Dear Reviewer,

We sincerely thank the reviewer for the comprehensive evaluation and the encouraging comments regarding the novelty and technical soundness of our proposed MIPV-NWP-PINNs framework.

We particularly appreciate the constructive suggestions concerning the clarity of the introduction, the generalizability of the PV forecasting model, and the physical interpretability of the PINNs parameters. We carefully considered all the comments and made detailed revisions to improve the manuscript in its scientific rigor and readability.

1. Reviewer Major Comment 1: In the abstract, the statement "raw Numerical Weather Prediction (NWP) outputs often fail to provide reliable forecasts because PV power is influenced by multiple coupled factors, including meteorological factors and photovoltaic modules" may be misleading. As currently written in the abstract, it may give the impression that PV power is directly predicted from NWP outputs. The authors are encouraged to clarify that NWP provides meteorological inputs, and that uncertainties arise from both atmospheric forecast errors and the subsequent PV power conversion process.

Response: We thank the reviewer for this comment. We have revised the abstract to explicitly clarify that NWP provides meteorological inputs and that uncertainties arise from two distinct sources: (1) inherent errors in NWP-derived meteorological variables (particularly solar irradiance), and (2) the complex nonlinear conversion process from meteorological inputs to PV power output.

Revised Text (Page 1, Lines 12-15): However, achieving reliable PV power forecasts remains challenging due to two primary sources of uncertainty: inherent errors in Numerical Weather Prediction (NWP)-derived meteorological variables, particularly solar irradiance, and the complex nonlinear conversion process from meteorological inputs to PV power output, which is influenced by both atmospheric conditions and PV module characteristics.

2. Reviewer Major Comment 2: The literature review presented in the Introduction is relatively lengthy and could be streamlined. In Section 1, it is suggested to reorganize this part around the core research problems by clearly summarizing: (1) the limited accuracy of NWP-derived GHI; (2) the "black-box" nature and lack of physical consistency in many PV power forecasting models; and (3) the potential of Physics-Informed Neural Networks (PINNs) to embed physical laws into the loss function and improve physical consistency. A more concise, problem-oriented review would improve readability and strengthen the overall motivation.

Response: We sincerely thank the reviewer for the constructive suggestion to streamline and reorganize the Introduction. In the revised manuscript, we have restructured Section 1 to center around the core research problems as suggested:

1. Limited accuracy of NWP-derived GHI: We highlighted the systematic overestimation issues in NWP models (e.g., WRF) due to cloud and aerosol misrepresentation, and the necessity of statistical post-processing.

2.'Black-box' nature of data-driven models: We explicitly discussed the limitations of pure deep learning approaches, emphasizing their lack of physical consistency despite their high accuracy.

3. Potential of Physics-Informed Neural Networks (PINNs): We introduced PINNs as a solution to bridge the gap between physical laws and data-driven efficiency, establishing the motivation for our proposed MIPV-NWP-PINNs framework.

We believe this problem-oriented restructuring significantly improves the readability and clearly articulates the motivation behind our work. We have added the revised introduction to **Pages 1-3, Lines 31-90**.

3. Reviewer Major Comment 3: While multi-site validation is conducted for GHI correction, the PV power forecasting results presented in Section 4 (e.g., the multi-horizon comparisons in Figures 12–15) are mainly based on a single station. The applicability of the proposed PINN-iTransformer to other regions, climatic conditions, and PV configurations should therefore be discussed more explicitly, or additional validation should be provided.

Response: We thank the reviewer for this constructive comment. To address this concern within available data constraints, we provide both theoretical justification and empirical supporting evidence. In addition, the uncertainties and limitations are explicitly discussed.

First, we validated the GHI correction model using 2025 data from the NASA POWER database. Application to WRF simulations at six different sites revealed that the proposed GHI correction model maintains its efficacy across independent spatial and temporal test cases.

Second, the physics-informed architecture embeds fundamental PV relationships (temperature-dependent efficiency, irradiance-power conversion) into the loss function, restricting the solution space to physically consistent predictions. This inductive bias is expected to enhance transferability, as the learned relationships reflect universal principles rather than site-specific correlations.

Finally, we conducted SHAP analysis to examine whether the model captures physically meaningful patterns (Figure 15). The results demonstrate: (i) an appropriate transition from temporal persistence at short horizons to meteorological dependence at longer horizons, and (ii) physically interpretable feature interactions consistent with PV thermodynamics.

