PINT: Physics-Informed Neural Time Series Models with Applications to Long-term Inference on WeatherBench 2m-Temperature Data

Keon Vin Park
Interdisciplinary Program in Artificial Intelligence
Seoul National University
Republic of Korea kbpark16@snu.ac.kr

Jisu Kim *
Department of Statistics
Seoul National University
Republic of Korea jkim82133@snu.ac.kr

Jaemin Seo †
Department of Physics
Chung-Ang University
Republic of Korea jseo@cau.ac.kr

Abstract

This paper introduces **PINT** (*Physics-Informed Neural Time Series Models*), a novel framework designed to integrate physical constraints into neural time series models, thereby enhancing their ability to capture complex dynamics in real-world datasets. To demonstrate its practical utility, we apply **PINT** to the ERA5 WeatherBench dataset, a widely-used benchmark for climate prediction, focusing on long-term forecasting of 2m-temperature data.

PINT leverages the *Simple Harmonic Oscillator Equation* as a physics-informed prior, incorporating its periodic dynamics into three popular neural architectures: RNN, LSTM, and GRU. The choice of the Simple Harmonic Oscillator Equation is motivated by its well-known analytical solutions (sine and cosine functions), which not only represent periodic dynamics but also enable rigorous evaluation of the performance improvements achieved through the incorporation of physics-informed constraints. By benchmarking against a linear regression baseline derived from the exact solutions of this equation, we quantify the added value of embedding physical principles in data-driven models.

Unlike traditional time series approaches that often rely on future observations for inference or training, **PINT** is designed for practical forecasting scenarios. Using only the first 90 days of observed data, the framework iteratively predicts the next two years, addressing challenges associated with limited or missing real-time updates.

Extensive experiments on the WeatherBench dataset showcase **PINT**'s ability to generalize to unseen data, accurately capture periodic trends, and align with underlying physical principles. This study highlights the potential of physics-informed neural time series models to bridge the gap between data-driven machine learning and the interpretability required for climate applications.

1 Introduction

Accurately modeling periodic dynamics in temporal data is a critical challenge in scientific modeling, with applications spanning physics, biology, and climate science. Real-world processes such as oscillatory systems, seasonal climate variations, and energy cycles exhibit strong periodic behaviors that demand models capable of leveraging domain-specific knowledge for reliable and interpretable predictions. While traditional machine learning approaches excel in short-term forecasting, their ability to generalize to long-term trends while adhering to underlying physical principles remains limited.

*Co-Corresponding authors. Email: jkim82133@snu.ac.kr
†Co-Corresponding authors. Email: jseo@cau.ac.kr

This paper introduces **PINT (Physics-Informed Neural Time Series Models)**, a framework designed to integrate physical constraints into neural architectures for robust handling of periodic patterns in temporal data. Central to the framework is the *Simple Harmonic Oscillator Equation*, a fundamental model of periodic dynamics with well-known analytical solutions (sine and cosine functions). By embedding these solutions as constraints, **PINT** enhances the ability of neural models to capture oscillatory behaviors and improves their interpretability. Furthermore, the availability of analytical solutions allows us to quantitatively benchmark the performance improvements achieved by incorporating physics-informed constraints.

To evaluate its practical applicability, we apply **PINT** to the ERA5 WeatherBench dataset Rasp et al. (2020), focusing on long-term forecasting of 2m-temperature data. Climate datasets, such as ERA5, are inherently periodic due to Earth's rotation, revolution, and seasonal cycles, making them ideal testbeds for physics-informed approaches. Unlike traditional numerical weather prediction (NWP) models, such as the Integrated Forecasting System (IFS) Bauer et al. (2021), which rely on solving complex partial differential equations (PDEs) for atmospheric simulations, **PINT** provides a scalable data-driven alternative with reduced computational overhead and improved interpretability.

In comparison to existing machine learning approaches like ClimODE Verma et al. (2024), ClimaX Nguyen et al. (2023), and FourCastNet Kurth et al. (2023), which focus on accurate short-term numerical predictions, **PINT** differentiates itself by focusing on long-term forecasting. By iteratively predicting extended future trends using only initial observations, PINT addresses challenges associated with cumulative errors and trend fidelity in long-term inference. Furthermore, by benchmarking against a linear regression baseline derived from the harmonic oscillator's analytical solutions, we ensure rigorous and interpretable evaluation of the proposed framework's performance.

