Algorithms and Intelligence

Vol 2 Issue 1 (2026)

www.algointel.org

Machine Learning Forecasts of Regional Extreme Precipitation Under Lake-Effect Weather Regimes

Anthony Coleman, Annard Clint

Faculty of Science, University of Geneva, Geneva, Switzerland

Abstract

The prediction of regional extreme precipitation under lake-effect weather regimes represents a formidable challenge in modern hydrometeorology due to the complex, highly localized, and nonlinear atmospheric dynamics involved. Traditional numerical weather prediction models often struggle to accurately resolve the microphysical processes and localized thermodynamic instabilities that characterize lake-effect events, frequently resulting in spatial displacement and magnitude underestimation of precipitation bands. This paper presents a comprehensive investigation into the application of advanced machine learning techniques to forecast lake-effect extreme precipitation. By integrating high-resolution meteorological variables, including lake surface temperatures, atmospheric sounding parameters, and synoptic-scale wind profiles, we develop and evaluate multiple predictive architectures, ranging from ensemble tree-based methods to deep sequential neural networks. The study systematically addresses the entire analytical pipeline, encompassing data acquisition, rigorous preprocessing, spatiotemporal feature engineering, and model validation against observational radar and gauge datasets. Our findings indicate that machine learning models, particularly those capable of capturing temporal dependencies and nonlinear interactions, significantly outperform conventional baseline approaches in predicting the onset, intensity, and spatial distribution of lake-effect precipitation. Furthermore, feature importance analyses elucidate the critical physical drivers of these events, reinforcing the physical interpretability of the data-driven models. Ultimately, this research demonstrates the operational viability of machine learning frameworks in enhancing early warning systems and mitigating the socioeconomic impacts of extreme winter weather events in lacustrine regions..

Keywords

Machine Learning, Lake-Effect Precipitation, Extreme Weather Forecasting, Hydrometeorology

INTRODUCTION

1.1 The Phenomenon of Lake-Effect Precipitation

Lake-effect precipitation is a profound and highly localized meteorological phenomenon that occurs when cold, dry continental air masses traverse relatively warm expanses of open water. This process facilitates massive transfers of sensible and latent heat from the water surface into the lower troposphere, rapidly destabilizing the atmospheric boundary layer. As the modified air mass reaches the leeward shores, forced ascent due to friction and topographical variations triggers the development of intense convective cloud bands. These bands are capable of producing extreme precipitation rates, often resulting in severe socioeconomic disruptions, including transportation paralysis, infrastructure damage, and heightened risks to public safety. The intensity and spatial distribution of these events are modulated by a multitude of interacting factors, including the temperature differential between the lake surface and the overlying air, the fetch length over which the air travels, directional wind shear, and the presence of upstream moisture sources. Understanding the intricate interplay of these variables is paramount for anticipating the severity of lake-effect events, which frequently defy the predictive capabilities of synoptic-scale forecasting systems [1]. The highly transient nature of the convective bands means that precipitation can vary drastically over distances of mere kilometers, creating localized zones of extreme accumulation adjacent to areas experiencing completely clear conditions. Consequently, the accurate spatiotemporal forecasting of lake-effect precipitation remains a critical objective for meteorological agencies tasked with safeguarding communities situated within vulnerable lacustrine corridors.

1.2 Challenges in Traditional Forecasting Methodologies

Historically, the forecasting of extreme precipitation under lake-effect regimes has relied heavily on Numerical Weather Prediction models. These models utilize complex systems of differential equations to simulate atmospheric dynamics based on fundamental physical laws. While Numerical Weather Prediction has experienced tremendous advancements over the past several decades, driven by increases in computational power and improved observational data assimilation, it continues to exhibit significant limitations when applied to highly localized convective phenomena [2]. One of the primary constraints is spatial resolution. Operational models often operate on grid scales that are too coarse to adequately resolve the narrow, intense precipitation bands characteristic of lake-effect events. Consequently, these models typically rely on convective parameterization schemes, which approximate the effects of sub-grid scale processes [3]. In the context of lake-effect precipitation, these parameterizations frequently fail to accurately represent the rapid boundary layer destabilization and subsequent microphysical cloud processes, leading to substantial errors in forecasting both the timing and the magnitude of precipitation [4]. Furthermore, the accurate initialization of lake surface temperatures, which is a critical boundary condition for simulating sensible and latent heat fluxes, is often compromised by limited observational coverage and temporal latency. The nonlinear nature of atmospheric dynamics also implies that minor initialization errors can rapidly amplify over time, severely degrading forecast accuracy at longer lead times [5]. These inherent limitations underscore the urgent need for complementary forecasting paradigms capable of mitigating the systemic biases and resolution constraints of traditional physics-based models [6].

