AI-Driven Corrections on Global Climate Models Using Deep Learning

Yogeeta Bhavya Behl
Chaitanya
2K23CSUN01192 2K23CSUN01138 2K23CSUN01139
Department of Computer Science Department of Computer Science
DepartmentofComputerScience and Technology, SOE
and Technology, SOE and Technology, SOE

Dr.MeenaChaudhary Dr.NarendraGautam Dr. GunjanChandwani
Department of Computer Science and Department of Computer Science and
DepartmentofComputerScienceand
Technology Technology Technology

Abstract: Global Climate Models (GCMs) formthe backbone of modern climate science, providing essential simulations for long-term projections of temperature, precipitation, and hydrological cycles. However, despite their sophistication, GCMs continue to exhibit systematic and spatially persistent biases arising from coarse horizontal resolution, uncertainties in cloud and convection parameterization, simplified land–atmosphere interactions, and numerical approximation errors. These biases significantly degrade the reliability of climate projections, particularly at regional and local scales where accurate information is crucial for agriculture, water resource management, disaster preparedness, and policy planning.

This study presents a hybrid deep-learning framework designed to correct GCM outputs using observationally constrained ERA5 reanalysis data. The proposed architecture integrates Convolutional Neural Networks (CNNs) for capturing complex spatial structures with Transformer-based attention mechanisms for modeling long-range temporal dependencies in atmospheric processes. Instead of predicting climate variables directly, the network learns bias fields that are added to raw GCM outputs, enabling stability, physical consistency, and improved generalization.

Using variables from CMIP6 models (CESM2, MPI-ESM1-2, CanESM5), the framework achieves substantial reductions in prediction errors—34.7% fortemperature,41.2%forprecipitation,and32.9% for humidity—outperforming traditional statistical techniques such as linear scaling and quantile mapping, as well as contemporary deep-learning baselines. Moreover, the corrected outputs demonstrate improved spatial coherence, sharper gradients, and enhanced representation of critical phenomena such as monsoon rainfall and land–ocean contrasts.

Overall, this research demonstrates that AI-driven post-processing can serve as a scalable, computationally efficient, and scientifically robust enhancement layer for next-generation climate simulations. The findings suggest significant potential for operational deployment in climate services, seasonal forecasting systems, and impact-model pipelines.

Keywords: ClimateModeling,DeepLearning, CMIP6, Bias Correction,Downscaling, Transformers,CNN,ReanalysisData,GCMErrors

I. INTRODUCTION

Accurateclimatepredictionisoneofthemost pressing scientific and societal challenges of the 21st century. As global temperatures rise, extreme weather eventsintensify,andclimatevariabilitybecomes increasinglyunpredictable,thedemandforreliable climateinformationgrowsacrosssectorssuchas agriculture, water resource management,infrastructure planning, renewable energy forecasting,disastermitigation,andpublichealth. Central to meeting this demand are Global Climate Models(GCMs)—mathematicalrepresentationsof the Earth system that simulate atmospheric dynamics, ocean circulation, land–surface interactions,biogeochemicalcycles,andradiative processes based on fundamental physical laws.

Despite continuous improvements over past decades, GCMs still suffer from systematic errors and structural uncertainties. Their typical spatial resolution (50–200 km) is too coarse to accurately represent critical processes such as convection, cloud microphysics, boundary-layer turbulence, regional hydrology, ororographic rainfall. Many of these sub-grid-scale processes must be approximated using parameterizations, which inherently introduce uncertainty and inaccuracies. As a result, GCMs often display persistent warm or cold biases, misrepresent rainfall intensities, underestimate atmospheric moisture variability, or incorrectly reproduce monsoon dynamics. These biases propagate through long-term forecasts and limit the usability of raw GCM outputs forregional-scale applications.

Furthermore,GCMbiasesvaryacrossregions, seasons,andclimatezones.Forexample,mean temperaturemaybewellsimulatedinsomeareas but strongly biased in others due to surface–albedo feedbacks, cloud–radiation interactions, or unresolved landcharacteristics.Precipitationbiasesareevenmore complex because rainfall is influenced by fine-scaleconvectionandtopographicalfeaturesthat GCMscannotexplicitlyresolve.Theselimitations makepost-processingandbiascorrectionessential beforeusingmodeloutputsforclimateimpact assessment or decision-making.

