IRSTI 37.23.31 https://doi.org/10.26577/jpcsit4120264

Zh. Baishemirov[1, 2, 3], D. Ospanova[2*], B. Amirgaliyev[2], S. Mukhambetzhanov[4]

[1]Kazakh-British Technical University, Almaty, Kazakhstan
[2]Astana IT University, Astana, Kazakhstan
[3]Narxoz University, Almaty, Kazakhstan
[4]Al-Farabi Kazakh National University, Almaty, Kazakhstan
*e-mail: 242982@astanait.edu.kz

COMPARATIVE ANALYSIS OF PHYSICS-INFORMED AND CONVENTIONAL LSTM AND RNN MODELS FOR TEMPERATURE FORECASTING USING ERA5 REANALYSIS DATA

Abstract. Climate change is one of the most serious modern problems affecting the Earth's atmosphere, as it causes a range of harmful effects worldwide. Due to the uneven nature of climate data, forecasting climate change is a challenging task today. Many previous studies in climate and machine learning have used recurrent neural networks (RNNs) and long short-term memory (LSTM) models to predict climate trends. Although these models are effective at identifying long-term trends in data, they often fail to satisfy physical laws such as energy conservation, mass balance, and thermodynamic principles. In this research, the aim was to develop oscillation-constrained RNN and LSTM models, in which an annual harmonic prior is incorporated into the loss function, and compare their performance with standard RNN and LSTM models. The study utilized data on air temperature at 2 meters above the surface for the cities of Astana, Almaty, and Shymkent for model training, validation, and testing. According to the results, physics-informed models achieved the lowest root mean square errors in Almaty (3.52 °C) and Shymkent (3.80 °C). RNN and LSTM models performed better in Astana (RMSE = 5.44 and 5.47 °C), where seasonal changes are relatively abrupt. These findings demonstrate that PINNs constrained by the annual harmonic oscillation provide more physically consistent forecasts for moderate climates, while conventional recurrent models perform better in locales with highly variable conditions.

Keywords: climate prediction, Physics-Informed Neural Networks, Long Short-Term Memory, Recurrent Neural Network, ERA5 reanalysis, temperature forecasting, Numerical Weather Prediction.

1. Introduction

The problem of climate change remains one of the greatest and most challenging issues in modern Earth system science, due to the chaotic and non-intuitive nature of the Earth's atmosphere. It is chaotic to an extent, where even the smallest change can yield an unexpected outcome over time. According to [1], global warming as a result of climate change can take place suddenly in a few decades, or even in a few years, in the form of a climate shock. As a result, weather prediction becomes highly unreliable, especially when events last more than a few days. The problem is exacerbated even further when the scale of the entire climate system is taken into account. Some events may happen in the span of a few minutes, while other events may take as long as a few decades to unfold. For instance, earlier research indicates a permanent shift of the Intertropical Convergence Zone towards the warmer hemisphere as a result of disturbances in the interhemispheric asymmetry [2]. Climatic events work in harmony, and disturbances at local scales can trigger responses at the global level.

Another problem encountered by researchers worldwide is data scarcity. Highly accessible locales account for the majority of the data, while less accessible locales, such as the Arctic, deep-ocean basins, and areas in developing countries, account for just a minuscule amount of data. For example, although Polar regions have an impact on global temperatures, they are not well studied because of their extreme conditions [3]. This leads to an incomplete picture of the global climate. As a result, a systematic bias is introduced into AI models and algorithms trained on these datasets.

Some weather prediction methods that were established in the past are outdated in the current meteorological scenario. While these traditional approaches have contributed greatly to humanity's progress in climatology, they are now incompatible with the scales of current climatic events. For instance, global climate models possess horizontal limits of 70 to 250 kilometers [4], which could be insufficient in encompassing numerous individual clouds, urban heat islands, and local topographic effects.

Hardware limitations of past methods have forced researchers to extrapolate data to a larger area, a process that can introduce statistical errors. Data extrapolation using AI and neural network approaches can lead to results that are not compatible with fundamental physical principles such as energy conservation, mass balance, or thermodynamics [5]. Therefore, predictive models that go beyond fundamental physics are prone to produce implausible and unreliable results. The following observation is crucial in the onset of PINN or Physics-Informed Neural Networks, which are especially practical in their compatibility with predefined physical datasets [6]. As introduced by Raissi, PINNs are neural networks designed to learn while adhering to physical laws governed by general nonlinear partial differential equations [7]. This study compares PINN efficiency in the prediction of air temperature with traditional Long Short-Term Memory and Recurrent Neural Network models. We hypothesize that incorporating physical laws via PINNs will result in more reliable and physically plausible forecasts of weather events compared to conventional approaches.