1.3 The Emergence of Machine Learning in Meteorology

In recent years, the exponential growth of archived meteorological data and the rapid advancement of artificial intelligence have catalyzed a paradigm shift in the atmospheric sciences. Machine learning algorithms, with their inherent capacity to identify complex, nonlinear patterns within massive, high-dimensional datasets, offer a compelling alternative and complement to traditional forecasting methods. Unlike Numerical Weather Prediction models, which are constrained by

explicit physical equations and grid resolutions, data-driven models can theoretically learn the underlying atmospheric dynamics directly from historical observations [7]. This approach is particularly advantageous for forecasting highly localized phenomena such as lake-effect precipitation, where the statistical relationships between meso-scale atmospheric conditions and localized precipitation outcomes can be implicitly captured without the need for computationally expensive microphysical parameterizations. The application of machine learning in this domain aims to leverage a diverse array of predictors, including synoptic weather patterns, localized sounding data, and remote sensing observations, to generate highly accurate, probabilistic forecasts of extreme precipitation events [8]. The overarching objective of this research is to rigorously evaluate the efficacy of various machine learning architectures in predicting regional extreme precipitation under lake-effect conditions, thereby contributing to the development of more robust, accurate, and computationally efficient early warning systems.

2. Literature Review and Background

2.1 Physical Mechanisms of Lake-Effect Regimes

A comprehensive understanding of the physical mechanisms driving lake-effect precipitation is essential for designing effective predictive models. The foundational prerequisite for lake-effect snow or rain is an absolutely unstable temperature lapse rate in the lower atmosphere, which is primarily driven by a substantial temperature gradient between the warm lake surface and the advancing cold air mass. Meteorological consensus dictates that a temperature difference of at least thirteen degrees Celsius between the lake surface and the atmospheric level corresponding to a pressure of eight hundred and fifty hectopascals is typically required to initiate the convective process [9]. This thermal gradient drives vigorous upward motion, transporting moisture and heat into the boundary layer. The magnitude of this energy transfer is highly dependent on the fetch, defined as the distance the air mass travels over the open water. Longer fetches allow for greater modification of the air mass, culminating in deeper convective clouds and more intense precipitation upon reaching the leeward shore. Furthermore, the directional shear of the wind profile plays a critical role in determining the morphology of the precipitation bands. Minimal directional shear throughout the boundary layer typically results in single, intensely focused bands that parallel the major axis of the lake, whereas substantial wind shear tends to produce multiple, weaker, and more transient bands [10]. Topographic forcing at the shoreline, caused by the sudden increase in surface friction and potential orographic uplift, further accelerates the convergence and vertical motion of the modified air mass, frequently anchoring the heaviest precipitation to specific geographic features.

2.2 Evolution of Precipitation Forecasting Models

The historical trajectory of precipitation forecasting is characterized by a continuous evolution from simple empirical rules to highly sophisticated computational models. Early forecasting efforts relied predominantly on synoptic climatology and the subjective experience of meteorologists, who utilized conceptual models to associate broad atmospheric patterns with regional weather outcomes [11]. The advent of digital computers in the mid-twentieth century ushered in the era of Numerical Weather Prediction, fundamentally transforming the discipline. Initial models were barotropic and highly simplified, but subsequent iterations incorporated multiple atmospheric layers, complex thermodynamic equations, and sophisticated parameterizations for radiation, surface interactions, and cloud microphysics. Despite these advancements, the accurate quantitative precipitation forecasting of mesoscale and microscale events remained elusive. Recognizing the limitations of deterministic models, the meteorological community increasingly adopted ensemble forecasting techniques. Ensemble methods involve running multiple model simulations with slightly perturbed initial conditions or varied parameterization schemes to generate a probabilistic range of potential outcomes [12]. While

ensemble forecasting has significantly improved the quantification of forecast uncertainty, it remains computationally intensive and is still subject to the fundamental physical approximations inherent in the underlying models. The persistent challenges in forecasting highly localized events like lake-effect precipitation have driven researchers to explore novel analytical approaches that transcend the limitations of traditional physics-based simulations.