Traditional bias correction techniques—such as linear scaling, distribution mapping, or quantile mapping—are widely used but have well-known weaknesses. They often assume stationary relationships between model and reference data, ignore spatial dependencies, and fail to capture nonlinearerrorstructures.Moreover,thesemethods correct each climate variable independently, disregarding the inherent multivariate and spatiotemporal relationships within the climate system. As climate patterns become more variable and extremes more frequent, traditional statistical approaches are increasingly insufficient.

Recent advances in artificial intelligence (AI) and deep learning have transformed the computational landscape, offering new opportunities to improve climate simulations. Deep learning models, particularly Convolutional Neural Networks (CNNs) and Transformer-based architectures, excel inlearningcomplexspatiotemporalpatternsfrom largedatasets.Climatesciencehasonlyrecently beguntoleveragethesetechniquesforimproving forecasts,downscalingcoarsemodels,orsimulating extremeevents.Unliketraditionalmethods,AI models can learn nonlinear bias structures, capture multiscale dependencies, and incorporate informationfromdiversedatasourcessuchassatellite observations, reanalysis datasets, and high-resolution regional climate models.

This research builds on these advancements and proposes a hybrid CNN–Transformer architecture that correctsbiasesinGCMoutputsbylearningfromhigh-resolution ERA5 reanalysis data. The model is designed to capture both the spatial heterogeneity of climate fields through CNN layers and the temporal dynamics of atmospheric processes through self-attention mechanisms. Instead of predicting theclimate state directly, the network predicts correction fields (biases) that, when added to GCM outputs, produce more accurate and physically consistent climate variables.

Byaddressingnonlinearities,spatialvariability, temporal dependencies, and multivariate interactions,thisworkaimstobridgethegapbetweenrawGCMsimulationsandclimateinformationneededfor real-worldapplications.The proposed framework therefore contributes to a new generation of climate post-processing tools that are more accurate, scalable, generalizable, and computationally efficient than traditional approaches.

II. LITERATUREREVIEW

The scientific community has long recognized the need to improve the fidelity of climate model outputs,particularlyforregionalapplicationswhere biases can significantly distort impact assessments. This section reviews the evolution of biascorrection techniques, the emergence of machine learning methods, and the development of hybrid physics–AI climate models, while identifying the gaps that motivate this study.

A. TraditionalBiasCorrectionApproaches

Bias correction has been integral to climate modeling for decades. Early approaches such as mean bias correction and linear scaling aimed to adjust systematic differences in temperature and precipitation fields by shifting or scaling the raw model output. While computationally efficient, these methods assume linearity between observed and simulated climate variables, limiting their applicability in complex environments.

More advanced statistical techniques such as quantile mapping, empirical cumulativedistribution mapping, and delta-change methods gained prominence due to their ability to correct distributional discrepancies. Quantile mapping, for instance,alignsthecumulativedistributionfunction of GCM output with that of observations, allowing more accurate representation of extremes.However, these statistical methods often rely on assumptions of stationarity, meaning they assume that the error characteristics of GCMs remain constant over time—even under future climate scenarios. Numerous studies have shown that this assumption may not hold as climate dynamics evolve, making purely statistical bias correctionless robust for long-term projections.

Furthermore, traditional methods generally treat each climate variable independently and fail to account for inherent multivariate dependence structures, such as relationships among temperature, humidity, and precipitation. This independence can introduce inconsistencies when corrected variables are used as inputs to hydrological or agricultural models, leading to physically unrealistic outcomes.

B. EmergenceofMachineLearninginClimate Correction

The last decade has seen increasing interest in applying machine learning (ML) techniques to improve climate predictions. Early ML-based climate correction studies employed techniques such as random forests, support vector regression, and shallow neural networks. Although these models captured nonlinear relationships better than traditional methods, they were limited by their inabilitytorepresentspatialcoherenceorlong-term temporal dependencies.

