

Research paper

Forecasting near-surface air temperature via SARIMA and LSTM: A regional time-series study



Muhammed M. Aksoy ^a, Md. Najmul Mowla ^b, Mehmet Bilgili ^c *, Engin Pinar ^c, Tahir Durhasan ^b, Davood Asadi ^b

^a Department of Energy Systems Engineering, Osmaniye Korkut Ata University, 80000, Osmaniye, Turkey

^b Department of Aerospace Engineering, Adana Alparslan Türkeş Science and Technology University, Sarıçam, 01250 Adana, Turkey

^c Department of Mechanical Engineering, Ceyhan Engineering Faculty, Çukurova University, Ceyhan, 01950, Adana, Turkey

ARTICLE INFO

Keywords:

Air temperature forecasting
SARIMA
LSTM
Climate change projection

ABSTRACT

Accurate modeling of near-surface air temperature (AT) trends is critical for assessing global and regional climate risks, particularly in light of the intensifying warming signals observed across the northern hemisphere and the tropics. This study proposes a robust and computationally efficient framework for forecasting near-surface AT across the global, the northern hemisphere, the southern hemisphere, and the tropics using two complementary time-series modeling techniques: seasonal autoregressive integrated moving average (SARIMA) and long short-term memory (LSTM) networks. The models are trained to capture both structured seasonal patterns and nonlinear temporal dynamics by leveraging the ERA5 reanalysis dataset (1970–2024) and incorporating preprocessing steps such as detrending and Z-score normalization. SARIMA consistently outperformed LSTM across most domains, particularly in the global region, achieving lower RMSE (0.0967 °C) and higher correlation ($R = 0.9975$), reflecting its superior capacity for linear and seasonal signal extraction. Quantitatively, SARIMA demonstrates 5%–10% lower RMSE and slightly higher correlation than LSTM across all domains, underscoring the statistical significance of its performance advantage. Projected near-surface AT anomalies by 2050 reveal a marked warming trend, with the SARIMA model estimating a global anomaly of +1.078 °C and a northern hemisphere anomaly of +1.474 °C, closely aligning with IPCC-reported trajectories and exceeding CMIP5 RCP4.5 projections. The findings underscore SARIMA's reliability for short- to mid-term near-surface AT forecasting and LSTM's potential for future hybrid modeling schemes. This work fills a critical methodological gap by integrating statistical rigor with scalable deep learning, offering enhanced fidelity for regional climate adaptation planning.

1. Introduction

Global climate change is among the most pressing scientific, environmental, and socioeconomic challenges of the 21st century (Baris-Tuzemen and Lyhagen, 2024; Varotsos et al., 2021). Driven by anthropogenic greenhouse gas (GHG) emissions, land-use changes, and natural variability, the Earth's climate system has undergone unprecedented transformations, especially since the mid-20th century (Infante-Amate and Aguilera, 2024; Avşar and Mowla, 2022). According to the World Meteorological Organization (WMO), the global average air temperature in 2023 reached approximately 1.45 ± 0.12 °C above pre-industrial levels, with disproportionate warming in the northern hemisphere and tropics, which are heating nearly four times faster than the global mean (Golitsyn and Varotsos, 2025; Jones, 2024). This accelerated warming intensifies feedback mechanisms such as albedo

loss, permafrost thaw, and ecosystem destabilization, all exacerbating the climate crisis (Wang et al., 2023; Varotsos and Efthathiou, 2015).

Among key indicators of climate change, near-surface air temperature (AT) is particularly critical in the polar and sub-polar regions, especially the tropics (Xu et al., 2023; Varotsos et al., 2013). Accurate projections of near-surface AT trends are essential to understanding complex climate feedback loops and enhancing the robustness of early warning systems, seasonal forecasting, and adaptation strategies in agriculture, energy, fisheries, water security, and infrastructure planning (Akperov et al., 2025).

The warming of the Earth's climate system is unequivocal (Savaş, 2025). Changes since the 1950s are unprecedented in the instrumental record and likely unmatched over millennia (Udy et al., 2024). This warming is driven by a growing radiative forcing imbalance caused by

* Corresponding author.

E-mail address: mbilgili@cu.edu.tr (M. Bilgili).

increasing atmospheric CO₂, CH₄, N₂O, and halogenated gases, which account for over 96% of total radiative forcing since 1750 (Bruhwiler et al., 2021). The effective radiative forcing (ERF) from well-mixed greenhouse gases reached 3.48 W m⁻² by 2023, of which CO₂ alone contributed 2.28 W m⁻² (Forster et al., 2024). These changes have resulted in the warming of the atmosphere and oceans, rising sea levels, and significant reductions in snow and ice cover (Ding et al., 2021; Chattopadhyay and Chattopadhyay, 2008).

The Earth's energy imbalance (EEI), the difference between absorbed solar radiation and outgoing longwave radiation, has increased markedly over time (Loeb et al., 2024; Li et al., 2024b). From an average of 0.65 W m⁻² during 1976–2023, the imbalance has risen to 0.96 W m⁻² in 2011–2023, reflecting heightened planetary heat retention, with ocean heat content accounting for over 90% of this excess energy (Von Schuckmann et al., 2023). These shifts have amplified the warming of surface air and ocean layers and intensified cryospheric loss (Cai et al., 2023).

Despite the objectives of the Paris Agreement to limit warming to well below 2 °C, preferably 1.5 °C, current trends are alarming (Meinshausen et al., 2022). The WMO's 2024 greenhouse gas bulletin reported record-breaking atmospheric concentrations of CO₂ (420 ppm), CH₄ (1,934 ppb), and N₂O (336.9 ppb), which are 151%, 265%, and 125% higher, respectively, than pre-industrial levels (Li et al., 2024a; World Meteorological Organization, 2024). The 28th UN climate conference (COP28) emphasized that these concentrations will continue to grow without immediate and aggressive reductions in GHG emissions, weakening the natural carbon sinks in terrestrial ecosystems and oceans and further accelerating global warming (World Meteorological Organization, 2023). The impacts of climate change are spatially heterogeneous and especially severe in high-latitude regions (Zhang et al., 2025). In the northern hemisphere, warming reshapes agricultural zones, melts glaciers, and alters water availability (Fallah et al., 2024).

While earth system models (ESMs) and coupled model intercomparison projects (CMIP5/CMIP6) provide valuable insights into long-term climate projections (Hamed et al., 2022; Hu et al., 2025; Winkler and Sierra, 2025), they often suffer from coarse resolution (Elshall et al., 2022; Fletcher et al., 2022), computational overhead (Acosta et al., 2023; Counillon et al., 2023), and limited capability to capture localized (Decharme et al., 2025), nonlinear (DelSole and Tippett, 2024), and abrupt changes (Matthes et al., 2025), particularly in tropics regions. Furthermore, physical models alone are often insufficient for real-time adaptation planning, especially when the climate system exhibits nonstationary behavior, abrupt feedback, and region-specific variability (Saleh et al., 2024). Thus, there is a growing recognition of the need for machine learning (ML)-based, computationally efficient hybrid models that can assimilate large-scale observational and reanalysis datasets (e.g., ERA5) and produce accurate projections of near-surface AT across various spatial scales, global, northern hemisphere, southern hemisphere, and tropic (Gardner and Dorling, 1998).

Motivated by the limitations of existing physical models and the need for real-time, scalable, and explainable predictive systems, this study proposes a robust deep learning (DL) framework for projecting near-surface AT under different shared socioeconomic pathways (SSPs). In contrast to traditional models, the proposed approach leverages spatiotemporal feature extraction capabilities of modern neural architectures to improve the fidelity of regional projections while reducing computational costs.

To address these challenges, we propose a comparative modeling framework that evaluates two fundamentally distinct time-series approaches, such as seasonal autoregressive integrated moving average (SARIMA) and long short-term memory networks (LSTM). SARIMA is a classical statistical method that excels at capturing linear dependencies and seasonal cycles in stationary climate series, offering transparent interpretability and efficient parameterization (Önder, 2024). In contrast, LSTM networks, a subset of recurrent neural networks (RNNs), are

designed to learn complex, nonlinear, and long-range temporal dependencies, making them suitable for nonstationary and highly dynamic time series (Mowla et al., 2025). By juxtaposing these models, we aim to assess whether traditional statistical approaches remain competitive in climate forecasting or whether modern DL architectures offer tangible advantages. This dual-model perspective highlights the trade-offs between interpretability, accuracy, and computational efficiency in the context of regional climate variability.

Recent studies have advanced the application of both statistical and ML approaches for climate and meteorological forecasting. For instance, using reanalysis data, Varotsos and Mazei (2019) analyzed hemispheric and global near-surface AT variability, highlighting the persistent long-term warming trend and regional disparities. Stachura et al. (2024) demonstrated that modern ML models can effectively capture nonlinear climate dynamics, yet emphasized model interpretability and computational cost challenges. Adeniran et al. (2024) applied SARIMA to regional meteorological forecasting, confirming its strength in capturing seasonality and linear trends. In contrast, Uluocak and Bilgili (2024) and Devi et al. (2023) explored LSTM and hybrid architectures for temperature prediction, reporting improved accuracy in nonlinear regimes but sensitivity to training data length and preprocessing. Comparative works such as Ray and Chattopadhyay (2021) underscored the trade-offs between statistical transparency and DL adaptability, while Durhasan et al. (2025) and Tran et al. (2021) stressed the importance of integrating large-scale climate drivers (e.g., ENSO, volcanic aerosols) into predictive frameworks. He et al. (2022) further emphasized that benchmarking against established climate scenarios (e.g., CMIP5/CMIP6) is critical for ensuring physical plausibility.

Despite these advances, few studies have conducted a coordinated, multi-domain evaluation of SARIMA and LSTM over multi-decadal periods, explicitly incorporating trend-reintroduction techniques to preserve physical interpretability, and benchmarking results against physically based scenarios. This study addresses this gap by leveraging ERA5 reanalysis data (1970–2024) across four spatial domains: the globe, the northern hemisphere, the southern hemisphere, and the tropics to compare the predictive skill, strengths, and weaknesses of SARIMA and LSTM. The novelty lies in integrating statistical and deep learning approaches within a unified framework, applying physically informed preprocessing, and validating projections against CMIP5 RCP4.5, offering methodological and interpretive advancements over existing literature.

Key contributions of this study include:

- Development of a unified framework for predicting near-surface AT across global, northern hemisphere, southern hemisphere, and tropics domains under multiple future scenarios.
- Integration of DL pipelines and statistical models with climate reanalysis datasets for fine-resolution projections.
- Evaluation of model robustness and accuracy against physical benchmarks and observed data for present and future periods.
- Detailed benchmarking of SARIMA and LSTM in reproducing historical temperature variability and projecting future trends across multiple spatial domains, validated against CMIP5 RCP4.5 scenarios.