The use of PINNs in long-term climate frameworks to increase weather forecast accuracy over long time horizons is what makes this research novel. Modern literature on climate prediction focuses mainly on short-term or localized frameworks, with little effort in applying PINN models. The study uses the ERA5 reanalysis dataset to validate this approach.

It is expected that PINNs will improve interpretability and forecast reliability for city-level climate datasets. Other climatic variables, such as precipitation, humidity, and wind speed, can also be predicted later, if this approach proves effective. This would help establish a unified system for long-term climate forecasting and environmental monitoring.

2. Literature Review

Numerical weather prediction (NWP) is a widely used concept in atmospheric studies. NWP models simulate future meteorological events by solving primitive equations of motion numerically [8]. The main advantage of NWP models is their physically based representations of fluid dynamics, thermodynamics, and radiative transfer. As a result, numerical models generate high-quality short-term weather prognoses for up to about one week [9]. Nevertheless, such models have high computational costs and limited resolution. Moreover, they overdepend on sub-grid processes such as convection and cloud microphysics, and accumulate errors rapidly due to chaotic dynamics [10]. Therefore, it is difficult to apply this model for long-term climate forecasts. Researchers employ Regional Climate Models (RCMs) along with dynamical downscaling to tackle this problem. This technique provides better spatial detail by using finer grids inside global models. While RCMs are effective, they require greater computational resources than global NWP methods and are impractical for frequent or long-term use [11].

To overcome these challenges, researchers developed statistical downscaling techniques. These methods can find relationships between large-scale and local climate variables [12]. Compared to dynamical downscaling, statistical downscaling provides more global detail and better computational efficiency. However, it depends on reliable and consistent observational records, which may become problematic as the climate changes [13,14].

The advent of machine learning (ML) algorithms has supplied the scientific world with another approach for climate prediction. 1998 marked the development of the first neural network models working with short-term precipitation prediction. In later years, non-deep learning methods were also introduced to enhance medium and long-term precipitation predictions. Models such as CGF, CycleGANs, DeepESD, and NNCAM have achieved superior accuracy compared to conventional physical models. They excelled in capturing temporal and spatial climate patterns, refining resolution, and ultimately reducing computational time [15]. Machine learning has also found its use in climate science in miscellaneous tasks such as tropical cyclone tracking, cloud classification, and air quality predictions [16 – 18].

The gradual growth of computational power and increasing scalability of climate datasets have led scientists to focus on more innovative models for capturing spatial and temporal dependencies in climate processes. Several deep learning models, such as Convolutional Neural Networks (CNN) have been employed to perform image super-resolution for enhancing coarse-resolution climate model outputs [19], while learning models such as

Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) were used in capture of temporal dependencies in climate and hydrological time series, as demonstrated by [20]. LSTM models effectively capture temporal trends and patterns, which enable accurate long-term predictions of noisy climate datasets. Transformer-based architectures have also proven to be useful in modeling long-range spatiotemporal interactions. For instance, Ramu et al. (2022) proposed a Transformer-based model for daily temperature forecasting, integrating a Spatial-Temporal Fusion Module, Hierarchical Graph Representation, and a Dynamic Temporal Graph Attention Mechanism to capture spatiotemporal dependencies and improve temporal feature extraction [21].

Although deep-learning algorithms offer many benefits, they often struggle to be physically consistent [22]. Slater et al. (2023) observe that data-driven models are usually not very good at forecasting extreme or novel events that have never occurred before and were not contained in historical data [23]. Besides, these models face the difficulty of optimizing high-dimensional multivariate outputs. These issues of deep learning algorithms have led to the development of Physics-Informed Neural Networks. PINNs use physical laws together with training data, which helps them make accurate predictions even when data is limited or sparse. PINNs help identify the main dynamics of a system and ensure the equations used remain consistent. This gives them an advantage over models that rely only on data, such as RNNs or LSTMs [6].