We emphasize that interpretability analysis provides supporting evidence of physical consistency but does not substitute for direct multi-site validation. This limitation is now explicitly stated in Section 5, with multi-site evaluation identified as a priority for future work. All changes are highlighted in blue in the revised manuscript (**Pages 21-22, Lines 437-467**) for the reviewer’s convenience.

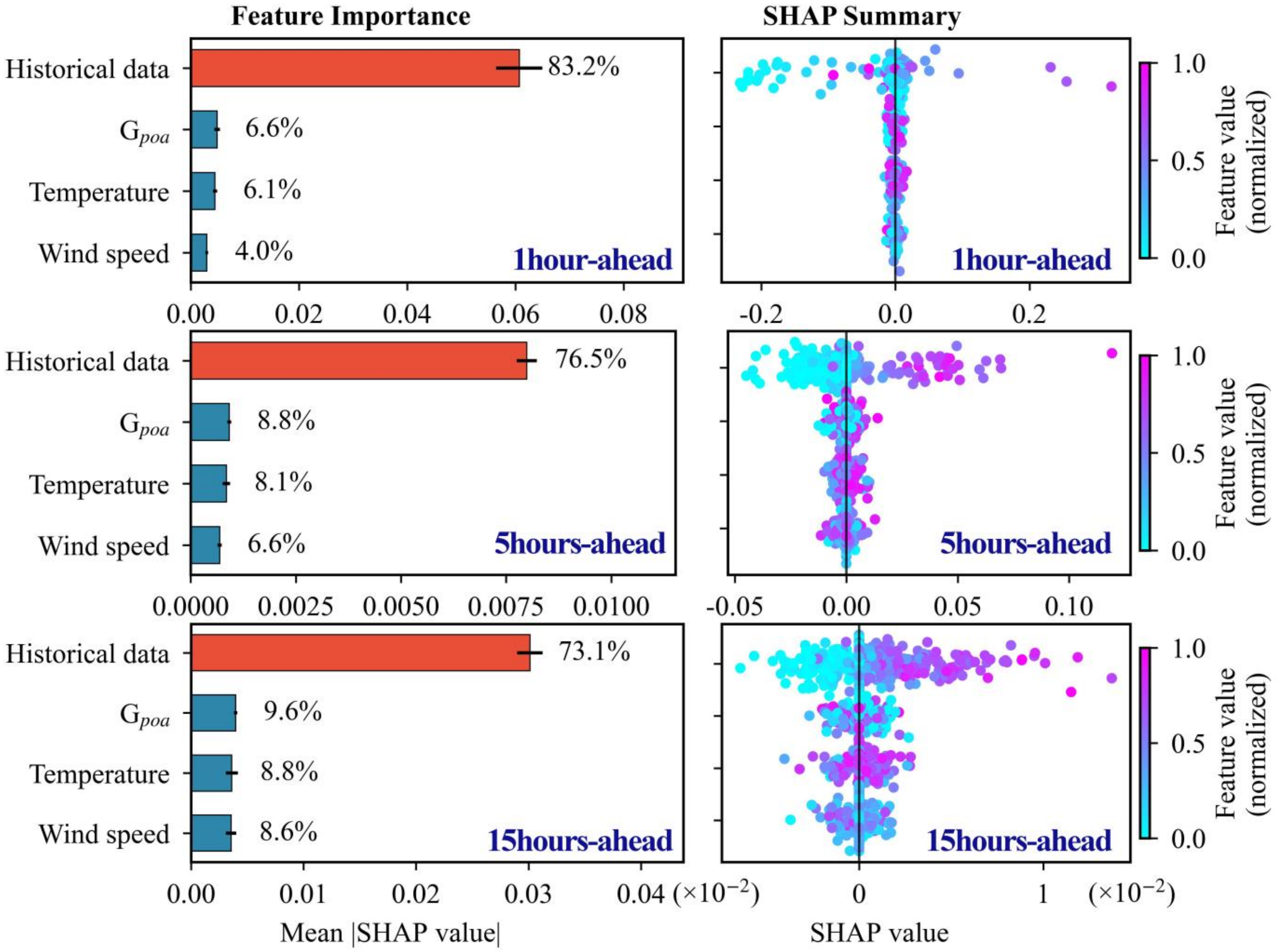

Figure 15: Left panels (top to bottom): distributions of feature importance at forecast horizons of 1, 5, and 15 h; right panels (top to bottom): SHAP beeswarm plots at forecast horizons of 1, 5, and 15 h.