This work addresses a critical gap in climate projection methodologies by fusing data-driven intelligence with physical interpretability, offering a scalable alternative to traditional ESMs for regional adaptation and risk assessment.

2. Materials and method

This section introduces the proposed SARIMA and LSTM models for accurately forecasting near-surface AT, accompanied by a comprehensive study area and dataset description. Additionally, the evaluation metrics used to assess the performance of the models are discussed.

Table 1
Hyperparameter optimization of SARIMA models for regional variability.

Hyperparameter	Global	Northern Hemisphere	Southern Hemisphere	Tropics
AR order (p)	1	1	1	1
Differencing order (d)	1	1	1	1
MA order (q)	1	1	1	0
Seasonal AR order (P)	1	1	0	1
Seasonal Differencing order (D)	1	1	1	1
Seasonal MA order (Q)	1	1	1	1
Seasonal period (s)	12	12	12	12
Constant	-5.7493e-05	0	0	0
AR(1) coefficient	0.93117	0.89195	0.89493	0.94571
SAR(12) coefficient	-0.1388	-0.15041	-	-0.15307
MA(1) coefficient	-0.39548	-0.45084	-0.38197	-
SMA(12) coefficient	-0.85253	-0.84793	-0.82469	-0.86115
Variance	0.010167	0.026422	0.014672	0.007384
AIC	-1139.9614	-513.5577	-902.6451	-1354.4335
BIC	-1113.2586	-491.3054	-884.7681	-1336.6316

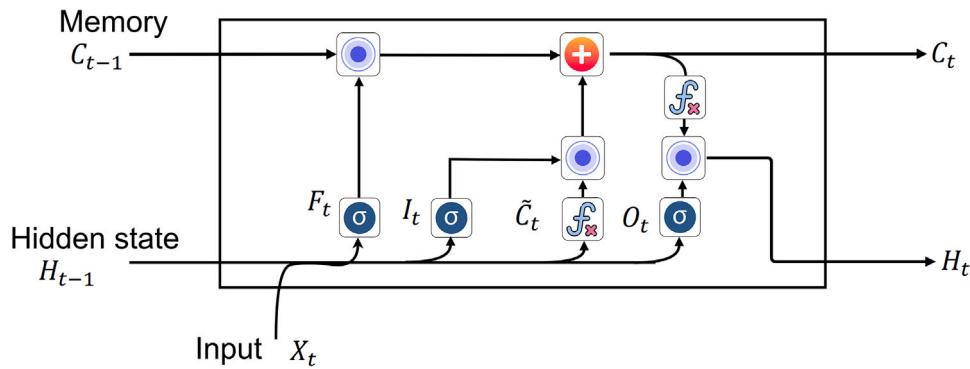


Fig. 1. LSTM cell architecture for near-surface AT prediction. The model leverages input (X_t) and previous hidden state (H_{t-1}) to update the cell state (C_t) and hidden output (H_t) through forget (F_t), input (I_t), and output (O_t) gates.

2.1. SARIMA model

The SARIMA model is an extension of the ARIMA model designed to handle seasonal data (Dubey et al., 2021). The general form of a SARIMA(p, d, q)(P, D, Q)_s model is expressed as:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^Dy_t = \Theta_Q(B^s)\theta_q(B)\epsilon_t, \quad (1)$$

where B is the backward shift operator such that $By_t = y_{t-1}$. The parameters p, d , and q represent the non-seasonal autoregressive, differencing, and moving average orders, respectively, while P, D , and Q denote their seasonal counterparts. The seasonal period is represented by s . The polynomials $\phi_p(B)$ and $\theta_q(B)$ are of order p and q , respectively, and $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are seasonal polynomials. The error term ϵ_t is assumed to be white noise.

2.1.1. Hyperparameter in SARIMA

Hyperparameter selection is crucial to capture temporal dependencies while mitigating overfitting (Hosseini et al., 2024). The model parameters are estimated using Maximum Likelihood Estimation (MLE) by optimizing the log-likelihood function (Boyko and O'Meara, 2024).

$$\mathcal{L}(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^n (y_t - \hat{y}_t)^2, \quad (2)$$

where θ represents the set of model parameters.

The selected SARIMA configurations for different regions are shown in Table 1. Explanations for each model's hyperparameters, coefficients, and selection criteria have been provided to ensure methodological clarity and scientific rigor.

The Global model features $p = 1, q = 1, P = 1$, and $Q = 1$, reflecting strong seasonal patterns. The Tropics model, with $p = 1, q = 0, P = 1$, and $Q = 1$, exhibited the highest AR(1) coefficient (0.94571), indicating persistent temporal correlation. The Northern Hemisphere

model includes $p = 1, q = 1, P = 1$, and $Q = 1$, with the highest MA(1) coefficient magnitude (-0.45084). Model performance metrics, including AIC and BIC, were used for selection, with the Tropics model showing the best fit.

2.2. LSTM model

LSTM networks, as shown in Fig. 1, constitute an advanced class of RNNs designed to address long-range dependencies in sequential data (Wang et al., 2024). Traditional RNNs suffer from vanishing and exploding gradient problems when processing extended sequences, leading to difficulties in capturing long-term dependencies (Waqas and Humphries, 2024; Mowla et al., 2025). LSTMs addressed these challenges through a specialized gating mechanism that regulates memory over time (Mowla et al., 2025).

The LSTM architecture is structured around a fundamental computational unit known as the LSTM cell, which integrates three primary gates: the input gate, the forget gate, and the output gate (Uluocak and Bilgili, 2024). The forget gate regulates retaining or discarding previous cell state information, effectively removing irrelevant or redundant details (Rahman et al., 2024; Mowla et al., 2025). Mathematically, the forget gate activation at time step t is computed as:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

where f_t represents the forget gate output, σ denotes the sigmoid activation function, W_f is the weight matrix, h_{t-1} corresponds to the hidden state from the previous time step, x_t is the current input, and b_f is the bias term.

The input gate modulates the incorporation of new information into the cell state, enabling the model to update its memory representation based on incoming input adaptively. It is

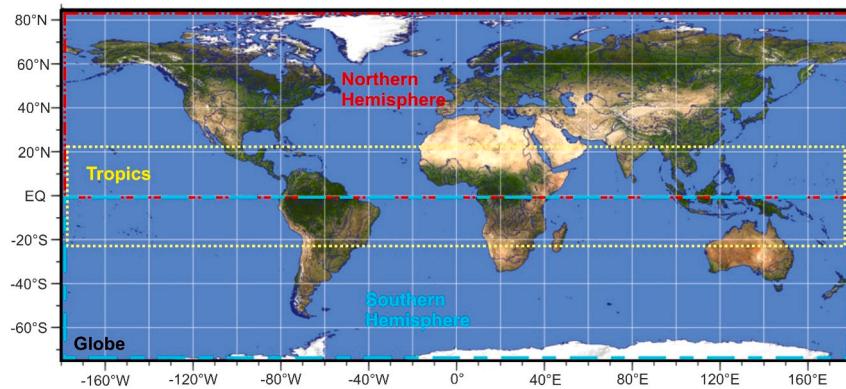


Fig. 2. Geographical boundaries of the study regions: the Globe (90° S– 90° N), Northern Hemisphere (0° – 90° N), Southern Hemisphere (0° – 90° S), and Tropics (23.5° N– 23.5° S).

Table 2
LSTM hyperparameters used in this study.

Hyperparameter	Value
Number of Hidden Units	50
Learning Rate	0.005 (piecewise drop $\times 0.2$ every 125 epochs)
Batch Size	Default (sequence-to-sequence)
Epochs	500
Gradient Threshold	1.0
Optimizer	Adam
Learn Rate Schedule	Piecewise

defined as:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (4)$$

where i_t denotes the input gate activation, and W_i , b_i represent the corresponding weight and bias parameters. The updated cell state is then computed as:

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (5)$$

C_t denotes the updated cell state, C_{t-1} is the previous cell state, and \tilde{C}_t represents the candidate cell state computed via a tanh activation function.

The output gate governs information propagation from the cell to the next hidden state, ensuring that only relevant features contribute to subsequent computations (Khala et al., 2025). It is expressed as:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

The hidden state h_t is then updated as:

$$h_t = o_t \tanh(C_t) \quad (7)$$

where o_t denotes the output gate activation, and W_o , b_o correspond to the associated parameters.

2.2.1. Hyperparameters in LSTM networks

The performance of LSTM networks is influenced by several hyperparameters that require careful selection to ensure model stability and generalization, as summarized in Table 2, which outlines key hyperparameters and their typical parameter ranges.

The number of hidden units determines the network's capacity to learn complex patterns (Kenne et al., 2024), while the learning rate η controls the step size in parameter updates, affecting convergence speed and stability (Wang et al., 2024). In standard gradient-based optimization, the parameter update at iteration t follows:

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t) \quad (8)$$

where θ_t represents the model parameters, and $\nabla L(\theta_t)$ is the gradient of the loss function.

The batch size impacts computational efficiency and gradient stability (Sekertekin et al., 2021), whereas the number of epochs dictates the extent of model training, requiring a balance to prevent underfitting or overfitting (Cho et al., 2022). Gradient clipping is applied to stabilize training by constraining the gradient norm:

$$\nabla L(\theta) = \frac{\nabla L(\theta)}{\max(1, \frac{\|\nabla L(\theta)\|}{\tau})} \quad (9)$$

where τ is the clipping threshold. Dropout regularization enhances generalization by randomly deactivating a fraction p of input units during training, represented as:

$$h_i = z_i \tilde{h}_i, \quad z_i \sim \text{Bernoulli}(1 - p) \quad (10)$$

where h_i is the output of a neuron, z_i is a binary mask sampled from a Bernoulli distribution, and p is the dropout rate.

The LSTM model was implemented in MATLAB (DL Toolbox) and trained with the Adam optimizer (Mowla et al., 2025). An 80% training and 20% testing split was used. Training data were standardized to zero mean and unit variance using training-set statistics (μ, σ) , which were applied to the test set. The network comprised a single LSTM layer with 50 hidden units, followed by a fully connected layer producing a scalar output and a regression loss. Training used 500 epochs, initial learning rate 0.005 with a piecewise drop factor of 0.2 every 125 epochs, and gradient threshold 1. Evaluation included (i) free-running multi-step forecasting via recursive prediction and (ii) rolling one-step forecasting with state updates. In both cases, outputs were de-standardized, and RMSE (shown in Eq. (20)) was computed against held-out targets.

