

Data-Driven Evaluation of Climate Projections Using Statistical and Deep-Learning Methods for Pavement ME Design

Mahtab Delfanazari, S.M.ASCE¹; Joshua Qiang Li, M.ASCE²; Musharraf Zaman, F.ASCE³; Ashik Ali, M.ASCE⁴; and Jason Furtado⁵

Abstract: Environmental conditions have a major influence on pavement design, performance, and long-term service life. In Oklahoma state, diverse and often extreme climatic conditions can greatly accelerate deterioration through cracking, rutting, moisture damage, and surface wear. To build more resilient transportation infrastructure, it is critical to integrate reliable climate projections into the mechanistic–empirical (ME) pavement design process. This study evaluates and ranks 10 of NASA’s NEX-GDDP-CMIP6 for Oklahoma under three shared socioeconomic pathways: SSP1-2.6 (low emissions/strong mitigation), SSP2-4.5 (intermediate stabilization/moderate mitigation), and SSP5-8.5 (high emissions/limited mitigation). Two complementary approaches were employed to rank model suitability. The first used traditional statistical metrics aggregated with the technique for order preference by similarity to ideal solution multi-decision-making method to assess how closely models reproduced observed temperature and precipitation. The second applied a Siamese long short-term memory neural network to evaluate temporal similarity between observed and simulated time series, capturing nonlinear patterns and seasonal dynamics that standard statistics may overlook. Together, these approaches address the limitations of using statistical metrics alone, which can be sensitive to daily variability or extreme values, by balancing accuracy with temporal consistency. Results indicated that the SSP5-8.5–CanESM5 model provided the strongest overall performance, followed by SSP2-4.5–FGOALS and SSP5-8.5–FGOALS. The framework not only identifies the most suitable models for pavement design in Oklahoma but also demonstrates a transferable methodology for climate-adaptive infrastructure planning in other regions. DOI: [10.1061/JPEODX.PVENG-2111](https://doi.org/10.1061/JPEODX.PVENG-2111). © 2026 American Society of Civil Engineers.

Introduction

Environmental conditions play a critical role in the design, performance, and long-term performance of pavement. Severe climate conditions—such as extreme temperatures, heavy rainfall, and freeze–thaw cycles—can induce various forms of pavement distress, including cracking, rutting, raveling, and moisture damage, ultimately compromising road safety and functionality (Hasan and Tarefder 2018; Hosseini et al. 2024; Li et al. 2012, 2013; Qiao 2015; You et al. 2015). These challenges are particularly relevant in regions such as Oklahoma, which experience diverse and extreme climatic conditions.

Integrating climate models into pavement design improves the accuracy of long-term performance assessments by accounting for both historical and projected climate data. This allows engineers to anticipate the impacts of temperature changes and precipitation patterns on pavement durability (Heitzman et al. 2011). As a result, climate-informed design leads to more resilient infrastructure that can better withstand changing environmental conditions (Heitzman et al. 2011; Li et al. 2012; You et al. 2015).

Global climate models (GCMs) are essential tools for investigating earth system processes, understanding climate change patterns, and informing adaptation strategies (Almazroui et al. 2021; Krantz et al. 2021; Parding et al. 2020; Pierce et al. 2009). Due to the complexity of climate conditions at a large scale, GCM simulations are associated with uncertainties stemming from natural climate variability, future socioeconomic developments, and differences in climatic modeling approaches (Almazroui et al. 2021; Brunner and Sippel 2023; Parding et al. 2020; Pierce et al. 2022). This has led to a significant body of work focused on evaluating GCM performance and selecting appropriate subsets of models for regional climate-related studies (Krantz et al. 2021; Parding et al. 2020; Pierce et al. 2022). Within CMIP6, future projections are commonly organized using shared socioeconomic pathways (SSPs), which represent alternative emissions trajectories: SSP1-2.6 (SSP126), a low-emission pathway with strong mitigation and substantial reductions in greenhouse gas emissions; SSP2-4.5 (SSP245), an intermediate stabilization pathway with moderate mitigation and midrange emissions; and SSP5-8.5 (SSP585), a high-emission pathway with limited mitigation leading to very high greenhouse gas emissions (Riahi et al. 2011).

¹Ph.D. Student, Dept. of Civil and Environmental Engineering, Oklahoma State Univ., Stillwater, OK 74078. ORCID: <https://orcid.org/0000-0002-6802-3714>. Email: mdelfan@okstate.edu

²Professor, School of Civil and Environmental Engineering, Oklahoma State Univ., Stillwater, OK 74078 (corresponding author). ORCID: <https://orcid.org/0000-0002-2632-7808>. Email: qiang.li@okstate.edu

³Professor, School of Civil Engineering and Environmental Science, Univ. of Oklahoma, Norman, OK 73019. Email: zaman@ou.edu

⁴Research Assistant Professor, School of Civil Engineering and Environmental Science, Univ. of Oklahoma, Norman, OK 73019. ORCID: <https://orcid.org/0000-0001-8552-3377>. Email: syed.a.ali@ou.edu

⁵Associate Professor, School of Civil Engineering and Environmental Science, Univ. of Oklahoma, Norman, OK 73019. Email: jfurtado@ou.edu

Note. This manuscript was submitted on October 13, 2025; approved on April 28, 2026; published online on August 16, 2026. Discussion period open until January 16, 2027; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Transportation Engineering, Part B: Pavements*, © ASCE, ISSN 2573-5438.

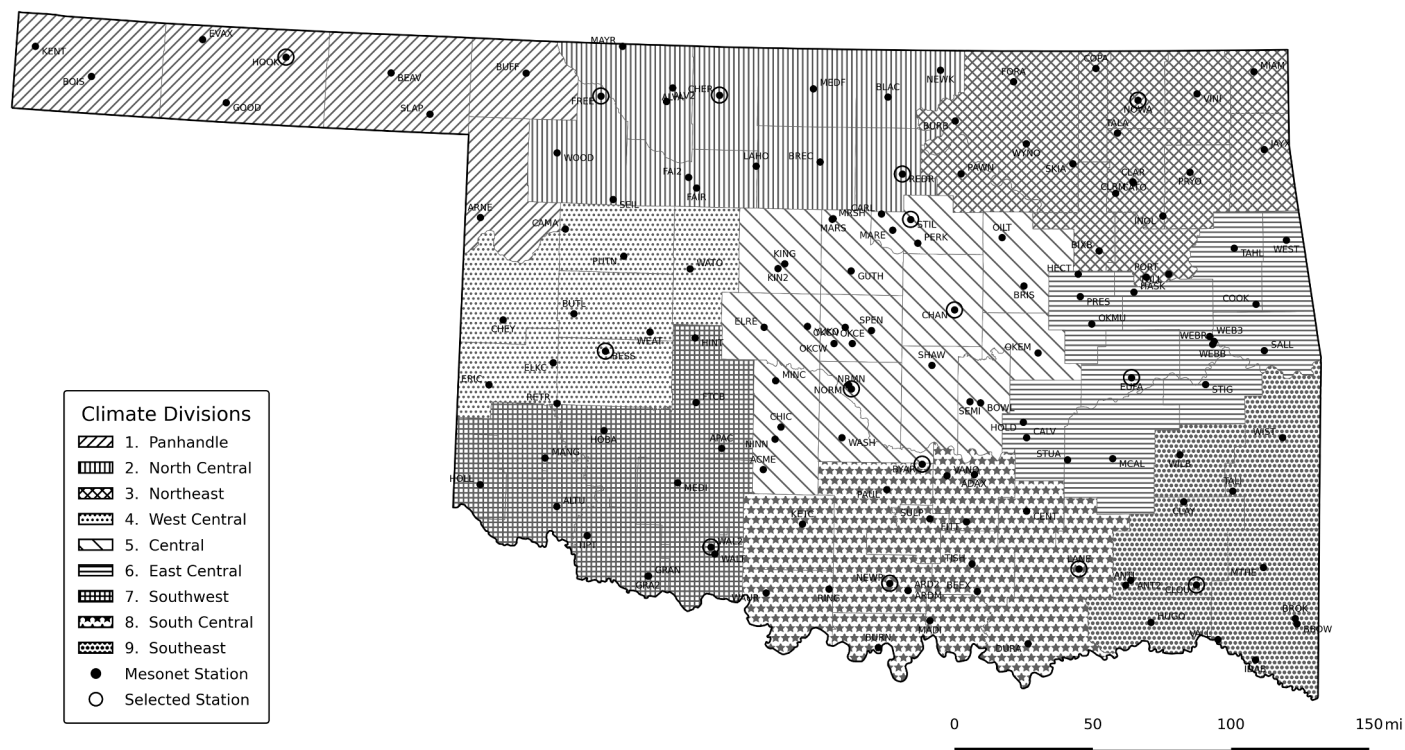


Fig. 1. Climate division of selected locations. (Map by authors.)