4. Reviewer Major Comment 4: In the description of the PINN-iTransformer framework, particularly in the subsection introducing the physics-informed loss and the first-order relaxation ODE, the physical meaning of the relaxation coefficient k is not sufficiently explained. In addition, the role of the physics-informed loss weight (λ) is introduced without further discussion. Clarifying whether k is constant or learnable, and providing a brief sensitivity or ablation analysis for λ, would strengthen the physical interpretability and robustness of the proposed approach.

Response: We thank the reviewer for these valuable suggestions and we have revised the methodology and results sections to provide a deeper explanation of $\boldsymbol{k}$ and $\boldsymbol{\lambda}$. As suggested, we have also added a comprehensive sensitivity analysis (now presented in the new Figure S2) to evaluate the impact of $\boldsymbol{\lambda}$ on model performance.

1. Clarification of the Relaxation Coefficient $\boldsymbol{k}$: We have revised Section 2 to explicitly state that $\boldsymbol{k}$ is a learnable parameter, not a fixed constant. Physically, it represents the adaptive coupling strength (or restoration rate) between the theoretical equilibrium ($\boldsymbol{P_{eq}}$) and the predicted power. As shown in the Figure S2(e), during training, does not fluctuate randomly but converges to a stable range of approximately 0.90–0.96. This indicates that the model successfully learns an intrinsic restoration rate, balancing the simplified physical theory with complex observed data patterns.

2. Sensitivity Analysis of the Physics-Informed Loss Weight $\boldsymbol{\lambda}$: We conducted a sensitivity analysis for $\boldsymbol{\lambda}$ ranging from 10^{-4} to 10^{0} to evaluate the trade-off between data fidelity and physical constraints. Based on the sensitivity analysis, the key findings are integrated into the manuscript:

(1). **Optimal Regime ($\boldsymbol{\lambda < 0.01}$)**: In this range, the physical loss acts as a soft regularizer. Both RMSE and MAE remain consistently below the baseline iTransformer (dashed lines in Figure S2), demonstrating that appropriate physical constraints improve generalization without overriding data-driven features.

(2). **Over-Constraint Regime ($\boldsymbol{\lambda > 0.05}$)**: Performance degrades significantly (underfitting) when $\boldsymbol{\lambda}$ exceeds 0.05. This confirms that excessive weighting forces the model to adhere too strictly to the simplified first-order ODE, which cannot fully capture the stochastic fluctuations inherent in real-world weather data.

All changes, including the detailed discussion of these results, are highlighted in blue in the revised manuscript (**Page 9, Lines 220-229** and **Pages 20-21, Lines 425-436**) for the reviewer's convenience. We believe these revisions have substantially improved the clarity of our description.