According to Feng et al. (2023), predictions in the data-sparse areas can be made more robust by imposing physical constraints [24]. Other studies show that combining physical models with LSTM improved accuracy in the validation period and reduced uncertainty in future flood forecasts [25]. PGnet, a physics-based deep learning model, improves tropospheric temperature predictions by using physical principles with generative neural networks and a guiding mask to enhance low-quality prediction areas [26]. However, most studies still focus on specific fields or small-scale cases, often looking at isolated physical processes rather than complex systems. Moreover, the ability of models to generalize data under the influence of changing climate conditions and scarce data is still insufficiently examined. Few studies discuss the scalability of PINNs for large-scale climatic conditions and high-dimensional datasets, as well as their integration into long-term monitoring frameworks. Such research gaps reflect the need for further studies on the application of PINNs in large-scale, data-driven climate prediction and environmental monitoring.

3. Materials and Methods

3.1 Dataset Selection and Preprocessing

ERA5 reanalysis dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF) was used in this study to train the models. The dataset integrates information from satellites, ground stations, and other observation systems. It was chosen for this research because it offers high spatial and temporal resolution. ERA5 also provides long-term, globally consistent, and reliable data.

In detail, we adopted the post-processed ERA5 daily statistics at single levels for 2008-2022. Air temperature at 2m above the land and sea surfaces was selected as the predicted variable. The dataset was divided into three parts: 2008–2018 for training, 2019–2020 for validation, and 2021–2022 for testing. All input variables were normalized to zero mean and unit variance using statistics computed from the training set, and these values were subsequently applied to the validation and test subsets.

Overlapping sliding windows with a stride of one day were then formed independently within each subset, ensuring that no input sequence or forecast horizon contained information from future periods. Using these sequences, a direct multi-step forecasting strategy was employed, where each input sequence of 365 consecutive daily temperature observations was used to predict the subsequent 90 daily temperature values in a single forward pass. This approach avoids recursive error accumulation associated with autoregressive inference and is well suited for seasonal and sub-seasonal temperature forecasting tasks.

Regional comparison was performed using climatic data from Astana, Almaty, and Shymkent, three major cities in Kazakhstan. The choice is also justified by uniquely distinct thermal footprints each city possesses, with continental sharp cold and seasonal contrasts in Astana, moderate and humid mountainous climate in Almaty, and warm semi-arid climate with relatively mild winters in Shymkent. Diversity of climates makes it possible to thoroughly evaluate how models perform in different climate conditions within one country.

Table 1. Summary of 2m temperature statistics by city

City	Coordinates	Mean (°C)	Std (°C)
Astana	latitude = 51.1694, longitude = 71.4491	3.37	14.60
Almaty	latitude = 43.238949, longitude = 76.889709	5.55	10.60
Shymkent	latitude = 42.3417 longitude = 69.5901	13.56	10.98

3.2. Models

3.2.1. LSTM

For this study, a Long Short-Term Memory (LSTM) neural network was employed to model and predict temporal variations in 2-meter air temperature. LSTM networks are a type of recurrent neural network (RNN) that can learn long-term patterns in sequential data. Hence, they are suitable for time-series forecasting tasks, such as climate and weather prediction.

Mean squared error (MSE) as the loss function and the Adam optimizer with a learning rate of 0.001 and 200 epochs were used in the model training process. To avoid overfitting and enhance generalization, we used early stopping and dropout regularization.

Table 2. Hyperparameters of the LSTM model

Hyperparameter	Value
Optimizer	Adam
Learning rate	0.001
Number of layers	2
Neurons per layer	64
Dropout	0.1
Number of epochs	200
Early stopping	Yes

3.2.2. RNN

Unlike the LSTM, which uses memory cells and gates, the vanilla RNN employs simple recurrent connections to capture short-term temporal dependencies in sequential data. Though more prone to vanishing gradients and worse at capturing long-term dependencies, the RNN is a lighter and simpler baseline computationally for time series forecasting applications such as weather and climate prediction.

The RNN model was also optimized with the Adam optimizer, using a learning rate of 0.001. There were 150 epochs of training. The RNN architecture offers a simpler approach to temperature forecasting than the LSTM architecture. Other model training parameters are listed in Table 3.

Table 3. Hyperparameters of the RNN model

Hyperparameter	Value
Optimizer	Adam
Weight initialization	Random
Learning rate	0.001
Number of layers	2
Neurons per layer	64
Dropout	0.1
Number of epochs	150
Early stopping	Yes

3.2.3. Physics-Informed models

While inspired by physics-informed approaches, the oscillation-constrained models do not encode full thermodynamic principles or conservation laws. The physics term used here reflects only the regular annual oscillation of temperature.