Key Contributions:

- We introduce **PINT**, a Physics-Informed Neural Time Series framework, integrating the *Simple Harmonic Oscillator Equation* as a constraint to capture periodic dynamics in temporal data.
- We demonstrate **PINT**'s applicability on the ERA5 WeatherBench dataset for long-term forecasting, showcasing its ability to model periodic trends inherent in real-world climate data.
- Unlike traditional RNN, LSTM, and GRU architectures, which also can perform autoregressive inference, **PINT** leverages embedded physical laws to deliver more robust and interpretable results for long-term forecasting.

2 METHODOLOGY

2.1 RECURRENT NEURAL NETWORKS (RNN)

Recurrent Neural Networks (RNN) are a foundational class of deep learning models designed to handle sequential data. By employing recurrent connections, RNN can capture temporal dependencies in data, making them suitable for time-series forecasting tasks. However, traditional RNN often suffer from the vanishing gradient problem, which hampers their ability to learn long-term dependencies Hopfield (1982). Despite this limitation, RNN provide a baseline for evaluating the effectiveness of more advanced architectures.

2.2 LONG SHORT-TERM MEMORY NETWORKS (LSTM)

Long Short-Term Memory Networks (LSTM) extend RNNs by introducing memory cells and gating mechanisms—namely, the input, forget, and output gates—that regulate the flow of information Hochreiter & Schmidhuber (1997). These gates enable LSTM to learn and retain long-term dependencies effectively, overcoming the vanishing gradient problem. Due to their ability to model complex temporal relationships, LSTM are widely used in climate forecasting tasks Qin et al. (2017).

2.3 GATED RECURRENT UNITS (GRU)

Gated Recurrent Units (GRU) simplify LSTM by merging the input and forget gates into a single update gate Cho et al. (2014). This reduction in complexity decreases the number of parameters while maintaining the ability to model long-term dependencies. GRU have shown comparable performance to LSTM in various sequential modeling tasks, including weather prediction Yu et al. (2017).

2.4 PHYSICS-INFORMED NEURAL TIME SERIES MODELS (PINT)

Physics-informed neural networks (PINN) incorporate domain knowledge by embedding physical constraints into the learning process Raissi et al. (2019). For example, in climate modeling, seasonal variations can be described by the harmonic oscillator equation:

$$u''(t) + \omega^2 u(t) = 0, \tag{1}$$

where $\omega = \frac{2\pi}{T}$ represents the angular frequency for a periodic signal with period T. Given that the period T for climate data corresponds to one year (365 days), ω is set to $\frac{2\pi}{365}$.

Notably, since the data has been standardized (mean centered and variance normalized), no offset value (constant term) is required. The standardization ensures that the mean of the data is zero, effectively aligning the harmonic oscillator equation with the standardized time series. Moreover, as shown in Figure 1, the framework combines data-driven loss with physics-based constraints to enhance model predictions.

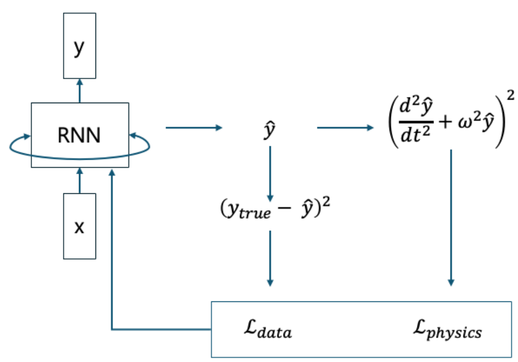

Figure 1: Structure of PINT utilizing physics knowledge based on Simple harmonic oscillator.

The corresponding physics-informed loss ensures that the model predictions align with known physical laws:

$$\mathcal{L}_{\text{physics}} = \frac{1}{N^2} \sum_{i=1}^{N} \left(u''(t_i) + \omega^2 u(t_i)\right)^2 . \tag{2}$$

The total loss function used during training combines the physics-informed loss with the standard data-driven loss, weighted by a hyperparameter λ_{physics}, which balances the influence of the two terms:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{data}} \mathcal{L}_{\text{data}} + \lambda_{\text{physics}} \mathcal{L}_{\text{physics}}, \tag{3}$$

where $\mathcal{L}_{\text{data}}$ represents the standard data loss (e.g., mean squared error between predictions and observations), λ_{data} is the data loss weight which is set to 1, and λ_{physics} is the physics loss weight. This combined loss formulation enables the model to leverage prior knowledge effectively while ensuring accurate predictions from the data.