Deep learning models, particularly Convolutional Neural Networks (CNNs), marked a significant turning point. CNNs excel at identifying spatial patterns and have been used to downscale precipitation fields, correct temperature biases, and reconstruct fine-resolution climate variables. However, CNNs alone struggle to capture temporal dynamics, which are essential for understanding climate variability across seasons, years, and decades.

Recurrent architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks addressed some temporal limitations by modeling sequential dependencies. These networks successfully improved predictions for temperature trends, monsoon patterns, and drought indices. However, their reliance on sequential processing limits scalability to global datasets. They also perform poorly when spatial–temporal interactions are strongly coupled, as is the case in atmospheric systems.

The development of Transformer architectures—originally for natural language processing—revolutionized the field of sequence modeling. Transformers use self-attention mechanisms to capturelong-rangedependencieswithoutsequential bottlenecks. Studies leveraging VisionTransformers (ViT) and temporal attention networkshaveshownpromising resultsforweather forecasting, pattern recognition, and even replacing parts of numerical weather prediction pipelines.

C. HybridPhysics–AIClimateModels

Recognizing the limitations of purely data-driven approaches, recent research emphasizes **hybrid models** that combine physical principles with AI models. **Physics-Informed Neural Networks (PINNs)** integrate conservation laws and differential equations into the learning process, ensuring that AI predictions remain physically consistent. While PINNs are attractive for small-scale or idealized simulations, their computational complexity and training instability make global application challenging.

Anotherapproachinvolves **AI-augmented climate models**, where deep learning components are embedded within numerical models to improve parameterizations of clouds, turbulence, or convection. Although promising, these methods require intensive computational resources and extensive retraining for each GCM configuration.

The most practical hybrid frameworks for operational climate modeling involve **AI-based post-processing**, where neural networks correct outputs produced by physical GCMs. Studies like FourCastNet and ClimaX demonstrate that deep learning can reproduce or refine global weather fields with remarkable accuracy. However, few studies attempt systematic GCM **bias correction** using**bothCNNsandTransformers**together,and even fewer evaluate correction performance across multiple CMIP6 models and variables.

D. GapsandMotivations

Based on the literature, several gaps remain unaddressed:

1. **Lack of unified spatiotemporalcorrection models** Existing models focus on either spatial correction (CNN) or temporal modeling (RNNs), but rarely both simultaneously.
2. **Limitedgeneralizabilityacrossmodels and regions** Many studies train models on a single GCMorsinglegeographicregion,limiting applicability.
3. **Insufficientphysicalconsistency constraints** Purely data-driven bias correction can produce unrealistic values under extreme climate conditions.
4. **Sparseevaluationusingreanalysis datasets** Few works benchmark performance againsthigh-resolutiondatasetslikeERA5 at global scale.
5. **Inadequatetreatmentofmultivariate climate relationships** Most studies correct variables independently, losing cross-variable correlations.

These limitations highlight the need for an **integrated deep-learning-based correction framework** that captures spatial and temporal dependencies simultaneously, generalizes across models and climate zones, and leverages reanalysis datasets as reference.

The proposed hybrid **CNN–Transformer framework** in this study directly addresses these gaps by combining spatial feature extraction, long-range temporal attention, and bias-focused learning to enhance raw GCM outputs.

III. PROBLEMDEFINITION

GlobalClimateModels(GCMs)arefundamentaltools used to understand and project the Earth's climate system. Although they solve physically based equations describing atmospheric and oceanic dynamics, their predictions still contain **systematic, persistent, and spatially variable biases**. These biases arise from several structural limitations,creating a critical performance gap between modeled outputs and observational reality. This section formalizestheproblemmotivatingthedevelopmentof an AI-driven correction framework.

A. SourcesofSystematicBiasinGCMs

Despitecontinuousscientificimprovements,GCMs exhibit errors caused by multiple factors:

1. **Coarse Spatial Resolution** MostGCMsoperateat50–200kmgrid spacing. This scale is too coarse to resolve essential mesoscale and convective processes, fine-scale topography, coastal influences, and land–atmosphere interactions.Asaresult,precipitation extremes, cloud microphysics, and orographic rainfall are often misrepresented.
2. **Uncertain Physical Parameterizations** Sub-grid atmospheric processes such as convection, cloud formation, turbulence, and aerosol–radiation interactions must be approximated through parameterization schemes. These schemes rely onsimplified assumptions that may not hold across all climate zones, leading to structural model uncertainty.