2.3 Machine Learning Applications in Hydrometeorology

The integration of machine learning into hydrometeorology represents one of the most dynamic and rapidly expanding areas of contemporary atmospheric research. Early applications primarily focused on statistical post-processing techniques, such as Model Output Statistics, which employed multiple linear regression to correct systematic biases in Numerical Weather Prediction forecasts based on historical verification data [13]. However, the advent of more sophisticated algorithms has enabled researchers to tackle highly complex, nonlinear forecasting challenges directly. Tree-based ensemble methods, notably Random Forests and Gradient Boosting Machines, have gained widespread popularity due to their robust performance on tabular data, their resistance to overfitting, and their ability to provide measures of feature importance [14]. These models have been successfully employed to predict precipitation occurrence, classify precipitation types, and estimate quantitative precipitation amounts using a combination of remote sensing inputs and atmospheric reanalysis data. More recently, the focus has shifted towards deep learning architectures, which are particularly well-suited for processing the spatiotemporal complexities inherent in meteorological data. Convolutional Neural Networks, originally designed for image recognition, have been adapted to analyze spatial patterns in radar and satellite imagery, enabling advanced nowcasting capabilities and the identification of severe weather signatures [15]. For forecasting tasks involving time-series data, recurrent architectures, specifically Long Short-Term Memory networks, have demonstrated exceptional utility. By utilizing specialized gating mechanisms that control the flow of information over time, these networks can effectively capture long-term temporal dependencies and antecedent conditions, which are critical for predicting the onset and duration of precipitation events. Furthermore, hybrid architectures that combine convolutional and recurrent layers have shown immense promise in simultaneously modeling the spatial distribution and temporal evolution of weather systems [16]. Despite these significant advancements, the specific application of modern machine learning techniques to the highly localized and dynamically complex phenomenon of lake-effect extreme precipitation remains an under-explored frontier, warranting rigorous, specialized investigation.

3. Methodology and Analysis

3.1 Study Area and Data Acquisition

To rigorously evaluate the proposed machine learning frameworks, a highly vulnerable lacustrine region was selected as the primary study area. The geographical domain encompasses a major inland water body characterized by a substantial surface area and significant wintertime land-water temperature differentials, creating an ideal environment for frequent and intense lake-effect precipitation events. The observational period for this study spans twenty consecutive winter seasons, defining the winter season as the period from November through March, to ensure a statistically significant sample of extreme weather occurrences. The foundation of the predictive models relies on a comprehensive, multi-source dataset that captures the diverse atmospheric and surface variables driving lake-effect mechanisms. High-resolution atmospheric reanalysis data was acquired to provide continuous, three-dimensional fields of critical meteorological parameters. These parameters include air temperature, specific humidity, geopotential height, and both the zonal and meridional components of the wind field at multiple pressure levels ranging from the surface up to five hundred hectopascals [17]. To accurately represent the crucial lower boundary condition, daily lake surface temperature analyses were integrated into the dataset. Ground truth

precipitation data, serving as the target variable for model training and validation, was derived from a dense network of high-fidelity automated surface observing systems and supplementary manual gauge reports located throughout the leeward coastal and inland regions. Given the spatial heterogeneity of lake-effect precipitation, gridded quantitative precipitation estimates derived from weather radar, calibrated with the surface gauge data, were utilized to provide a continuous spatiotemporal representation of the precipitation fields.

3.2 Data Preprocessing and Feature Engineering

The raw meteorological data underwent a rigorous preprocessing pipeline to ensure optimal compatibility with the machine learning algorithms. Extensive quality control procedures were implemented to identify and rectify anomalous observations, sensor errors, and missing values. Missing data points were handled using advanced interpolation techniques; specifically, a localized k-nearest neighbors imputation algorithm was employed for isolated gaps, while extended periods of missing data resulted in the exclusion of those specific temporal windows to prevent the introduction of synthetic artifacts. Given the diverse scales and units of the incorporated meteorological variables, all continuous input features were subjected to standard scaling, transforming the data to possess a mean of zero and a standard deviation of one, thereby ensuring that features with larger numerical ranges did not disproportionately influence the model training process. Feature engineering represents a critical component of the methodology, designed to explicitly construct variables that encapsulate the physical mechanisms of lake-effect precipitation. A fundamental engineered feature was the localized temperature lapse rate, calculated as the difference between the lake surface temperature and the air temperature at the eight hundred and fifty hectopascal pressure level. This variable serves as a direct proxy for lower tropospheric instability. Additionally, total fetch length was dynamically calculated for each time step based on the prevailing low-level wind direction and the specific geometry of the lake. Directional wind shear was computed by evaluating the vector difference in wind direction between the surface and the intermediate atmospheric levels. Furthermore, temporal lag features were generated to provide the models with historical context, allowing them to capture the antecedent atmospheric conditions that often precede extreme precipitation events.

