e-ISSN : 2722-4368

Cross-Regime Transfer Learning for Dekadal Rainfall Prediction in Indonesia Using a Transformer Encoder

Tonny Wahyu Aji[1], Marzuki Sinambela[2], Hapsoro Agung Nugroho[3], Edward Trihadi[4], Tonni Limbong[5]
[1,2,3,4,]State College of Meteorology, Climatology, and Geophysics (STMKG), Tangerang, Indonesia
[5]Santo Thomas Catholic University, Medan, Indonesia
Email : tonny.wahyuaji@gmail.com[1], sinambela.m@gmail.com[2], hapsoro.agung@stmkg.ac.id[3], etrihadi@stmkg.ac.id[4], tonni.budidarma@gmail.com[5]

***Abstract**. Rainfall prediction across regions with different rainfall regimes remains challenging because a model trained in one climatic domain can fail when applied to another domain. This study evaluates transfer learning for dekadal rainfall prediction using a Transformer encoder trained on 10 rainfall sites in West Java and adapted to three target regions representing different rainfall patterns: Kupang (monsoonal), Padang (equatorial), and Ambon (local). Three transfer strategies were evaluated: zero-shot inference, head-only fine-tuning, and staged unfreezing. The source model used 22 meteorological and climate-index features, site embedding, a 36-step input window, and H1-H6 forecasting horizons. Zero-shot transfer produced negative R^2 on all targets, indicating substantial domain shift. Staged unfreezing consistently improved the zero-shot model, reducing RMSE by 49.83% in Kupang, 9.13% in Padang, and 24.76% in Ambon. On the raw millimeter scale, the final transferred model obtained RMSE values of 69.30 mm, 168.36 mm, and 92.91 mm for Kupang, Padang, and Ambon, respectively. Compared with simple baselines, the transferred model outperformed persistence on all targets but outperformed climatology only in Ambon. These results indicate that transfer learning is useful for adapting deep rainfall models to new regions, but reliable deployment still requires target-domain fine-tuning, raw-scale evaluation, and comparison against simple climatological baselines.*

Keyword : Transfer learning; rainfall prediction; Transformer encoder; domain adaptation; dekadal rainfall

INTRODUCTION

Rainfall forecasting is a critical component of climate-risk management in agriculture, water resources, and disaster preparedness, especially in tropical countries where livelihoods remain strongly exposed to rainfall variability. Rainfall information is particularly important where agriculture depends on rain-fed water supply and where shifts in rainfall timing or intensity can disrupt food production and operational planning [1], [2]. In Indonesia, this need is reinforced by the importance of water management for sustaining agricultural productivity, including in dryland systems where short growing seasons and uneven rainfall distribution constrain cropping decisions [3]. Evidence from operational forecast applications also shows that sub-monthly and dekadal information can support in-season early warning and agricultural decision-making when forecast skill emerges early enough before harvest [4]. More generally, rainfall forecast skill can change with temporal aggregation, region, and lead time, which makes the evaluation scale an important part of forecast design [5].

Rainfall forecasting in Indonesia is difficult because the country does not have a single homogeneous rainfall regime. Recent analysis over the Indonesian region identifies three dominant rainfall signals: monsoonal, equatorial, and local rainfall patterns, each with distinct spatial distributions and temporal behaviorr [6]. Monsoonal patterns dominate much of southern Indonesia, equatorial patterns are found around parts of Sumatra and Kalimantan, and local patterns are prominent around Maluku and nearby islands [6]. These regime differences are modulated by ocean-atmosphere interactions involving local sea-surface temperature, the Intertropical Convergence Zone, ENSO, El Nino Modoki, and the Indian Ocean Dipole, with impacts that vary by season and region [7]. Indonesian rainfall is also shaped by Maritime Continent coastal processes and a strong diurnal cycle, making local convection sensitive to coastline geometry and land-sea thermal contrast [8]. Long-term Indonesia-wide analyses further show spatial heterogeneity in rainfall variability and

e-ISSN : 2722-4368

extremes, including high interannual variability and increasing extreme-rainfall tendencies in parts of eastern Indonesia [9].