Integrating this combined loss term into the training process improves interpretability and generalization, especially for tasks requiring the enforcement of physical consistency, such as climate forecasting.

2.5 Linear Regression under Physics's Law

Linear Regression serves as a fundamental statistical technique used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to observed data. In the context of climate modeling, a simple harmonic equation is often sufficient for representing seasonal cycles:

$$x(t) = \beta_1 \cos(\omega t) + \beta_2 \sin(\omega t), \tag{4}$$

where $\omega = \frac{2\pi}{T}$. For this study, the period T is set to 365 days, resulting in $\omega = \frac{2\pi}{365}$. This formulation mirrors the basic structure of a simple harmonic oscillator, providing a straightforward yet effective baseline for comparing more sophisticated machine learning approaches.

Notably, since the dependent variable $x(t)$ has been standardized (mean centered and variance normalized), no offset value (constant term) is required in this formulation. The standardization ensures that the mean of the data is zero, thereby eliminating the need for a constant intercept term.

2.6 Autoregressive Inference for Long-Term Forecasting

In the context of climate modeling, autoregressive inference is a practical approach for long-term forecasting. Unlike standard predictive models that rely on complete input sequences for each prediction step, autoregressive models iteratively use their own predictions as inputs for subsequent forecasts. This allows them to extend forecasting horizons without requiring intermediate observations.

Figure 2 illustrates the autoregressive inference workflow in the context of Recurrent Neural Networks (RNN). In each step:

- The model takes an input vector x, which consists of observed data for the initial time steps.
- The model predicts the next set of values $\hat{y}$ for the forecasting horizon.
- The predicted values $\hat{y}$ are concatenated with the original input x, forming a new input sequence for the next iteration.

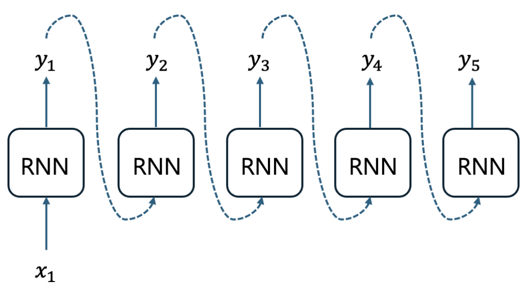

Figure 2: Workflow of the autoregressive inference for long-term forecasting.

2.6.1 Performance Metrics

To evaluate the models, two key performance metrics are used:

- **Root Mean Square Error (RMSE)**: This measures the average magnitude of the prediction errors, defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \tag{5}$$

  where y_i are the observed values and $\hat{y}_i$ are the predicted values. A lower RMSE indicates better model accuracy.

- **Correlation Coefficient (CORR)**: This quantifies the strength and direction of the linear relationship between the observed and predicted values, defined as:

$$\text{CORR} = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2 \sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2}}, \tag{6}$$

where $\bar{y}$ and $\bar{\hat{y}}$ are the means of the observed and predicted values, respectively. A CORR value closer to 1 indicates better alignment between predictions and observations.

3 Data

To evaluate the proposed framework, we utilize the ERA5 dataset from WeatherBench, a comprehensive real-world time series dataset specifically designed for benchmarking weather forecasting models.

3.1 ERA5 WeatherBench

The ERA5 dataset, derived from the European Centre for Medium-Range Weather Forecasts (ECMWF), offers a global atmospheric reanalysis that has been widely utilized in climatic studies Rasp et al. (2020). For this study, we specifically focus on the *2m_temperature (t2m)* variable from WeatherBench, which records the air temperature at two meters above the ground. The *t2m* variable is particularly suitable for our analysis as it reflects near-surface atmospheric conditions, making it highly relevant for climate modeling and forecasting applications.

We selected three cities with distinct climatic conditions—Seoul (South Korea), Washington, D.C. (United States), and Beijing (China)—to ensure the robustness and applicability of our model across different geographic and environmental contexts. The data for each city is spatially averaged around the nearest grid point to the city's coordinates, ensuring that local climatic conditions are adequately captured.