Several studies have assessed how well GCMs reproduce historical mean climate and the variability of temperature and precipitation in specific regions, such as the Upper Colorado River Basin (Pierce et al. 2022), the Northeastern US (Karmalkar et al. 2019), or the Indian subcontinent (Rahman and Pekkat 2024). Although CMIP6 performance has been evaluated at continental scales in North America, these broad findings are not necessarily transferable to Oklahoma due to strong regional variability in model skill. Commonly used metrics for climatic evaluation include root mean square error (RMSE), correlation coefficients (R), normalized root mean square error, and skill score (Krantz et al. 2021; Parding et al. 2020; Pierce et al. 2022; Rahman and Pekkat 2024). These statistical metrics cannot capture the nonlinear continuous temporal patterns in climate data variability.

In recent years, deep learning, particularly long short-term memory (LSTM) networks, a type of recurrent neural network (RNN), has gained significant attention in hydrology and climate science (Jeong et al. 2020; Kratzert et al. 2018; Ouma et al. 2022; Yokoo et al. 2022). LSTMs have demonstrated strong performance in capturing long-term dependencies and complex temporal features within time series data (Connor et al. 1994; Hou et al. 2022; Kratzert et al. 2018; Ouma et al. 2022; Tealab 2018). Siamese long short-term memory (S-LSTM) networks have been implemented across climate science, hydrology, environmental monitoring, and transportation systems due to their ability to model long-term dependencies and compare temporal sequences. In climate and weather forecasting, LSTMs have enhanced prediction accuracy for wind speed and rainfall using attention mechanisms and transfer learning (Sha and Guha 2023). S-LSTMs offer key advantages such as automatic feature learning, scalability, robustness to data noise, and efficient parameter sharing, making them a powerful tool for comparative analysis and prediction in climate and environmental systems (Connor et al. 1994; Hou et al. 2022; Kamineni et al. 2018; Tealab 2018; Zhang et al. 2024). There is limited research combining statistical and deep-learning approaches to evaluate GCM performance at a regional scale, and, to our knowledge, no

comprehensive GCM evaluation has been conducted specifically for Oklahoma. This work addresses that gap by providing an Oklahoma-specific, station-based assessment.

Objectives

Before incorporating projected climate data into engineering applications such as pavement design, it is essential to identify the models that best represent regional conditions. Although previous studies have evaluated climate model performance at continental scales in North America (Almazroui et al. 2021; Fan et al. 2020; Thorarinsdottir et al. 2020), their broad findings may not be directly transferable to Oklahoma due to strong regional variability in model skill. In addition, some pavement-focused studies have selected GCMs based on worst-case pavement response rather than agreement with local observations (Hosseini et al. 2024), which supports risk-based design but differs from accuracy-focused climate evaluation. Consequently, a systematic, observation-based CMIP6 assessment tailored to Oklahoma has not been explicitly established, and pavement design practice in Oklahoma still relies largely on stationary historical climate data that may be inadequate for long-term resilience under climate change.

To address this gap, this study evaluates and ranks 10 CMIP6 climate models, with each model analyzed under all three SSPs using Oklahoma Mesonet observations. Two complementary evaluation approaches were applied: (1) traditional statistical metrics aggregated using the technique for order preference by similarity to ideal solution (TOPSIS); and (2) a S-LSTM framework that compares multivariate time series and captures nonlinear temporal behavior. Unlike prior studies that rely mainly on statistical metrics, the S-LSTM provides an additional perspective by evaluating similarity in temporal evolution alongside magnitude-based performance. Although centered on Oklahoma, the proposed framework can be adapted to other regions and climate-sensitive infrastructure applications.

Table 1. Selected climate models

No.	Climate models	Institution/country	Method
1	GFDL-ESM4	Geophysical Fluid Dynamics Laboratory/United States of America	ESM
2	CanESM5	Centre for Climate Modelling and Analysis/Canada	ESM
3	CNRM-ESM2-1	Centre National de Recherches Météorologiques and the European Centre for Advanced Research and Training in Scientific Computing/France	ESM
4	MPI-ESM1-2-HR	Max Planck Institute for Meteorology/Germany	ESM
5	MRI-ESM2-0	Meteorological Research Institute/Japan	ESM
6	NorESM2-MM	Norwegian Meteorological Institute/Norway	ESM
7	EC-Earth3-Veg-LR	A European community Earth System Model/Europe-wide	ESM
8	EC-Earth3		GCM
9	ACCESS_CM2	Community Climate and Earth System Simulator/Australia	GCM
10	FGOALS-g3	Institute of Atmospheric Physics and Academy of Sciences/China	GCM

Climate Data Sources

Two primary climate data sources were utilized for this analysis: high-resolution observational records from the Oklahoma Mesonet network and model projected outputs from NASA's NEX-GDDP-CMIP6 archive. The Mesonet observations were used as the reference benchmark for evaluating the selected climate models. The evaluation period was defined as 2015–2023, which is the longest common overlap between the scenario-based NASA NEX-GDDP-CMIP6 datasets used in this study and the available Mesonet record at the time of analysis. To evaluate the climate models, temperature and precipitation were selected as the primary variables due to their fundamental roles in environmental processes and infrastructure applications. Their accurate representation is critical for credible climate impact assessments (Almazroui et al. 2021; Thrasher et al. 2012; Wood et al. 2004).

Oklahoma Mesonet Field Observed Data

The Oklahoma Mesonet is a high-quality, statewide environmental monitoring network consisting of automated weather stations across the state. Jointly developed by Oklahoma State University and the University of Oklahoma in the early 1990s, the network is now maintained by the Oklahoma Climatological Survey. Its primary purpose is to deliver accurate, high-resolution weather data in near real-time to support applications in agriculture, emergency response, hydrology, and atmospheric sciences.

Each Mesonet station transmits data every 15 min to a central processing center, making it one of the most responsive and comprehensive weather monitoring systems in the United States. Stations are equipped with standardized instruments. The Oklahoma Mesonet exemplifies how localized data collection can inform decision-making and enhance resilience to weather and climate extremes (Mesonet 2025).

In this study, climatic data from fifteen Oklahoma Mesonet stations were selected to provide broad spatial coverage across the state's major climatic variability, while ensuring sufficient data completeness for the 2015–2023 evaluation period. The geographic locations and corresponding Mesonet climate divisions of the selected stations are shown in Fig. 1 (Mesonet 2025).

Global Climate Model Projection Data

This study utilized climate projections from the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) in the Coupled Model Intercomparison Project Phase 6 (CMIP6). These datasets provide high-resolution ($0.25^\circ \times 0.25^\circ$) daily

projections of key climate variables up to the year 2100 and have undergone rigorous bias correction and spatial downscaling, making them well-suited for regional climate assessments (Thrasher et al. 2022).