This climatic diversity creates a direct modeling problem. A model trained at one location can learn statistical relationships that do not transfer cleanly to regions with different rainfall-generating mechanisms. In time-series forecasting, this issue is related to the common assumption that source and target distributions are sufficiently similar, even though real-world series often have time-varying characteristics and distributional differences [10]. In geophysical applications, direct transfer across space is especially difficult because climatic processes, topography, and predictor-predictand relationships vary geographically [11]. Data availability compounds the problem: high-quality observations are unevenly distributed, and vulnerable regions are not always the regions with the best observational coverage [11]. Indonesian applications similarly note that sparse ground observations motivate the use of satellite products and data-efficient modeling strategies [12].

Deep learning has become increasingly prominent for hydrological and rainfall forecasting because it can model nonlinear and multiscale temporal dependencies that are difficult for conventional approaches to capture. A recent review in Earth system science concludes that deep learning is promising for spatiotemporal forecasting while still facing challenges in generalizability, interpretability, and uncertainty estimation [13]. Transformer architectures are relevant in this context because self-attention enables sequence models to learn long-range dependencies without relying on recurrent computation [14]. In hydrology, Transformer-based models have shown competitive performance in streamflow and rainfall-runoff forecasting compared with persistence, LSTM, GRU, and other baselines [15], [16]. These findings suggest that attention-based encoders are suitable candidates for rainfall prediction tasks requiring long temporal context.

Transfer learning offers a practical route for adapting such models to target regions with limited local data or different rainfall regimes. In hydrology, transferring knowledge from data-rich to data-sparse basins has become an important problem because observations are concentrated in relatively few well-monitored locations [17]. Recent studies show that Transformer- or Informer-based transfer frameworks can improve forecasting accuracy in target domains, especially when limited target-domain fine-tuning is available [18]. Fine-tuned regional models can outperform untuned models, indicating that pretraining may capture general temporal structure that must still be specialized using local data [19]. Related work on geographic generalization also shows that domain shift can substantially reduce out-of-domain performance and that domain adaptation or fine-tuning can recover part of the lost skill [11].

Despite this progress, a gap remains for Indonesian dekadal rainfall prediction. Existing studies establish the diversity of Indonesian rainfall regimes, while transfer-learning studies in hydrology are still dominated by streamflow, flood forecasting, downscaling, or non-Indonesian basins rather than dekadal rainfall transfer across contrasting Indonesian regimes [6], [18]. Benchmark design is also important because simple references such as climatology and persistence can remain competitive in seasonal precipitation forecasting and should not be omitted when evaluating deep models [20], [21]. Another unresolved issue is extreme-event behavior: heavy-rainfall verification becomes harder as event thresholds increase, so evaluation should consider metrics sensitive to rare-event detection and tail errors, such as CSI, ETS, or related contingency-based scores [22].

Accordingly, this study investigates whether a Transformer encoder trained on rainfall data from 10 sites in West Java can be transferred to three contrasting Indonesian rainfall regimes represented by Kupang, Padang, and Ambon. The study evaluates three adaptation strategies: zero-shot inference, head-only fine-tuning, and staged unfreezing. This design tests how much transferable rainfall structure exists across monsoonal, equatorial, and local regimes, and how much target-specific adaptation is required. By focusing on dekadal rainfall prediction within Indonesia, this study connects attention-based transfer learning with the operational need for data-efficient forecasting across heterogeneous tropical rainfall regimes.

e-ISSN : 2722-4368

METHOD

Research Design

This study used a quantitative experimental design to evaluate transfer learning for dekadal rainfall prediction. A source model was trained on combined multi-site data from West Java and then evaluated and adapted to three out-of-domain target regions. The transfer protocol consisted of zero-shot inference, head-only fine-tuning, and staged unfreezing. All final results were reported on the 2022-2025 test set.