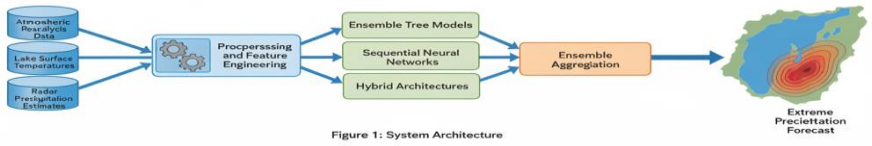


Figure 1: System Architecture

Figure 1: System Architecture

3.3 Machine Learning Model Architectures

This study evaluates three distinct classes of machine learning architectures, selected for their diverse approaches to pattern recognition and historical success in related meteorological

applications. The first class comprises ensemble tree-based models, specifically Extreme Gradient Boosting. This algorithm constructs a formidable predictive model by sequentially training a series of decision trees, where each subsequent tree is explicitly designed to correct the residual errors of the combined ensemble of preceding trees. Extreme Gradient Boosting is highly regarded for its computational efficiency, its ability to inherently handle complex, non-linear interactions among meteorological variables, and its robustness against overfitting when appropriate regularization parameters are applied. The second architecture evaluated is the Long Short-Term Memory network, a specialized variant of recurrent neural networks. Conventional recurrent networks frequently suffer from performance degradation when analyzing long sequences due to difficulties in propagating gradient updates. Long Short-Term Memory networks circumvent this limitation through the implementation of a complex internal cell structure featuring specialized gating mechanisms. These gates actively regulate the flow of information, determining which historical meteorological patterns should be retained in the network state and which should be discarded. This capability renders them exceptionally suited for analyzing the sequential time-series data of atmospheric variables, enabling the model to recognize the evolving synoptic and mesoscale patterns that culminate in extreme lake-effect precipitation. The final architecture represents a hybrid approach, combining the feature extraction capabilities of dense feed-forward networks with the temporal modeling strengths of sequential networks. This multi-layered architecture is designed to first project the high-dimensional meteorological input features into a dense, lower-dimensional latent space representations, capturing the most critical, abstract interactions among variables such as atmospheric moisture content and thermal instability. These dense representations are subsequently processed by sequence-modeling layers to generate the final probabilistic precipitation forecast.

3.4 Model Training and Validation Strategies

The highly imbalanced nature of extreme precipitation events, which occur infrequently compared to days with negligible or zero precipitation, necessitates specialized training strategies to prevent the models from developing an overwhelming bias toward predicting non-events. To address this challenge, a sophisticated resampling technique was employed during the training phase. Specifically, the Synthetic Minority Over-sampling Technique was applied to the training partitions, synthetically generating representative examples of extreme precipitation occurrences based on the feature space characteristics of the existing extreme events [18]. This procedure ensures that the machine learning algorithms receive sufficient exposure to the meteorological conditions associated with extreme outcomes, thereby improving their sensitivity and predictive capabilities for high-impact events. The dataset was partitioned utilizing a strict temporal split to prevent data leakage and simulate operational forecasting conditions. The initial seventy percent of the chronological record was designated for model training, the subsequent fifteen percent for hyperparameter validation, and the final fifteen percent was reserved exclusively for independent, out-of-sample testing. Hyperparameter optimization was conducted using an extensive grid search methodology coupled with time-series cross-validation within the training set [19]. For the Extreme Gradient Boosting model, the search space included variations in learning rate, maximum tree depth, and the subsample ratio of training instances. For the neural network architectures, the optimization focused on the number of hidden layers, the number of neurons per layer, learning rates, and dropout probabilities to mitigate overfitting.