Ten GCMs were selected from the NEX-GDDP-CMIP6 archive based on their availability of essential parameters for infrastructure applications—particularly those required by the AASHTOWare Pavement ME Design software—as illustrated in NASA's technical documentation (Thrasher et al. 2022). The chosen models represent a diverse set of modeling institutions and include two typical climatic model types: the general circulation models, which simulate large-scale physical processes, and the Earth system models (ESMs) which further incorporate biogeochemical cycles such as vegetation and carbon dynamics. Details of the climate models, their originating institutions, and data derivation methods are provided in Table 1. Each model provides projections under three SSPs, which reflect varying degrees of global mitigation and adaptation strategies to changing climatic patterns.

Climate Models Evaluation

To identify the climate models most suitable for predicting future conditions in Oklahoma, two complementary evaluation approaches were used. Prior to computing all metrics and training the S-LSTM, Mesonet records were quality-screened by converting missing/error markers (e.g., –998, –999) to NaN, aligning both datasets to a common daily time index (2015–2023), and removing any remaining unmatched or missing entries (limited at the selected stations). For spatial matching, each Mesonet station was paired with the NASA NEX-GDDP-CMIP6 time series extracted from the nearest model grid cell to the station latitude/longitude (nearest-neighbor selection).

First, traditional statistical metrics—including the RMSE, mean absolute error (MAE), correlation coefficient (R), uncertainty coefficient (U), Z score, skill score (SS), and Taylor skill score (TS)—were calculated to quantify agreement between model outputs and observed data. The calculated metrics are defined in Eqs. (1)–(8). While these metrics remain essential, they are limited in evaluating overall agreement, without fully capturing sequential dependencies in time series. Previous work emphasized that mean-state and variability metrics can miss important aspects of temporal changes (Gleckler et al. 2008). To address this limitation, S-LSTM networks can effectively capture nonlinear temporal dependencies by learning and comparing latent representations of paired time series (Zhang et al. 2024). This complementary approach evaluates not only error magnitude but also how well models reproduce observed

Table 2. Statistical metrics for temperature

Climate model	SSP	RMSE (°C)	MAE (°C)	U (°C)	R	TS	SS
GFDL-ESM4	SSP126	6.77	5.22	0.17	0.78	0.774	0.66
GFDL-ESM4	SSP245	6.59	5.09	0.17	0.78	0.776	0.70
GFDL-ESM4	SSP585	6.69	5.22	0.17	0.80	0.790	0.59
CanESM5	SSP126	6.97	5.38	0.18	0.78	0.780	0.45
CanESM5	SSP245	6.66	5.17	0.17	0.80	0.794	0.52
CanESM5	SSP585	6.59	5.06	0.17	0.78	0.775	0.64
CNRM-ESM2-1	SSP126	6.52	5.13	0.17	0.80	0.797	0.51
CNRM-ESM2-1	SSP245	6.71	5.28	0.17	0.79	0.790	0.50
CNRM-ESM2-1	SSP585	6.35	5.00	0.16	0.80	0.791	0.63
MPI-ESM1-2-HR	SSP126	6.89	5.36	0.18	0.77	0.766	0.58
MPI-ESM1-2-HR	SSP245	6.86	5.36	0.17	0.78	0.773	0.55
MPI-ESM1-2-HR	SSP585	6.76	5.28	0.17	0.78	0.782	0.57
MRI-ESM2-0	SSP126	6.75	5.25	0.17	0.78	0.776	0.64
MRI-ESM2-0	SSP245	7.22	5.68	0.18	0.78	0.774	0.46
MRI-ESM2-0	SSP585	6.76	5.27	0.17	0.79	0.779	0.61
NorESM2-MM	SSP126	6.34	4.96	0.16	0.80	0.797	0.62
NorESM2-MM	SSP245	6.22	4.77	0.16	0.80	0.800	0.65
NorESM2-MM	SSP585	6.39	4.99	0.16	0.80	0.800	0.62
EC-Earth3-Veg-LR	SSP126	6.80	5.25	0.17	0.77	0.765	0.58
EC-Earth3-Veg-LR	SSP245	6.98	5.54	0.17	0.79	0.785	0.46
EC-Earth3-Veg-LR	SSP585	6.77	5.28	0.17	0.77	0.766	0.70
EC-Earth3	SSP126	6.83	5.32	0.17	0.78	0.779	0.44
EC-Earth3	SSP245	6.86	5.29	0.17	0.77	0.767	0.56
EC-Earth3	SSP585	7.04	5.59	0.18	0.78	0.777	0.40
ACCESS-CM2	SSP126	6.37	4.93	0.16	0.80	0.793	0.66
ACCESS-CM2	SSP245	6.70	5.25	0.17	0.79	0.789	0.58
ACCESS-CM2	SSP585	6.40	4.92	0.16	0.80	0.790	0.64
FGOALS-g3	SSP126	6.36	4.90	0.17	0.79	0.790	0.74
FGOALS-g3	SSP245	6.20	4.77	0.16	0.80	0.798	0.69
FGOALS-g3	SSP585	6.35	4.88	0.16	0.79	0.787	0.68

temporal dynamics. By jointly learning temperature–precipitation patterns across three SSPs, the S-LSTM supports more robust climate-model selection for Oklahoma.

Statistical Metrics

To quantitatively assess the performance of the selected climate models in reproducing observed patterns over Oklahoma, a suite of widely recognized statistical metrics was employed. These metrics were selected for their ability to capture different aspects of model performance ranging from accuracy and consistency to correlation and bias, thus enabling a reliable and comprehensive evaluation.

All metrics were computed using daily data over a 9-year common period (2015–2023) at 15 locations across Oklahoma. Observational Mesonet data served as the reference baseline against which each model's predictions were evaluated under three SSPs (SSP1-2.6, SSP2-4.5, and SSP5-8.5). Each metric captures different aspects of model performance: RMSE emphasizes the magnitude of errors, R highlights linear temporal coherence (but can miss nonlinear dependencies), and the U statistic reflects agreement in amplitude patterns but can be sensitive to outliers and may mask variability or timing errors. In addition to these classical metrics, multivariate methods such as the TS and Z score were calculated as well. TS combines correlation, variability, and error into a single index, while the use of Z scores allows absolute comparison of errors across dissimilar variables, helping to identify the most

suitable climate models (Kwon et al. 2018; Pierce et al. 2022; Taylor 2001). These statistical metrics ensure balanced evaluation and support objective climate model selection for region-specific applications, while they have been used commonly to assess the climate model in previous studies (Almazroui et al. 2021; Karmalkar et al. 2019; Thorarinsdottir et al. 2020)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{\text{pred},i} - y_{\text{obs},i})^2} \quad (1)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_{\text{pred},i} - y_{\text{obs},i}| \quad (2)$$

$$U = \frac{\sum_i^n (y_{\text{pred},i} - y_{\text{obs},i})^2}{\sqrt{\sum_i^n (y_{\text{pred},i})^2} + \sqrt{\sum_i^n (y_{\text{obs},i})^2}} \quad (3)$$

$$R = \frac{\sum_i^n (y_{\text{pred},i} - \bar{y}_{\text{pred}})(y_{\text{obs},i} - \bar{y}_{\text{obs}})}{\sqrt{\sum_i^n (y_{\text{pred},i} - \bar{y}_{\text{pred}})^2} \sqrt{\sum_i^n (y_{\text{obs},i} - \bar{y}_{\text{obs}})^2}} \quad (4)$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (5)$$