Dataset and Study Area

The dataset was derived from ERA5-Land and processed into dekadal time series. The full data period was 1990-2025. The temporal split used 1990-2018 for training, 2019-2021 for validation, and 2022-2025 for testing. The source domain consisted of 10 sites in West Java, whereas the transfer targets were Kupang, Padang, and Ambon.

Table 1. Study area and data split

Role	Site	Domain or rainfall pattern	Latitude	Longitude	Period	Split role
Source	Bogor	West Java source	-6.5539	106.7440	1990-2025	Source training
Source	Bandung	West Java source	-6.8836	107.5973	1990-2025	Source training
Source	Bekasi	West Java source	-6.2383	106.9756	1990-2025	Source training
Source	Cirebon	West Java source	-6.7561	108.5397	1990-2025	Source training
Source	Cianjur	West Java source	-6.8200	107.1408	1990-2025	Source training
Source	Ciasem	West Java source	-6.3541	107.6626	1990-2025	Source training
Source	Pangandaran	West Java source	-7.6969	108.6568	1990-2025	Source training
Source	Cibeureum	West Java source	-7.0394	107.4959	1990-2025	Source training
Source	Ciberes	West Java source	-6.3760	107.5830	1990-2025	Source training
Source	Ciawi	West Java source	-6.6623	106.8528	1990-2025	Source training
Target	Kupang	Monsoonal	-10.1679	123.6705	1990-2025	Transfer target
Target	Padang	Equatorial	-0.5452	100.2977	1990-2025	Transfer target
Target	Ambon	Local	-3.6559	128.1895	1990-2025	Transfer target

Variables and Preprocessing

The prediction target was total precipitation, converted to millimeters and aggregated by summation at the dekadal scale. The full feature scheme contained 22 predictors, including seasonal-cycle components, meteorological variables, relative humidity, QBO indices, MJO indices, rainfall lags, and climate-index lags. The target was transformed using $\sqrt{x + c},\quad c = 0.25$, while input features were standardized using training-set statistics.

Source Model

The source model was a Transformer encoder-only forecaster with site embedding. Based on

e-ISSN : 2722-4368

the Bayesian Optimization artifact for the combined Transformer model, the best source configuration used 2 layers, d_model = 64, 6 attention heads, a window size of 36, dropout of 0.35, batch size of 32, learning rate of 5e-5, and weight decay of 0.001. The model produced multi-horizon predictions from H1 to H6.

Table 2. Source Transformer configuration and source-domain performance.

Component	Value
Architecture	Transformer encoder-only
Source domain	Combined West Java, 10 sites
Input features	22
Site embedding	Yes, 12D
Window size	36
Forecast horizon	H1-H6
HPO source	Bayesian Optimization
Parameters	107,006
FLOPs	3,705,216
Source RMSE mean H1-H6	2.9934
Source MAE mean H1-H6	2.3687
Source R2 mean H1-H6	0.6790
Source S_robust	7.8766

Transfer Learning Strategies

Three transfer strategies were evaluated.

1. Zero-shot inference. The source model was directly applied to the target domain without weight updates. This strategy tested direct cross-region generalization.
2. Head-only fine-tuning. The encoder, site embedding, input projection, positional encoding, and Transformer blocks were frozen. Only the prediction head was fine-tuned on the target data.
3. Staged unfreezing. The model was first adapted with a trainable prediction head, then the last encoder block was unfrozen, and finally the full model was unfrozen. In the e30 experiments, the staged-unfreezing schedule used 15 epochs for the last-block stage followed by 30 epochs for the full-model stage. Checkpoint selection used validation RMSE, while final reporting used the test set.