2.3. Geographical coverage and data sources

This study investigates the historical variability and projected near-surface AT trends across four distinct geographical domains: the Globe, the Northern Hemisphere, the Southern Hemisphere, and the Tropics. As illustrated in Fig. 2, the spatial extents of these regions are defined as follows: the globe spans from 90° S to 90° N; the northern hemisphere, from 0° to 90° N; the southern hemisphere, from 0° to 90° S; and the tropics, from 23.5° N to 23.5° S.

Monthly near-surface (2-meter) AT data were acquired from the *Climate Reanalyzer* platform, a publicly accessible and widely recognized resource developed by the Climate Change Institute at the University of Maine (Climate Change Institute, University of Maine, 2024). These data were accessed on January 5, 2025, via <https://climatereanalyzer.org>. The platform is frequently employed in climate-related studies due to its robust data visualization tools and utility in scenarios with limited direct observational records.

The AT data used in this study are derived from the ERA5 reanalysis dataset, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 provides high-resolution global climate data

Table 3

Statistical characteristics of monthly near-surface AT data (°C).

Statistics	Globe	Northern Hemisphere	Southern Hemisphere	Tropics
Minimum	11.5720	7.3480	10.0520	23.8690
Maximum	16.9530	22.3120	16.7820	26.8030
Mean	14.2257	15.0808	13.4700	25.2187
Median	14.2405	15.4410	13.4575	25.1655
Standard Deviation	1.4247	4.6215	1.9302	0.4908
Skewness	-0.0271	-0.0436	0.0460	0.0281
Kurtosis	1.6465	1.5265	1.5647	2.8603

at a spatial resolution of $0.5^\circ \times 0.5^\circ$, supporting detailed assessments at both global and regional scales. The dataset spans from January 1970 to November 2024, offering a consistent and comprehensive temporal record suitable for rigorous trend analysis and climate modeling.

The statistical summary presented in Table 3 outlines the fundamental characteristics of the monthly near-surface AT data used for modeling global, northern hemisphere, southern hemisphere, and tropical projections. This dataset, stratified across four climatologically distinct regions, global, northern hemisphere, southern hemisphere, and tropics, captures central tendencies and higher-order statistical moments necessary for constructing data-driven, generalizable prediction models under diverse climatic regimes.

Globally, the near-surface AT data show a mean value of 14.23 °C with a relatively narrow standard deviation of 1.42 °C, indicative of moderate temporal variability on a planetary scale. This stability, however, masks the spatial heterogeneity apparent at regional levels. The Northern Hemisphere, a key region for tropics-focused modeling, exhibits a broader temperature range (7.35 °C to 22.31 °C) and significantly higher variability (standard deviation: 4.62 °C). The slight negative skewness (-0.0436) and kurtosis of 1.53 suggest a near symmetric, platykurtic distribution, with reduced probability mass in the tails, indicating fewer extreme temperature occurrences relative to a Gaussian distribution.

The Southern Hemisphere demonstrates a lower mean AT of 13.47 °C and reduced spread (standard deviation: 1.93 °C), consistent with the dominance of oceanic thermal regulation and lower seasonal amplitude in this region.

In contrast, the Tropics exhibit a tightly clustered distribution, characterized by the highest mean (25.22 °C), minimal variability (standard deviation: 0.49 °C), and near-Gaussian features, reflected in skewness of 0.0281 and kurtosis approaching 3 (2.86).

2.4. Data preprocessing

To enhance the accuracy of our forecasting models, we implemented a comprehensive data preprocessing framework specifically tailored for time-series analysis.

2.4.1. Data partitioning

The dataset was divided into training and testing subsets to facilitate model evaluation. Specifically, 80% of the data was allocated for training, with the remaining 20% reserved for testing. Let N denote the total number of observations; the number of training samples, N_{train} , is calculated as:

$$N_{\text{train}} = \lfloor 0.80 \times N \rfloor \quad (11)$$

The training and testing datasets are defined as:

$$\begin{aligned} X_{\text{train}} &= \{X_1, X_2, \dots, X_{N_{\text{train}}}\} \\ X_{\text{test}} &= \{X_{N_{\text{train}}+1}, X_{N_{\text{train}}+2}, \dots, X_N\} \end{aligned} \quad (12)$$

where X_t represents the near-surface AT measurement at time step t .

2.4.2. Data normalization

Z-score normalization was applied to reduce scale effects and improve numerical stability. The mean (μ) and sample standard deviation (σ) were computed from the training segment only:

$$\begin{aligned} \mu &= \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} X_i, \\ \sigma &= \sqrt{\frac{1}{N_{\text{train}} - 1} \sum_{i=1}^{N_{\text{train}}} (X_i - \mu)^2}. \end{aligned} \quad (13)$$

The same parameters were then applied to standardize both inputs and targets in the training and test sets:

$$\begin{aligned} X_{\text{train},\text{norm}} &= \frac{X_{\text{train}} - \mu}{\sigma}, \\ X_{\text{test},\text{norm}} &= \frac{X_{\text{test}} - \mu}{\sigma}. \end{aligned} \quad (14)$$

This approach ensures zero mean and unit variance for the training set while preventing data leakage into the test set.

2.4.3. Detrending for time-series modeling

The dataset was detrended before training to enhance the effectiveness of time-series forecasting models. This step removed long-term trends, allowing the models to focus on fluctuations around a stable mean (Allan et al., 2022). The dataset was detrended before training to enhance the effectiveness of time-series forecasting models. A centered simple moving average (SMA) filter was selected due to its transparency, low computational cost, and ability to isolate short- to mid-term variability while preserving the dominant long-term climate signal. The detrending was applied only during the training phase; after forecasting, the removed trend was reintroduced to the predicted residuals (Eq. (17)) to reconstruct the full series. This ensures that essential low-frequency climate signals, including anthropogenic warming trends, remain intact.

Trend Estimation: The trend component, T_t , was estimated using a moving average filter:

$$T_t = \frac{1}{m} \sum_{i=t-m/2}^{t+m/2} X_i \quad (15)$$

where m represents the window size.

Detrending Transformation: The detrended series, X_{det} , was obtained as:

$$X_{\text{det}} = X_t - T_t \quad (16)$$

Reconstruction after Forecasting:

After obtaining the forecasted values, the trend component was reintroduced:

$$\hat{X}_t = \hat{X}_{\text{det}} + T_t \quad (17)$$

This process ensured that the final forecasts accurately reflected both short-term variations and long-term trends in the near-surface AT.

This process ensured that the final forecasts accurately reflected short-term variations and long-term trends in near-surface AT. Physically, the reintroduced trend represents the persistent, large-scale warming signal detected in the ERA5 reanalysis dataset (Hersbach et al., 2020; Soci et al., 2024), which is driven by a combination of

anthropogenic greenhouse gas emissions, aerosol forcing, and natural low-frequency variability (e.g., Pacific Decadal Oscillation, Atlantic Multidecadal Oscillation). This approach is consistent with established climate time-series decomposition practices, where the trend component captures externally forced climate change while the residual represents internal variability (Von Storch and Zwiers, 2002; Yao et al., 2025). By restoring this trend after forecasting the detrended residuals, we retain the full climatic context, preserving the long-term radiative imbalance and ocean heat uptake influences (Forster et al., 2024) while allowing the models to focus on predictable oscillatory and interannual variability components during training. This ensures that the final projections remain physically consistent with the observed climate system and relevant for long-term adaptation planning (Forster et al., 2024; Masson-Delmotte et al., 2021).

LSTM model: Sliding window sequence generation. For the LSTM model, the time series was transformed into input–output pairs using a sliding-window approach, enabling the network to learn temporal dependencies effectively (Wu et al., 2022).

- Input Sequence Formation

Given a sequence of normalized observations, each input sequence, X_{seq} , was structured as:

$$X_{\text{seq}} = \{X_t, X_{t+1}, \dots, X_{t+k-1}\} \quad (18)$$

where k denotes the sequence length used as input to the model.

2.5. Evaluation metrics

To thoroughly assess the predictive performance of the SARIMA and LSTM models, we utilize four commonly accepted statistical metrics, such as mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and the correlation coefficient (R). Each metric provides distinct insights into the model's accuracy and error distribution, enabling a comprehensive evaluation of predictive reliability.

The mathematical expressions for these metrics are defined as follows:

- **MAE:** Measures the average absolute difference between predicted and actual values (Yao et al., 2022).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (19)$$

- **RMSE:** Gives greater weight to larger errors by squaring the differences, making it more sensitive to outliers (Kodama et al., 2024).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (20)$$

- **MAPE:** Represents the error as a percentage of actual values, enabling direct interpretability (Bao et al., 2024).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|\hat{y}_i - y_i|}{y_i} \right) \times 100 \quad (21)$$

- **Correlation Coefficient (R):** Quantifies the strength of the linear relationship between predicted and actual values, ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation) (Mateo-Benito et al., 2024).

$$R = \left(1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \right)^{\frac{1}{2}} \quad (22)$$

Here, \hat{y}_i represents predicted values, y_i denotes actual observed values, \bar{y}_i is the mean of the observed values, and n is the total number of observations.

Table 4

Performance comparison of SARIMA and LSTM models for AT forecasting across climate zones.

Region	Model	MAPE (%)	MAE (°C)	RMSE (°C)	R
Global	SARIMA	0.5160	0.0745	0.0967	0.9975
	LSTM	0.5348	0.0774	0.1025	0.9973
Northern Hemisphere	SARIMA	0.8841	0.1175	0.1499	0.9995
	LSTM	0.9475	0.1300	0.1663	0.9993
Southern Hemisphere	SARIMA	0.6852	0.0919	0.1132	0.9983
	LSTM	0.6796	0.0918	0.1158	0.9982
Tropics	SARIMA	0.2452	0.0628	0.0781	0.9829
	LSTM	0.2782	0.0713	0.0874	0.9786

3. Result and discussion

This study comprehensively investigates the spatiotemporal variability of near-surface AT across four distinct geographical domains: the globe, the northern hemisphere, the southern hemisphere, and the tropics. A linear detrending procedure was applied to ensure that the long-term trends in the monthly near-surface AT data did not obscure the underlying variability and periodic behavior. The resulting residuals provide a more accurate representation of short- to mid-term fluctuations in atmospheric temperature patterns.

3.1. Model performance evaluation

Table 4 presents the comparative performance of two predictive models, SARIMA and LSTM, across four climate zones using standard error metrics such as MAPE, MAE, RMSE, and the R.