Table 3. Statistical metrics for precipitation

Climate model	SSP	RMSE (mm)	MAE (mm)	U (mm)	R	TS	SS
GFDL-ESM4	SSP126	8.77	3.47	0.68	0.02	0.023	0.78
GFDL-ESM4	SSP245	8.57	3.50	0.67	0.03	0.024	0.75
GFDL-ESM4	SSP585	8.15	3.23	0.68	0.00	0.000	0.84
CanESM5	SSP126	7.29	3.20	0.66	0.01	0.012	0.83
CanESM5	SSP245	7.37	3.24	0.66	0.01	0.012	0.82
CanESM5	SSP585	7.26	3.14	0.65	0.05	0.033	0.81
CNRM-ESM2-1	SSP126	8.05	3.24	0.67	0.03	0.026	0.83
CNRM-ESM2-1	SSP245	7.93	3.17	0.67	0.02	0.016	0.85
CNRM-ESM2-1	SSP585	8.30	3.42	0.68	0.01	0.006	0.78
MPI-ESM1-2-HR	SSP126	8.72	3.49	0.68	0.02	0.013	0.79
MPI-ESM1-2-HR	SSP245	8.18	3.24	0.67	0.02	0.020	0.82
MPI-ESM1-2-HR	SSP585	8.30	3.27	0.67	0.02	0.018	0.82
MRI-ESM2-0	SSP126	8.60	3.56	0.68	0.01	0.014	0.76
MRI-ESM2-0	SSP245	7.76	3.03	0.67	0.02	0.018	0.88
MRI-ESM2-0	SSP585	8.39	3.46	0.67	0.01	0.015	0.78
NorESM2-MM	SSP126	8.52	3.51	0.67	0.02	0.025	0.76
NorESM2-MM	SSP245	8.81	3.66	0.67	0.03	0.029	0.71
NorESM2-MM	SSP585	8.51	3.48	0.67	0.03	0.025	0.77
EC-Earth3-Veg-LR	SSP126	8.63	3.51	0.67	0.03	0.025	0.77
EC-Earth3-Veg-LR	SSP245	8.03	3.16	0.67	0.01	0.016	0.82
EC-Earth3-Veg-LR	SSP585	8.81	3.59	0.68	0.02	0.018	0.75
EC-Earth3	SSP126	8.87	3.50	0.69	0.00	0.004	0.80
EC-Earth3	SSP245	9.10	3.67	0.68	0.02	0.017	0.74
EC-Earth3	SSP585	9.05	3.61	0.68	0.01	0.006	0.76
ACCESS-CM2	SSP126	8.47	3.68	0.67	0.01	0.008	0.73
ACCESS-CM2	SSP245	8.25	3.45	0.68	0.00	0.000	0.79
ACCESS-CM2	SSP585	8.13	3.47	0.67	0.01	0.009	0.77
FGOALS-g3	SSP126	7.98	3.55	0.67	0.00	0.000	0.75
FGOALS-g3	SSP245	7.75	3.37	0.66	0.03	0.030	0.78
FGOALS-g3	SSP585	7.69	3.37	0.66	0.03	0.026	0.79

$$TS = 4(1 + R)/[(\sigma_{\text{pred}}/\sigma_{\text{obser}} + \sigma_{\text{obser}}/\sigma_{\text{pred}})^2(1 + R_0)] \quad (6)$$

where “ n ” = total number of data points; “ i ” = time index; “ $y_{\text{pred},i}$ ” = model’s predicted data; “ $y_{\text{obser},i}$ ” = real observed data; $(\bar{y}_{\text{pred}})$ and $(\bar{y}_{\text{obser}})$ = average of observed and predicted parameters over the “ n ” series; “ R ” = correlation coefficient; “ R_0 ” = maximum correlation achievable; and “ σ_{pred} ” and “ σ_{obser} ” = standard deviations of predicted and observed series.

The Z-score method was used to evaluate how well climate models reproduce seasonal observations by normalizing model-observation differences against observed variability (Pierce et al. 2022). A SS derived from Z scores indicates model accuracy, with a score of 1 representing perfect agreement

$$z_{\text{score}}(x, y) = \frac{(\langle \text{Pred}(x, y, t) \rangle - \langle \text{obser}(x, y, t) \rangle)}{\sigma(\text{obser}(x, y, t))} \quad (7)$$

$$ss = 1 - RMSE(z_{\text{score}}) \quad (8)$$

where angle brackets $\langle \rangle$ = temporal averaging of seasonal means; “ σ ” = standard deviation over time; and “ $\text{pred}(x, y, z)$ ” and “ $\text{obser}(x, y, z)$ ” = time series of seasonal mean data in the model and observations, respectively.

For each location, first, different metrics were calculated, and then the values were averaged across all locations. The calculated

metric values are presented in Tables 2 and 3 for temperature and precipitation data, respectively. Note that this evaluation was conducted daily for temperature and precipitation, which show much greater variability than annual averages. As a result, statistical metrics can yield larger error values, and the presence of extremes (especially in precipitation) tends to lower both R and U scores.

TOPSIS Approach

To identify the most suitable climate models for regional analysis in Oklahoma, the TOPSIS was employed. TOPSIS is a widely used multicriteria decision-making method that ranks alternatives based on their geometric proximity to an ideal solution—minimizing the distance from the best-performing option while maximizing the distance from the worst (Behzadian et al. 2012; Tzeng and Huang 2011).

In this study, the alternatives are the various climate models, and the criteria are performance indices used to assess model skill in simulating both climate variables of temperature and precipitation (Raju and Nagesh Kumar 2015; Zamani and Berndtsson 2019). The TOPSIS method following the steps shown in Fig. 2 was used to determine the positive ideal and negative ideal solutions. All TOPSIS criteria were assigned equal weights in the primary analysis (RMSE, MAE, R , TS, and SS each = 0.20), and a sensitivity test with reduced weight for SS (seasonal) produced the same rankings. Relative closeness coefficients were then calculated for

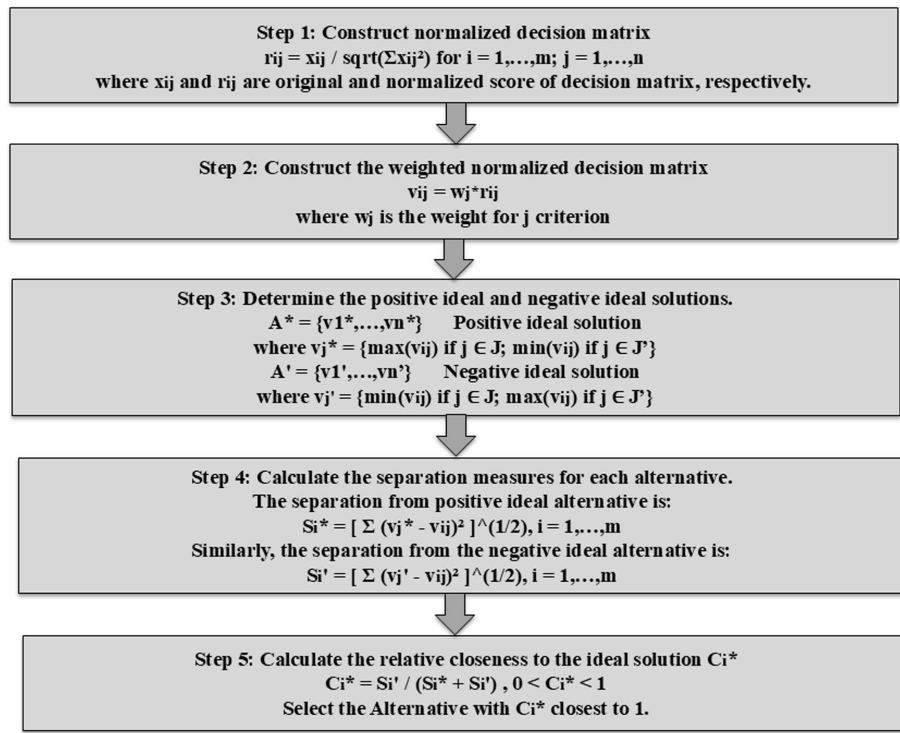


Fig. 2. TOPSIS method procedure.

each model to enable the ranking of models based on their overall performance (Behzadian et al. 2012). Rankings were conducted for a combined assessment of both variables of temperature and precipitation together as shown in Table 4.

