World Journal of Environmental Biosciences
Available Online at: **www.environmentaljournals.org**

Volume 15, Issue 1: 97-105

https://doi.org/10.51847/bO2tYhjWMD

ISSN 2277- 8047

Vectorial Feature Engineering of Temporal Variables in GRU-based Machine Learning Models for Temperature Analysis in Puno

Leonel Coyla-Idme[1], Fred Torres-Cruz[1]*, Ernesto Nayer Tumi-Figueroa[1], Elqui-Yeye Pari-Condori[1], Leonid Aleman Gonzales[1], Jose Pánfilo Tito Lipa[1]

[1]Department of Statistical Engineering and Informatics, National University of the Altiplano, Puno, Peru.

ABSTRACT

Accurate prediction of daily mean temperature in high-altitude environments remains challenging due to pronounced thermal variability, strong seasonal oscillations, and data limitations typical of remote meteorological stations. This study proposes a vectorial feature engineering strategy for temporal variables within a Gated Recurrent Unit (GRU) framework, applied to the Puno station (3,827 m a.s.l.) in the Peruvian Altiplano. The dataset comprises 7,279 daily observations (2003–2024), preprocessed through a four-stage pipeline: sentinel value detection, dual-consensus outlier removal (interquartile range and monthly Z-score), cubic spline imputation, and Min–Max normalization. The core contribution is the cyclic sine–cosine encoding of calendar variables (month, day), projecting discrete temporal indices onto the unit circle in $\mathbb{R}^2$. This representation removes artificial discontinuities at calendar boundaries (e.g., December–January), which persist under conventional linear scaling. The proposed architecture—GRU (64), SpatialDropout1D (0.2), GRU (32), Dense (16, ReLU), Dense (1)—is trained using 30-day sliding windows for one-step-ahead forecasting. A controlled ablation study compares the proposed approach against an identical baseline using normalized but non-cyclic temporal features. On the test set (n = 1,088; 2021–2024), cyclic encoding yields consistent improvements: MAE decreases from 0.0663 to 0.0655 (+1.17%), RMSE from 0.0836 to 0.0829 (+0.82%), MAPE from 13.20% to 12.63% (+4.32%), and R^2 increases from 0.6705 to 0.6758 (+0.80%). The denormalized error (0.81 °C) remains within standard measurement uncertainty, confirming the effectiveness of cyclic temporal encoding as an inductive bias without increasing model complexity.

Keywords: *Feature engineering, Gated recurrent unit, Temperature prediction, Cyclic encoding, Machine learning, Altiplano climate*

Corresponding author: Fred Torres Cruz
e-mail ✉ ftorres@unap.edu.pe
Received: 18 December 2025
Accepted: 28 March 2026

INTRODUCTION

The prediction of surface air temperature is a fundamental task in climate science, with direct implications for agriculture, water resource management, public health, and disaster risk reduction (Gu *et al*., 2024). In high-altitude tropical environments such as the Peruvian altiplano, characterized by elevations exceeding 3,800 m above sea level, intense solar radiation, and pronounced diurnal thermal amplitudes, accurate temperature forecasting is particularly critical. The region surrounding Lake Titicaca sustains agricultural systems and pastoral livelihoods that are acutely sensitive to frost events and thermal anomalies (Garreaud *et al*., 2009; Khan *et al*., 2024; Conti *et al*., 2025, Loganathan *et al*., 2025). Despite this socioeconomic importance, the meteorological station network in the altiplano remains sparse, and many stations suffer from data quality issues, including missing records, sentinel-coded values, and instrumental artifacts (Vuille *et al*., 2008; SENAMHI, 2023; Snodin & McCrossen, 2024; Raza *et al*., 2025). These constraints motivate the development of data-driven prediction models capable of extracting reliable thermal patterns from imperfect observational records.

Traditional approaches to temperature prediction have relied on numerical weather prediction (NWP) models, which solve the governing equations of atmospheric dynamics on discretized grids (Schultz *et al*., 2021). While NWP models achieve high accuracy at synoptic scales, their performance degrades for site-specific local predictions at individual stations, particularly in regions of complex topography where sub-grid-scale processes dominate the thermal signal (Rasp *et al*., 2020). Statistical methods, including autoregressive integrated moving-average (ARIMA) models and multiple linear regression, offer computationally efficient alternatives but assume stationarity and linearity, assumptions frequently violated by climate time series exhibiting nonlinear seasonal interactions and long-range dependencies (Lim & Zohren, 2021; Lee *et al*., 2025).

The emergence of deep learning has fundamentally transformed time series forecasting by enabling the automatic extraction of hierarchical temporal features from raw observational data (LeCun *et al*., 2015; Torres *et al*., 2021; Lee *et al*., 2025). Recurrent neural networks (RNNs) are well-suited to sequential data, as they maintain an internal hidden state that evolves. However, standard RNNs suffer from the vanishing gradient problem, which prevents effective learning of dependencies spanning more than a few dozen time steps (Bengio *et al*., 1994; Csep *et al*., 2024; Njoroge & Odhiambo, 2025). The Long Short-Term Memory (LSTM) network addressed this limitation through gating mechanisms that regulate information flow