The architecture of the PINN models mirrors the baseline structures, consisting of two hidden layers with 64 neurons each, followed by a fully connected dense output layer. The tanh activation function was used for hidden units to capture nonlinear temperature dynamics, and a dropout rate of 0.1 was applied to reduce overfitting. The total loss function combines a data-driven loss with a physics-based loss term:

$$L_{total} = \lambda_{data} L_{data} + \lambda_{phys} L_{phys} \quad (1)$$

where L_{data} is the mean squared error between observed and predicted temperatures, and L_{phys} can be defined as [27]:

$$L_{phys} = \frac{1}{N}\sum_{i=1}^{N} (u''(t_i) + \omega^2 u(t_i))^2 \quad (2)$$

where $\omega = \frac{2\pi}{T}$ is the angular frequency and T is the period. In this study, T corresponds to one year, giving $\omega = \frac{2\pi}{365}$. This formulation captures annual periodicity but does not explicitly model subannual variability, phase shifts between locations, or thermodynamic energy exchanges.

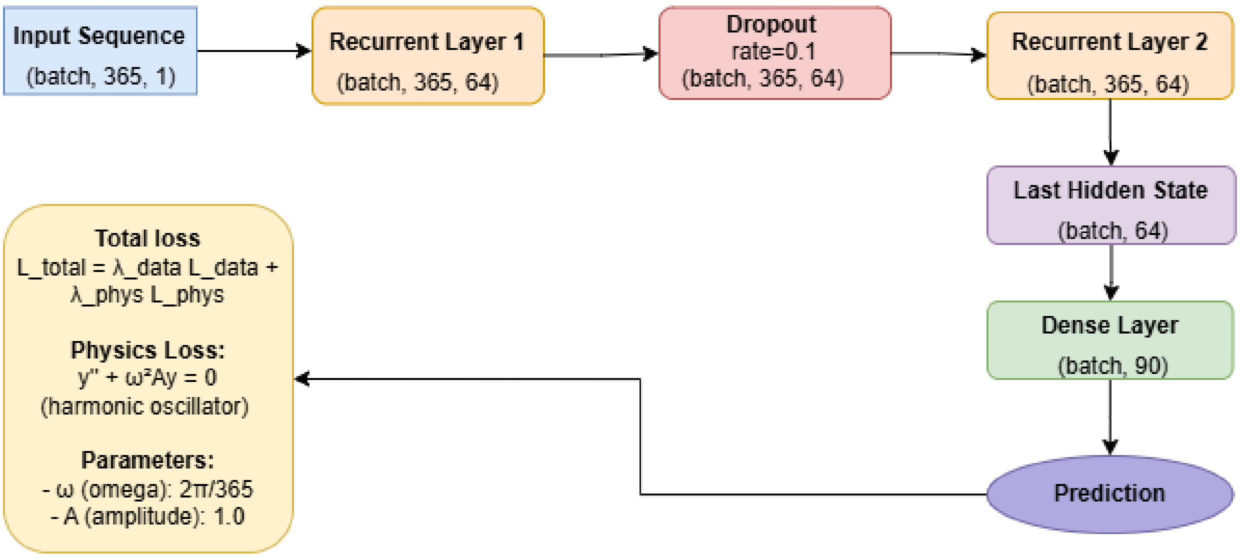

Figure 1. Architecture of the Physics-Informed Neural Network models

To control the loss balance, weighting coefficients ($\lambda_{data} = 1$, $\lambda_{phys} = 0.0001$) were applied. The model was trained using the Adam optimizer with a learning rate of 0.001. The training process was conducted over 500 epochs.

Table 4. Hyperparameters of the PINN models

Hyperparameter	**Value**
Optimizer	Adam
Weight initialization	Random
Learning rate	0.001
Hidden layers	2
Neurons per layer	64
Dropout	0.1
Number of epochs	500
Early stopping	Yes
Physics loss weight	0.0001

4. Results and Discussion

To compare the temperature dynamics of the three cities, a spectral analysis was performed. The power spectral density (PSD) is plotted on a log-log scale to visualize the relationship between frequency and temperature variance. The analysis was conducted on temperature residuals, which represent the fluctuations that remain after the predictable seasonal cycles have been removed. Temperature cycles for Astana, Almaty, and Shymkent are shown in blue, orange, and green, respectively, on the graph.

