporal samples equally without volatility-based weighting. This setup effectively stabilizes the latent space and preserves temporal coherence but lacks sensitivity to highly uncertain periods caused by wildfire-induced irradiance fluctuations. In the second configuration, the original boundary is kept, meaning the latent variables follow a static Gaussian prior without adaptive regularization, while VWVL is introduced to modulate the learning process

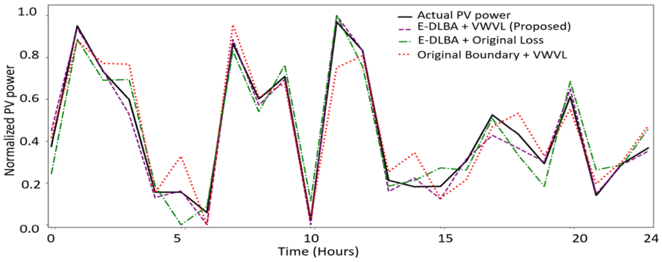

Fig. 2. Time-series visualization of actual vs. forecasted PV power values obtained from the proposed and two comparative models (ablated variants) on February 14, 2019.

979-8-3315-8138-1

TABLE I. RESULTS FROM ABLATION ANALYSIS OF THE PROPOSED METHODOLOGY

Model	MAE		MAPE (%)	
	IS	OoS	IS	OoS
E-DLBA + Original Loss	24.5	26.2	15.8	17.1
Original Boundary + VWVL	26.4	27.6	17.2	18.8
E-DLBA + VWVL (Proposed)	19.1	20.4	12.8	14.2

according to local volatility. This variant improves robustness against noise and uncertainty, enabling the model to focus more effectively on volatile regions, yet the absence of E-DLBA may reduce the temporal coherence of latent trajectories and smoothness in temporal transitions. Finally, the full configuration, which combines the proposed E-DLBA and VWVL components, demonstrates the strongest overall performance in terms of both metrics, as clearly reflected in Table I. Furthermore, significant differences are seen between the results of the proposed model and its two simpler versions in Table I, further illustrating the effectiveness of the proposed E-DLBA and VWVL components. Together, these components enable the proposed deep GenAI model to achieve both stable latent evolution and precise reconstruction under varying uncertainty conditions.

B. Ablation Analysis Based on Input Features

The details of the scenarios examined in this ablation analysis based on input features and the results obtained from this analysis are summarized in Table II. This analysis focuses exclusively on the proposed enriched set of input features. In Table II, ✓ denotes inclusion of an input feature, whereas × denotes its exclusion. As confirmed by this Table, the proposed enriched set of input features achieves the lowest error across both evaluation metrics, whereas the exclusion of any proposed feature increases the forecasting error.

IV. CONCLUSIONS

This paper introduces an energy-guided generative framework for PV power forecasting under wildfire smoke conditions. The model leverages an enriched set of meteorological features, including wildfire-related aerosol indicators, to enhance contextual awareness. By integrating E-DLBA and VWVL, the proposed deep GenAI framework adaptively regulates latent boundaries and dynamically adjusts the loss function of the learning process based on local volatility. Ablation analyses demonstrate that the proposed framework achieves superior forecasting accuracy compared to its ablated variants, highlighting the effectiveness of both the proposed components and input features. Future work will extend the proposed framework toward multi-site and spatial–temporal forecasting with improved interpretability.

TABLE II. RESULTS FROM ABLATION ANALYSIS OF THE ENRICHED INPUT FEATURE SET

Scenarios	Features						MAE	MAPE (%)
	PM10	SO_2	*NO*	NO_2	NH_3	O_3	*OoS*	*OoS*
1	✓	✓	✓	✓	✓	✓	20.4	14.2
2	✓	✓	✓	✓	✓	×	22.6	17.8
3	✓	✓	✓	✓	×	✓	22.5	17.6
4	✓	✓	✓	×	✓	✓	22.7	17.9
5	✓	✓	×	✓	✓	✓	22.4	17.5
6	✓	×	✓	✓	×	✓	23.2	18.6
7	×	✓	✓	✓	✓	✓	22.5	17.7

REFERENCES

[1] S. Vahedi, *et al.*, "Wildfire and power grid nexus in a changing climate," *Nature Reviews Electrical Engineering*, vol. 2, pp. 225–243, 2025.

[2] M. S. Miah, R. Shah, N. Amjady, T. Surinkaew and S. Islam, "Resiliency assessment and enhancement of renewable dominated edge of grid under high-impact low-probability events—A review," *IEEE Trans. Ind. Appl.*, vol. 60, no. 5, pp. 7578-7598, Sept.-Oct. 2024.

[3] D. L. Donaldson, D. M. Piper and D. Jayaweera, "Temporal solar photovoltaic generation capacity reduction from wildfire smoke," *IEEE Access*, vol. 9, pp. 79841-79852, 2021.