across time (Hochreiter & Schmidhuber, 1997; Sak *et al*., 2014; Ganea *et al*., 2024; Petchesi *et al*., 2025). The Gated Recurrent Unit (GRU), proposed by Cho *et al*. (2014), simplified the LSTM architecture by merging the cell state and hidden state into a single representation vector, reducing the parameter count while achieving comparable or superior performance on multiple benchmark tasks (Cho *et al*., 2014; Chung *et al*., 2014; Zar *et al*., 2024; Yu *et al*., 2025). In the context of meteorological prediction, both LSTM and GRU architectures have demonstrated competitive accuracy for temperature (Salman *et al*., 2015; Karevan & Suykens, 2020; Mickevičius *et al*., 2024; Yılmazer & Altinok, 2024), wind speed (Zheng *et al*., 2023), and precipitation forecasting (Hewage *et al*., 2020). Zhang *et al*. (2018) applied LSTM networks to hydrological time series (Zhang *et al*., 2023; Jagsi *et al*., 2025), reporting that recurrent architectures outperformed traditional regression models when sufficient temporal context was provided. Reichstein *et al*. (2019) further argued that deep learning models offer a complementary paradigm to physics-based Earth system models, capable of capturing complex nonlinear interactions that resist analytical formulation.

Despite these advances, a critical and often overlooked aspect of deep learning for climate time series is the representation of temporal input variables. Most existing studies encode calendar variables month, day-of-year, and hour as raw integers or linearly normalized scalars, implicitly imposing an ordinal structure that introduces artificial discontinuities at calendar boundaries (Lim & Zohren, 2021; Dupont & Lefevre, 2024; Kowalski *et al*., 2024). For instance, representing the month as m/12 maps December to 1.0 and January to 0.083, creating a Euclidean distance of 0.917 between temporally adjacent months. This misalignment between input-space geometry and temporal proximity forces the neural network to expend representational capacity learning that numerically distant values are temporally contiguous, a task that is neither guaranteed to succeed nor computationally efficient. Kazemi *et al*. (2019) proposed Time2Vec, a learnable temporal encoding that demonstrated performance improvements across multiple forecasting benchmarks. Wang *et al*. (2022) conducted a comparative study of temporal encoding methods for energy prediction, reporting that trigonometric representations consistently outperformed linear normalization (Wang *et al*., 2022). However, the specific application of cyclic sine-cosine encoding to high-altitude climate prediction has received limited attention in the literature.

The feature engineering stage has been recognized as a decisive factor in the performance of deep learning prediction systems. Yan *et al*. (2022) demonstrated that a multi-stage feature engineering and selection framework combining feature crossing and reinforcement learning-based selection significantly improved axle temperature prediction accuracy when paired with a bidirectional GRU network (Yan *et al*., 2022). Their results confirm that the quality and structure of input features can influence prediction accuracy as much as the choice of network architecture itself. SpatialDropout1D, originally introduced by Tompson *et al*. (2015) for convolutional architectures, has been adapted for recurrent networks as a regularization mechanism that drops entire feature channels along the temporal dimension rather than individual elements (Tompson *et al*., 2015). This strategy preserves the temporal coherence of dropout patterns and has been shown to be more effective than standard dropout for sequential data (Srivastava *et al*., 2014; Elamin *et al*., 2023; Pardo-Zamora & Castellano-Rioja, 2024). The Adam optimizer (Kingma & Ba, 2015) provides adaptive learning rate adjustment and has become the de facto standard for training recurrent networks due to its robustness to hyperparameter selection.

In this context, the present study addresses the following research question: Does the vectorial transformation of temporal variables through cyclic sine-cosine encoding improve the predictive performance of GRU networks for daily mean temperature prediction in the Peruvian altiplano? To answer this question, we propose a complete analytical pipeline comprising: (1) a four-stage data cleaning protocol with dual-consensus outlier detection and cubic spline imputation (Tukey, 1977; de Boor, 1978) (2) vectorial feature engineering that projects month and day onto orthogonal sine-cosine components on the unit circle; (3) a deep GRU architecture with SpatialDropout1D regularization trained using sliding windows of 30 days; and (4) a controlled ablation experiment that isolates the contribution of the cyclic encoding by comparing against an identical model receiving raw normalized temporal variables. The main contributions of this work are threefold:

(a) A dual-consensus outlier detection method combining IQR and monthly climate Z-Score, specifically designed for high-altitude climate data where legitimate temperature extremes (frosts, heat events) must be preserved.

(b) Empirical evidence that vectorial sine-cosine encoding of temporal variables improves GRU prediction accuracy by up to +4.32% in MAPE without increasing model complexity or parameter count.

(c) A reproducible end-to-end analytical pipeline validated on 22 years of daily temperature data from the Puno altiplano station, achieving a denormalized prediction error of 0.81 deg C within the measurement uncertainty of standard meteorological instrumentation.

MATERIALS AND METHODS

The experimental framework proposed in this study comprises four sequential stages: data cleaning, vectorial feature engineering, supervised sequence construction, and deep recurrent modeling. The raw dataset was obtained from the meteorological station of Puno, Peru (latitude -15.84 deg, longitude -70.02 deg, elevation 3,827 m a.s.l.), and contains N = 7,279 daily observations spanning from January 1, 2003, to December 31, 2024. Three thermal variables were recorded: mean temperature (T_mean), mean maximum temperature (T_max), and mean minimum temperature (T_min), all expressed in degrees Celsius.