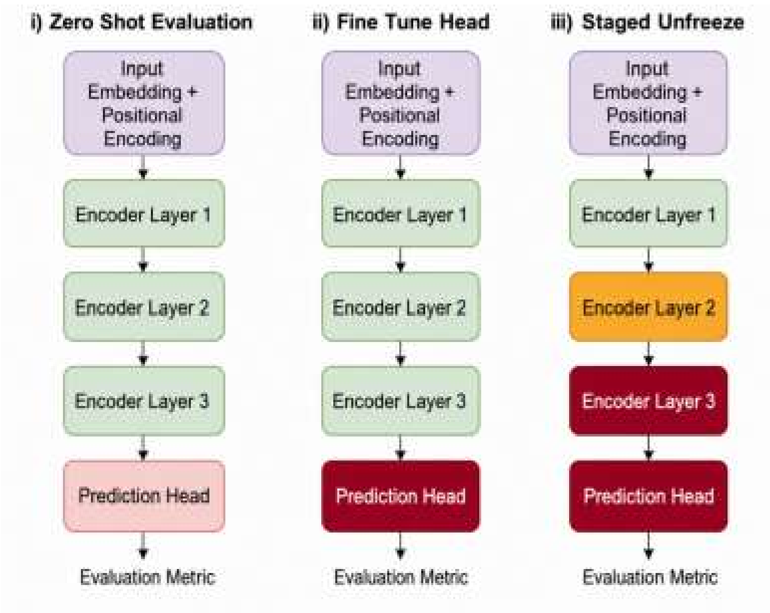

Figure 1. Transfer Learning Strategies

e-ISSN : 2722-4368

Table 3. Transfer strategies and selected runs.

Target	Selected run	Fine-tuning setting
Kupang	sweep_e30_lr1e4	ft_epochs = 30, learning rate 1e-4
Padang	sweep_e30_lr5e4	ft_epochs = 30, learning rate 5e-4
Ambon	sweep_e30_lr5e4	ft_epochs = 30, learning rate 5e-4

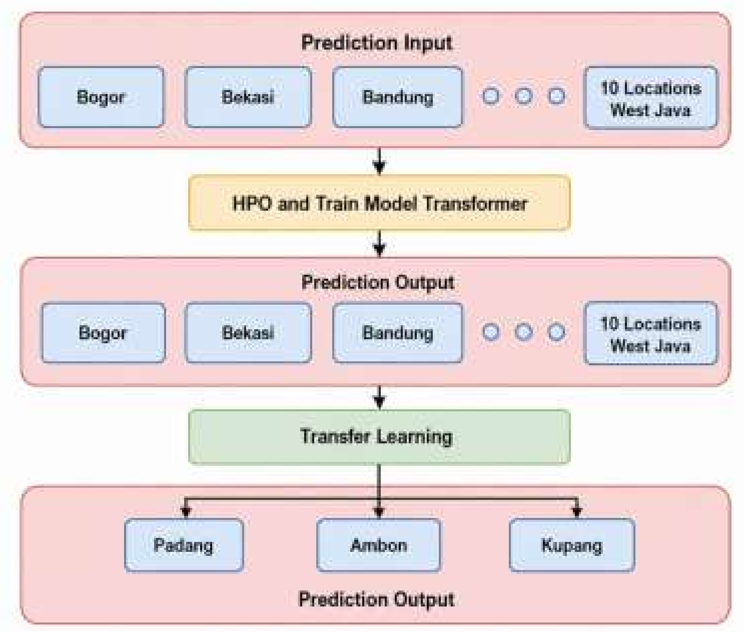

Figure 2. Transfer Learning Scheme

Evaluation Metrics

Model performance was evaluated using RMSE, MAE, R2, CVaR95, and CSI@P95. RMSE and MAE measured average error magnitude, R^2 measured the proportion of target variance explained by the model, CVaR95 measured the mean absolute error in the largest 5% of errors, and CSI@P95 evaluated extreme-event detection using the 95th-percentile rainfall threshold. The transferred model was also compared with climatology and persistence baselines on the raw millimeter scale.