Two performance metrics were used to evaluate the models. Root Mean Square Error (RMSE) is the square root of the average of squared differences between predicted and actual values; it gives greater weight to larger errors. Mean Absolute Error (MAE) measures the average absolute difference between predictions and actual values. Table 5 reports the values of these metrics.

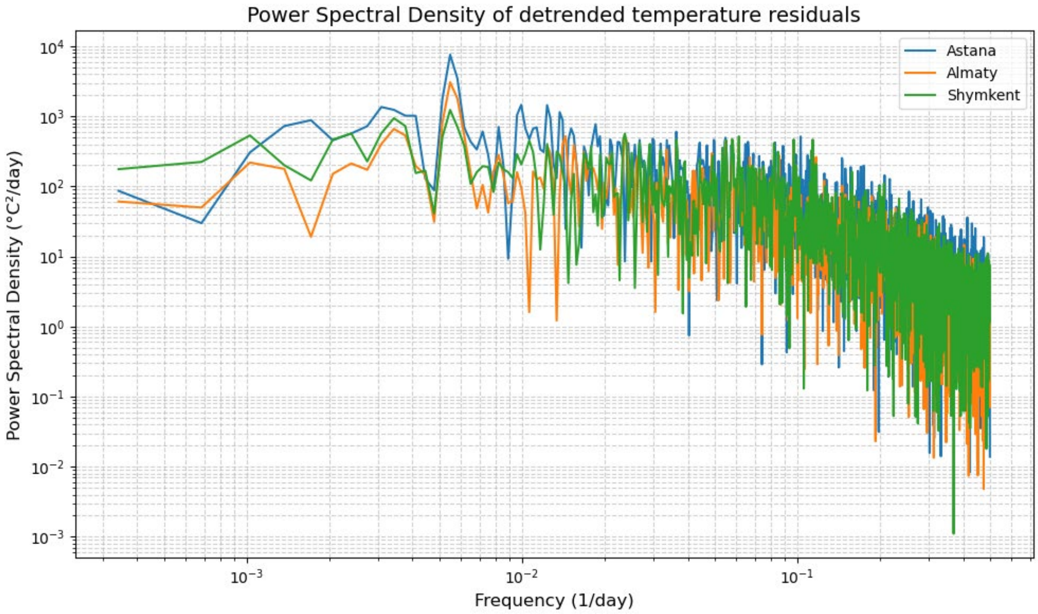

Figure 2. Power Spectral Density of temperature variations by city

Table 5. RNN, LSTM, and PINNs Performance Comparison

City	Model	RMSE (°C)	MAE(°C)
Astana	LSTM	5.4708	4.3581
	Physics-Informed LSTM	5.6639	4.5192
	RNN	5.4449	4.3247
	Physics-Informed RNN	5.5432	4.4241
Almaty	LSTM	3.6076	2.8777
	Physics-Informed LSTM	3.5230	2.8194
	RNN	3.8271	3.0705
	Physics-Informed RNN	3.7238	2.9830
Shymkent	LSTM	3.9618	3.1057
	Physics-Informed LSTM	3.7976	2.9931
	RNN	3.8853	3.0313
	Physics-Informed RNN	4.1836	3.2728

4.1. Astana

Table 1 demonstrates that Astana's standard deviation (14.60 °C) is the highest among all the cities presented. According to Figure 2, Astana consistently exhibits the highest spectral power. This indicates that Astana's temperature residuals contain more variance. Astana has the largest RMSE and MAE across all trained models, with values of RMSE ranging from 5.44 °C (RNN) to 5.66 °C (Physics-Informed LSTM).

It is noteworthy that pure data-driven time series models performed better in Astana than Physics-Informed Neural Networks (PINNs), contradicting our initial hypothesis. In fact, the vanilla RNN recorded the smallest RMSE and slightly better results than the LSTM and all PINN models. This can be attributed to the ability of conventional recurrent models to flexibly learn and adapt to rapid temporal fluctuations in historical data. The oscillation-constrained models impose a regular seasonal pattern, which may limit their ability to capture abrupt or extreme variations in Astana's temperature. As a result, their performance on rapidly changing weather conditions appears more constrained.

4.2. Almaty

In the case of Almaty, the RMSE ranged from 3.52 °C (Physics-Informed LSTM) to 3.83 °C (RNN), and both RMSE and MAE were lower than those observed in Astana. A relatively average and stable climate is reflected in the city's relatively low temperature standard deviation (10.6 °C), which is lower than in Astana and Shymkent. The log-log PSD analysis of residuals confirms this stability; Almaty's curve generally sits lower than Astana's, particularly at high frequencies.