The SARIMA model demonstrates higher predictive accuracy in all evaluated domains, outperforming the LSTM model in most metrics. For instance, across the Global dataset, SARIMA yields a MAPE of 0.5160%, outperforming LSTM (0.5348%), alongside a lower RMSE (0.0967°C) and a marginally higher correlation ($R = 0.9975$). These results imply that SARIMA is particularly effective in modeling linear trends and seasonality in structured climatological time series.

In the tropical region, where temperature variability is relatively minimal, SARIMA achieves notable accuracy again with an MAPE of 0.2452% and R of 0.9829, compared to LSTM's MAPE of 0.2782% and R of 0.9786. This marginal yet consistent superiority is especially critical in low-variance regions, where model sensitivity to subtle shifts in temperature patterns becomes paramount.

Interestingly, performance differences between SARIMA and LSTM are less pronounced in the Southern Hemisphere. The models exhibit nearly identical MAE values (SARIMA: 0.0919°C vs. LSTM: 0.0918°C), suggesting that linear and nonlinear approaches are comparably robust in less volatile climatic zones. However, SARIMA retains a slight edge in RMSE and correlation, further affirming its reliability in capturing both central tendency and distributional spread.

While the correlation coefficients (R) between SARIMA and LSTM models appear marginally different (e.g., $\Delta R < 0.005$), even such small improvements are meaningful in climatological time-series modeling, given the long temporal span (1970–2024), slight enhancements in correlation translate to improved temporal coherence and reduced error propagation across forecasting horizons, moreover, lower RMSE and MAE values consistently support these correlation improvements across all regions, reinforcing SARIMA's practical forecasting advantage. Significantly, SARIMA exhibits its most significant advantage over LSTM in the tropics, with a larger R gap (0.9829 vs. 0.9786) and markedly lower RMSE (0.0781 vs. 0.0874). The tropics exhibit relatively low interannual variability and strong cyclical behavior, which align closely with SARIMA's strengths in modeling consistent autoregressive structures. In contrast, LSTM's linear learning mechanisms may over-smooth temperature fluctuations or require more data to generalize effectively in such regimes. These observations suggest that SARIMA is particularly

Table 5

Observed and projected Near-surface AT trends and anomalies for the Globe, Hemispheres, and Tropics (1970–2050).

Variable	Year	Global	Northern Hemisphere	Southern Hemisphere	Tropics
Near-surface AT (°C)	1970 Obs	13.80	14.50	13.20	24.92
	2000 Obs	14.12	15.01	13.33	25.04
	2023 Obs	14.97	16.04	14.00	25.86
	2050 SARIMA	15.45	16.76	14.22	26.19
	2050 LSTM	15.37	16.72	14.11	26.11
Near-surface AT Anomaly (°C) relative to the 1991–2020	1970 Obs	-0.5755	-0.7886	-0.3627	-0.4277
	2000 Obs	-0.2525	-0.2754	-0.2309	-0.3042
	2023 Obs	+0.6008	+0.7584	+0.4442	+0.5105
	2050 SARIMA	+1.0784	+1.4742	+0.6639	+0.8435
	2050 LSTM	+0.9958	+1.4327	+0.5530	+0.7597
Near-surface AT Trend (°C/decade)	1970–2000 Obs	+0.177	+0.229	+0.124	+0.163
	2001–2023 Obs	+0.247	+0.317	+0.176	+0.186
	2024–2050 SARIMA	+0.191	+0.291	+0.102	+0.132
	2024–2050 LSTM	+0.183	+0.287	+0.098	+0.156

Table 6

Annual near-surface AT anomalies (°C) for 2025 and 2050, based on the 1991–2020 climatological baseline. Comparison between CMIP5 RCP 4.5, SARIMA, and LSTM projections.

Region	2025			2050		
	CMIP5 RCP 4.5	SARIMA (this study)	LSTM (this study)	CMIP5 RCP 4.5	SARIMA (this study)	LSTM (this study)
Global	0.498	0.6352	0.5726	1.087	1.0784	0.9958
Northern Hemisphere	0.614	0.7618	0.7498	1.340	1.4742	1.4327
Southern Hemisphere	0.382	0.4287	0.3213	0.834	0.6639	0.5530
Tropics	0.468	0.5134	0.1839	1.002	0.8435	0.7597

effective in the tropics, where the climate system exhibits stronger periodicity and lower noise, allowing it to outperform more flexible but less stable DL models.

3.2. Observed and projected near-surface AT trends (1970–2050)

Table 5 synthesizes observed and modeled Near-surface AT trends from 1970 to 2050 to assess long-term climate evolution. Both the observed data and model projections consistently depict a warming trajectory, with the Northern Hemisphere exhibiting the most significant increase, from 14.50 °C in 1970 to 16.04 °C in 2023, with further projections reaching 16.76 °C (SARIMA) and 16.72 °C (LSTM) by 2050.

The global near-surface AT anomaly shifts from a negative baseline in 1970 (-0.5755 °C) to a pronounced positive anomaly in 2023 (+0.6008 °C), signifying anthropogenically induced warming. Notably, the SARIMA model forecasts a more aggressive anomaly growth (+1.0784 °C by 2050), particularly in the Northern Hemisphere, where the anomaly is projected at +1.4742 °C, compared to +1.4327 °C by the LSTM model. This suggests that SARIMA may exhibit heightened sensitivity to recent accelerations in near-surface AT anomalies, aligning well with IPCC-documented patterns of intensified hemispheric warming.

Despite slight differences in absolute values, both models converge in recognizing the Tropics as the least affected zone by future warming in terms of anomalies, with projections of +0.8435 °C (SARIMA) and +0.7597 °C (LSTM). This agrees with the climatological understanding that persistent oceanic heat fluxes and convective equilibrium partially buffer equatorial near-surface ATs.

While the observed near-surface AT trend during 2001–2023 is steeper than the model-projected rates for 2024–2050, this does not indicate a decline in global warming. Instead, it reflects the presence of short-term climatic accelerations driven by intensified anthropogenic forcing, polar amplification, and episodic natural phenomena such as El Niño events (Trenberth, 1997) superimposed on the long-term trend. The SARIMA and LSTM projections smooth out such high-frequency variability due to their autoregressive and memory-based architectures, emphasizing the structural component of climate evolution. As shown in Figs. 6 and 10, both models forecast continued warming across all domains through 2050, with global anomalies exceeding +1.0 °C. The slight moderation in projected trends is therefore a result of statistical averaging, not a signal of reduced climate risk.

3.3. Model comparison with CMIP5 rcp4.5 scenario

Table 6 presents a comparative analysis of annual surface temperature anomalies projected for 2025 and 2050 using the SARIMA and LSTM models developed in this study, alongside the CMIP5 RCP 4.5 scenario. The baseline period for anomaly calculation is 1991–2020. This comparative evaluation provides insights into the predictive capacity of data-driven approaches relative to established climate model scenarios.

The results reveal a consistent increase in projected anomalies across all regions, with both SARIMA and LSTM models generally predicting higher anomalies than the CMIP5 RCP 4.5 scenario by 2025. Specifically, the SARIMA model projects a global anomaly of 0.6352 °C in 2025, compared to 0.5726 °C for the LSTM model and 0.498 °C under CMIP5 RCP 4.5. A similar pattern is observed across hemispheres and in the tropics, indicating that data-driven models may offer a more sensitive response to short-term warming signals, possibly due to their direct reliance on historical patterns without embedded physical parameterizations.

By 2050, the SARIMA model continues to project higher warming than LSTM in all regions, with the Northern Hemisphere showing the most pronounced divergence. SARIMA estimates an anomaly of 1.4742 °C in the Northern Hemisphere by 2050, compared to 1.4327 °C from LSTM and 1.340 °C from CMIP5 RCP 4.5. This suggests that SARIMA may better capture the nonlinear acceleration of warming, particularly in higher-latitude regions. Conversely, in the Southern Hemisphere, where warming trends are generally more moderate, both models tend to underpredict relative to CMIP5, with LSTM showing a notably lower anomaly of 0.5530 °C compared to 0.6639 °C (SARIMA) and 0.834 °C (CMIP5).

In the tropics, the SARIMA model again projects the highest anomaly (0.8435 °C) by 2050, followed by LSTM (0.7597 °C) and CMIP5 (1.002 °C). Notably, the LSTM model's projections for the tropics in 2025 are significantly lower than both SARIMA and CMIP5, indicating a potential underfitting or lag in capturing subtropical variability, which could be attributed to the temporal dependency structure in LSTM's recurrent architecture.

These findings highlight that SARIMA, while more simplistic in structure, provides competitive, if not superior, projections compared

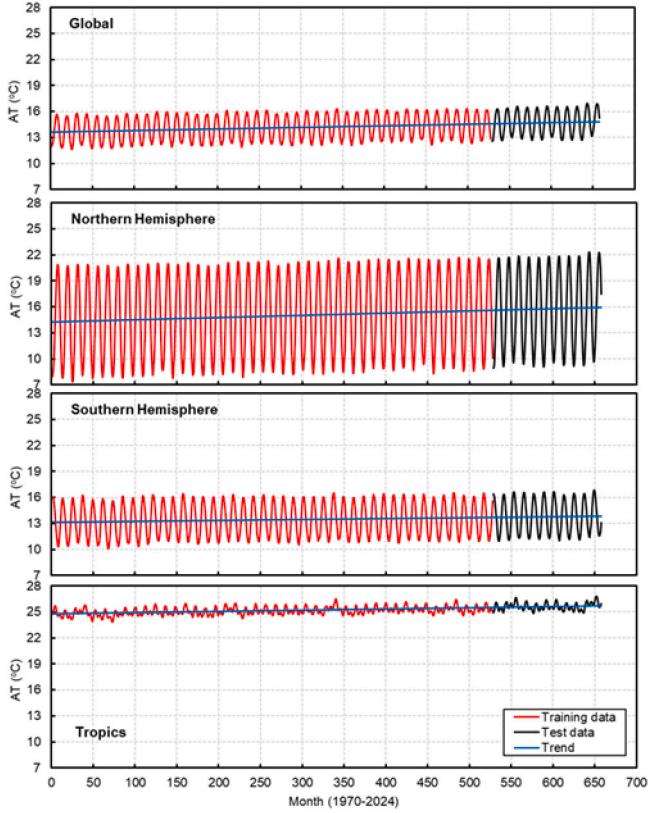


Fig. 3. Partitioning of the AT dataset into training and testing subsets, along with the corresponding linear trend line.

to the DL model, for instance, LSTM, especially in regions where the warming trend exhibits linear or seasonal periodicity. Moreover, both data-driven models offer complementary perspectives to traditional climate models, particularly for near-term climate projections and regional adaptation planning.