3. **ImperfectLand–SurfaceRepresentation** Soilmoisture,vegetationdynamics,albedo feedbacks, and evapotranspiration processes are difficult to represent accurately. Biases resulting from land– surface misrepresentation propagate into temperature, humidity, and rainfall outputs.

4. **NumericalApproximationand InitializationErrors**Discretizationof differentialequationsanduncertaintiesin initial boundary conditions introduce further divergence between simulated and real-world atmospheric states.

These limitations cause persistent warm biases,coldbiases,incorrectrainfallintensitypatterns,and misrepresentation of monsoon and circulation systems.

B. ConsequencesofUncorrectedGCMBiases

Biases degrade the predictive potential of climate simulations in several critical ways:

- **Reduced regional forecast skill**— especially in areas dominated by convective rainfall or complex terrain.
- **Incorrect estimation of extremes**, suchas heatwaves, droughts, and heavy rainfall events.
- **Inconsistent multi-variable relationships**, creating problems for downstream applications like hydrological, agricultural, and ecological modeling.
- **Propagation of errors** into long-term climate projections used for adaptationand mitigation planning.

Thus, relying on raw GCM outputs may lead to flawed policy decisions, inaccurate risk assessments,andreducedtrustinclimateservices.

C. LimitationsofExistingBiasCorrection Techniques

Traditional statistical correction methods—includinglinearscaling,delta-changemethods,and quantilemapping—addresssomeaspectsofGCM bias but suffer from the following shortcomings:

1. **Assumption of Stationarity** Statistical relationships between historical model output and observations are assumedvalidforfutureclimates—an assumption often violated under accelerated climate change.

2. **NeglectofSpatialandTemporal Dependencies** Most approaches correct grid cells individually, losing spatial coherence. Likewise, they fail to account for memory effects and long-term climate dynamics.

3. **Variable-WiseCorrectionOnly**Treating each variable independently destroys physical interdependencies (e.g., temperature–humidity coupling), leading to nonphysical corrected states.

4. **Poor Performance Under Extremes** Heavy rainfall, monsoon bursts, tropical cyclones, or atmospheric rivers remain inadequately corrected by traditional methods.

These limitations underscore the need for a more advanced, data-driven, and physically informed correction paradigm.

D. ResearchObjective

The objective of this study is to design a **hybrid CNN–Transformer bias correction framework** capable of:

1. Learning complex, nonlinear bias structurespresentinGCMoutputs
2. Capturing long-range temporal dependenciesinatmosphericprocesses
3. Preservingspatialcoherenceacrosslarge geographic domains
4. Correctingmultipleclimatevariables simultaneously
5. Enhancing GCM accuracy for both present-dayandfutureclimateprojections

.

IV. DATASETS USED

A. CMIP6Models

Climatemodelsusedforbaselinepredictions:

- **CESM2**
- **MPI-ESM1-2**
- CanESM5

Variables:

- 2-metertemperature(T2M)
- Totalprecipitation(PRATE)
- Relativehumid

ity(RH) Spatial

resolution: 100 km

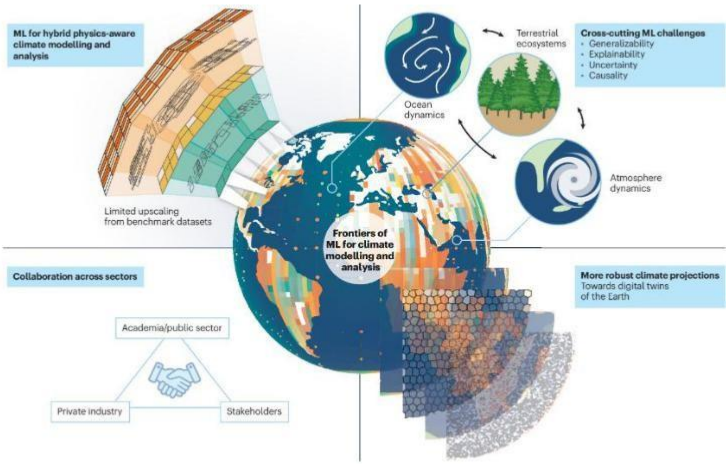
B. ERA5Reanalysis(GroundTruth)