[4] Q. -T. Phan, K. -C. Zhan, Y. -Y. Hsu, Y. -K. Wu and P. Mandal, "Next-generation PV power forecasting: Advances in AI model architectures, sky imaging, weather classification, and data validation strategies," in *IEEE Industry Applications Society Annual Meeting*, Taipei, Taiwan, 2025, pp. 1-6.

[5] P. Di Leo, A. Ciocia, G. Malgaroli, and F. Spertino, "Advancements and challenges in photovoltaic power forecasting: A comprehensive review," *Energies*, vol. 18, no. 8, p. 2108, 2025.

[6] M. Gupta, A. Arya, U. Varshney, J. Mittal, and A. Tomar, "A review of PV power forecasting using machine learning techniques," *Progress in Engineering Science*, vol. 2, no. 1, p. 100058, 2025.

[7] S. D. Gilletly, N. D. Jackson, and A. Staid, "Evaluating the impact of wildfire smoke on solar photovoltaic production," *Applied Energy*, vol. 348, p. 121303, 2023.

[8] S. M. Treacy and A. B. Moura, "The effects and financial impacts of wildfire smoke on solar photovoltaic power production in Alberta, Canada," *REM- Research in Economics and Mathematics*, Working Paper No. 0373, Mar. 2025.

[9] A. J. Ali, L. Zhao and M. H. Kapourchali, "Data-driven-based analysis and modeling for the impact of wildfire smoke on PV systems," *IEEE Trans. Ind. Appl.*, vol. 60, no. 2, pp. 2076-2084, Mar.-Apr. 2024.

[10] F. Bartram, B. Yuan, and K. M. Zhang, "Predicting solar photovoltaic generation impacted by severe wildfire smoke," *Environmental Research Letters*, vol. 20, no. 6, p. 064015, May 2025.

[11] R. Rastgoo, N. Amjady, S. Lin and S. M. Muyeen, "Ultra-short-term solar power prediction using sky image sequences by a residual vision reformer," *IEEE Trans. Sustain. Energy*, vol. 16, pp. 2972-2988, 2025.

[12] M. Oliveira, C. Delerue-Matos, M. C. Pereira, and S. Morais, "Environmental particulate matter levels during 2017 large forest fires and megafires in the center region of Portugal: A public health concern?," *International Journal of Environmental Research and Public Health*, vol. 17, no. 3, p. 1032, Feb. 2020.

[13] D. A. Jaffe, *et al.*,"Wildfire and prescribed burning impacts on air quality in the United States," *Journal of the Air & Waste Management Association*, vol. 70, no. 6, pp. 583–615, Jun. 2020.

[14] J. Shim, S. Park, and D. Song, "Impact of particulate matter (PM10, PM2.5) on global horizontal irradiance and direct normal irradiance in urban areas," *Building and Environment*, vol. 271, p. 112610, 2025.

[15] V. Sultan, D. Coleman, E. Robles, and M. Van, "Assessing the impact of air quality on solar energy production," *International Journal of Advanced and Applied Sciences*, vol. 12, pp. 1–6, 2025.

[16] SunBrush Mobil GmbH, "Removal of ammonia deposits," *SunBrush Mobil.* Available: https://shorturl.at/UeDQW. [Accessed: Oct. 22, 2025].

[17] R. Rastgoo, N. Amjady, R. Shah and S. M. Muyeen, "A diffusion-based probabilistic ultra-short-term solar power prediction using the sky image sequences," *IEEE Access*, vol. 13, pp. 149532-149553, 2025.

[18] H. Zhao, F. Liu, L. Li and C. Luo, "A novel softplus linear unit for deep convolutional neural networks," *Applied Intelligence*, vol. 48, no. 7, pp. 1707-1720, 2018.

[19] "NSRDB: National Solar Radiation Database ," [Online]. Available: https://nsrdb.nrel.gov/data-viewer (Accessed Oct. 20, 2025).

[20] NSW Department of Planning & Environment, "Air Quality Data Download Facility," [Online]. Available: https://www.airquality.nsw.gov.au/air-quality-data-services/data-download-facility (Accessed Oct. 20, 2025).

[21] National Aeronautics and Space Administration, "NASA POWER: Data Access Viewer," [Online]. Available: https://power.larc.nasa.gov/data-access-viewer/ (Accessed Oct. 20, 2025).

[22] Australian Bureau of Meteorology, "Climate data online," [Online]. Available: http://www.bom.gov.au/climate/data/ (Accessed Oct. 19, 2025).

[23] "Daily available capacity and planned outages," Powercor Distribution Annual Planning Report, 2025. [Online]. Available: https://dapr.powercor.com.au/. (Accessed Oct. 24, 2025).

[24] C. A. T. Cortes *et al.*, "Analysis of wildfire visualization systems for research and training: Are they up for the challenge of the current state of wildfires?," *IEEE Trans. Vis. Comput. Graph.*, vol. 30, no. 7, pp. 4285-4303, July 2024.