RESULT AND DISCUSSION

Zero-Shot Transfer Performance

Zero-shot transfer produced negative R^2 on all targets: -0.5140 for Kupang, -0.1674 for Padang, and -0.3203 for Ambon. This result indicates that the West Java source model could not be directly generalized to the target regions without adaptation. Zero-shot RMSE on the transformed scale was also high, reaching 6.7237 in Kupang, 5.7404 in Padang, and 4.9939 in Ambon.

Transfer Strategy Comparison

Staged unfreezing produced the best RMSE on all targets. Compared with zero-shot inference, RMSE decreased by 49.83% in Kupang, 9.13% in Padang, and 24.76% in Ambon. The largest improvement occurred in Kupang, whereas the smallest improvement occurred in Padang.

Table 4. Transfer performance on transformed scale

Site	Pattern	Zero-shot RMSE	Head-only RMSE	Staged RMSE	RMSE reduction (%)	Zero-shot R^2	Staged R^2	Zero-shot CVaR95	Staged CVaR95
Kupang	Monsoonal	6.7237	3.8381	3.3730	49.83	-0.5140	0.6190	12.0135	7.6842
Padang	Equatorial	5.7404	5.2190	5.2164	9.13	-	0.0361	13.9080	12.7238

Site	Pattern	Zero-shot RMSE	Head-only RMSE	Staged RMSE	RMSE reduction (%)	Zero-shot R^2	Staged R^2	Zero-shot CVaR95	Staged CVaR95
						0.1674			
Ambon	Local	4.9939	4.0172	3.7572	24.76	- 0.3203	0.2527	10.3566	8.9989

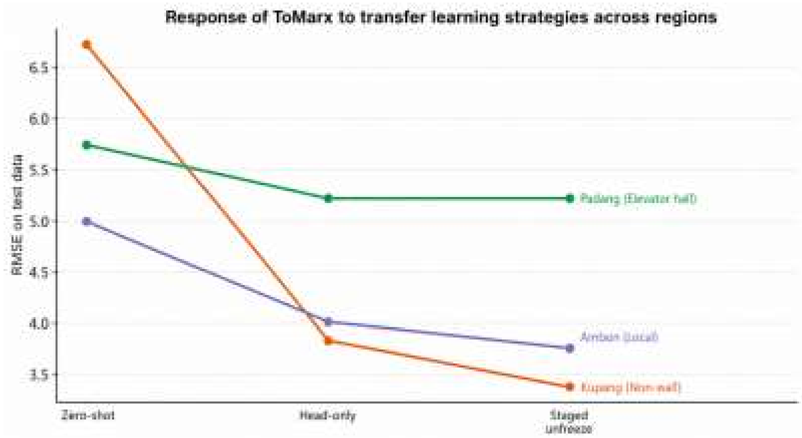

Figure 3. RMSE comparison among transfer strategies

Robustness and Tail-Risk

CVaR95 decreased after staged unfreezing on all targets. The absolute CVaR95 reduction was 4.3293 in Kupang, 1.1842 in Padang, and 1.3577 in Ambon. Thus, staged unfreezing reduced not only average error but also the risk of large tail errors. However, CSI@P95 was 0.0000 on all targets, so the model should not be claimed to detect extreme rainfall events reliably.