3.4. Trend analysis and predictive evaluation of SARIMA and LSTM

Fig. 3 presents the monthly near-surface AT data from 1970 to 2024 across four spatial domains: global, northern, southern, and tropics. Each series is partitioned into training (red) and testing (black) subsets, with a superimposed linear trend line (blue). The plots reveal a consistent warming signal in all regions. Globally, AT increases from approximately 12.0 °C to 15.5 °C, with a moderate seasonal amplitude of 2.5 °C. The Northern Hemisphere exhibits the most pronounced seasonal variation, ranging from 10.0 °C to 19.5 °C, driven by land-dominant surfaces and amplified radiative forcing. The Southern Hemisphere shows a more modest rise (from 11.0 °C to 15.5 °C), with seasonal amplitude constrained to 4–5 °C due to oceanic thermal regulation. The Tropics display a narrow temperature range (24.0 °C to 26.5 °C), minimal seasonality (2.5 °C), and visibly erratic short-term fluctuations.

Fig. 4 shows each region's corresponding detrended AT series, generated by subtracting the linear component estimated using a moving average filter, as defined in Eq. (15). This preprocessing step yields stationary residuals emphasizing short- to mid-term variability by eliminating long-term warming trends. The detrended global and southern hemisphere series exhibit moderate cyclical behavior within ±2.5 °C. The northern hemisphere retains high-amplitude, seasonally structured residuals, with variations up to ±5 °C, while the Tropics demonstrate

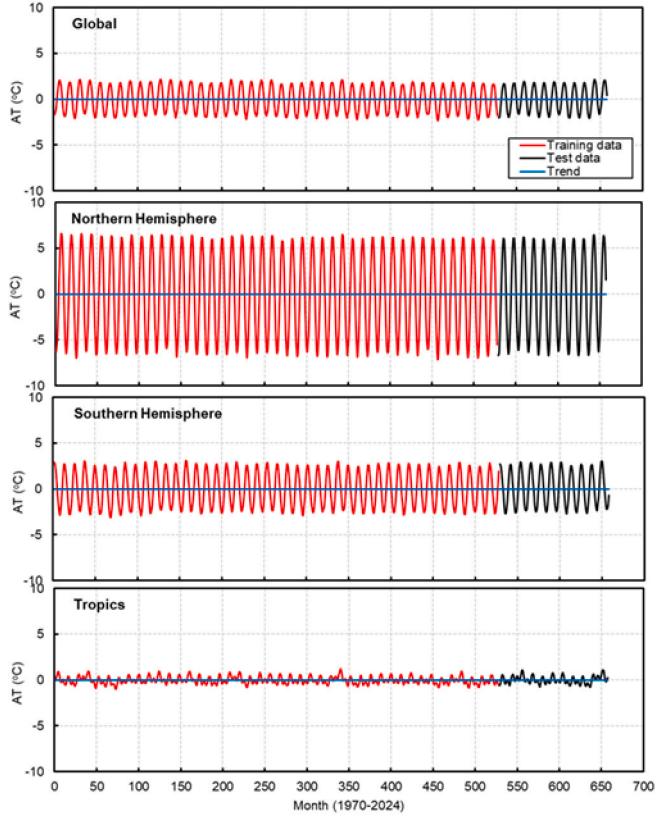


Fig. 4. Detrended AT data for various regions, highlighting underlying variability after removing the linear trend.

low-amplitude, high-frequency noise bounded within ±1 °C. The absence of strong seasonality in the tropics confirms the dominance of convective equilibrium and thermal inertia.

Besides, **Fig. 5** compares the observed AT series from 2014 to 2024 with predictions generated by SARIMA and LSTM models across four spatial domains. In the global domain, observed temperatures fluctuate between 13.0 °C and 16.8 °C with seasonal amplitudes of 3.8 °C, which both models closely replicate. The Northern hemisphere shows the largest variation, ranging from 10.0 °C to 19.5 °C; SARIMA better tracks the seasonal extremes, reflecting its capacity to model strong periodicity. In the southern hemisphere, temperatures span 12.5 °C to 16.5 °C, with both models yielding near-identical performance. The Tropics exhibit irregular fluctuations between 24.5 °C and 26.7 °C, where SARIMA more accurately captures high-frequency variability, while LSTM provides smoother transitions.

Fig. 6 presents long-term projections (2024–2050) of AT using SARIMA and LSTM models for four geographic domains. All regions show a consistent warming trajectory with preserved seasonal oscillations. The Global AT increases from 15.5 °C to 17.2 °C, with a seasonal amplitude of 3.5 °C. The Northern Hemisphere exhibits the steepest rise, from 13.0 °C to above 22.0 °C, with seasonal amplitudes reaching nearly 9 °C. The Southern Hemisphere shows a more gradual increase from 13.5 °C to 17.0 °C, maintaining seasonal oscillations within a 3.5 °C band. In the Tropics, where baseline temperatures remain high (25.0 °C to 26.5 °C), both models project a 1.0 °C increase by 2050, with minimal seasonal fluctuation. SARIMA consistently forecasts slightly higher temperatures than LSTM across all regions, with the gap becoming more noticeable post-2040, particularly in the Tropics and the Northern Hemisphere. This reflects SARIMA's sensitivity to long-term autocorrelated trends, while LSTM's smoother output reflects its nonlinear sequence learning mechanism. The green trend lines confirm

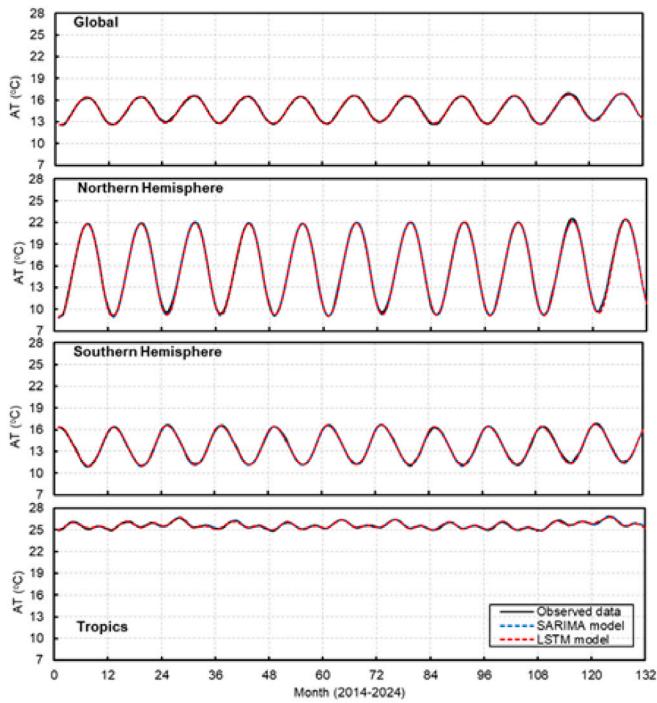


Fig. 5. Comparison between observed AT data and model outputs from SARIMA and LSTM in the testing process.

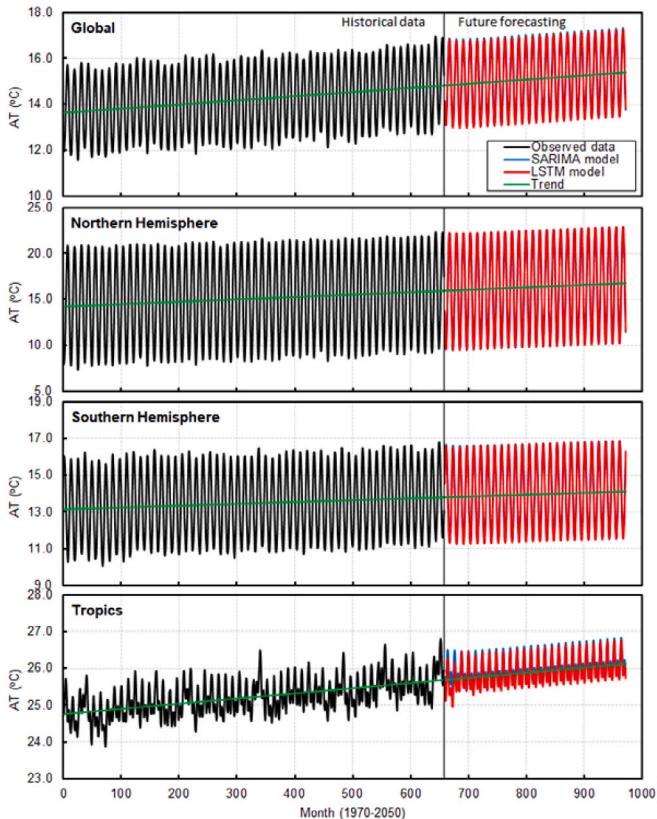


Fig. 6. Future near-surface AT projections for the Globe, Hemispheres, and Tropics using SARIMA and LSTM models.

a persistent upward trajectory, especially in mid- and high-latitude zones. These projections align with the numerical estimates in [Table 6](#) and emphasize the complementary strengths of both models.

The predictive performance of the SARIMA and LSTM models is assessed in [Fig. 7](#), which compares predicted versus observed AT across four spatial domains. Both models follow the 1:1 reference line closely, with most predictions falling within a $\pm 2.5\%$ error margin, indicating high accuracy. In the global and southern hemisphere regions, predictions are tightly clustered across the observed AT ranges ($13\text{--}17\text{ }^{\circ}\text{C}$ and $11\text{--}16.5\text{ }^{\circ}\text{C}$, respectively), with SARIMA exhibiting slightly lower dispersion. The northern hemisphere displays near-perfect alignment over a broader range ($9.5\text{--}23.5\text{ }^{\circ}\text{C}$), reflecting model robustness under pronounced seasonal variation. A higher spread is observed in the tropics ($24.2\text{--}27.5\text{ }^{\circ}\text{C}$), especially at the upper range, where LSTM marginally underestimates peak values.

[Fig. 8](#) presents notched box plots comparing the observed and modeled near-surface AT distribution across four spatial domains. SARIMA and LSTM accurately reproduce the observed central tendency and spread, with median values tightly aligned in all cases. The notches, representing the 95% confidence interval of the median, show substantial overlap, indicating no statistically significant deviation between observed and predicted medians. In the global and southern hemisphere domains, SARIMA and LSTM display comparable performance, with minimal divergence in interquartile range and symmetric whiskers. In the northern hemisphere, which exhibits the largest AT variability ($11\text{--}23\text{ }^{\circ}\text{C}$), SARIMA matches both the median and interquartile range more precisely. Similarly, in the tropics ($25\text{--}26.7\text{ }^{\circ}\text{C}$), where temperature variability is lower, SARIMA demonstrates marginally closer adherence to observed distributions.