Siamese Long Short-Term Memory Deep-Learning Method

LSTM networks are a special type of RNN to capture long-term dependencies in sequences (Costanti et al. 2024; Krantz et al. 2021; Yokoo et al. 2022). They use a cell state and three gates (input, forget, output) to control information flow. When the LSTM block receives the input time series vector $x(s)$ at time “ s ,” it updates the cell state $c(s)$ and generates the hidden states $h(s)$. At the next step ($s + 1$), the hidden states $h(s)$ at time s are replaced into the LSTM block together with the input time series vector $x(s + 1)$. The overall process can be expressed as Eqs. (9)–(13) (Yokoo et al. 2022).

The input gate (g_i) writes new content, the forget gate (g_f) removes irrelevant past information, and the output gate (g_o) exposes part of the cell state as hidden (Ouma et al. 2022; Roy et al. 2022)

$$g_i(s) = \sigma(W_{ii}x(s) + b_{ii} + W_{hi}h(s - 1) + b_{hi}) \quad (9)$$

$$g_f(s) = \sigma(W_{if}x(s) + b_{if} + W_{hf}h(s - 1) + b_{hf}) \quad (10)$$

$$g_o(s) = \sigma(W_{io}x(s) + b_{io} + W_{ho}h(s - 1) + b_{ho}) \quad (11)$$

$$c(s) = g_f(s) \otimes c(s - 1) + g_i(s) \otimes \tanh(W_{ic}x(s) + b_{ic} + W_{hc}h(s - 1) + b_{hc}) \quad (12)$$

$$h(s) = g_o(s) \otimes \tanh(c(s)) \quad (13)$$

where “ x ” = input vector; and “ c ” and “ h ” = cell state and hidden state, respectively. The symbols “ W ” and “ b ” refer to the weight

matrices and bias vectors, which serve as the trainable parameters of the LSTM model. Each weight and bias include two subscripts: the first indicates whether it connects to the input or hidden state, and the second specifies the corresponding gate. The function “ σ ” represents the sigmoid activation, and “ $\otimes$ ” denotes the Hadamard product.

Siamese neural networks employ two identical, weight-sharing branches to learn a similarity (distance) between paired inputs by encoding each input into an embedding and comparing the resulting embeddings with a distance/similarity function (Fig. 3) (Pontes et al. 2018; Shih et al. 2017; Zhang et al. 2024). S-LSTM extends this framework by using LSTMs as the twin encoders to transform sequential inputs (e.g., time series or text) into fixed-length representations, which are then compared to quantify semantic or behavioral similarity (Pontes et al. 2018; Ruan et al. 2022; Shih et al. 2017; Yih et al. 2011). Sharing parameters across the twin LSTMs reduces computational cost and promotes consistent learning (Kamini et al. 2018; Shih et al. 2017).

In our study, we implemented S-LSTM networks to evaluate the similarity between observed climate data from 15 Oklahoma Mesonet stations and daily projections from 10 selected NASA CMIP6 climate models, each under three SSPs. The architecture consists of two identical LSTM branches with shared weights, where each branch receives a multivariate input time series—either observed or modeled daily temperature and precipitation data over the period 2015–2023—and encodes it into a fixed-length feature vector. Before model input, the observational and modeled data are temporally aligned and Z score normalized (per variable) to ensure comparability across stations and variables. The Siamese-LSTM model was trained for a maximum of 50 epochs with early stopping (patience = 5) to prevent overfitting, and batch size of 8 was used. Training was monitored using training/validation loss curves (Fig. 4).

For training, model–observation pairs are labeled using an independent baseline distance computed directly from the

Table 4. Climate models ranking via TOPSIS

Climate model	SSP	Score	Rank
FGOALS-g3	SSP245	0.74	1
FGOALS-g3	SSP585	0.69	2
CanESM5	SSP585	0.66	3
NorESM2-MM	SSP585	0.62	4
CNRM-ESM2-1	SSP126	0.62	4
NorESM2-MM	SSP126	0.60	6
NorESM2-MM	SSP245	0.59	7
CanESM5	SSP245	0.58	8
ACCESS-CM2	SSP585	0.57	9
CNRM-ESM2-1	SSP585	0.55	10
CNRM-ESM2-1	SSP245	0.54	11
ACCESS-CM2	SSP126	0.53	12
FGOALS-g3	SSP126	0.50	13
GFDL-ESM4	SSP585	0.49	14
GFDL-ESM4	SSP245	0.48	15
MPI-ESM1-2-HR	SSP585	0.47	16
EC-Earth3-Veg-LR	SSP245	0.45	17
MPI-ESM1-2-HR	SSP245	0.45	17
CanESM5	SSP126	0.44	19
MRI-ESM2-0	SSP245	0.44	19
MRI-ESM2-0	SSP585	0.43	21
GFDL-ESM4	SSP126	0.42	22
ACCESS-CM2	SSP245	0.41	23
EC-Earth3-Veg-LR	SSP126	0.39	24
MRI-ESM2-0	SSP126	0.36	25
EC-Earth3-Veg-LR	SSP585	0.36	25
EC-Earth3	SSP245	0.29	27
MPI-ESM1-2-HR	SSP126	0.29	27
EC-Earth3	SSP126	0.28	29
EC-Earth3	SSP585	0.19	30

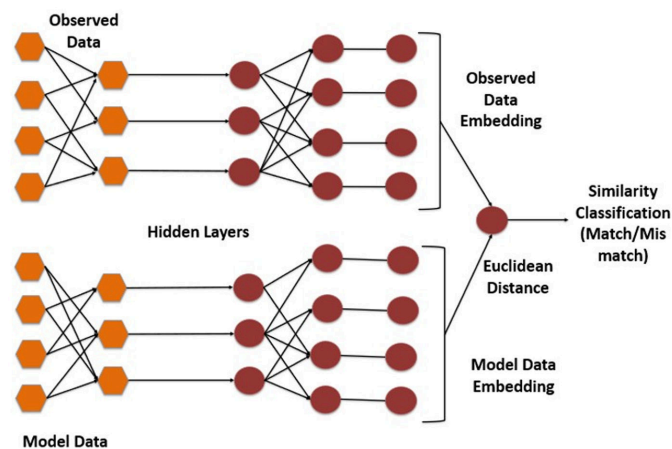


Fig. 3. Structure of the S-LSTM network.

Z-score-normalized input series: the top- k closest pairs are assigned label 1 (similar) and the bottom- k pairs label 0 (dissimilar). This labeling provides supervised contrast for learning, while the S-LSTM similarity is computed from the learned embeddings. The Siamese-LSTM learns long-term temporal dependencies by encoding observed and modeled sequences into a latent representation, and similarity is measured as the Euclidean distance between these embeddings. Lower distance values indicate a better match to the

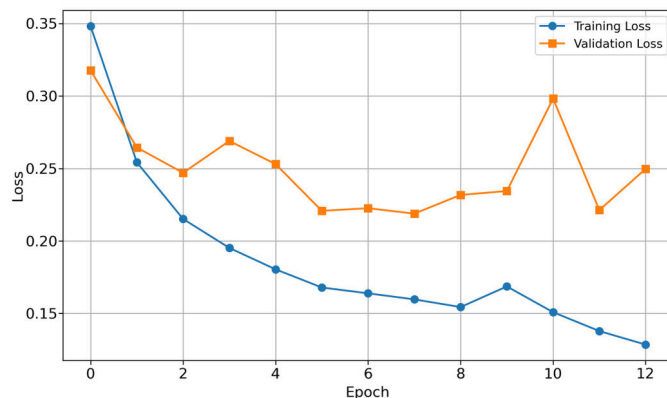


Fig. 4. S-LSTM training/validation loss curve.

observed climate series. The climate models' rankings are summarized in Table 5. This architecture not only preserves the sequential nature of climate data but also enables robust model evaluation by learning high-level temporal patterns rather than relying solely on statistical pointwise comparisons. The shared weights ensure consistent feature extraction across both inputs, and the use of LSTMs provides superior handling of long-term dependencies often found in climate time series. This method complements traditional statistical metrics and offers a novel approach for ranking