To distinguish legitimate extreme values from instrumental errors, we implemented a consensus-based detection scheme that combines two complementary methods.

The first method employs the interquartile range (IQR). For each variable x, the acceptance interval is defined as:

$$\mathcal{I}_{IQR} = [Q_1 - 1.5 \cdot \text{IQR},\ Q_3 + 1.5 \cdot \text{IQR}] \quad (1)$$

where $\text{IQR} = Q_3 - Q_1$ and Q_k denotes the k-th quartile of the sample distribution.

The second method uses the monthly climate Z-Score. For each observation x_i recorded in month m:

$$z_{i,m} = \frac{x_i - \bar{x}_m}{\sigma_m} \quad (2)$$

where $\bar{x}_m$ and σ_m represent the mean and standard deviation of observations corresponding to month m. Monthly segmentation is essential in the altiplano context, where thermal amplitude varies significantly between the dry season (May–September) and the wet season (October–April). A value is classified as an outlier if $|z_{i,m}| > 3$.

The consensus criterion establishes that a value is confirmed as atypical only when both methods detect it independently. This conservative strategy reduces the risk of removing legitimate extreme values, particularly relevant in the altiplano, where severe frosts and heat waves constitute real meteorological phenomena. The protocol identified 52 confirmed outliers: 7 in T_{mean} and 45 in T_{max}.

Imputation via cubic spline interpolation

Missing values (original, sentinel, and removed outliers) were imputed using temporal cubic spline interpolation. Given the vector of temporal positions $\{t_j\}$ with valid values $\{y_j\}$, we construct the spline function $S(t)$ that minimizes:

$$\sum_{j=1}^{n} \big(y_j - S(t_j)\big)^2 + \lambda \int [S''(t)]^2 dt \quad (3)$$

where λ controls the trade-off between data fidelity and curve smoothness. We used scipy.interpolate.UnivariateSpline with degree $k = 3$ and smoothing factor $s = 2n$, where n is the number of valid observations.

The choice of cubic spline over simpler methods (linear interpolation, moving average) is justified by its ability to preserve the seasonal curvature of thermal series. A total of 1,059 values were imputed: 787 in T_{mean}, 180 in T_{max}, and 92 in T_{min}.

To ensure stable numerical convergence in the neural network, all thermal variables were normalized to the interval [0,1] using the Min-Max transformation:

$$\hat{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (4)$$

Normalization parameters (x_{min}, x_{max}) were preserved to enable subsequent denormalization of predictions. **Table 1** summarizes the ranges for each variable.

Table 1. Min-Max normalization parameters

Variable	x_{min} (°C)	x_{max} (°C)
T_{mean}	4.50	16.80
T_{max}	10.80	22.90
T_{min}	**−4.80**	9.40

Temporal variables Month and Day exhibit an inherent discontinuity when represented as integers: the model perceives December (12) and January (1) as distant values, when they are in fact adjacent in the annual cycle. This problem is exacerbated in the Puno altiplano context, where the transition between the dry and wet seasons falls precisely on the year boundary. To resolve this discontinuity, we apply cyclic encoding using trigonometric functions. Each temporal variable v with period P is transformed into a pair of orthogonal components:

$$v_{sin} = sin\left(\frac{2\pi . v}{P}\right) \quad (5)$$

$$v_{cos} = cos\left(\frac{2\pi . v}{P}\right) \quad (6)$$

For month (P=12):

$$Month_{sin} = sin\left(\frac{2\pi . m}{12}\right) \quad (7)$$

$$Month_{cos} = cos\left(\frac{2\pi . v}{12}\right) \quad (8)$$

For day (P=31):

$$Day_{sin} = sin\left(\frac{2\pi . m}{31}\right) \quad (9)$$

$$Day_{cos} = cos\left(\frac{2\pi . v}{31}\right) \quad (10)$$

This transformation projects each temporal instant onto a unit circle in $\mathbb{R}^2$, ensuring that the Euclidean distance between representations faithfully reflects cyclic temporal proximity. Both components (sine and cosine) are required to obtain a bijective representation: the sine alone cannot distinguish months symmetric about the solstice. After transformation, the original Month and Day variables are removed from the input space. The resulting feature vector comprises six dimensions:

$$\mathbf{x}_t = \left[\text{Month}_{\text{sin}}, \text{Month}_{\text{cos}}, \text{Day}_{\text{sin}}, \text{Day}_{\text{cos}}, \hat{T}_{\text{max}}, \hat{T}_{\text{min}}\right] \quad (11)$$

The prediction problem is formulated as a supervised regression task with a sliding window. For each time step i, the input sequence and its corresponding target value are constructed as:

$$\boldsymbol{X}_i = [\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_{i+W-1}] \in \mathbb{R}^{W \times F} \quad (12)$$

$$y_i = \hat{T}_{mean,\, i+W} \quad (13)$$

where $W = 30$ is the window size (days), and $F = 6$ is the feature vector dimensionality. The resulting input tensor has a dimension $\mathbb{R}^{N' \times 30 \times 6}$, with $N' = N - W$. The selection of $W = 30$ corresponds to the monthly temporal scale of altiplano climate patterns, providing the model with sufficient information to capture intra-monthly trends and seasonal transitions.