3.2 Data Processing and Sequence Modeling

The temperature data from WeatherBench is split into training, validation, and test sets covering different years. Moreover, descriptive statistics for each is shown in Table **??**:

- **Training Data:** 2008-2012
- **Validation Data:** 2013-2015
- **Test Data:** 2016-2018

Each dataset includes daily temperature averages, which are used to construct input sequences of 90 days. These sequences are utilized to predict the subsequent 30 days of temperatures, allowing the model to capture both short-term fluctuations and seasonal trends. This sequence modeling approach is designed to test the model's ability to leverage past climate data to forecast future conditions effectively.

4 Results

The predictive performance of the models was evaluated using two key metrics: Root Mean Square Error (RMSE) and correlation coefficient (CORR). RMSE provides a measure of the average magnitude of errors between predicted and observed temperatures, offering a direct comparison of model accuracy. CORR evaluates the strength and direction of the linear relationship between predictions and actual values, reflecting the models' ability to capture temporal trends in the data. Figures 4, 5 illustrate the time series comparisons for Seoul, Beijing, and Washington-DC, respectively, highlighting the performance of different models across these cities. Furthermore, Estimated Linear Regression Beta Coefficients for each city are presented in Table 1.

This section evaluates the performance of different models, focusing on four key comparisons: (1) Choosing a Hyperparameter: physics loss weight, (2) RNN-family models vs. Physics-Informed

Table 1: Linear Regression Beta Coefficients

City	$\hat{\beta}_1$(cosine)	$\hat{\beta}_2$(sine)
Seoul	-1.2038	-0.5845
Beijing	-1.3088	-0.3113
Washington-DC	-1.0197	-0.8799

Counterparts, (3) A Best Model vs. Linear Regression as a baseline, and (4) general observations across datasets.

4.1 Choosing a Hyperparameter: Physics Loss Weight

The physics loss weight was validated for each city using five different hyperparameter values, ranging from 10^{-1} to 10^{-5}, decreasing by a factor of 10 at each step. The optimal value for each city was selected based on the lowest RMSE. As a result, the best-performing physics loss weight was determined to be 10^{-5} for Seoul and 10^{-3} for both Beijing and Washington, D.C.

To illustrate this selection, the following figure presents a comparison of the Physics-Informed LSTM model in Seoul using physics loss weights of 10^{-3}, 10^{-4}, and 10^{-5} (see Figure 3).

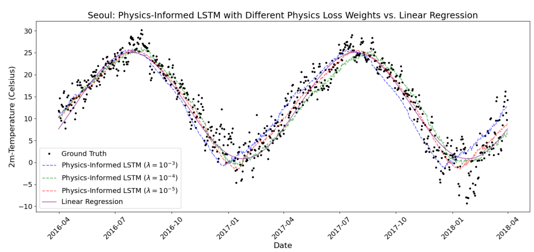

Figure 3: Seoul: Physics-Informed LSTM with different physics loss weights

4.2 Comparison 1: RNN-family Models vs. Physics-Informed Counterparts

Physics-Informed Neural Networks (PINNs) were developed to enhance the accuracy and interpretability of RNN-family models by embedding physical constraints inspired by the harmonic oscillator equation. Among the tested models, the Physics-Informed LSTM consistently delivered the best performance across the three cities analyzed, underscoring the value of incorporating physical principles into machine learning models. Results are summarized in Table 2, Table 3, and Table 4. A comprehensive visual comparison is presented in Figure 4 and Figure 6.

Table 2: Seoul Model Performance Summary, when $\lambda_{\text{physics}} = 10^{-5}$

Model	RMSE	RMSE (Physics)	CORR	CORR (Physics)
RNN	3.9329	**2.8231 (-1.1098)**	0.9033	**0.9509 (+0.0476)**
LSTM	3.9597	2.9413 (-1.0184)	0.9036	0.9492 (+0.0456)
GRU	2.9262	2.9467 (+0.0205)	0.9491	0.9475 (-0.0016)

Table 3: Beijing Model Performance Summary, when $\lambda_{\text{physics}} = 10^{-3}$

Model	RMSE	RMSE (Physics)	CORR	CORR (Physics)
RNN	4.2696	4.6542 (+0.3846)	0.9416	0.9297 (-0.0119)
LSTM	4.3549	**3.7866 (-0.5683)**	0.9439	**0.9545 (+0.0106)**
GRU	5.7860	5.6230 (-0.163)	0.8982	0.8987 (+0.0005)

- **Seoul: The Physics-Informed RNN Outperforms All Models** In Seoul, which exhibited strong seasonal trends, the Physics-Informed RNN achieved the best performance:

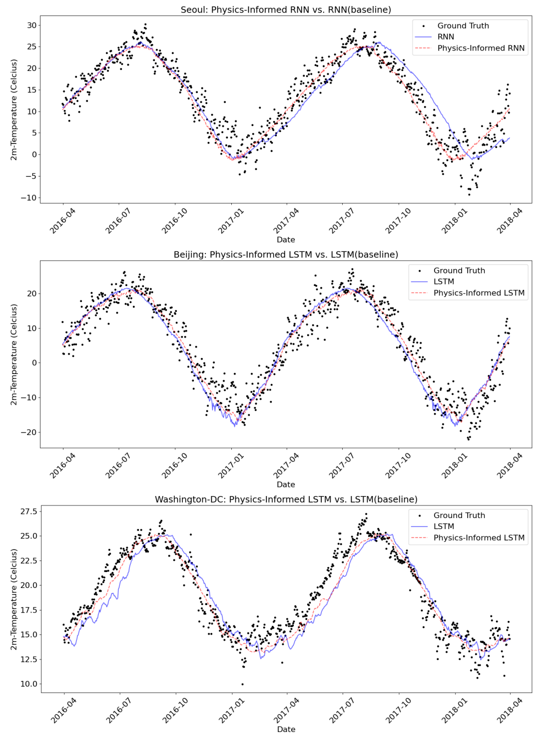

Figure 4: A Comparison for RNN-family Models vs. Physics-Informed Counterparts across cities.

Table 4: Washington-DC Model Performance Summary, when $\lambda_{\text{physics}} = 10^{-3}$

Model	RMSE	RMSE (Physics)	CORR	CORR (Physics)
RNN	4.0391	2.4875 (-1.5516)	0.5126	0.8204 (+0.3078)
LSTM	1.8667	**1.3907 (-0.4760)**	0.9040	**0.9485 (+0.0455)**
GRU	2.3930	1.5689 (-0.8241)	0.8442	0.9374 (+0.0932)

  - **CORR:** 0.9509
  - **RMSE:** 2.8231

  The Physics-Informed LSTM performed similarly but with slightly higher RMSE, emphasizing the effectiveness of RNN-based architectures for periodic datasets.

- **Beijing: Physics-Informed LSTM Shows the Superiority** In Beijing, the Physics-Informed LSTM delivered the best balance between RMSE and CORR:
  - **CORR:** 0.9545
  - **RMSE:** 3.7866

  The Physics-Informed LSTM demonstrated superior performance across all metrics, highlighting its ability to effectively capture temporal patterns and trends in the data.

- **Washington-DC: Physics-Informed LSTM Excels** In Washington-DC, the Physics-Informed LSTM significantly outperformed all other models:
  - **CORR:** 0.9475
  - **RMSE:** 1.3907

  These results highlight the LSTM's ability to model complex dependencies and the added benefit of embedding physical constraints, even in datasets with relatively stable temporal variations.

4.3 Comparison 2: A Best Model vs. Linear Regression (Baseline)

Linear Regression served as a baseline model, leveraging harmonic oscillator-based seasonal trends for interpretable predictions. Unlike deep learning models, Linear Regression required only the training set (2008–2012) and avoided the need for a validation set. Despite its simplicity, it performed competitively in Beijing and Washington-DC, highlighting the effectiveness of simpler models in certain contexts. For visual comparisons, see Figure 5.

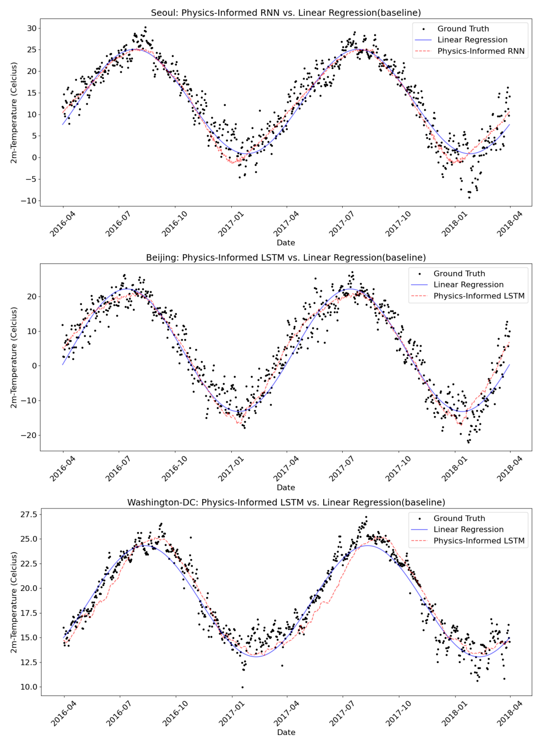

Figure 5: A Comparison for Best Model per city vs. Linear Regression (baseline).