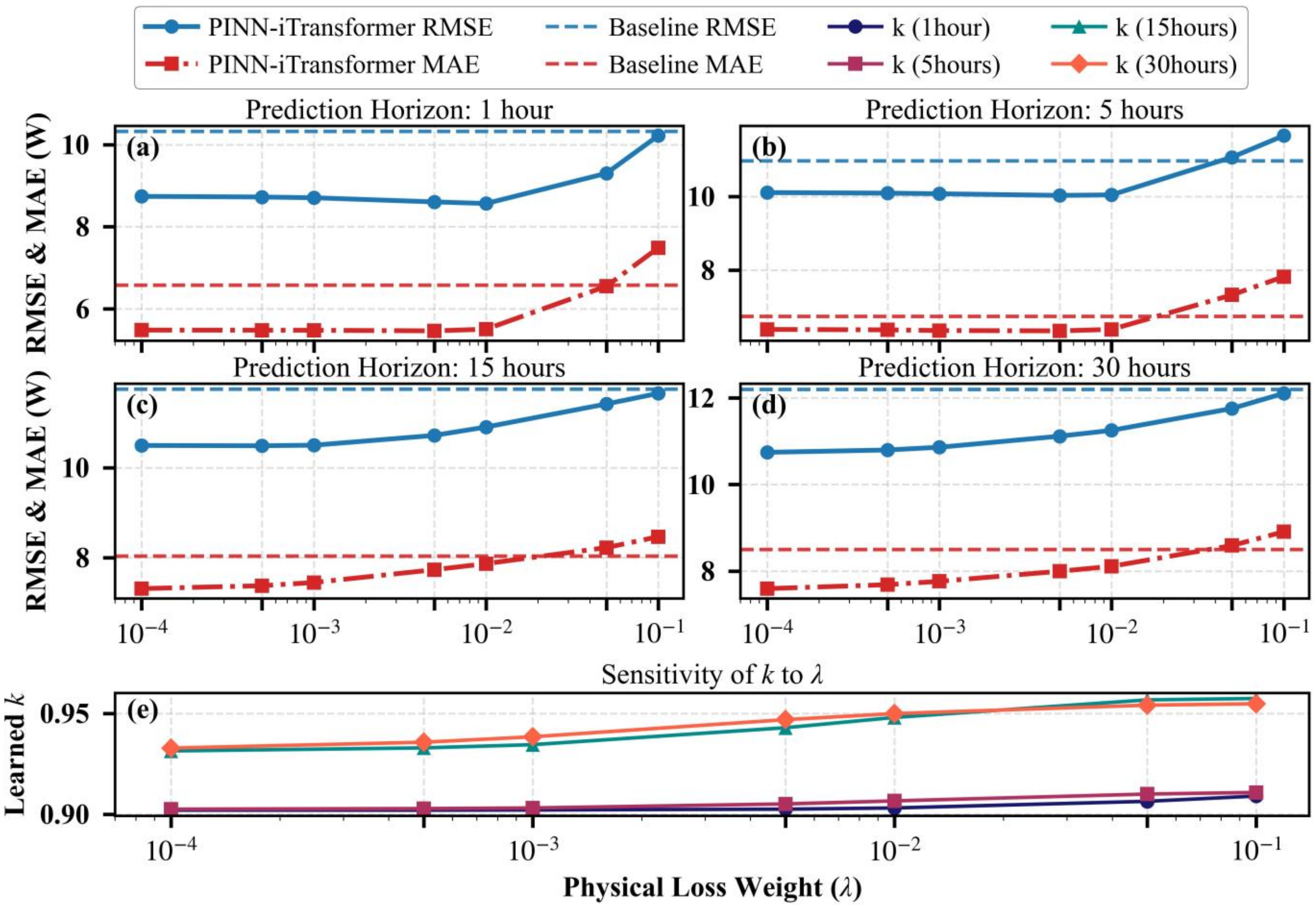

Figure S2: (a)–(d) Variation curves of RMSE and MAE with the physical constraint parameter λ at forecast horizons of 1, 5, 15, and 30 h, respectively; (e) variation curve of the relaxation coefficient k with λ.

5. Minor Comment 1: Although the manuscript emphasizes the importance of aerosol and cloud effects on GHI in the Introduction, the QM-TPA-LSTM model described in the GHI correction section remains purely statistical. This limitation should be more clearly acknowledged and discussed, particularly in relation to physical interpretability.

Response: We thank the reviewer for this insightful observation. We have added a new paragraph at the end of Section 3 that clearly states that the current QM-TPA-LSTM model is a purely statistical approach that does not explicitly incorporate physical variables, such as aerosol optical depth or cloud properties. We discuss that while the model effectively reduces NWP biases, it lacks direct physical interpretability, and we outline potential future improvements, including the integration of satellite-derived aerosol and cloud products.

Revised Text (Page 16, Lines 362-369): It should be acknowledged that the current QM-TPA-LSTM model operates as a purely statistical post-processing approach without explicitly incorporating physical variables such as aerosol optical depth, cloud fraction, or cloud optical thickness. While this statistical framework effectively reduces systematic NWP biases and captures temporal patterns in GHI errors, it lacks direct physical interpretability regarding the underlying atmospheric processes. The model learns implicit relationships between NWP outputs and observed GHI, but cannot explicitly attribute correction magnitudes to specific physical factors such as aerosol scattering or cloud attenuation. Future work could enhance physical interpretability by integrating satellite-derived aerosol and cloud products as additional input features, or by developing physics-guided correction schemes that explicitly model the radiative transfer processes affected by atmospheric constituents.