Table 5. Climate model ranking via S-LSTM

Climate model	SSP	Distance	Rank
CNRM-ESM2-1	SSP585	0.19	1
CanESM5	SSP585	0.20	2
EC-Earth3-Veg-LR	SSP585	0.25	3
FGOALS-g3	SSP585	0.27	4
FGOALS-g3	SSP245	0.27	4
NorESM2-MM	SSP126	0.29	6
CNRM-ESM2-1	SSP126	0.29	6
MRI-ESM2-0	SSP126	0.30	8
NorESM2-MM	SSP245	0.32	9
EC-Earth3-Veg-LR	SSP126	0.34	10
ACCESS-CM2	SSP585	0.36	11
CanESM5	SSP126	0.37	12
MRI-ESM2-0	SSP585	0.38	13
GFDL-ESM4	SSP126	0.40	14
NorESM2-MM	SSP585	0.42	15
ACCESS-CM2	SSP245	0.42	15
EC-Earth3-Veg-LR	SSP245	0.43	17
CanESM5	SSP245	0.47	18
MPI-ESM1-2-HR	SSP245	0.49	19
FGOALS-g3	SSP126	0.53	20
CNRM-ESM2-1	SSP245	0.54	21
MRI-ESM2-0	SSP245	0.55	22
GFDL-ESM4	SSP245	0.58	23
ACCESS-CM2	SSP126	0.61	24
EC-Earth3	SSP126	0.70	25
MPI-ESM1-2-HR	SSP585	0.77	26
EC-Earth3	SSP245	0.96	27
EC-Earth3	SSP585	1.12	28
GFDL-ESM4	SSP585	1.41	29
MPI-ESM1-2-HR	SSP126	1.49	30

and validating climate models based on their temporal similarity to real observations.

Discussions of Results

This study used two complementary methods to evaluate 10 CMIP6 climate models for Oklahoma and their suitability for climate-informed pavement design. The first method applied statistical metrics (RMSE, MAE, correlation, Taylor skill score, Z score), aggregated with TOPSIS, to rank models against observed temperature and precipitation. Temperature and precipitation were evaluated separately due to their different statistical properties. The second method used a Siamese-LSTM to compare multivariate time series of observed and modeled data. After standardizing both variables, the S-LSTM learned temporal similarity patterns and generated an integrated similarity score. Unlike the statistical approach, it captures nonlinear trends and seasonal dynamics simultaneously. It is important to note that road surface performance can be sensitive to extreme heat and heavy precipitation, and GCMs may differ in their ability to represent these events. In this study, daily-scale model evaluation is conducted, because the projection dataset used in this work provides future information at daily resolution, and our objective is to identify models that best represent baseline regional climate statistics for Oklahoma.

A comparison of the rankings of climate models derived from the S-LSTM and TOPSIS-based evaluations is presented in Fig. 5. The rankings are performance-based, and differences are not attributed to model resolution. Each point represents one model under a SSP, with lower rank values indicating better performance. The red dashed line represents the 45° line of perfect agreement, where a model would receive identical rankings from both evaluation methods. Several key patterns emerge from the plot:

- **High Agreement Zone:** Models such as SSP585–CanESM5, SSP245–FGOALS-g3, and SSP585–FGOALS-g3 are clustered near the lower-left corner and lie close to the 45° line, indicating best rankings and strong agreement between both methods. These models demonstrate consistent performance in terms of both statistical accuracy and temporal pattern similarity, making them top candidates for application in climate-informed infrastructure design.
- **Discrepancy Zone:** Some models, such as SSP585–EC-Earth3, SSP245–EC-Earth3, and SSP126–MPI-ESM1-2-HR, appear in the upper-right region. These models received high (i.e., worse) rankings by both methods. This pattern suggests these models may not capture Oklahoma’s regional climate conditions well.
- **S-LSTM-Favored Models:** A few models such as SSP585–CNRM-ESM2-1 and SSP585–EC-Earth3-Veg-LR received noticeably better rankings from the S-LSTM approach compared to TOPSIS. This implies these models may better capture

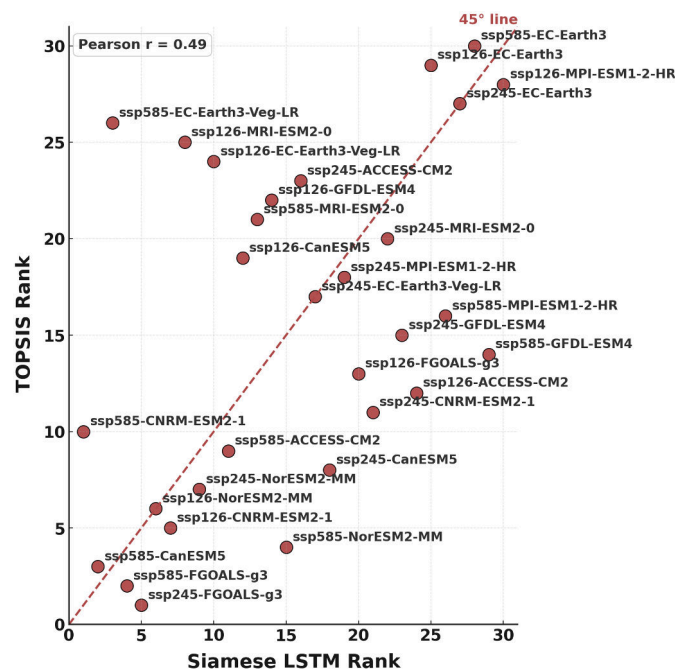


Fig. 5. Comparison of TOPSIS and S-LSTM climate model rankings.

temporal dynamics and intervariable relationships, even if their pointwise accuracy is less optimal.

- **TOPSIS-Favored Models:** Conversely, SSP585–NorESM2-MM and SSP245–CanESM5 performed better under TOPSIS but were penalized by the S-LSTM evaluation, suggesting that their temporal behavior deviates from observed sequences despite producing statistically accurate individual values.

The moderate correlation ($r = 0.49$) suggests that the two methods highlight different but complementary aspects of model performance. The S-LSTM emphasizes temporal structure, while the statistical approach captures pointwise agreement. Models that score well in both—those located near the origin and diagonal—emerge as the most reliable for pavement design, where both magnitude and time-dependent climate behavior are important.

Mechanistic–Empirical Pavement Design Analysis

To assess how the selection of climate-model inputs may influence mechanistic–empirical (ME) pavement predictions, a focused pavement ME sensitivity analysis was conducted. AASHTOWare Pavement ME requires climate inputs at an hourly time step; an hourly weather-generation procedure was applied following Hiruta et al. (2022) to convert the daily climate series into hourly inputs. The analysis was performed for a flexible pavement with climate data for EUFA station, providing a site-specific demonstration of

how differences among climate inputs can translate into differences in predicted distress.

All nonclimatic design inputs (pavement structure, materials, traffic, and design settings) were held constant and were obtained from the long-term pavement performance (LTPP) database (Table 6) (Elkins et al. 2017). Thus, the only factor varied across simulations was the climate dataset. Representative CMIP6 model–scenario inputs were selected from the three best- and worst-performing GCMs identified by state-wide ranking framework in this paper. The selected GCMs included CanESM5–SSP585, FGOALS-g3–SSP245, FGOALS-g3–SSP585, EC-Earth3–SSP245, EC-Earth3–SSP585, and MPI-ESM1-2-HR–SSP126.

The results (Table 7) indicate that climate inputs can meaningfully influence predicted pavement distress, with the most pronounced effect observed for AC thermal cracking. Across the sensitivity runs, thermal cracking consistently controlled the reliability outcome and exhibited substantial variability across the tested climate datasets, including differences relative to the Mesonet-driven case. In contrast, bottom-up fatigue cracking and rutting remained low and showed comparatively minor changes across model inputs. These outcomes support that differences in extreme-temperature behavior can lead to different distress predictions.