In this city, Physics-Informed LSTM achieved the best performance, having a slight edge over LSTM and RNN. The physical constraints enable the model to capture regular, predictable seasonal cycles and improve generalization. In contrast, the standard RNN exhibited lower accuracy. The Physics-Informed LSTM also outperformed its standard version. These observations imply that PINNs offer an advantage in areas where seasonal patterns are moderate because the physical knowledge allows the accurate replication of predictable variations in temperature.

4.3. Shymkent

Shymkent has mid-range results: RMSE values equal to 3.80 °C for the Physics-Informed LSTM and 4.18 °C for the Physics-Informed RNN. Its temperature standard deviation is 10.98 °C, showing less seasonal amplitude than Astana but more irregular short-term variability than Almaty. This is visually represented in the log-log PSD plot, where the green curve (Shymkent) rises above Almaty's (orange) in the higher frequency range.

The oscillation-constrained LSTM showed slightly lower RMSE than the conventional LSTM, while the oscillation-constrained RNN performed comparably to the standard RNN. These findings suggest that the effectiveness of oscillation-constrained models may depend on local climatic conditions and the chosen network architecture, although the differences are small and should be interpreted cautiously.

5. Conclusion

This study evaluated the performance of Physics-Informed Neural Networks relative to standard LSTM and RNN models in forecasting 2-meter air temperature for Astana, Almaty, and Shymkent. All the models were trained to capture temporal patterns in daily 2-meter air temperature using the ERA5 reanalysis dataset. The results suggest that local climate features strongly influence how well physical restrictions can be incorporated. Physics-informed recurrent models constrained by a simple harmonic prior perform better when the target temperature dynamics are dominated by smooth seasonal variability. In climates characterized by high-frequency and abrupt temperature changes, such constraints may limit model flexibility and reduce predictive accuracy.

Future research will focus on including more meteorological parameters, such as wind speed and humidity, to support more comprehensive climate monitoring. It would also be useful to compare Physics-Informed models with other forecasting techniques, such as CNN-LSTM hybrids or Transformer-based models, to provide a broader context for model performance. Also, the examination of more cities and regions in the current study would test the models under various climate scenarios. Finally, future work will incorporate statistical significance testing and uncertainty quantification, such as bootstrap confidence intervals or year-wise performance variability, to evaluate the robustness of observed differences in model accuracy.

Funding

This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No.BR24992852 "Intelligent models and methods of Smart City digital ecosystem for sustainable development and the citizens' quality of life improvement").

Author Contributions

Conceptualization, Z.B., S.M. and D.O.; Methodology, Z.B. and D.O.; Software, D.O.; Validation, Z.B. and D.O.; Formal Analysis, D.O.; Investigation, D.O.; Resources, Z.B.; Data Curation, D.O.; Writing – Original Draft Preparation, D.O.; Writing – Review & Editing, Z.B. , S.M. and B.A.; Visualization, D.O.; Supervision, Z.B. and B.A.; Funding Acquisition, Z.B. and B.A.

Conflicts of Interest

The authors declare no conflict of interest.

References:

1. L. Bornmann, R. Haunschild, K. Boyack, W. Marx, and J. C. Minx, "How relevant is climate change research for climate change policy? An empirical analysis based on Overton data," *PLoS ONE*, vol. 17, no. 9, p. e0274693, Sep. 2022, https://doi.org/10.1371/journal.pone.0274693

2. A. Mamalakis, J.T. Randerson, J. Yu, M.S. Pritchard, G. Magnusdottir, P. Smyth, P.A. Levine, S. Yu, and E. Foufoula-Georgiou, "Zonally contrasting shifts of the tropical rain belt in response to climate change," *Nature Climate Change*, vol. 11, no. 2, pp. 143–151, Jan. 2021, https://doi.org/10.1038/s41558-020-00963-x

3. G. C. Smith *et al.*, "Polar Ocean observations: A critical gap in the observing system and its effect on environmental predictions from hours to a season," *Frontiers in Marine Science*, vol. 6, Aug. 2019, https://doi.org/10.3389/fmars.2019.00429