The GRU processes the temporal sequence through three equations that regulate information flow:

$$Z_t = \sigma(w_z[\mathbf{h}_{t-1}, x_t] + b_z) \quad (14)$$

$$r_t = \sigma(w_r[\mathbf{h}_{t-1}, x_t] + b_r) \quad (15)$$

$$\tilde{\mathbf{h}}_t = \tanh(w_h[r_t \odot [\mathbf{h}_{t-1}, x_t] + b_h)] \quad (16)$$

$$\mathbf{h}_t = (1 - z_t) \odot \boldsymbol{h}_{t-1} + z_t \odot \tilde{\mathbf{h}}_t \quad (17)$$

where z_t is the update gate, r_t the reset gate, σ the sigmoid function, and $\odot$ the Hadamard product. The update gate controls how much information from the previous state is retained, while the reset gate determines which portion of the prior state contributes to the candidate.

The deep architecture comprises five sequential layers:

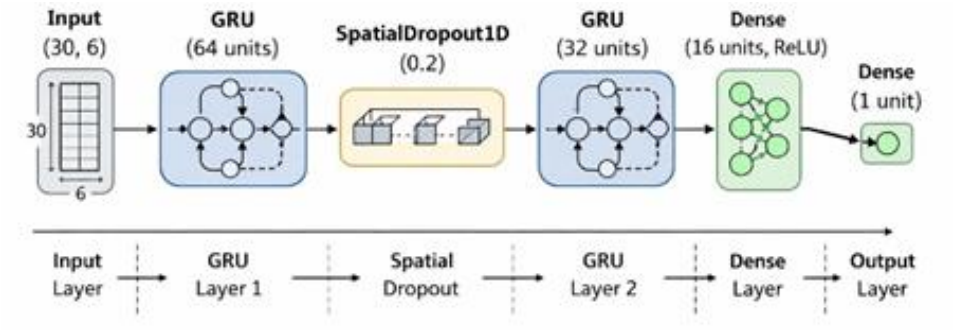

Figure 1. GRU Deep Architecture for mean Temperature prediction.

The first GRU layer with 64 units (return_sequences=True) extracts low-level temporal patterns and propagates the full sequence of hidden states. The SpatialDropout1D layer with rate $p = 0.2$ applies regularization by dropping entire feature channels along the temporal dimension, which is more effective than standard dropout for sequential data. The second GRU layer with 32 units compresses the sequence into a fixed-length vector representation. The final dense layers perform nonlinear regression toward the scalar prediction. **Figure 1** illustrates the complete model topology.

The model is optimized using Adam with learning rate $\alpha = 0.001$ and MSE loss function:

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (18)$$

Data are split temporally without shuffling to respect the causal structure of the series: 70% training ($n = 5\,074$), 15% validation ($n = 1\,087$), and 15% test ($n = 1\,088$). Training uses mini-batches of size 32 for a maximum of 200 epochs, with two control mechanisms:

Early Stopping: halts training if validation loss does not improve for 15 consecutive epochs, restoring the weights from the best epoch.

Model Checkpoint: stores the model with the lowest observed validation loss.

To quantify the effect of vectorial feature engineering, a second GRU model with identical architecture is trained using temporal variables in their raw form, linearly normalized: $\text{Month}' = m/12$ and $\text{Day}' = d/31$. This normalization preserves the cyclic discontinuity and serves as the experimental baseline.

Predictive performance is evaluated on the test set using four complementary metrics:

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

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

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

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

MAE quantifies the average absolute error in normalized units. RMSE disproportionately penalizes large errors, which is useful for detecting aberrant predictions. MAPE expresses the error as a relative percentage, facilitating practical interpretation. The R^2 coefficient measures the proportion of variance explained by the model relative to trivial mean prediction.

RESULTS AND DISCUSSION

Table 2 summarizes the transformations applied at each protocol stage. The process started from 963 missing or anomalous values and produced a complete dataset without compromising the temporal structure of the series. Post-processing verification confirmed strict monotonicity on the time axis and total absence of null values.

Table 2. Data cleaning protocol summary

Stage	T_{mean}	T_{max}	T_{min}
Original nulls	780	114	69

Sentinel (−999)	0	21	23
Outliers (consensus)	7	45	0
Total imputed	**787**	**180**	**92**

The dual consensus method proved its value for the $T_{\min}$ variable: while the climate Z-Score detected 33 suspicious values, the IQR method identified none. Since the interquartile range of minimum temperatures naturally spans from −5.5 to 12.9 °Ca wide interval consistent with nocturnal altiplano frostsno extreme value simultaneously exceeded both thresholds. This result validates the conservative consensus strategy for avoiding the removal of legitimate climate extremes.

Table 3 presents the evaluation metrics for both models on the test set ($n = 1\,088$ days, period 2021–2024).

Table 3. Test set performance comparison

Metric	With FE	Without FE	Improvement
MAE	0.0655	0.0663	+1.17%
RMSE	0.0829	0.0836	+0.82%
MAPE	12.63%	13.20%	+4.32%
$\boldsymbol{R^2}$	0.6758	0.6705	$+\mathbf{0.80}\%$

The model with vectorial feature engineering outperforms the control model across all four evaluated metrics. The most pronounced improvement is observed in MAPE (+4.32%), indicating that sine-cosine cyclic encoding consistently reduces the relative percentage error. In absolute terms, the MAE of 0.0655 in normalized scale translates to a mean error of $0.0655 \times (16.80 - 4.50) = 0.81\,^{\circ}\text{C}$ in the original temperature scale, an acceptable margin for climate monitoring applications in the altiplano.