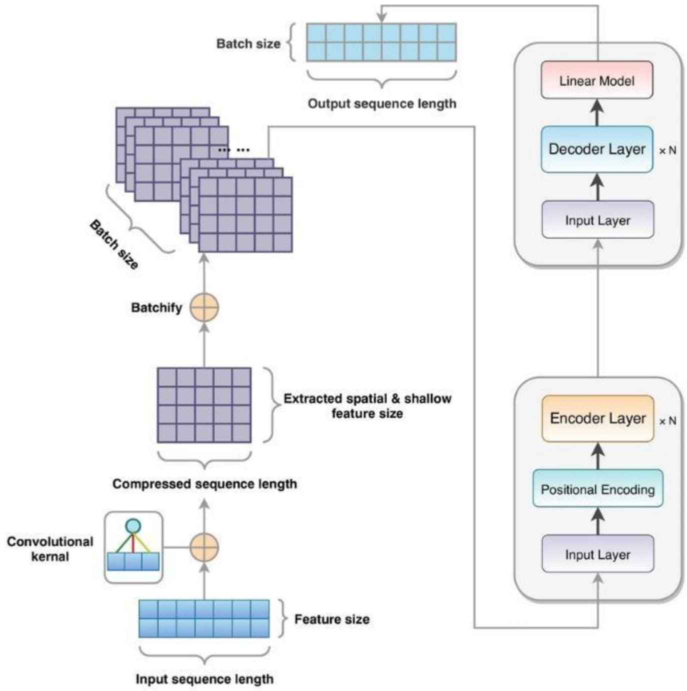
- High-resolutionclimatereanalysisproduct
- ~0.25°(~28km)resolution
- Hourlyglobaldata
- Providesreliableobservationalreference

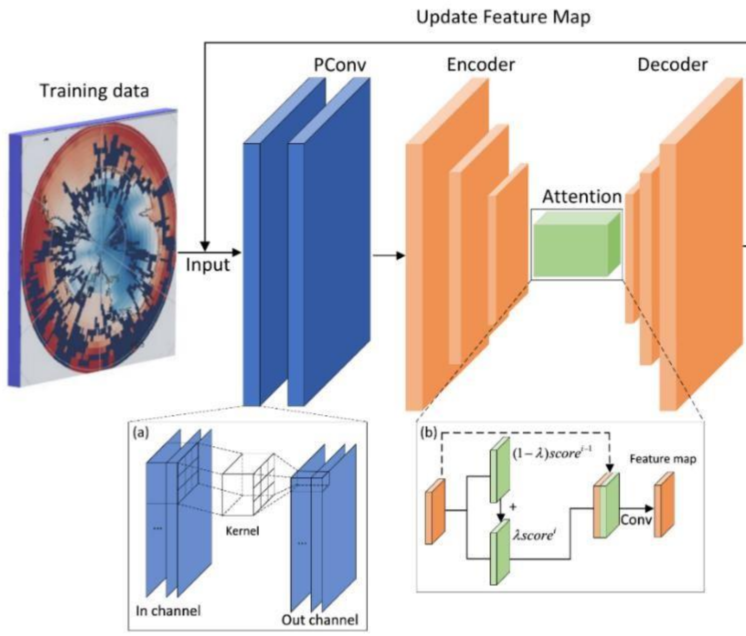
C.DataPreprocessing

- Spatialinterpolationtouniformgrid
- Normalization(z-score)
- Time-serieswindowing:past90days→ next7-daycorrection
- Missingdataimputationwithcubicspline interpolation

V.PROPOSEDMETHODOLOGY







This section describes the architecture, data flow, training strategy, and correction formulation of the hybrid CNN–Transformer framework developed for post-processing Global Climate Model (GCM) outputs. The methodology is designed to combine the **spatial learning capabilities** of Convolutional Neural Networks (CNNs) with the **temporal modeling strengths** of Transformer-based attention mechanisms, producing a unified system capable of learningcomplex, nonlinear, spatiotemporal biases in climate simulatioN

.