Conclusions and Future Work

This study proposed a dual-method framework to select CMIP6 models for climate-resilient pavement design in Oklahoma. Traditional statistical metrics were used alongside a Siamese-LSTM to evaluate accuracy and temporal patterns in daily temperature and precipitation. SSP585–CanESM5 ranked highest, followed by SSP245–FGOALS and SSP585–FGOALS, indicating strong suitability for infrastructure applications. The integration of statistical and deep-learning methods offers a robust, transferable framework for climate model evaluation. These findings suggest that these models are particularly well-suited for supporting future climate projections in engineering applications, including infrastructure planning and pavement design.

Moreover, the study highlights the strength of combining traditional statistical techniques with deep-learning models in evaluating climate datasets. The proposed framework not only enhances model selection for Oklahoma but also provides a transferable methodology that can be adapted to other regions and infrastructure sectors that rely heavily on accurate and locally relevant climate projections. Consistent with this, a focused AASHTOWare Pavement ME sensitivity analysis showed that using different climate inputs can lead to materially different distress predictions—most notably for AC thermal cracking—highlighting the practical importance of selecting locally reliable GCM datasets. Future research should focus on expanding the framework to include additional climatic variables (e.g., humidity, wind speed) and on coupling selected GCMs with mechanistic–empirical pavement design tools

Table 6. Pavement characterization

Input	Layer type	Material type	Thickness (in.)	Values	Modulus (psi)
AADTT	—	—	—	1,268	—
PG	—	—	—	64–22	—
Layer 1	Flexible	Asphalt concrete	10	—	—
Layer 2	No stabilized	A-1-a	10.3	—	40,000
Layer 3	Subgrade	A-6	4	—	14,000

Table 7. Pavement analysis results

Distress type	Target	Mesonet	CanESM5 (SSP585)	FGOALS-g3 (SSP245)	FGOALS-g3 (SSP585)	EC-Earth3 (SSP585)	MPI-ESM1-2- HR (SSP126)	EC-Earth3 (SSP245)
Terminal IRI (in./mile)	172	161.09	158.61	149.24	151.58	154.62	160.89	161.81
Permanent deformation (in.)	0.75	0.27	0.26	0.25	0.25	0.26	0.26	0.26
AC bottom-up fatigue cracking (%lane area)	25	1.45	1.45	1.45	1.45	1.45	1.46	1.45
AC thermal cracking (ft/mile)	1,000	1,590.96	2,425.19	1,399.29	1,672.31	2,011.52	1,606.13	2,769.92
AC top-down fatigue cracking (%lane area)	25	13.54	13.66	13.54	13.54	13.62	13.57	13.62
Permanent deformation—AC only (in.)	0.25	0.11	0.11	0.10	0.10	0.11	0.10	0.10

to directly assess their impact on long-term pavement performance for more locations.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

Mahtab Delfanazari: Data curation; Formal analysis; Methodology; Writing – original draft. Joshua Qiang Li: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing – review and editing. Musharraf Zaman: Conceptualization; Methodology; Writing – review and editing. Ashik Ali: Data curation; Methodology; Validation; Writing – review and editing. Jason Furtado: Conceptualization; Data curation; Writing – review and editing.

References

- Almazroui, M., et al. 2021. “Projected changes in temperature and precipitation over the United States, Central America, and the Caribbean in CMIP6 GCMs.” *Earth Syst. Environ.* 5 (1): 1–24. <https://doi.org/10.1007/s41748-021-00199-5>.
- Behzadian, M., S. Khanmohammadi Otaghsara, M. Yazdani, and J. Ignatius. 2012. “A state-of-the-art survey of TOPSIS applications.” *Expert Syst. Appl.* 39 (17): 13051–13069. <https://doi.org/10.1016/j.eswa.2012.05.056>.
- Brunner, L., and S. Sippel. 2023. “Identifying climate models based on their daily output using machine learning.” *Environ. Data Sci.* 2 (Jan): e22. <https://doi.org/10.1017/eds.2023.23>.
- Connor, J. T., R. D. Martin, and L. E. Atlas. 1994. “Recurrent neural networks and robust time series prediction.” *IEEE Trans. Neural Networks* 5 (2): 240–254. <https://doi.org/10.1109/72.279188>.
- Costanti, F., I. Cappelli, A. Fort, E. G. Ceroni, and M. Bianchini. 2024. “LSTM-based siamese networks for fault detection in meteorological time series data.” In *Proc., 2024 IEEE Int. Conf. on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering*, 906–911. New York: IEEE.
- Elkins, G. E., T. Thompson, B. Ostrom, A. Simpson, and B. Visintine. 2017. *Long-term pavement performance information management system user guide*. Washington, DC: USDOT, Federal Highway Administration.
- Fan, X., Q. Duan, C. Shen, Y. Wu, and C. Xing. 2020. “Global surface air temperatures in CMIP6: Historical performance and future changes.” *Environ. Res. Lett.* 15 (10): 104056. <https://doi.org/10.1088/1748-9326/abb051>.
- Gleckler, P. J., K. E. Taylor, and C. Doutriaux. 2008. “Performance metrics for climate models.” *J. Geophys. Res. Atmos.* 113 (D6): D06104. <https://doi.org/10.1029/2007JD008972>.
- Hasan, M. A., and R. A. Tarefder. 2018. “Development of temperature zone map for mechanistic empirical (ME) pavement design.” *Int. J. Pavement Res. Technol.* 11 (1): 99–111. <https://doi.org/10.1016/j.ijprt.2017.09.012>.
- Heitzman, M., D. Trimm, G. Takle, D. Herzmann, and D. D. Truax. 2011. *Development of climate data input files for the mechanistic-empirical pavement design guide (MEPDG)*. Rep. No. FHWA/MS-DOT-RD-11-232. Jackson, MS: Mississippi DOT.
- Hiruta, Y., N. N. Ishizaki, S. Ashina, and K. Takahashi. 2022. “Hourly future climate scenario datasets for impact assessment of climate change considering simultaneous interactions among multiple meteorological factors.” *Data Brief* 42 (Jun): 108047. <https://doi.org/10.1016/j.dib.2022.108047>.
- Hosseini, F., M. Nasimifar, N. Sivanewaran, and A. Golalipour. 2024. “Mutual impacts of changing climate and flexible pavement performance considering resilience and sustainable aspects.” *J. Cleaner Prod.* 460 (Jul): 142594. <https://doi.org/10.1016/j.jclepro.2024.142594>.
- Hou, J., Y. Wang, J. Zhou, and Q. Tian. 2022. “Prediction of hourly air temperature based on CNN-LSTM.” *Geomatics Nat. Hazards Risk* 13 (1): 1962–1986. <https://doi.org/10.1080/19475705.2022.2102942>.
- Jeong, J., E. Park, H. Chen, K.-Y. Kim, W. Shik Han, and H. Suk. 2020. “Estimation of groundwater level based on the robust training of recurrent neural networks using corrupted data.” *J. Hydrol.* 582 (Mar): 124512. <https://doi.org/10.1016/j.jhydrol.2019.124512>.
- Kaminen, A., M. Shrivastava, H. Yenala, and M. Chinnakotla. 2018. “Siamese LSTM with convolutional similarity for similar question retrieval.” In *Proc., 2018 Int. Joint Symp. on Artificial Intelligence and Natural Language Processing*, 1–7. New York: IEEE.
- Karmalkar, A. V., J. M. Thibeault, A. M. Bryan, and A. Seth. 2019. “Identifying credible and diverse GCMs for regional climate change studies—Case study: Northeastern United States.” *Clim. Change* 154 (3): 367–386. <https://doi.org/10.1007/s10584-019-02411-y>.
- Krantz, W., D. Pierce, N. Goldenson, and D. Cayan. 2021. “Memorandum on evaluating global climate models for studying regional climate change in California.” Accessed August 1, 2025. https://www.energy.ca.gov/sites/default/files/2022-09/20220907_CDAWG_MemoEvaluating_GCMs_EPC-20-006_Nov2021-ADA.pdf.
- Kratzert, F., D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger. 2018. “Rainfall-runoff modelling using long short-term memory (LSTM) networks.” *Hydrol. Earth Syst. Sci.* 22 (11): 6005–6022. <https://doi.org/10.5194/hess-22-6005-2018>.
- Kwon, S.-H., J. Kim, K.-O. Boo, S. Shim, Y. Kim, and Y.-H. Byun. 2018. “Performance-based projection of the climate-change effects on precipitation extremes in East Asia using two metrics.” *Int. J. Climatol.* 39 (4): 2324–2335. <https://doi.org/10.1002/joc.5954>.
- Li, Q., L. N. O. Mills, S. McNeil, and N. Attoh-Okine. 2012. *Exploring impact of climate change on pavement performance and design*. Rep. No. 12-2096. Washington, DC: Transportation Research Board.
- Li, Q. J., L. Mills, S. McNeil, and N. O. Attoh-Okine. 2013. “Integrating potential climate change into the mechanistic-empirical