The $R^2 = 0.6758$ value indicates that the model explains 67.6% of the daily mean temperature variance. The remaining 32.4% can be attributed to factors not included in the input vector: relative humidity, precipitation, cloud cover, wind speed, and large-scale climate teleconnections.

The training process converged at epoch 4, where the lowest validation loss was recorded (val_loss = 0.0064). The Early Stopping mechanism halted training at epoch 19 after 15 epochs without improvement, automatically restoring optimal weights. The learning curves reveal three relevant observations **(Figure 2)**. First, both models achieve rapid convergence (fewer than 5 epochs), suggesting that the GRU architecture has sufficient capacity to capture thermal patterns without requiring prolonged training. Second, the gap between training and validation loss remains narrow throughout the epochs, confirming that the combination of SpatialDropout1D ($p = 0.2$) and Early Stopping provides effective regularization without requiring additional techniques. Third, the model with feature engineering exhibits slightly faster and more stable convergence than the control model, evidencing that cyclic representation facilitates optimization by eliminating discontinuities in the input space.

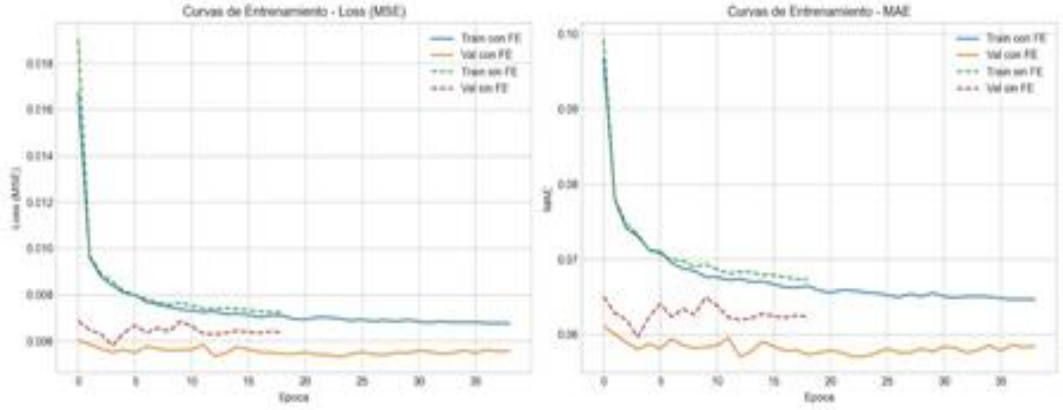

Figure 2. Learning Curves

The residual analysis ($e_i = y_i - \hat{y}_i$) of the feature-engineered model reveals an approximately symmetric distribution centered at zero, with mean $\bar{e} = -0.001$ and standard deviation $\sigma_e = 0.083$. The proximity of the mean to zero confirms the absence of systematic bias in predictions: the model neither consistently overestimates nor underestimates see **Figure 3**.

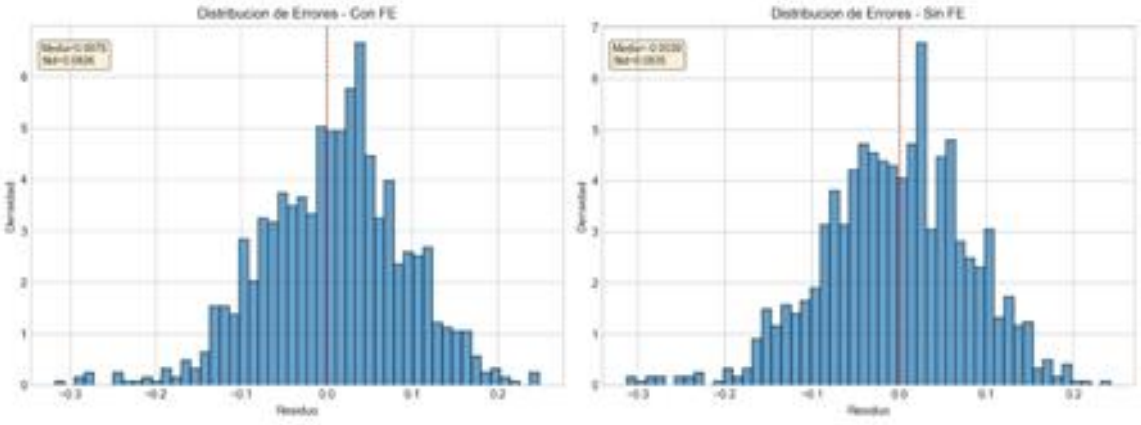

Figure 3. Error Distribution

In **Figure 4**. the scatter plot between actual and predicted values shows alignment with the identity line ($y = \hat{y}$), with higher concentration in the central temperature range ($0.3 < \hat{T}_{\text{mean}} < 0.7$) and increasing dispersion toward the extremes. This behavior is consistent with the lower observation frequency at extreme temperature values, which reduces the model's ability to generalize in those regions of the space.