Table 5: A Comparison of the Best Model and Linear Regression (2016–2018)

City	Year	Model	RMSE	CORR
Seoul	2016–2018	Linear Regression	3.0015	0.9463
		Physics-Informed RNN	**2.8231**	**0.9509**
Beijing	2016–2018	Linear Regression	4.0187	0.9514
		Physics-Informed LSTM	**3.7866**	**0.9545**
W-DC	2016–2018	Linear Regression	**1.2065**	**0.9627**
		Physics-Informed LSTM	1.3907	0.9485

- **Seoul: Physics-Informed RNN Outperforms Baseline** The Physics-Informed LSTM significantly outperformed Linear Regression:
  - **RMSE:** 2.8231 (vs. 3.0015 for Linear Regression)
  - **CORR:** 0.9509 (vs. 0.9463 for Linear Regression)

  These results underscore the advantage of combining data-driven approaches with physical constraints in datasets with pronounced seasonal trends.
- **Beijing: Physics-Informed LSTM Outperforms Baseline** In Beijing, The Physics-Informed LSTM slightly outperformed Linear Regression :
  - **RMSE:** 3.7866
  - **CORR:** 0.9545

This result highlights the flexibility of LSTM models in datasets dominated by complex patterns.

- **Washington-DC: Physics-Informed LSTM Matches Linear Regression** In Washington-DC, the Physics-Informed LSTM achieved comparable performance to Linear Regression:
  - **RMSE:** 1.3907 (vs. 1.2065 for Linear Regression)
  - **CORR:** 0.9485 (vs. 0.9627 for Linear Regression)

  The results demonstrate that the Physics-Informed LSTM effectively captures stable trends while maintaining physical interpretability.

5 Conclusion and Future Work

This study highlights the effectiveness of integrating physics-based constraints into neural network architectures for long-term forecasting tasks. The proposed framework, **PINT (Physics-Informed Neural Time Series Models)**, has been demonstrated on climate datasets, specifically for predicting near-surface air temperature (*t2m*), showcasing its ability to capture complex temporal dynamics and long-term trends. However, **PINT** is designed as a *general-purpose physics-informed framework* that can be extended to various domains beyond climate prediction, offering significant potential in fields with underlying periodic or physical dynamics, such as energy forecasting, finance, and healthcare.

The results indicate that **PINT** delivers superior performance compared to traditional recurrent neural networks (RNNs) for datasets with pronounced seasonal patterns. At the same time, simpler models like Linear Regression remain competitive for datasets with stable, well-defined seasonal trends. These findings underscore the importance of selecting models based on the nature and complexity of the dataset, balancing data-driven learning with physical interpretability.

Future Directions

To further extend and enhance the utility of **PINT**, several promising directions for future research are identified:

- **Multivariate Forecasting:** Expanding the framework to incorporate multivariate time series forecasting could offer richer insights. Variables like humidity, precipitation, and wind speed in climate science—or other interdependent factors in different domains—should be integrated to capture nonlinear interactions and dependencies.
- **Adaptive Physics Loss:** Developing adaptive weighting mechanisms for the physics-informed loss terms will improve the trade-off between physical constraints and data-driven learning. This will enhance the generalizability of **PINT** across diverse datasets and temporal resolutions.
- **Regional and Global Evaluation:** For climate science, future studies should evaluate **PINT** across a broader range of geographic regions, including polar, tropical, and arid zones, to validate the model's robustness under varying climatic regimes. Similarly, evaluations in other domains should address diverse use cases to demonstrate the framework's adaptability.

Acknowledgements

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Grant No. RS-2024-00346024). It was also partly supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) [NO.RS-2021-II211343, Artificial Intelligence Graduate School Program (Seoul National University)]. In addition, it was partly supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT). (0769-20240034)

REFERENCES

P. Bauer, A. Thorpe, and G. Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525:47–55, 2021.

K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734, Doha, Qatar, 2014.