4. C. E. Iles, R. Vautard, J. Strachan, S. Joussaume, B. R. Eggen, and C. D. Hewitt, "The benefits of increasing resolution in global and regional climate simulations for European climate extremes," *Geoscientific Model Development*, vol. 13, no. 11, pp. 5583–5607, Nov. 2020, https://doi.org/10.5194/gmd-13-5583-2020

5. T. Beucler, M. Pritchard, S. Rasp, J. Ott, P. Baldi, and P. Gentine, "Enforcing analytic constraints in neural networks emulating physical systems," *Physical Review Letters*, vol. 126, no. 9, Mar. 2021, https://doi.org/10.1103/PhysRevLett.126.098302

6. Z. K. Lawal, H. Yassin, D. T. C. Lai, and A. C. Idris, "Physics-Informed Neural Network (PINN) Evolution and Beyond: A systematic literature review and bibliometric analysis," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 140, Nov. 2022, https://doi.org/10.3390/bdcc6040140

7. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, Nov. 2018, doi: 10.1016/j.jcp.2018.10.045. Available: https://doi.org/10.1016/j.jcp.2018.10.045

8. F. Baer, "Numerical weather prediction," in *Advances in Computers*, 2000, pp. 91–157. https://doi.org/10.1016/S0065-2458(00)80017-0

9. M. G. Schultz *et al.*, "Can deep learning beat numerical weather prediction?," *Philosophical Transactions of the Royal Society a Mathematical Physical and Engineering Sciences*, vol. 379, no. 2194, p. 20200097, Feb. 2021, https://doi.org/10.1098/rsta.2020.0097

10. G. Flato, N. Gillett, V. Arora, A. Cannon, and J. Anstey, "Modelling future climate change," Jan. 2019. https://doi.org/10.4095/327808

11. C. Tebaldi and R. Knutti, "Evaluating the accuracy of climate change pattern emulation for low warming targets," *Environmental Research Letters*, vol. 13, no. 5, p. 055006, Apr. 2018, https://doi.org/10.1088/1748-9326/aabef2

12. F. Sun, A. Mejia, S. Sharma, P. Zeng, and Y. Che, "Evaluating the Credibility of Downscaling: Integrating Scale, Trend, Extreme, and Climate Event into a Diagnostic Framework," *Journal of Applied Meteorology and Climatology*, vol. 59, no. 9, pp. 1453–1467, Aug. 2020, https://doi.org/10.1175/JAMC-D-20-0078.1

13. A. Doury, S. Somot, S. Gadat, A. Ribes, and L. Corre, "Regional climate model emulator based on deep learning: concept and first evaluation of a novel hybrid downscaling approach," *Climate Dynamics*, vol. 60, no. 5–6, pp. 1751–1779, Jul. 2022, https://doi.org/10.1007/s00382-022-06343-9

14. A. Hernanz, C. Correa, M. Domínguez, E. Rodríguez-Guisado, and E. Rodríguez-Camino, "Comparison of machine learning statistical downscaling and regional climate models for temperature, precipitation, wind speed, humidity and radiation over Europe under present conditions," *International Journal of Climatology*, vol. 43, no. 13, pp. 6065–6082, Jul. 2023, https://doi.org/10.1002/joc.8190

15. L. Chen, B. Han, X. Wang, J. Zhao, W. Yang, and Z. Yang, "Machine Learning Methods in Weather and Climate Applications: a survey," *Applied Sciences*, vol. 13, no. 21, p. 12019, Nov. 2023, https://doi.org/10.3390/app132112019

16. G. Accarino, D. Donno, F. Immorlano, D. Elia, and G. Aloisio, "An ensemble machine learning approach for tropical cyclone localization and tracking from ERA5 Reanalysis data," *Earth and Space Science*, vol. 10, no. 11, Nov. 2023, https://doi.org/10.1029/2023EA003106

17. B. Guo, F. Zhang, W. Li, and Z. Zhao, "Cloud classification by Machine learning for Geostationary Radiation ImageR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–14, Jan. 2024, https://doi.org/10.1109/TGRS.2024.3353373

18. A. Rowley and O. Karakuş, "Predicting air quality via multimodal AI and satellite imagery," *Remote Sensing of Environment*, vol. 293, p. 113609, May 2023, https://doi.org/10.1016/j.rse.2023.113609

19. Y. Yasuda, R. Onishi, and K. Matsuda, "Super-resolution of three-dimensional temperature and velocity for building-resolving urban micrometeorology using physics-guided convolutional neural networks with image inpainting techniques," *Building and Environment*, vol. 243, p. 110613, Jul. 2023, https://doi.org/10.1016/j.buildenv.2023.110613