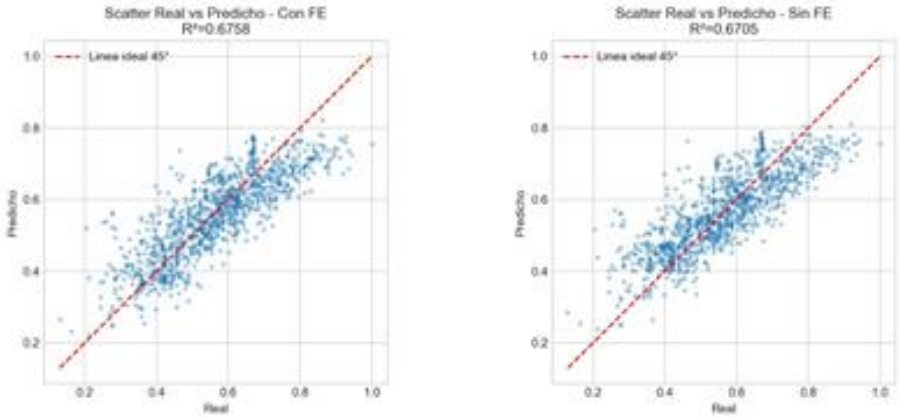

Figure 4. Real Vs Predicted Values.

The specific contribution of cyclic encoding manifests most clearly in the MAPE, where the improvement reaches 4.32%. This result admits a direct interpretation: when the model receives Month and Day as normalized scalar values, the December–January transition generates an abrupt jump from 1.0 to 0.083 in the Month variable (corresponding to $12/12 \rightarrow 1/12$). The GRU network must implicitly learn that these numerically distant values are contiguous in the temporal domain, a task that consumes representational capacity without guaranteed success.

Sine-cosine encoding eliminates this computational burden. In the vectorial representation, December and January map to adjacent points on the unit circle: $(\sin(2\pi \cdot 12/12), \cos(2\pi \cdot 12/12)) = (0,1)$ and $(\sin(2\pi \cdot 1/12), \cos(2\pi \cdot 1/12)) = (0.50,0.87)$, with a Euclidean distance of 0.52the smallest between consecutive months. This design frees network capacity to model more complex thermal patterns.

The moderate but consistent improvement across all metrics suggests that vectorial feature engineering serves as an effective inductive bias: it does not increase model complexity (the parameter count remains invariant), but structures the input space to facilitate learning of seasonal dependencies.

CONCLUSION

In conclusion, this study shows that accurate one-day-ahead prediction of the daily mean temperature is achievable, even in data-scarce, high-altitude contexts, by combining a recurrent architecture with a minimal yet physically plausible input set (thermal variables + time). The denormalized MAE of 0.81 °C positions the model's error within the range typically reported by recent data-driven forecasting work in climate and environmental applications, where performance is strongly conditioned by input richness and station instrumentation (Maslyakova *et al.*, 2023; Pardo-Zamora & Castellano-Rioja, 2024; Loganathan *et al.*, 2025; Wang *et al.*, 2025). The R^2 = 0.6758 is consistent with evidence that explained variance generally increases when models can exploit broader multivariate drivers (e.g., additional meteorological variables), which are often unavailable in real deployments (Jdi & Falih, 2024; Leadbeatter & Tjaya, 2024; Salem *et al.*, 2025; Wang *et al.*, 2025). Methodologically, the key takeaway is that the sine–

cosine (cyclic) temporal encoding yields a systematic gain over the baseline across all evaluation metrics without adding parameters or computational burden—an advantage aligned with the broader move toward practical, robust feature–model combinations in applied forecasting pipelines (Al-Mubarak *et al*., 2024; Gu *et al*., 2024; Bona *et al*., 2025; Lin *et al*., 2025). Because the transformation is simple, low-risk, and generalizable to any periodic time variable, it is especially suitable for operational monitoring systems in under-instrumented regions; moreover, its benefit should be even more pronounced at higher sampling frequencies or for variables with stronger diurnal periodicity, where raw integer time encodings introduce sharper discontinuities (Loganathan *et al*., 2025; Wang *et al*., 2025).

ACKNOWLEDGMENTS: The authors acknowledge the Servicio Nacional de Meteorología e Hidrología del Perú for providing the meteorological data used in this study.

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: This research was funded through the FEDU Competitive Research Projects 2025 under Resolution R.R. N° 1951-2025-R-UNA. No additional external funding was received.

ETHICS STATEMENT: None

REFERENCES

Al-Mubarak, A. M., Alkhaldi, F. A., Alghamdi, A. A., Almahmoud, M. A., & Alghamdi, F. A. (2024). Awareness and clinical competency of dental students in crown lengthening procedures. *Asian Journal of Periodontics and Orthodontics, 4*, 42–51. doi:10.51847/r5cLVpz1UT

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks, 5*(2), 157–166.

Bona, C., Camacho-Alonso, F., Vaca, A., & Llorente-Alonso, M. (2025). Oral biofilm control in patients using orthodontic aligners: Evidence from a systematic review. *Asian Journal of Periodontics and Orthodontics, 5*, 33–42. doi:10.51847/siIhUaqfip

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using an RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Conti, A., Ricci, L., & Esposito, M. (2025). In vitro evaluation and in vivo pharmacokinetic assessment of intranasal tadalafil nanocrystals. *Annals of Pharmaceutical Practice and Pharmacotherapy, 5*, 94–111. doi:10.51847/8J1kgy8yW6

Csep, A. N., Voiţă-Mekereş, F., Tudoran, C., & Manole, F. (2024). Understanding and managing polypharmacy in the aging population. *Annals of Pharmaceutical Practice and Pharmacotherapy, 4*, 17–23. doi:10.51847/VdKr0egSln

de Boor, C. (1978). *A practical guide to splines*. Springer-Verlag.