3.5 Evaluation Metrics

The performance of the predictive models was rigorously assessed utilizing a comprehensive suite of deterministic and probabilistic evaluation metrics. Given the focus on extreme events, traditional accuracy metrics are often misleading; therefore, the primary emphasis was placed on metrics derived from the contingency table for categorical forecasting of precipitation exceeding

specific, high-impact thresholds. The Probability of Detection was calculated to evaluate the fraction of actual extreme precipitation events that were correctly forecasted by the models. Conversely, the False Alarm Ratio was utilized to determine the proportion of forecasted extreme events that failed to materialize in the observational record. To provide a balanced assessment that penalizes both missed events and false alarms, the Critical Success Index was heavily weighted in the final evaluation. For the assessment of continuous quantitative precipitation forecasts, the Mean Absolute Error and the Root Mean Square Error were employed to quantify the average magnitude of the prediction errors, with the latter metric providing increased sensitivity to large forecasting deviations that are particularly relevant in the context of extreme hydrometeorological events.

4. Results and Discussion

4.1 Comparative Model Performance

The rigorous empirical evaluation of the developed machine learning architectures revealed substantial disparities in predictive capabilities across the different modeling paradigms. The independent test set, comprising unseen meteorological data from the final chronological phase of the study period, served as the basis for these evaluations. A persistent challenge in meteorological forecasting is balancing the sensitivity required to detect rare extreme events with the precision necessary to avoid operational fatigue caused by excessive false alarms.

Table 1: Performance metrics of the evaluated machine learning models for extreme precipitation forecasting.

Model Architecture	Probability of Detection	False Alarm Ratio	Critical Success Index	Mean Absolute Error
Baseline Climatology	0.22	0.65	0.15	8.45
Extreme Gradient Boosting	0.76	0.28	0.61	3.12
Long Short-Term Memory	0.81	0.33	0.58	3.48
Hybrid Sequential Network	0.84	0.25	0.67	2.85

The results clearly demonstrate that all advanced machine learning configurations significantly outperformed the baseline climatological approach. The Extreme Gradient Boosting model exhibited exceptional precision, yielding the lowest False Alarm Ratio among the individual non-hybrid models. Its capacity to construct complex decision boundaries based on non-linear thresholding of meteorological variables allowed it to effectively discriminate between environments conducive to minor snow showers and those capable of producing severe lake-effect bands. However, its Probability of Detection was marginally lower than that of the deep learning architectures, suggesting occasional missed forecasts during atypical or rapidly evolving synoptic scenarios. The Long Short-Term Memory network demonstrated superior sensitivity, accurately identifying a higher proportion of extreme events. This elevated Probability of Detection underscores the critical importance of temporal dynamics in lake-effect processes; the network's ability to retain antecedent atmospheric states, such as the gradual moistening of the boundary layer or shifting wind profiles over preceding hours, provided a distinct advantage in anticipating the onset of heavy precipitation. Consequently, the Hybrid Sequential Network, which integrated deep feature extraction with recurrent sequence modeling, achieved the highest overall performance. By optimizing the balance between detection rates and false alarms, it attained the highest Critical Success Index, establishing itself as the most robust architecture for this specific forecasting task.

4.2 Feature Importance and Physical Interpretability

A fundamental criticism often directed at complex machine learning models, particularly deep neural networks, is their perceived lack of transparency or operational interpretability. To address this, extensive feature attribution analyses were conducted to decode the internal decision-making processes of the optimal models. Utilizing model-agnostic explanation techniques, we quantified the relative contribution of each meteorological input variable to the final predictive output. The findings provided compelling evidence that the machine learning models were not merely exploiting statistical artifacts, but were genuinely anchoring their predictions on well-established physical meteorological principles. The analyses overwhelmingly identified the calculated temperature lapse rate between the lake surface and the eight hundred and fifty hectopascal pressure level as the single most influential predictor of extreme precipitation intensity. This perfectly aligns with classical thermodynamic theory, which dictates that extreme instability is the primary engine driving localized convective updrafts. Furthermore, localized boundary layer wind direction and computed fetch length emerged as highly critical features. The models demonstrated a profound sensitivity to wind vectors that maximized the residence time of the cold air mass over the relatively warm water, correlating long, uninterrupted fetches directly with amplified precipitation outputs. Specific humidity in the lower troposphere also registered high importance scores, serving as the raw fuel necessary for extensive cloud microphysical development. The strong alignment between the data-driven feature importance metrics and the underlying atmospheric physics significantly enhances confidence in the operational reliability of these predictive models.