6. Minor Comment 2: While improved PV power forecasting accuracy is demonstrated in the results section, the manuscript does not discuss how these forecasts could be used in downstream applications such as grid operation or energy management. A brief discussion in the concluding section would enhance the applied value of the study.

Response: We appreciate this valuable suggestion. We have added a new paragraph in the Conclusions section discussing the practical implications of improved PV forecasting for downstream applications. Specifically, we discuss how forecasts at different time scales support various grid operation tasks: day-ahead forecasts for unit commitment and market participation, intra-day forecasts for energy storage scheduling, and ultra-short-term forecasts for automatic generation control. We also note the potential for uncertainty quantification to support risk-aware decision-making.

Revised Text (Page 19, Lines 412-418): Furthermore, the improved multi-scale PV power forecasting demonstrated in this study has significant implications for power system operations. As emphasized in previous reviews (Antonanzas et al., 2016; Iheanetu, 2022), forecasts at different time scales serve distinct operational purposes: day-ahead forecasts (24-120 h) support unit commitment decisions and electricity market participation; intra-day forecasts (6-24 h) facilitate energy storage scheduling and real-time balancing; and ultra-short-term forecasts (1-6 h) are critical for automatic generation control and frequency regulation. These operational benefits underscore the practical value of integrating physics-informed approaches into PV forecasting systems.

7. Minor Comment 3: In the PV power forecasting analysis, additional investigation of prediction errors—such as SHAP analysis or other interpretability methods—would be beneficial for understanding the sources of model error and the relative importance of input variables.

Response: We thank the reviewer for this insightful suggestion. Following your recommendation, we have integrated a SHAP analysis into our revised manuscript. This analysis quantifies the contribution of each input variable across different forecast horizons (1h, 5h, and 15h) and provides a deeper look into the model's internal logic.

Summary of changes in the manuscript:

1. We added a detailed discussion in Section 5 exploring how the model transitions from relying on temporal persistence to physical meteorological signals as the forecast horizon increases.

2. We included a SHAP summary plot and a feature importance transition diagram to visually demonstrate these findings.

More detailed revision can refer to the response to major Comment 3.

8. Minor Comment 4: Throughout the manuscript, please ensure consistent terminology, particularly for forecasting horizons (e.g., "6 h" vs. "6-hour" vs. "6 hours") and model names (e.g., PINN-iTransformer vs. PINN–iTransformer)."

Response: We sincerely apologize for the inconsistencies in terminology and formatting in the original manuscript. We have conducted a thorough, line-by-line review of the entire manuscript to ensure consistency. Specifically, we have implemented the following standardizations:

1.We have standardized the usage of time units.

2.We have also checked other units and abbreviations to ensure they adhere to a uniform style throughout the paper.

All changes have been highlighted in the revised manuscript (or tracked via the new version).

9. Minor Comment 5: In the methodological sections and equations, several symbols (e.g., k, λ, and variables related to PV power and irradiance) are introduced without being clearly defined at first occurrence. Providing clearer definitions or a concise summary of symbols would improve readability."

Response: We appreciate the reviewer's comment regarding the clarity of our mathematical notation. We have thoroughly revised Section 2 (Methodology) to ensure that every symbol—including the physical parameters (k, λ) and the variables for PV power and solar irradiance—is explicitly defined upon its first appearance in the equations. Specific improvements made: In-text Definitions: We have added immediate definitions following each equation. We believe these changes significantly improve the clarity of the technical framework.

10. Minor Comment 6: In the multi-panel figures presenting PV forecasting results, some fonts and legends are relatively small. Improving font size and legend placement would enhance clarity.

Response: We thank the reviewer for this practical suggestion. In response, we have systematically optimized all figures in the manuscript (especially the multi-panel plots in the Results and Discussion sections) to ensure high clarity and professional presentation. The following specific improvements have been made:

1. The font sizes for all axis labels, tick marks, and subplot titles have been increased to ensure they remain legible when scaled to the journal's column width.

2. Legends have been resized and strategically repositioned to avoid overlapping with data curves, and in some multi-panel figures, we have used unified legends to reduce visual clutter and maximize the plotting area.

3. We also adjusted the line widths and marker sizes in the forecasting plots to ensure distinct visibility between the observed power and the predicted results from different models.

These updates have been applied to Figures 13-14. We believe the revised figures now meet the high standards for publication in GMD.