[Fig. 9](#) presents Taylor diagrams evaluating model performance across four spatial domains using standard deviation, correlation coefficient, and centered RMSE. In all panels, the SARIMA model lies closer to the observed reference point than LSTM, indicating stronger variance and temporal correlation agreement. In the global, southern hemisphere, and tropics, SARIMA exhibits slightly higher correlation values (0.98–0.99) and smaller deviations in standard deviation relative to LSTM. In the northern hemisphere, where interannual variability is greatest, both models achieve high correlation (0.99), but SARIMA maintains a lower RMSE, reflecting better alignment with observed dynamics. These findings are consistent with [Table 4](#), reinforcing SARIMA's robustness in reproducing regional temperature variability's amplitude and phase characteristics.

[Fig. 10](#) illustrates annual near-surface AT trends from 1970 to 2050 across the globe, northern hemisphere, southern hemisphere, and tropics. All regions exhibit a clear warming trajectory, with model forecasts closely aligned with historical observations. The northern hemisphere shows the most pronounced increase, rising from approximately $14.5\text{ }^{\circ}\text{C}$ in the 1970s to over $16.5\text{ }^{\circ}\text{C}$ by 2050, driven by extensive landmass and stronger radiative forcing. The global mean AT increases from $13.8\text{ }^{\circ}\text{C}$ to $15.5\text{ }^{\circ}\text{C}$, while the southern hemisphere warms from $13.0\text{ }^{\circ}\text{C}$ to $14.5\text{ }^{\circ}\text{C}$. The tropics remain the warmest region, increasing from $24.5\text{ }^{\circ}\text{C}$ to $26.5\text{ }^{\circ}\text{C}$. SARIMA and LSTM models effectively replicate historical trends, with SARIMA demonstrating slightly better alignment in post-2024 projections. The green trend lines indicate a continued linear warming pattern, while SARIMA and LSTM capture mild nonlinear deviations. These results affirm the robustness of both models in capturing long-term climatic behavior. In addition, [Fig. 11](#) presents annual AT anomaly projections from SARIMA and LSTM models for the globe, northern hemisphere, southern hemisphere, and tropics. Both models effectively capture historical warming patterns and exhibit comparable future trajectories through 2050. While the overall trends are consistent across regions, SARIMA projects slightly higher anomaly increases beyond 2024, particularly in the northern hemisphere and tropics. This subtle divergence highlights SARIMA's heightened sensitivity to long-term autoregressive structures. LSTM, though closely aligned, yields smoother forecasts.

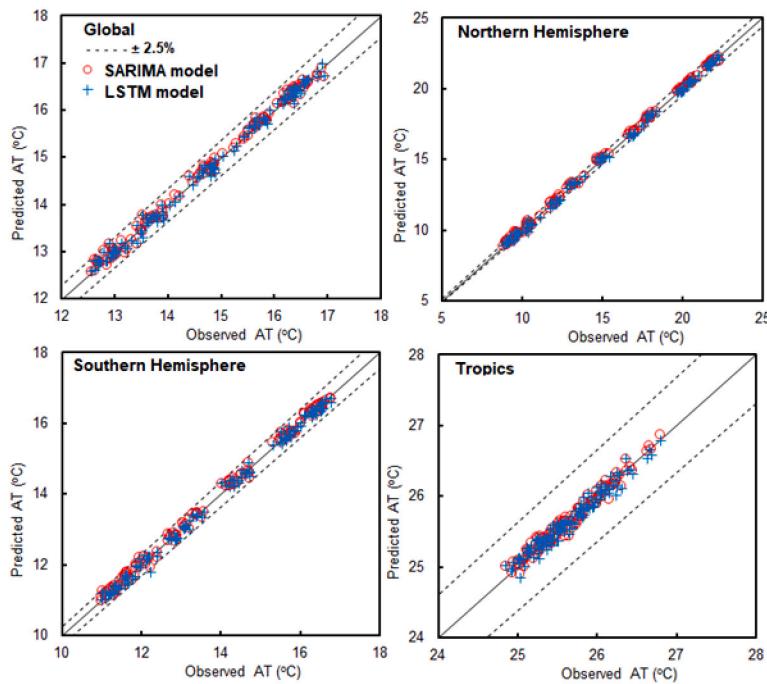


Fig. 7. Comparison of predicted and observed AT values using SARIMA and LSTM models.

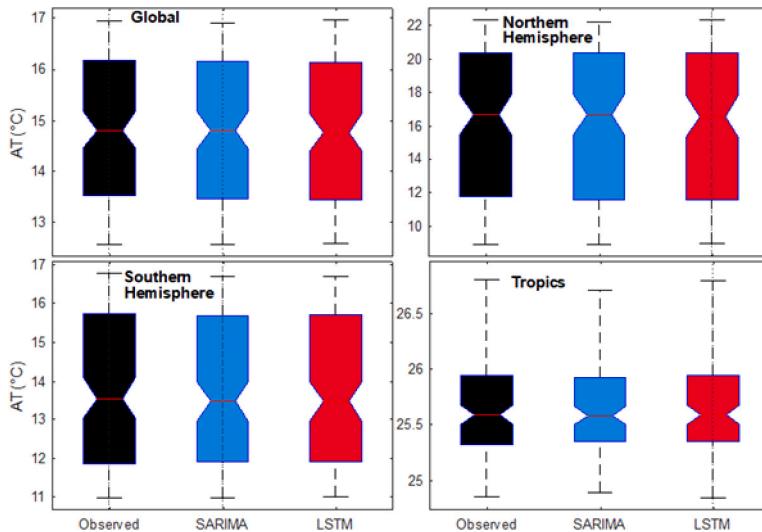


Fig. 8. Box plot comparison of observed AT with SARIMA and LSTM model outputs across different regions.

4. Conclusion

This study compared SARIMA and LSTM models for projecting near-surface air temperature (AT) across the globe, the northern hemisphere, the southern hemisphere, and the tropics using ERA5 reanalysis data (1970–2024) with projections to 2050. Both models reproduced the persistent global warming trend, with the northern hemisphere warming fastest and the tropics exhibiting the highest mean AT but lowest interannual variability. SARIMA consistently outperformed LSTM in regions dominated by strong seasonal and linear components, achieving lower MAPE, MAE, and RMSE values and closer alignment with observed and detrended trends. Projections for 2050 indicate a global AT anomaly of +1.078 °C and +1.474 °C for the northern hemisphere relative to 1991–2020, consistent with CMIP5 RCP4.5 scenarios.

The framework's main strengths include SARIMA's high predictive accuracy, strong alignment with observed trends, and computational efficiency, while LSTM demonstrated adaptability in modeling complex temporal patterns despite lower accuracy in this dataset. Limitations include LSTM's underperformance relative to SARIMA, the relatively small dataset size for deep learning, and the assumption of historical stationarity, which may hinder the capture of unprecedented climate shifts.

To address these limitations, future research should investigate hybrid modeling strategies that integrate SARIMA's seasonal and autoregressive strengths with LSTM's nonlinear capabilities. Expanding datasets, incorporating exogenous climate drivers (e.g., ENSO, volcanic aerosols, radiative forcing anomalies), and extending projections under multiple Shared Socioeconomic Pathways (SSPs) would enhance robustness. Finally, applying explainable AI techniques is recommended

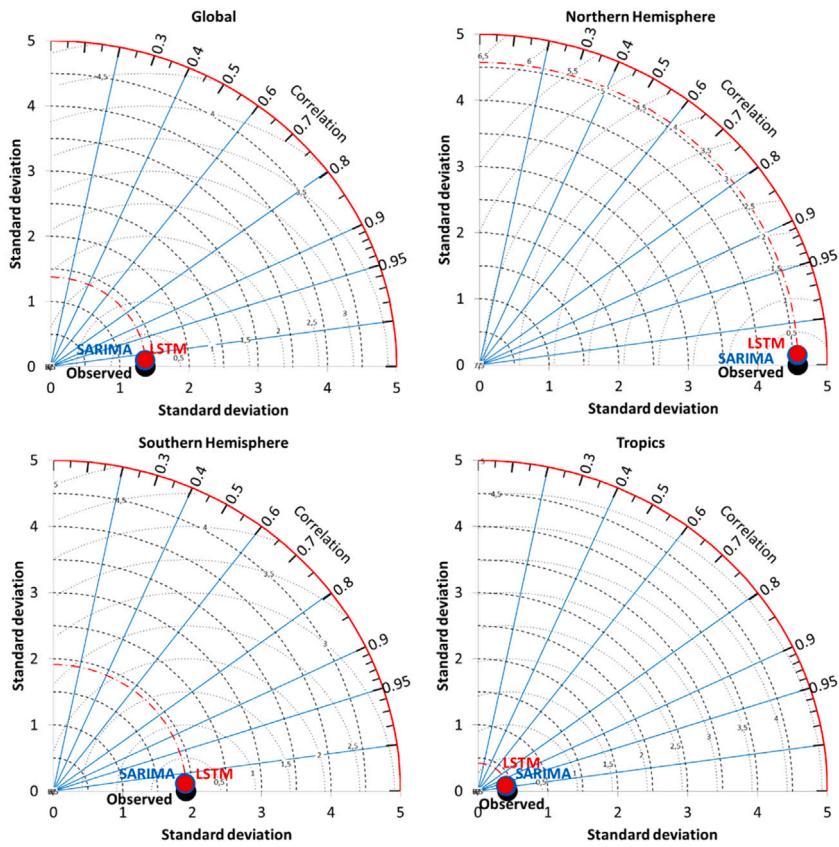


Fig. 9. Taylor diagram illustrating the performance of SARIMA and LSTM models in replicating observed AT across regions.