S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.

Thorsten Kurth, Shashank Subramanian, Peter Harrington, Jaideep Pathak, Morteza Mardani, David Hall, Andrea Miele, Karthik Kashinath, and Anima Anandkumar. FourCastNet: Accelerating global high-resolution weather forecasting using adaptive fourier neural operators. In *Proceedings of the Platform for Advanced Scientific Computing Conference (PASC '23)*, pp. 1–11. Association for Computing Machinery, 2023. doi: 10.1145/3592979.3593412. URL `https://doi.org/10.1145/3592979.3593412`.

Tung Nguyen, Johannes Brandstetter, Ashish Kapoor, Jayesh K Gupta, and Aditya Grover. Climax: A foundation model for weather and climate. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 25904–25938. PMLR, 23–29 Jul 2023. URL `https://proceedings.mlr.press/v202/nguyen23a.html`.

L. Qin, G. Song, and H. Ma. A dual-stage attention-based recurrent neural network for time series prediction. *IEEE Transactions on Neural Networks and Learning Systems*, 29(10):4901–4914, 2017.

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*, 378:686–707, 2019.

S. Rasp, M. Pritchard, and P. Gentine. Weatherbench: A benchmark dataset for data-driven weather forecasting. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1747–1769, 2020.

Yogesh Verma, Markus Heinonen, and Vikas Garg. Climode: Climate and weather forecasting with physics-informed neural odes. In *The Twelfth International Conference on Learning Representations (ICLR 2024)*, 2024. URL `https://openreview.net/forum?id=xuY33XhEGR`.

H. Yu, Z. Wu, S. Wang, and et al. Spatiotemporal recurrent neural networks: A deep learning framework for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*, 19(7): 2347–2356, 2017.

Table 6: Hyperparameters configuration of the RNN family used.

Hyperparameter	**Configuration**
Optimizer	Adam
Weight initialization	Random
Learning rate	0.001
Activation function	tanh
Dropout rate	0.1
Number of hidden layers	2
Number of neurons per layer	64
Sequence length	90
Prediction length	30
Training epochs	1000
Batch size	Full batch
Early stopping	No

Table 7: Hyperparameters configuration of the PINN used.

Hyperparameter	**Configuration**
Optimizer	Adam
Weight initialization	Random
Learning rate	0.001
Activation function	tanh
Data loss weight	1
Physics loss weight	10^{-1} to 10^{-5}
Regularization weight	-
Number of hidden layers	2
Number of neurons per layer	64
Dropout rate	0.1
Training epochs	1000
Early stopping	No
Batch size	Full batch

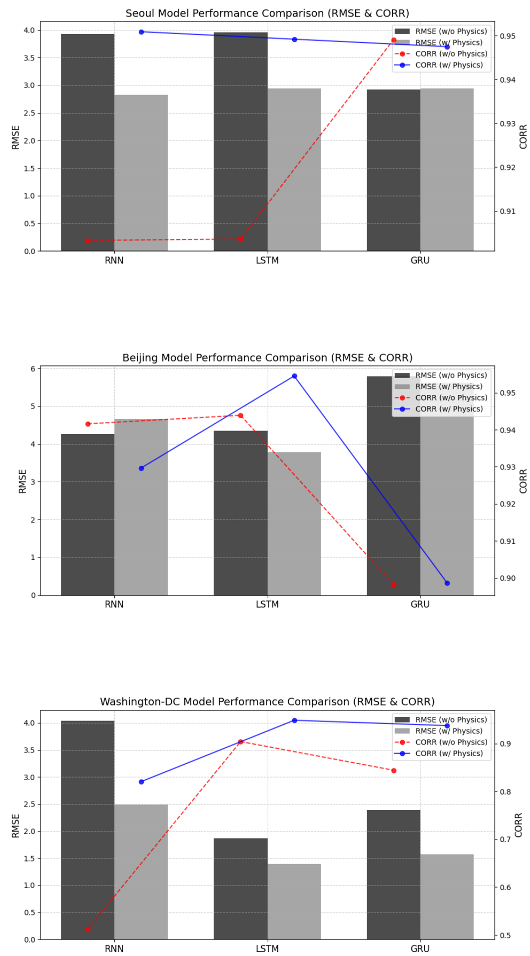

Figure 6: Bar plots for RNN-family Models vs. Physics-Informed Counterparts across cities.