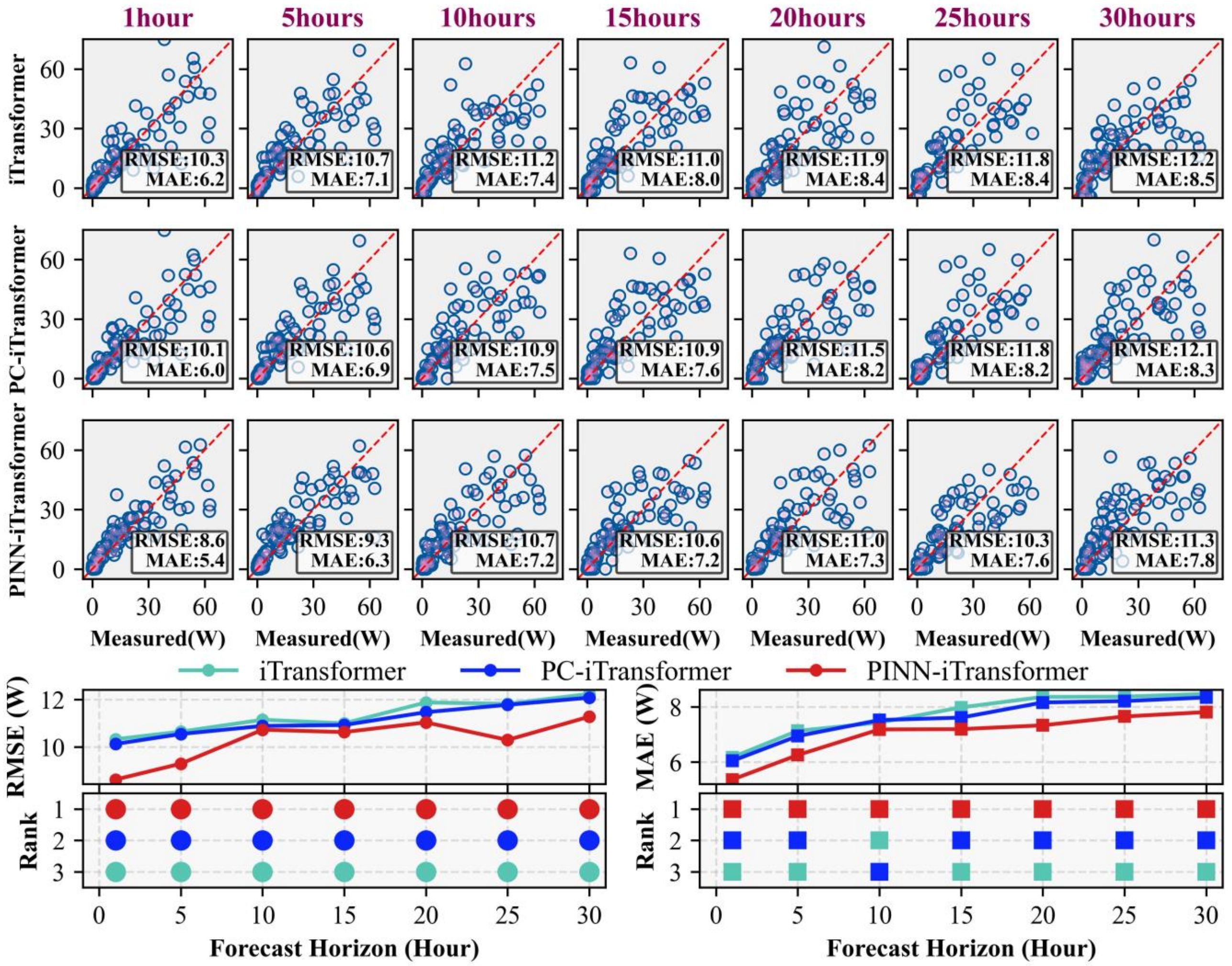

Figure 13: Multi-step PV power forecasting performance using a 5-day WRF forecast.