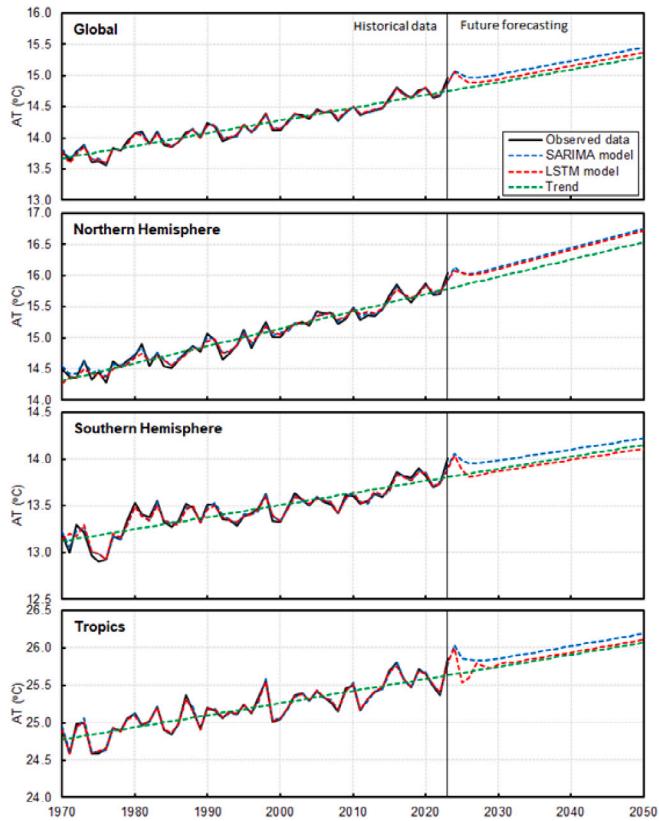


Fig. 10. Annual AT trends for the Globe, Northern Hemisphere, Southern Hemisphere, and Tropics.

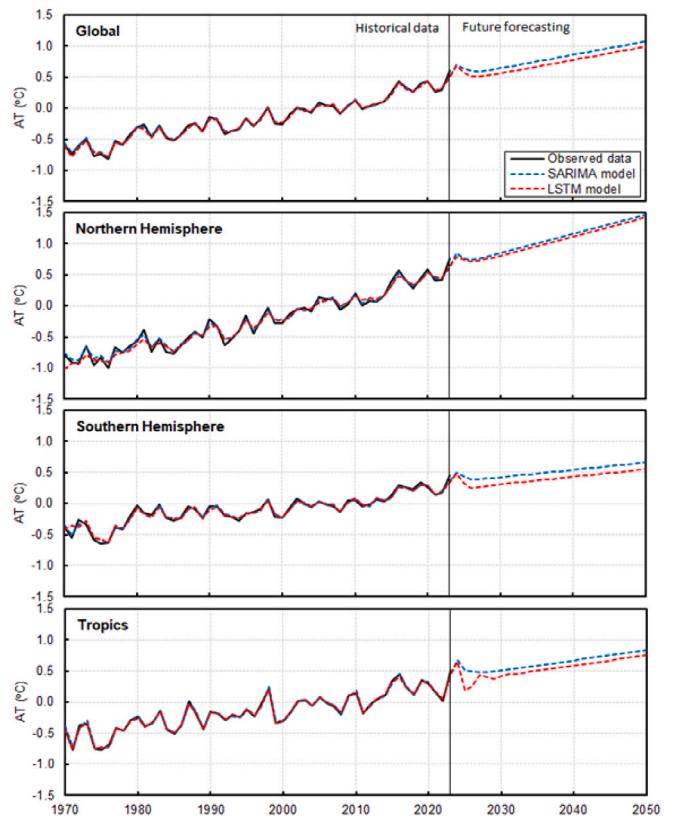


Fig. 11. Annual AT anomaly projections from SARIMA and LSTM models for the Globe, Northern Hemisphere, Southern Hemisphere, and Tropics.

to improve model transparency and support climate adaptation and policy decisions.

CRediT authorship contribution statement

Muhammed M. Aksoy: Writing – review & editing, Writing – original draft, Validation, Methodology, Data curation, Conceptualization. **Md. Najmul Mowla:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Data curation, Conceptualization. **Mehmet Bilgili:** Writing – review & editing, Data curation, Conceptualization. **Engin Pinar:** Writing – review & editing, Conceptualization. **Tahir Durhasan:** Writing – review & editing, Conceptualization. **Davood Asadi:** Writing – review & editing, Conceptualization.

Funding

The authors did not receive support from any organization for the submitted work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The dataset used in this study is publicly available from the Climate Reanalyzer platform provided by the Climate Change Institute at the University of Maine. The data can be accessed [Climate Change Institute, University of Maine \(2024\)](#).

References

- Acosta, M., Palomas, S., Tourigny, E., 2023. Balancing EC-Earth3 improving the performance of EC-earth CMIP6 configurations by minimizing the coupling cost. *Earth Space Sci.* 10 (8), e2023EA002912.
- Adeniran, I.A., Nazeer, M., Wong, M.S., Chan, P.-W., 2024. An improved machine learning-based model for prediction of diurnal and spatially continuous near surface air temperature. *Sci. Rep.* 14 (1), 27342.
- Akperov, M., Zhang, W., Koenigk, T., Eliseev, A., Semenov, V.A., Mokhov, I.I., 2025. Projected changes in near-surface wind speed in the Arctic by a regional climate model. *Polar Sci.* 43, 101162.
- Allan, R.P., Willett, K.M., John, V.O., Trent, T., 2022. Global changes in water vapor 1979–2020. *J. Geophys. Res.: Atmospheres* 127 (12), e2022JD036728.
- Avşar, E., Mowla, M.N., 2022. Wireless communication protocols in smart agriculture: A review on applications, challenges and future trends. *Ad Hoc Networks* 136, 102982.
- Bao, N., Peng, K., Yan, X., Lu, Y., Liu, M., Li, C., Zhao, B., 2024. Towards interpreting machine learning models for understanding the relationship between vegetation growth and climate factors: A case study of the Anhui Province, China. *Ecol. Indic.* 167, 112636.
- Baris-Tuzemen, O., Lyhagen, J., 2024. Revisiting the role of climate change on crop production: evidence from Mediterranean countries. *Environ. Dev. Sustain.* 1–14.
- Boyko, J.D., O'Meara, B.C., 2024. Dentist: Quantifying uncertainty by sampling points around maximum likelihood estimates. *Methods Ecol. Evol.* 15 (4), 628–638.
- Bruhwiler, L., Basu, S., Butler, J.H., Chatterjee, A., Dlugokencky, E., Kenney, M.A., McComiskey, A., Montzka, S.A., Stanitski, D., 2021. Observations of greenhouse gases as climate indicators. *Clim. Change* 165 (1), 12.
- Cai, W., Gao, L., Luo, Y., Li, X., Zheng, X., Zhang, X., Cheng, X., Jia, F., Purich, A., Santoso, A., et al., 2023. Southern Ocean warming and its climatic impacts. *Sci. Bull.* 68 (9), 946–960.
- Chattopadhyay, S., Chattopadhyay, G., 2008. Identification of the best hidden layer size for three-layered neural net in predicting monsoon rainfall in India. *J. Hydroinformatics* 10 (2), 181–188.
- Cho, D., Yoo, C., Son, B., Im, J., Yoon, D., Cha, D.-H., 2022. A novel ensemble learning for post-processing of NWP Model's next-day maximum air temperature forecast in summer using deep learning and statistical approaches. *Weather. Clim. Extrem.* 35, 100410.
- Climate Change Institute, University of Maine, 2024. Climate Reanalyzer. URL <https://climatereanalyzer.org>. (Accessed 12 January 2024).
- Counillon, F., Keenlyside, N., Wang, S., Devilliers, M., Gupta, A., Koseki, S., Shen, M.-L., 2023. Framework for an ocean-connected supermodel of the earth system. *J. Adv. Model. Earth Syst.* 15 (3), e2022MS003310.
- Decharme, B., Costantini, M., Colin, J., 2025. A simple approach to represent irrigation water withdrawals in Earth system models. *J. Adv. Model. Earth Syst.* 17 (4), e2024MS004508.
- DelSole, T., Tippett, M.K., 2024. Tuning Earth system models without integrating to statistical equilibrium. *J. Adv. Model. Earth Syst.* 16 (12), e2024MS004230.
- Devi, R.R., Banerjee, S., Chattopadhyay, S., 2023. Dempster-Shafer and LSTM based analysis and forecasting of total ozone data. *Remote. Sens. Lett.* 14 (10), 1034–1043.
- Ding, Y., Mu, C., Wu, T., Hu, G., Zou, D., Wang, D., Li, W., Wu, X., 2021. Increasing cryospheric hazards in a warming climate. *Earth-Sci. Rev.* 213, 103500.
- Dubey, A.K., Kumar, A., García-Díaz, V., Sharma, A.K., Kanhaiya, K., 2021. Study and analysis of SARIMA and LSTM in forecasting time series data. *Sustain. Energy Technol. Assessments* 47, 101474.
- Durhasan, T., Pinar, E., Ulucak, I., Bilgili, M., 2025. Future forecast of global mean surface temperature using machine learning and conventional time series methods. *Theor. Appl. Climatol.* 156 (1), 17.
- Elshall, A.S., Ye, M., Kranz, S.A., Harrington, J., Yang, X., Wan, Y., Maltrud, M., 2022. Earth system models for regional environmental management of red tide: Prospects and limitations of current generation models and next generation development. *Environ. Earth Sci.* 81 (9), 256.
- Fallah, B., Didovets, I., Rostami, M., Hamidi, M., 2024. Climate change impacts on Central Asia: Trends, extremes and future projections. *Int. J. Climatol.* 44 (10), 3191–3213.
- Fletcher, C.G., McNally, W., Virgin, J.G., King, F., 2022. Toward efficient calibration of higher-resolution Earth System Models. *J. Adv. Model. Earth Syst.* 14 (7), e2021MS002836.
- Forster, P.M., Smith, C., Walsh, T., Lamb, W.F., Lamboll, R., Hall, B., Hauser, M., Ribes, A., Rosen, D., Gillett, N.P., et al., 2024. Indicators of Global Climate Change 2023: annual update of key indicators of the state of the climate system and human influence. *Earth Syst. Sci. Data* 16 (6), 2625–2658.
- Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron)— review of applications in the atmospheric sciences. *Atmos. Environ.* 32 (14–15), 2627–2636.
- Golitsyn, G., Varotsos, C., 2025. The air temperature scaling effect. In: *The Stochastic Nature of Environmental Phenomena and Processes*. Springer Nature Switzerland, Cham, pp. 181–191.
- Hamed, M.M., Nashwan, M.S., Shahid, S., bin Ismail, T., Wang, X.-J., Dewan, A., Asaduzzaman, M., 2022. Inconsistency in historical simulations and future projections of temperature and rainfall: A comparison of CMIP5 and CMIP6 models over Southeast Asia. *Atmos. Res.* 265, 105927. <http://dx.doi.org/10.1016/j.atmosres.2021.105927>.
- He, Y., Chen, C., Li, B., Zhang, Z., 2022. Prediction of near-surface air temperature in glacier regions using ERA5 data and the random forest regression method. *Remote. Sens. Appl.: Soc. Environ.* 28, 100824.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146 (730), 1999–2049.
- Hosseini, F., Prieto, C., Álvarez, C., 2024. Hyperparameter optimization of regional hydrological LSTMs by random search: A case study from Basque Country, Spain. *J. Hydrol.* 643, 132003.
- Hu, Z., Liang, X., Bian, C., Min, X., Yue, Z., Ran, Y., 2025. Future precipitation projections and model evaluation in the Hengduan Mountains based on CMIP6. *J. Environ. Manag.* 382, 125378. <http://dx.doi.org/10.1016/j.jenvman.2025.125378>.
- Infante-Amate, J., Aguilera, E., 2024. Beyond fossil fuels: Considering land-based emissions reshapes the carbon intensity of modern economic growth (Spain, 1860–2017). *Hist. Methods: A J. Quant. Interdisciplinary Hist.* 57 (4), 226–241.
- Jones, R.N., 2024. Evidence and explanation for the 2023 global warming anomaly. *Atmosphere* 15 (12), 1507.
- Kenne, A.D., Toure, M., Logamou Seknewna, L., Ketsemen, H.L., 2024. Subseasonal prediction of summer temperature in West Africa using artificial intelligence: A case study of Senegal. *Int. J. Intell. Syst.* 2024 (1), 8869267.
- Khala, M., El Yanboiy, N., Elabbassi, I., Eloutassi, O., Halimi, M., El Hassouani, Y., Messaoudi, C., 2025. Enhancing machine learning model for early warning in PV plants: air temperature prediction informed by power temperature coefficient. *J. Supercomput.* 81 (2), 394.
- Kodama, K.M., Kourkchi, E., Longman, R.J., Lucas, M.P., Bateni, S.M., Huang, Y.-F., Kagawa-Viviani, A., Mclean, J., Cleveland, S.B., Giambelluca, T.W., 2024. Mapping daily air temperature over the Hawaiian Islands from 1990 to 2021 via an optimized piecewise linear regression technique. *Earth Space Sci.* 11 (1), e2023EA002851.
- Li, Z., Li, Q., Chen, T., 2024a. Record-breaking high-temperature outlook for 2023: An assessment based on the China global merged temperature (cmst) dataset. *Adv. Atmospheric Sci.* 41 (2), 369–376.
- Li, X., Li, Q., Wild, M., Jones, P., 2024b. An intensification of surface Earth's energy imbalance since the late 20th century. *Commun. Earth & Environ.* 5 (1), 644.
- Loeb, N.G., Ham, S.-H., Allan, R.P., Thorsen, T.J., Meyssignac, B., Kato, S., Johnson, G.C., Lyman, J.M., 2024. Observational assessment of changes in Earth's energy imbalance since 2000. *Surv. Geophys.* 1–27.

- Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., et al., 2021. Climate change 2021: the physical science basis. Contrib. Work. Group I to Sixth Assess. Rep. the Intergov. Panel Clim. Chang. 2 (1), 2391.
- Mateo-Beneito, A., Florescu, G., Tátosová, J., Carter, V.A., Chiverrell, R., Heiri, O., Vasiliev, I., Kuosmanen, N., Kuneš, P., 2024. Multi-proxy temperature and environmental reconstruction during the Late Glacial and Early Holocene in the Bohemian Forest, Central Europe. *Quat. Sci. Rev.* 331, 108647.
- Matthes, H., Damseaux, A., Westermann, S., Beer, C., Boone, A., Burke, E., Decharme, B., Genet, H., Jafarov, E., Langer, M., et al., 2025. Advances in permafrost representation: Biophysical processes in earth system models and the role of offline models. *Permafro. Periglac. Process.*
- Meinshausen, M., Lewis, J., McGlade, C., Gütschow, J., Nicholls, Z., Burdon, R., Cozzi, L., Hackmann, B., 2022. Realization of Paris Agreement pledges may limit warming just below 2 °C. *Nature* 604 (7905), 304–309.
- Mowla, M.N., Asadi, D., Masum, S., Rabie, K., 2025. Adaptive hierarchical multi-headed convolutional neural network with modified convolutional block attention for aerial forest fire detection. *IEEE Access* 13, 3412–3433. <http://dx.doi.org/10.1109/ACCESS.2024.3524320>.
- Onder, G.T., 2024. Comparative time series analysis of SARIMA, LSTM, and GRU models for global SF6 emission management system. *J. Atmos. Sol.-Terr. Phys.* 265, 106393.
- Rahman, M.M., Joha, M.I., Nazim, M.S., Jang, Y.M., 2024. Enhancing IoT-based environmental monitoring and power forecasting: A comparative analysis of AI models for real-time applications. *Appl. Sci.* (2076-3417) 14 (24).
- Ray, S.N., Chattopadhyay, S., 2021. Analyzing surface air temperature and rainfall in univariate framework, quantifying uncertainty through Shannon entropy and prediction through artificial neural network. *Earth Sci. Informatics* 14 (1), 485–503.
- Saleh, M.A., Rasel, H., Ray, B., 2024. A comprehensive review towards resilient rainfall forecasting models using artificial intelligence techniques. *Green Technol. Sustain.* 100104.
- Savaş, Y., 2025. Examining the link between load capacity and income inequality: evidence from OECD countries. *Environ. Dev. Sustain.* 1–19.
- Sekertekin, A., Bilgili, M., Arslan, N., Yıldırım, A., Celebi, K., Ozbek, A., 2021. Short-term air temperature prediction by adaptive neuro-fuzzy inference system (ANFIS) and long short-term memory (LSTM) network. *Meteorol. Atmos. Phys.* 133, 943–959.
- Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., et al., 2024. The ERA5 global reanalysis from 1940 to 2022. *Q. J. R. Meteorol. Soc.* 150 (764), 4014–4048.
- Stachura, G., Ustrnul, Z., Sekula, P., Bochenek, B., Kolonko, M., Szczęch-Gajewska, M., 2024. Machine learning based post-processing of model-derived near-surface air temperature—A multimodel approach. *Q. J. R. Meteorol. Soc.* 150 (759), 618–631.
- Tran, T.T.K., Bateni, S.M., Ki, S.J., Vosoughifar, H., 2021. A review of neural networks for air temperature forecasting. *Water* 13 (9), 1294.
- Trenberth, K.E., 1997. The definition of El Niño. *Bull. Am. Meteorol. Soc.* 78 (12), 2771–2778. [http://dx.doi.org/10.1175/1520-0477\(1997\)078<2771:TDOENO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2).
- Udy, D.G., Vance, T.R., Kiem, A.S., Holbrook, N.J., Abram, N., 2024. Australia's 2019/20 Black Summer fire weather exceptionally rare over the last 2000 years. *Commun. Earth & Environ.* 5 (1), 317.
- Uluocak, I., Bilgili, M., 2024. Daily air temperature forecasting using LSTM-CNN and GRU-CNN models. *Acta Geophys.* 72 (3), 2107–2126.
- Varotsos, C.A., Efstatihou, M.N., 2015. Symmetric scaling properties in global surface air temperature anomalies. *Theor. Appl. Climatol.* 121 (3), 767–773.
- Varotsos, C.A., Efstatihou, M.N., Cracknell, A.P., 2013. On the scaling effect in global surface air temperature anomalies. *Atmospheric Chem. Phys.* 13 (10), 5243–5253.
- Varotsos, C.A., Mazei, Y.A., 2019. Future temperature extremes will be more harmful: A new critical factor for improved forecasts. *Int. J. Environ. Res. Public Heal.* 16 (20), 4015.
- Varotsos, C.A., Mazei, Y., Saldaev, D., Efstatihou, M., Voronova, T., Xue, Y., 2021. Nowcasting of air pollution episodes in megacities: a case study for Athens, Greece. *Atmospheric Pollut. Res.* 12 (7), 101099.
- Von Schuckmann, K., Minière, A., Gues, F., Cuesta-Valero, F.J., Kirchengast, G., Adusumilli, S., Straneo, F., Ablain, M., Allan, R.P., Barker, P.M., et al., 2023. Heat stored in the Earth system 1960–2020: where does the energy go? *Earth Syst. Sci. Data* 15 (4), 1675–1709.
- Von Storch, H., Zwiers, F.W., 2002. Statistical Analysis in Climate Research. Cambridge University Press.
- Wang, S., Foster, A., Lenz, E.A., Kessler, J.D., Stroeve, J.C., Anderson, L.O., Turetsky, M., Betts, R., Zou, S., Liu, W., et al., 2023. Mechanisms and impacts of Earth system tipping elements. *Rev. Geophys.* 61 (1), e2021RG000757.
- Wang, H., Zhang, J., Yang, J., 2024. Time series forecasting of pedestrian-level urban air temperature by LSTM: Guidance for practitioners. *Urban Clim.* 56, 102063.
- Waqas, M., Humphries, U.W., 2024. A critical review of RNN and LSTM variants in hydrological time series predictions. *MethodsX* 102946.
- Winkler, A.J., Sierra, C.A., 2025. Towards a new generation of impulse-response functions for integrated earth system understanding and climate change attribution. *Geophys. Res. Lett.* 52 (8), e2024GL112295. <http://dx.doi.org/10.1029/2024GL112295>.
- World Meteorological Organization, 2023. COP28 concludes with historic agreement to try to tackle the climate crisis. Press Release, 13 December 2023, <https://wmo.int/news/media-centre/cop28-concludes-historic-agreement-try-tackle-climate-crisis>.
- World Meteorological Organization, 2024. Greenhouse gas concentrations surge again to new record in 2023. Press Release, 28 October 2024, <https://wmo.int/news/media-centre/greenhouse-gas-concentrations-surge-again-new-record-2023>.
- Wu, H., Liang, Y., Zuo, J., 2022. Human-inspired spatiotemporal feature extraction and fusion network for weather forecasting. *Expert Syst. Appl.* 207, 118089.
- Xu, M., Tian, W., Zhang, J., Screen, J.A., Zhang, C., Wang, Z., 2023. Important role of stratosphere-troposphere coupling in the Arctic mid-to-upper tropospheric warming in response to sea-ice loss. *Npj Clim. Atmospheric Sci.* 6 (1), 9.
- Yao, M., Tang, H., Huang, G., 2025. Roles of external forcing and internal variability in winter precipitation changes over Central Asia. *Earth's Futur.* 13 (7), e2025EF006064.
- Yao, R., Wang, L., Huang, X., Cao, Q., Peng, Y., 2022. A method for improving the estimation of extreme air temperature by satellite. *Sci. Total Environ.* 837, 155887.
- Zhang, H., Hu, Z., Chen, X., Li, J., Zhang, Q., Zheng, X., 2025. Global greening major contributed by climate change with more than two times rate against the history period during the 21st century. *Global Change Biol.* 31 (3), e70126.