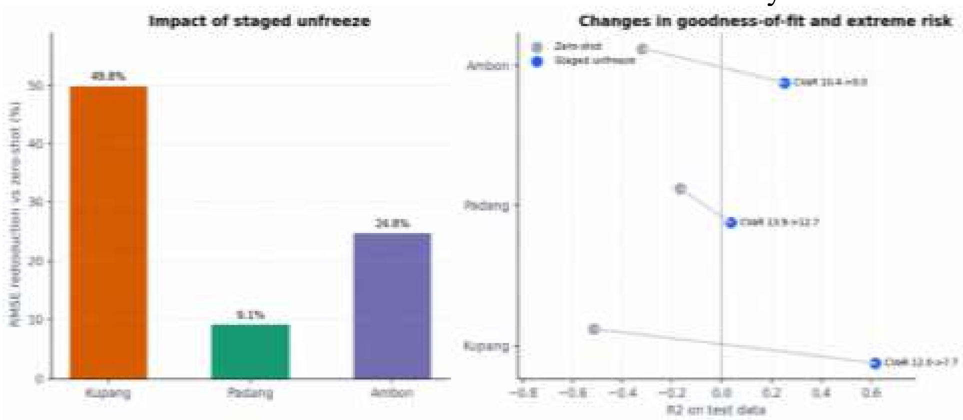

Figure 4. Transfer gain for RMSE, R2, and CVaR95

Final Evaluation on Raw Millimeter Scale

Raw-scale evaluation is important because transformed-scale metrics can obscure the physical magnitude of rainfall errors. On the raw millimeter scale, final staged-unfreeze RMSE was 69.2977 mm for Kupang, 168.3559 mm for Padang, and 92.9109 mm for Ambon. Raw CVaR95 showed that large-error risk remained substantial, especially in Padang, where CVaR95 reached 499.0328 mm.

Table 5. Final staged-unfreeze performance on raw millimeter scale

Site	Pattern	RMSE (mm)	MAE (mm)	R^2 raw	CVaR95 (mm)
Kupang	Monsoonal	69.2977	43.6268	0.4844	204.5188
Padang	Equatorial	168.3559	119.7087	0.0481	499.0328
Ambon	Local	92.9109	62.2198	0.2031	290.3573

e-ISSN : 2722-4368

Baseline Comparison

The transferred ToMarx model outperformed persistence on all targets, but it did not consistently outperform climatology. In Kupang and Padang, climatology produced lower RMSE than ToMarx. In Ambon, ToMarx produced lower RMSE than climatology. This comparison shows that transfer learning improved adaptation relative to zero-shot inference, but the resulting deep model still needs to be evaluated against simple reference forecasts.

Table 6. Baseline comparison on raw millimeter scale

Site	Pattern	Method	RMSE (mm)	MAE (mm)	R^2 raw
Kupang	Monsoonal	ToMarx	69.2939	43.6268	0.4844
Kupang	Monsoonal	Climatology	65.6406	41.3486	0.5374
Kupang	Monsoonal	Persistence	110.7194	68.4671	-0.3163
Padang	Equatorial	ToMarx	168.3492	119.7021	0.0481
Padang	Equatorial	Climatology	164.3111	116.2401	0.0932
Padang	Equatorial	Persistence	220.0303	162.9617	-0.6261
Ambon	Local	ToMarx	92.8936	62.2154	0.2032
Ambon	Local	Climatology	96.8877	62.0820	0.1332
Ambon	Local	Persistence	127.4420	87.3312	-0.4998