Dupont, H., & Lefevre, M. A. (2024). Patient-taught workshop to develop Calgary-Cambridge communication skills in hospital pharmacy residents: Implementation and outcomes. *Annals of Pharmaceutical Education, Safety, Public Health Advocacy, 4*, 185–191. doi:10.51847/Z4Tey6oMKA

Elamin, S. M., Redzuan, A. M., Aziz, S. A. A., Hamdan, S., Masmuzidin, M. Z., & Shah, N. M. (2023). Educational impact on glycemic outcomes among children and adolescents diagnosed with type 1 diabetes. *Journal of Medical Sciences Interdisciplinary Research, 3*(1), 41–64. doi:10.51847/s5KgRZ9e1O

Ganea, M., Horvath, T., Nagy, C., Morna, A. A., Pasc, P., Szilagyi, A., Szilagyi, G., Sarac, I., & Cote, A. (2024). Rapid method for microencapsulation of Magnolia officinalis oil and its medical applications. *Special Journal of Pharmacognosy and Phytochemistry Biotechnology, 4*, 29–38. doi:10.51847/UllqQHbfeC

Garreaud, R. D., Vuille, M., Compagnucci, R., & Marengo, J. (2009). Present-day South American climate. *Palaeogeography, Palaeoclimatology, Palaeoecology, 281*(3–4), 180–195.

Gu, T., Zhang, Y., & Wang, L. (2024). TCN-BiGRU: A hybrid deep learning architecture for enhanced temperature time series forecasting. In *2024 6th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI)* (pp. 590–594). doi:10.1109/RICAI64321.2024.10911112

Hewage, P., Trovati, M., Pereira, E., & Behera, A. (2020). Temporal convolutional neural (TCN) network for effective weather forecasting using time-series data from the local weather station. *Soft Computing, 24*, 16453–16482.

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

Jagsi, R., Lee, J., Roselin, D., Ira, K., & Williams, J. (2025). Do U.S. medical schools follow medical associations' recommendations on paid parental leave for faculty? *Annals of Pharmaceutical Education, Safety, Public Health Advocacy, 5*, 1–11. doi:10.51847/r117In8wdi

Jdi, H., & Falih, N. (2024). Leveraging transformer models for enhanced temperature forecasting: A comparative analysis in the Beni Mellal region. *Indonesian Journal of Electrical Engineering and Computer Science, 36*(3), 1694–1700. doi:10.11591/ijeecs.v36.i3.pp1694-1700

Karevan, Z., & Suykens, J. A. K. (2020). Transductive LSTM for time-series prediction: An application to weather forecasting. *Neural Networks, 125*, 1–9.

Khan, T. M., Tahir, H., Adil, Q., Baig, M. R., Jaber, A. A. S., Khaliel, A. M., & Mohammed, Z. M. (2024). A three-decade overview of female-specific cancers in Malaysia: A thorough examination. *Asian Journal of Current Research in Clinical Cancer, 4*(2), 5–18. doi:10.51847/LIdazW7afN

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.

Kowalski, T. W., Reis, L. B., Andreis, T. F., Ashton-Prolla, P., & Rosset, C. (2024). Rare co-occurrence of two mutational variants in NF1: Molecular testing reveals diagnostic surprises. *Journal of Medical Sciences Interdisciplinary Research, 4*(2), 20–29. doi:10.51847/H2qQlZTYO7

Leadbeatter, D., & Tjaya, K. C. (2024). Human rights and bioethical principles in correctional settings: A systematic review of the evidence. *Asian Journal of Ethics in Health and Medicine, 4*, 97–106. doi:10.51847/wSNBedLrGt

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.

Lee, Y. T., Tan, Y. J., & Oon, C. E. (2025). An overview of targeted therapy applications in cancer treatment. *Asian Journal of Current Research in Clinical Cancer, 5*(1), 30–35. doi:10.51847/P55dZHZAF2

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A, 379*(2194), 20200209.

Lin, D., Xu, X., Liu, K., Wu, T., Wang, X., & Zhang, R. (2025). Interpretable data-driven urban building energy modeling considering the inter-building effect. *Building and Environment, 274*, 112688. doi:10.1016/j.buildenv.2025.112688

Loganathan, E., S, N., Yoganathan, A., Kalyanasundaram, P., Anitha, P., & Geetha, R. (2025). A comprehensive review of artificial intelligence approaches for climate change prediction using remote sensing data. In *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 420–424). doi:10.1109/ICSCDS65426.2025.11166962

Maslyakova, A. R., Magomedova, S. A., Romantsov, I. N., Nurbagandov, S. M., Bulovin, M. N., & Podobin, O. R. (2023). Evaluation of the anticancer potential of selenium nanoparticles. *Archives of International Journal Cancer Allied Sciences, 3*(2), 41–47. doi:10.51847/POXX7HfEZo