20. M. E. Makkaoui, A. Dalli, and K. Elbaamrani, "A Comparative Study of RNN and DNN for Climate Prediction," *2024 International Conference on Global Aeronautical Engineering and Satellite Technology (GAST)*, pp. 1–6, May 2024, https://doi.org/10.1109/GAST60528.2024.10520748

21. T. B. Ramu, R. Kocherla, G.N.V.G. Sirisha, V.L. Chetana, P.V. Sagar, R. Balamurali, and N. Boddu, "Transformer based models with hierarchical graph representations for enhanced climate forecasting," *Scientific Reports*, vol. 15, no. 1, Jul. 2025, https://doi.org/10.1038/s41598-025-07897-4

22. K. Kashinath et al., "Physics-informed machine learning: case studies for weather and climate modelling," *Philosophical Transactions of the Royal Society a Mathematical Physical and Engineering Sciences*, vol. 379, no. 2194, p. 20200093, Feb. 2021, https://doi.org/10.1098/rsta.2020.0093

23. L. J. Slater et al., "Hybrid forecasting: blending climate predictions with AI models," *Hydrology and Earth System Sciences*, vol. 27, no. 9, pp. 1865–1889, May 2023, https://doi.org/10.5194/hess-27-1865-2023

24. D. Feng, H. Beck, K. Lawson, and C. Shen, "The suitability of differentiable, physics-informed machine learning hydrologic models for ungauged regions and climate change impact assessment," *Hydrology and Earth System Sciences*, vol. 27, no. 12, pp. 2357–2373, Jun. 2023, https://doi.org/10.5194/hess-27-2357-2023.

25. W. Liu, T. Yang, F. Sun, H. Wang, Y. Feng, and M. Du, "Observation-Constrained projection of global flood magnitudes with anthropogenic warming," *Water Resources Research*, vol. 57, no. 3, Jan. 2021, https://doi.org/10.1029/2020wr028830.

26. Z. Chen, J. Gao, W. Wang, and Z. Yan, "Physics-informed generative neural network: an application to troposphere temperature prediction," *Environmental Research Letters*, vol. 16, no. 6, p. 065003, May 2021, https://doi.org/10.1088/1748-9326/abfde9

27. K. V. Park, J. Kim, and J. Seo, "PINT: Physics-Informed Neural Time Series Models with Applications to Long-term Inference on WeatherBench 2m-Temperature Data," *arXiv (Cornell University)*, Feb. 2025, https://doi.org/10.48550/arXiv.2502.04018.

Information about Authors:

Zharasbek Baishemirov is a lead researcher at Kazakh British Technical University (Almaty, Kazakhstan, e-mail: z.baishemirov@kbtu.kz), Astana IT University and Narxoz University, recognized for his extensive experience in the fields of industry, research, and higher education. He has held key research positions at the above-mentioned universities, where his work has primarily focused on mathematical modeling, mathematics, machine learning, and artificial intelligence. Dr. Baishemirov has a strong academic foundation, having completed his studies at leading institutions such as the Abai Kazakh National Pedagogical University and other renowned universities.

Dina Ospanova is a promising junior scientist and graduate student at Astana IT University (Astana, Kazakhstan, e-mail: 242982@astanait.edu.kz), specializing in data science and machine learning. As a dedicated researcher, Dina Ospanova is actively involved in various projects that apply advanced technologies to solve real-world challenges. Her academic focus includes predictive modeling, intelligent routing systems, and data-driven decision-making for urban and environmental applications.

Beibut Amirgaliyev is a distinguished researcher at Astana IT University (Astana, Kazakhstan, e-mail: beibut.amirgaliyev@astanait.edu.kz) and is recognized for his contributions to academia and industry. He holds a PhD in Computer Science and serves as a Professor at Astana IT University, focusing on research areas such as machine learning and computer vision. Dr. Amirgaliyev has published numerous papers on automatic number plate recognition and solar collector systems, with his work cited by over 200 researchers.

Saltanbek Mukhambetzhanov is an assistant-professor at Al-Farabi Kazakh National University (Almaty, Kazakhstan, e-mail: saltanbek.kaznu@gmail.com). He has over 40 years of experience in applied mathematics field.

Submission received: 11 November, 2025.
Revised: 18 December, 2025.
Accepted: 25 December, 2025.