- based pavement design." *Can. J. Civ. Eng.* 40(12): 1173–1183. <https://doi.org/10.1139/cjce-2012-0465>.
- Mesonet. 2025. "History of the Mesonet." Accessed August 1, 2025. <https://www.mesonet.org/about/history-of-mesonet>.
- Ouma, Y. O., R. Cheruyot, and A. N. Wachera. 2022. "Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: Case study of Nzoia hydrologic basin." *Complex Intell. Syst.* 8(1): 213–236. <https://doi.org/10.1007/s40747-021-00365-2>.
- Parding, K. M., et al. 2020. "GCMeval—An interactive tool for evaluation and selection of climate model ensembles." *Clim. Serv.* 18 (Apr): 100167. <https://doi.org/10.1016/j.cliser.2020.100167>.
- Pierce, D. W., T. P. Barnett, B. D. Santer, and P. J. Gleckler. 2009. "Selecting global climate models for regional climate change studies." *Proc. Natl. Acad. Sci. U.S.A.* 106(21): 8441–8446. <https://doi.org/10.1073/pnas.0900094106>.
- Pierce, D. W., D. R. Cayan, J. Goodrich, T. Das, and A. Munévar. 2022. "Evaluating global climate models for hydrological studies of the upper Colorado River Basin." *JAWRA J. Am. Water Resour. Assoc.* 58(5): 709–734. <https://doi.org/10.1111/1752-1688.12974>.
- Pontes, E. L., S. Huet, A. C. Linhares, and J.-M. Torres-Moreno. 2018. "Predicting the semantic textual similarity with Siamese CNN and LSTM." Preprint, submitted October 24, 2018. <https://arxiv.org/abs/1810.10641>.
- Qiao, Y. 2015. "Flexible pavements and climate change: Impact of climate change on the performance, maintenance, and life-cycle costs of flexible pavements." Ph.D. thesis, Dept. of Civil Engineering, Univ. of Nottingham.
- Rahman, A., and S. Pekkat. 2024. "Identifying and ranking of CMIP6-global climate models for projected changes in temperature over Indian subcontinent." *Sci. Rep.* 14(1): 3076. <https://doi.org/10.1038/s41598-024-52275-1>.
- Raju, K. S., and D. Nagesh Kumar. 2015. "Ranking general circulation models for India using TOPSIS." *J. Water Clim.* 6(2): 288–299. <https://doi.org/10.2166/wcc.2014.074>.
- Riahi, K., S. Rao, V. Krey, C. Cho, V. Chirkov, G. Fischer, G. Kindermann, N. Nakicenovic, and P. Rafaj. 2011. "RCP 8.5—A scenario of comparatively high greenhouse gas emissions." *Clim. Change* 109(1): 33. <https://doi.org/10.1007/s10584-011-0149-y>.
- Roy, D., T. Ishizaka, C. K. Mohan, and A. Fukuda. 2022. "Detection of collision-prone vehicle behavior at intersections using Siamese interaction LSTM." *IEEE Trans. Intell. Transp. Syst.* 23(4): 3137–3147. <https://doi.org/10.1109/TITS.2020.3031984>.
- Ruan, G., D. S. Kirschen, H. Zhong, Q. Xia, and C. Kang. 2022. "Estimating demand flexibility using Siamese LSTM neural networks." *IEEE Trans. Power Syst.* 37(3): 2360–2370. <https://doi.org/10.1109/TPWRS.2021.3110723>.
- Sha, R., and T. Guha. 2023. "Climate time series prediction with deep learning and LSTM." In *Proc., 2023 4th Int. Conf. on Smart Electronics and Communication*. New York: IEEE.
- Shih, C.-H., B.-C. Yan, S.-H. Liu, and B. Chen. 2017. "Investigating Siamese LSTM networks for text categorization." In *Proc., 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conf.*, 641–646. New York: IEEE.
- Taylor, K. E. 2001. "Summarizing multiple aspects of model performance in a single diagram." *J. Geophys. Res.: Atmos.* 106(D7): 7183–7192. <https://doi.org/10.1029/2000JD900719>.
- Tealab, A. 2018. "Time series forecasting using artificial neural networks methodologies: A systematic review." *Future Comput. Inf. J.* 3(2): 334–340. <https://doi.org/10.1016/j.fcij.2018.10.003>.
- Thorarindottir, T. L., J. Sillmann, M. Haugen, N. Gissibl, and M. Sandstad. 2020. "Evaluation of CMIP5 and CMIP6 simulations of historical surface air temperature extremes using proper evaluation methods." *Environ. Res. Lett.* 15(12): 124041. <https://doi.org/10.1088/1748-9326/abc778>.
- Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy. 2012. "Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping." *Hydrol. Earth Syst. Sci.* 16(9): 3309–3314. <https://doi.org/10.5194/hess-16-3309-2012>.
- Thrasher, B., W. Wang, A. Michaelis, F. Melton, T. Lee, and R. Nemani. 2022. "NASA global daily downscaled projections, CMIP6." *Sci. Data* 9(1): 262. <https://doi.org/10.1038/s41597-022-01393-4>.
- Tzeng, G.-H., and J.-J. Huang. 2011. *Multiple attribute decision making: Methods and applications*. Boca Raton, FL: CRC Press.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier. 2004. "Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs." *Clim. Change* 62(1): 189–216. <https://doi.org/10.1023/B:CLIM.0000013685.99609.9e>.
- Yih, W., K. Toutanova, J. C. Platt, and C. Meek. 2011. "Learning discriminative projections for text similarity measures." In *Proc., 15th Conf. on Computational Natural Language Learning*, edited by S. Goldwater and C. Manning, 247–256. Stroudsburg, PA: Association for Computational Linguistics.
- Yokoo, K., K. Ishida, A. Ercan, T. Tu, T. Nagasato, M. Kiyama, and M. Amagasaki. 2022. "Capabilities of deep learning models on learning physical relationships: Case of rainfall-runoff modeling with LSTM." *Sci. Total Environ.* 802 (Jan): 149876. <https://doi.org/10.1016/j.scitotenv.2021.149876>.
- You, Z., X. Yang, J. Hiller, D. Watkins, and J. Dong. 2015. *Improvement of Michigan climatic files in pavement ME design*. Houghton, MI: Michigan Technological Univ.
- Zamani, R., and R. Berndtsson. 2019. "Evaluation of CMIP5 models for west and southwest Iran using TOPSIS-based method." *Theor. Appl. Climatol.* 137(1): 533–543. <https://doi.org/10.1007/s00704-018-2616-0>.
- Zhang, S., Y. Yang, K. Xie, J. Gao, Z. Zhang, Q. Niu, G. Wang, Z. Che, L. Mu, and S. Jia. 2024. "Spatial-temporal Siamese convolutional neural network for subsurface temperature reconstruction." *IEEE Trans. Geosci. Remote Sens.* 62 (Jan): 1–16. <https://doi.org/10.1109/TGRS.2023.3348653>.