Mickevičius, I., Astramskaitė, E., & Janužis, G. (2024). A systematic review of the implant success rate following immediate implant placement in infected sockets. *Journal of Current Research in Oral Surgery, 4*, 20–31. doi:10.51847/PcPJL1v1XF

Njoroge, E., & Odhiambo, S. (2025). Elucidating the therapeutic mechanisms of Agrimonia pilosa Ledeb. extract for acute myocardial infarction via network pharmacology and experimental validation. *Pharmaceutical Sciences and Drug Design, 5*, 48–63. doi:10.51847/eZOWCUj80m

Pardo-Zamora, F., & Castellano-Rioja, G. (2024). Liquid biopsy in oral cancer diagnosis: A narrative review of emerging diagnostic tools. *Archives of International Journal Cancer Allied Sciences, 4*(1), 1–6. doi:10.51847/CcaLqtzvoN

Petchesi, C. D., Kozma, K., Iuhas, A. R., Hodisan, R., & Jurca, A. D. (2025). Co-occurrence of Beckwith-Wiedemann syndrome and familial long QT syndrome type I: A case report. *Interdisciplinary Research in Medical Sciences Special, 5*(1), 17–22. doi:10.51847/ihFGrsCY5a

Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. (2020). WeatherBench: A benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems, 12*(11), e2020MS002203.

Raza, S., Khan, A., Mehmood, F., & Farooq, U. (2025). Nationwide implementation of essential pharmacogenomic testing in the Netherlands: A decision-analytic model of lives saved and cost-effectiveness. *Special Journal of Pharmacognosy and Phytochemistry Biotechnology, 5*, 39–49. doi:10.51847/PUWEymkYkk

Sak, H., Senior, A., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large-scale acoustic modeling. In *Interspeech Proceedings* (pp. 338–342).

Salem, H. M., Watanabe, S., & Chang, A. H. (2025). Ethical concerns in managing anorexia nervosa: A content analysis of ethics consultation records. *Asian Journal of Ethics in Health and Medicine, 5*, 25–35. doi:10.51847/oHEI6FgL3V

Salman, A. G., Kanigoro, B., & Heryadi, Y. (2015). Weather forecasting using deep learning techniques. *Procedia Computer Science, 72*, 281–285.

Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., & Leufen, L. H. (2021). Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A, 379*(2194), 20200097.

SENAMHI. (2023). *Climas del Perú: Mapa de clasificación climática nacional*. Servicio Nacional de Meteorología e Hidrología del Perú.

Snodin, D. J., & McCrossen, S. D. (2024). Regulatory considerations of pharmaceutical impurities with emphasis on genotoxic impurities. *Pharmaceutical Sciences and Drug Design, 4*, 1–15. doi:10.51847/ck2yogXhAS

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*(1), 1929–1958.

Tompson, J., Goroshin, R., Jain, A., LeCun, Y., & Bregler, C. (2015). Efficient object localization using convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 648–656).

Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep learning for time series forecasting: A survey. *Big Data, 9*(1), 3–21.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Vuille, M., Francou, B., Wagnon, P., Juen, I., Kaser, G., Mark, B. G., & Bradley, R. S. (2008). Climate change and tropical Andean glaciers: Past, present, and future. *Earth-Science Reviews, 89*(3–4), 79–96.

Wang, Y., Cao, Y., Qu, X., Wang, M., Wang, Y., & Zhang, C. (2025). A review of the application of machine learning techniques in thermal error compensation for CNC machine tools. *Measurement, 243*, 116341. doi:10.1016/j.measurement.2024.116341

Wang, Y., Li, C., & Peng, L. (2022). A comparative study of temporal feature encoding methods for energy consumption prediction. *Applied Energy, 316*, 119082.

Yan, G., Bai, Y., Yu, C., & Yu, C. (2022). A multi-factor driven model for locomotive axle temperature prediction based on multi-stage feature engineering and a deep learning framework. *Machines, 10*(9), 759.

Yılmazer, E., & Altinok, A. (2024). Innovative approaches to delivering mindfulness-based stress reduction (MBSR) in cancer care: Improving access and engagement. *International Journal of Social Psychological Aspects of Healthcare, 4*, 1–12. doi:10.51847/4u9e1ZvfMS

Yu, M., Ma, Y., Han, F., & Gao, X. (2025). Effectiveness of mandibular advancement splint in treating obstructive sleep apnea: A systematic review. *Journal of Current Research in Oral Surgery, 5*, 25–32. doi:10.51847/AInSXrD9rc

Zar, H. J., Moore, D. P., Andronikou, S., Argent, A. C., Avenant, T., Cohen, C., Green, R. J., Itzikowitz, G., Jeena, P., Masekela, R., et al. (2024). Principles of diagnosis and treatment in children with acute pneumonia. *Interdisciplinary Research in Medical Sciences Special, 4*(2), 24–32. doi:10.51847/4RVz1Zxy4h

Zhang, Y., Pan, A., Wang, J., Pan, X., Chen, J., Li, H., Cao, A., & Liu, G. (2023). Assessing the role of play therapy in easing anxiety and despair in children with cancer. *International Journal of Social Psychological Aspects of Healthcare, 3*, 40–48. doi:10.51847/S7vZ2lgmuc

Zheng, Z., Chen, H., & Luo, X. (2023). Spatial-temporal recurrent neural network for wind power forecasting. *Energy Conversion and Management, 275*, 116478.