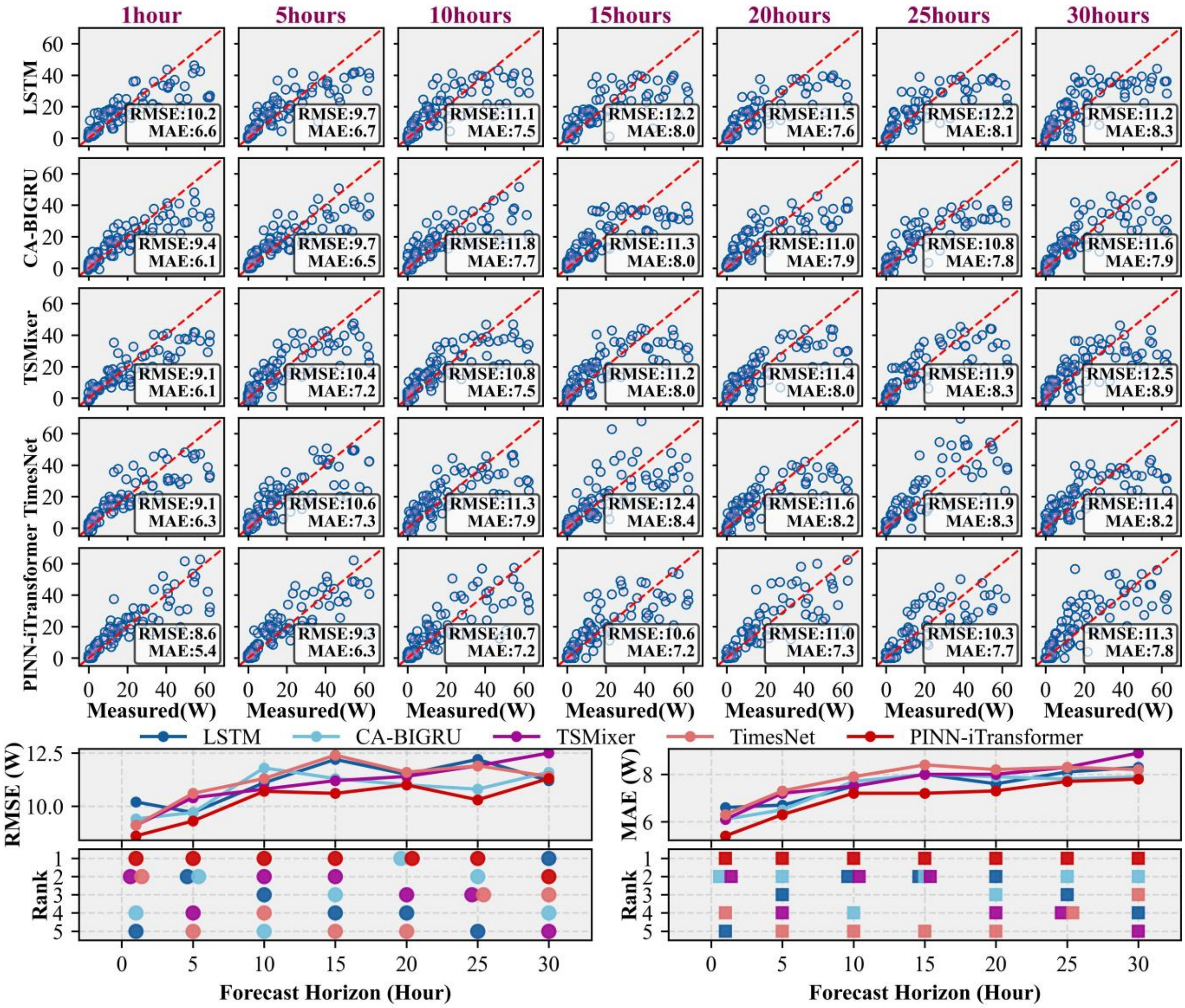

Figure 14: Performance comparison of different models for multi-step forecasting.

11. Minor Comment 7: Please ensure that physical units (e.g., W/m^2, kW, MW) are consistently formatted throughout the manuscript and that spacing between numbers and units follows journal conventions.

Response: We thank the reviewer for pointing out these formatting details. We have conducted a comprehensive audit of the entire manuscript to ensure that all physical units and their spacing adhere to the Copernicus Publications and GMD journal guidelines.

We hope that the revisions and explanations provided adequately address the reviewer's concerns. We believe that these improvements have significantly strengthened the manuscript, making it suitable for publication.

Thank you once again for your time and effort in reviewing our work.

Sincerely,

The Authors