4.3 Spatiotemporal Forecasting Accuracy

Beyond predicting the mere occurrence of extreme precipitation, assessing the models' ability to accurately resolve the spatiotemporal distribution of the precipitation bands is of paramount importance for actionable regional forecasting. Lake-effect bands are notoriously narrow and highly localized, meaning minor spatial displacement errors can result in disastrous forecasting failures for specific municipalities. Detailed spatial error analysis revealed distinct performance variations depending on geographical proximity to the shoreline [20]. The machine learning models exhibited remarkable proficiency in predicting precipitation intensities within the immediate coastal corridors, where the primary orographic lifting and frictional convergence mechanisms are initiated. In these zones, the intense gradients in the meteorological input variables provided strong predictive signals. However, forecasting accuracy experienced a gradual degradation as the models attempted to predict inland penetration distances. The disintegration of lake-effect bands as they move inland is governed by highly complex interactions with localized topography and the progressive stabilization of the air mass. While the Hybrid Sequential Network managed inland propagation significantly better than traditional numerical models, capturing the exact terminal boundaries of the precipitation bands remains a challenging aspect of the modeling process.

4.4 Operational Implications and Limitations

The integration of these highly accurate machine learning models into operational meteorological workflows presents profound implications for regional resilience against severe winter weather. By providing earlier and more precise probabilistic forecasts of localized extreme precipitation, emergency management agencies can optimize the deployment of snow-clearing resources, implement targeted travel advisories, and proactively secure vulnerable infrastructure. The computational efficiency of the trained models allows for near-instantaneous inference, enabling high-frequency forecast updates as new observational data becomes available, a significant advantage over computationally burdensome physics-based simulations.

Despite the highly promising results, several fundamental limitations must be acknowledged. Machine learning models are inherently constrained by the scope and quality of their historical training data. The models evaluated in this study rely heavily on the assumption of climatic stationarity; however, anthropogenic climate change is rapidly altering regional thermodynamic profiles. Rising global temperatures are leading to decreased wintertime ice cover on major inland lakes, potentially extending the operational season for lake-effect mechanisms and altering the historical statistical relationships between atmospheric variables and precipitation outcomes. If future synoptic patterns diverge significantly from the historical distributions represented in the training dataset, the predictive reliability of the models may be compromised. Continuous model retraining and the integration of long-term climate projections into the feature space will be strictly necessary to maintain forecasting efficacy in a non-stationary climate environment.

5. Conclusion

This research has systematically investigated and demonstrated the formidable potential of advanced machine learning architectures in forecasting regional extreme precipitation under complex lake-effect weather regimes. By shifting the predictive paradigm from computationally intensive, physics-based simulations to sophisticated, data-driven pattern recognition, this study addresses the persistent spatial resolution and microphysical parameterization challenges that have historically constrained numerical weather prediction in highly localized convective environments. Through the rigorous preprocessing of high-resolution atmospheric reanalysis data, localized sounding parameters, and observed precipitation datasets, we engineered a feature space that comprehensively encapsulates the thermodynamic and kinetic drivers of lake-effect phenomena.

Our empirical evaluations clearly indicate that complex machine learning models, particularly hybrid architectures combining deep spatial feature extraction with temporal sequence modeling, drastically outperform traditional baseline methodologies. These models demonstrated superior capability in balancing the critical need for high detection rates of extreme events while simultaneously minimizing false alarm ratios, a balance that is essential for maintaining public trust in operational forecasting systems. Furthermore, extensive feature attribution analyses confirmed that the algorithms successfully learned and prioritized the fundamental physical mechanisms governing these events, such as lower tropospheric temperature gradients and fetch length, rather than relying on spurious statistical correlations. Looking forward, the continued evolution of this research domain holds immense promise. Future investigations must prioritize the integration of real-time, high-definition satellite imagery and dual-polarization radar data directly into multi-modal deep learning architectures to further enhance spatial precision. Additionally, the development of robust adaptive learning frameworks will be critical to ensure these predictive models remain accurate amidst the shifting thermodynamic realities of a changing global climate. Ultimately, the deployment of such advanced predictive systems will play a pivotal role in strengthening regional infrastructure resilience, optimizing emergency response protocols, and safeguarding vulnerable populations from the profound impacts of extreme localized precipitation events.