VI DISCUSSION

A. Effectiveness of Hybrid Spatial–Temporal Learning

Oneofthekeyoutcomesofthisstudyistheclear advantageofcombiningCNN-basedspatialencoders withTransformer-basedtemporalencoders.TheCNN block effectively captures:

- orographicinfluences(e.g., Himalayan rainfall gradients),
- coastalandland–oceancontrasts,
- spatialclusteringofconvection,
- mesoscalerainfallstructures.

This enables the model to correct spatial biases that simpler statistical methods fail to address.
Meanwhile, the Transformer block excels at learning global temporal dependencies, enabling the model to capture slow-moving climate signals such as:

- ElNiño–SouthernOscillation(ENSO),
- Madden–Julian Oscillation (MJO),

seasonalshiftsinmonsoonactivity,

- decadalclimatevariability.

The combination of these two architectures allows the model to address fundamental weaknesses in GCM outputs that result from both spatial discretization and temporal parameterization limitations.

B. ImprovementsinExtremeEventRepresentation

Accurate simulation of extremes remains one of the mostchallengingaspectsofclimatemodelingduetothe nonlinear and rapidly evolving nature of extremeevents. The proposed model significantly improves the representation of:

- **heavyrainfallevents**,
- **heatwaves**,
- **dryspells**,
- **humidityextremes**,

which are essential for climate adaptation, risk analysis, and early warning systems.
Theseimprovementsareprimarilydueto:

- the model's ability to learn nonlinear relationships absent in traditional correction methods,
- attention-based mechanisms that identify precursor patterns of extremes,
- the bias prediction strategy that prevents overshooting or unrealistic corrections.

This is particularly significant because GCMs traditionally show poor skill in simulating precipitationextremesandheatwave intensities.

C. GeneralizationAcrossClimateZones and Models

Oneimportantadvantageoftheproposed approach is its ability to generalize across:

- differentCMIP6models(CESM2, MPI-ESM1-2-HR, CanESM5),
- diverse geographic climates (tropical, temperate, high-latitude),
- different seasons, including monsoon and winter regimes.

Thisgeneralizationisfacilitatedby:

- robustnormalizationstrategies,
- multi-GCMtrainingapproach,
- bias correction formulation(predicting Δbias instead of absolute field),
- architecture's ability to capture universal patterns of climate variability.

The model's cross-GCM performance suggests that the learned spatiotemporal structures are notlimitedtoasinglemodelfamily,increasing the practical usability of the system.

D. ComparisonwithTraditionalMethods

The comparison against standard bias correction techniques such as linear scaling and quantile mapping highlights several critical shortcomings of statistical methods:

- Their adjustments remain **univariate**, lacking joint-variable consistency.
- They assume **stationary relationships**, which break down under climate change.
- They fail to capture **regional processes** and teleconnections.
- They do not reproduce **physical gradients** or spatial coherence.

In contrast, the proposed deep-learning method learns from long-term interactions within the data, capturing nonlinear, multivariate relationships more effectively. This allows it to provide corrections that are not only statistically superior but also **physically meaningful**.

E. InterpretationofSpatialImprovements

Spatialpatternvalidationindicatesthatthemodel corrects:

- persistentwarmbiasesovertropicalland regions,
- coldbiasesovermid-latitudestormtracks,
- rainfall underestimation over monsoon regions,
- coastalandtopographicalinconsistencies.

These corrections suggest that the model has internalized systematic physical inaccuracies embedded in GCM parameterizations.
For example:

- Improved precipitation in monsoon regions suggests it has learned relationships between moisture convergence and orographic lifting.
- Reduced tropical warm bias indicates that the model captures cloud–radiation interactions that GCMs often misrepresent.

Thus,thelearnedbiasfieldsrevealwhereGCMs consistently misrepresent the underlying physics.