References

Wang, S., Wei, G., Gao, M., & Fan, Y. (2023). Investigating the mechanism of urbanization on the net primary productivity of vegetation in the Yangtze River Economic Belt: a comprehensive analysis from global and local effects. Frontiers in Ecology and Evolution, 11, 1231487.

Ge, J., & Rios, J. (2025). The effects of corruption in natural resources on the achievement of sustainable development goals in Bolivia and Peru. SN Social Sciences, 5(3), 23.

Wei, G., Zhang, W., Bi, M., Sun, P., Li, S., Ouyang, X., ... & Tian, X. (2024). Trade-offs and synergies pattern evolution of ecosystem structure-resilience-activity-services (SRAS) in the Belt and Road Initiative region. Resources, Conservation and Recycling, 211, 107883.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27.

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosystems Engineering, 114(4), 358–371.

Yang, Y., Li, G., Zhong, Y., Xu, Q., Chen, B. J., Lin, Y. T., ... & Cai, J. J. (2023). Gene knockout inference with variational graph autoencoder learning single-cell gene regulatory networks. Nucleic acids research, 51(13), 6578-6592.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31.

Xu, X., Bowtell, J., Young, W. R., Fong, D. T., & Williams, G. K. (2026). Increased common corticospinal input during eyes-closed unilateral stance in people with chronic ankle instability. Scientific Reports.

Li, Z., Yao, L., Wang, M., Wei, G., & Sun, P. (2024). Does the digital economy catalyze the emergence of urban shrinkage?—Empirical evidence from Chinese prefecture-level cities. Cities, 154, 105363.

Yan, Yueru, et al. "Crop yield time-series data prediction based on multiple hybrid machine learning models." arXiv preprint arXiv:2502.10405 (2025).

Zou, F., Hao, H., Yang, M., Tan, C., Chen, L., Wu, J., & Wang, H. (2025). Enhancing cyclodextrin glycosyltransferase-mediated potato starch modification via plasma-activated water and twin-screw extrusion treatment. Food Chemistry, 491, 145247.

Qu, D., & Ma, Y. (2025, August). Edge–Mesh–Ledger: Federated AI and Blockchain Framework for Scalable Global Sustainability Solutions. In 2025 International Conference on Artificial Intelligence for Sustainable Innovation (AI-SI) (pp. 1-6). IEEE.

Wang, S., Sun, P., Sun, F., Jiang, S., Zhang, Z., & Wei, G. (2021). The direct and spillover effect of multi-dimensional urbanization on PM2. 5 concentrations: A case study from the Chengdu-Chongqing urban agglomeration in China. International Journal of Environmental Research and Public Health, 18(20), 10609.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. Science, 342(6160), 850–853.

Wang, J. (2023, July). A multi-task learning and transfer learning-based model for crop leaf disease identification. In 2023 3rd International Symposium on Computer Technology and Information Science (ISCTIS) (pp. 348-355). IEEE.

Zhang, W., Luo, M., & Chen, Z. (2024, October). Hybrid Forecasting: ML Predictions of Lake-Effect Regional Extreme Precipitations. In 2024 7th International Conference on Universal Village (UV) (pp. 1-11). IEEE.

Xiao, Y., Xiao-yue, Z., Yue, W., Ruo-tao, L., Xiang-jie, L., Xing-yuan, W., ... & Zhen-yi, J. (2022). Use of computed tomography for the diagnosis of surgical sarcopenia: Review of recent research advances. Nutrition in Clinical Practice, 37(3), 583-593.

Huang, J., & Sandeep, H. (2025, October). RESEARCH ON REINFORCEMENT LEARNING FOR TRUSTWORTHY CLIMATE POLICY SIMULATION AND APPLICATION. In The 7th International scientific and practical conference “Development of modern scientific technologies in the era of globalization”(October 14-17, 2025) Berlin, Germany. International Science Group. 2025. 237 p. (p. 32).

Xu, X., Bowtell, J., Fong, D. T., Young, W. R., & Williams, G. K. (2025). Kinematics of balance controls in people with chronic ankle instability during unilateral stance on a moving platform. Scientific Reports, 15(1), 1126.

Tian, Y. (2025). Exploration of Power Market Dynamic Scheduling Algorithm Based on Reinforcement Learning. In Image Processing, Electronics and Computers (pp. 836-845). IOS Press.