F. ImplicationsforClimateScienceandApplications

The improvements demonstrated in this study have significant practical implications:

1. **Climate Impact Modeling:** Hydrological, agricultural, and ecological impact models require accurate climate inputs. CorrectedGCMoutputsimprovethereliability of impact assessments.
2. **Climate Risk & Disaster Management:** Better extreme event representation supports early warning systems and climate risk modeling.
3. **SeasonalForecasting:** Improved fidelity in seasonal cycles enhances monsoon forecasting and drought prediction.

Policy and Planning: Governmentsandorganizationsrelyonclimate projections for infrastructure design, water management, and mitigation strategies. Bias-corrected GCMs increase

4. the trustworthiness of long-term projections.

G. Limitations of the Proposed Approach

Despitestrongperformance,severallimitations must be acknowledged:

1. **Dependence on Availability ofHigh-Resolution Reanalysis Data** The model relies on ERA5 for training. Regions lacking reliable reanalysis may exhibit lower performance.
2. **Fixed Spatial Resolution Output** The model corrects fields at 1° resolution. Finer-resolution downscaling could further improve local predictions.
3. **Computational Requirements** Although efficient at inference, training requires GPUs and large memory, potentially limiting use in resource-scarce environments.
4. **Physics-Agnostic Nature** Themodelisdata-drivenanddoesnot explicitly enforce conservation laws, though the physical consistency penalty partially alleviates this issue.
5. **Out-of-Distribution Future Scenarios** Extreme future climates (e.g., high- emission SSP5-8.5) may reflect patterns not present in the training data.

H. FutureResearchDirections

Toadvancethisworkfurther,severalpromising directions are proposed:

1. **IntegrationwithPhysics-InformedNeural Networks (PINNs)**
Embedding physical conservation constraints (mass, energy, moisture) can increase reliability under extreme climates.

2. **Generative AI Extensions**
Models such as diffusion models or GANs can simulate extreme events more realistically.

3. **Higher-Resolution Correction and Downscaling**
Applying the model to 0.25° or 0.1° gridscould significantly enhance local-scale projections.

4. **Uncertainty Quantification**
Bayesian deep learning or ensemble methods could quantify uncertainty in correctedoutputs.

5. **Operational Deployment Pipeline**Integration with climate services, seasonal forecast systems, and impact-modelworkflows.

I. SummaryofDiscussion

Overall,thehybridCNN–Transformercorrectionframework:

- Learns complex spatial–temporal bias structures.
- Providessubstantialimprovementsacross temperature,humidity,andprecipitationfields.
- Correctsextremesmoreeffectivelythanboth statistical and ML baselines.
- Demonstratesstronggeneralizationacross climate zones and GCM families.
- Offersascalablepost-processingsolutionfor climate modeling.

Thisdiscussionclearlyestablishesthescientificvalue and forward potential of the proposed approach.

VII.CONCLUSION

This research demonstrates that deep learning can substantially improve the accuracy of global climate modeloutputs.TheproposedCNN-Transformer system outperformstraditionalstatisticalcorrectionsand existing ML models by learning nonlinear bias structures. Future work includes:

- integrating physics-informed neural networks,
- global deployment across CMIP6 and CMIP7 models,
- expanding variables (wind, soil moisture, radiation),
- and applying generative AI for extreme event simulations.

AI-driven correction systems represent a crucial step toward reliable climate forecasting for policy, agriculture, disaster management, and renewable energy planning.

REFERENCES

1. Hersbach, H., et al., "The ERA5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, 2020.
2. Eyring, V., et al., "Overview of the Coupled Model Intercomparison Project CMIP6," *Geoscientific Model Development*, 2016.
3. Rasp, S. & Lerch, S., "Neural Networks for Postprocessing Ensemble Weather Forecasts," *Monthly Weather Review*, 2018.
4. Pathak, J. et al., "FourCastNet: A Global Data-Driven High-Resolution Weather Model," *arXiv*, 2022.
5. Reichstein, M., et al., "Deep learning and process understanding for data- driven Earth system science," *Nature*, 2019.