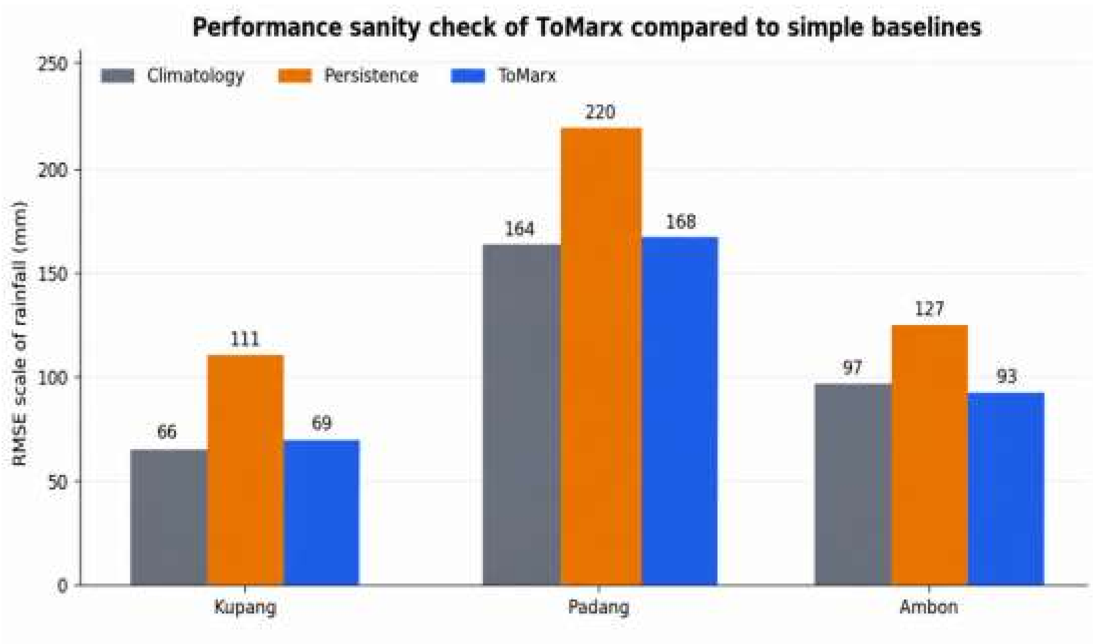

Figure 5. Baseline comparison on raw millimeter scale

Visual Diagnostic

The observed-versus-predicted scatter plot indicates underestimation at high rainfall values. This behavior is consistent with the large raw-scale CVaR95, particularly in Padang. The H1 time series shows that the model followed part of the general temporal dynamics, but the predictions were smoother than the observations and did not always capture rainfall peaks.

e-ISSN : 2722-4368

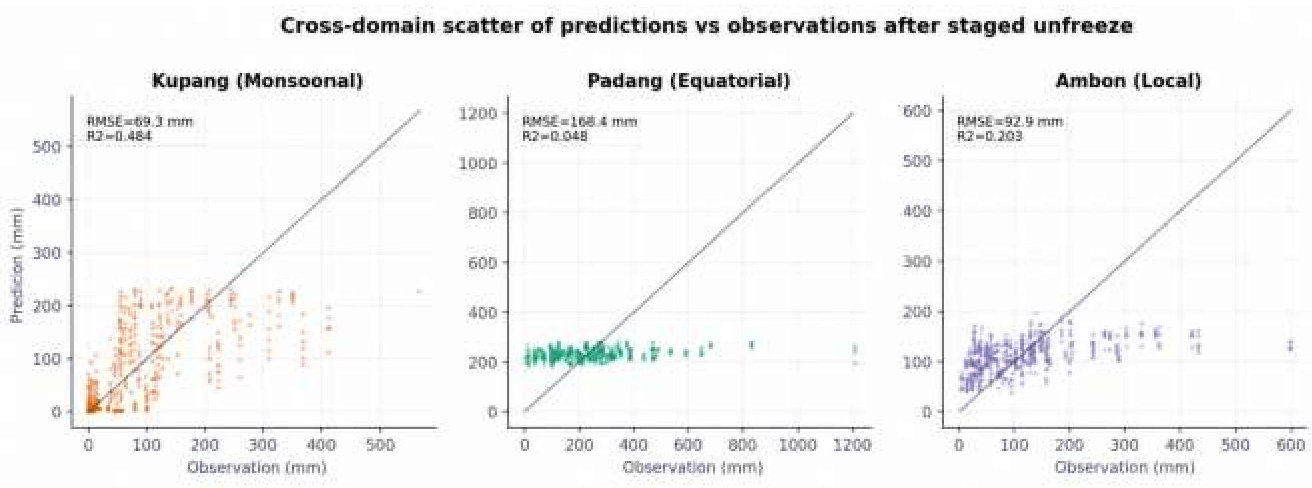

Figure 6. Observed vs predicted rainfall on raw millimeter scale

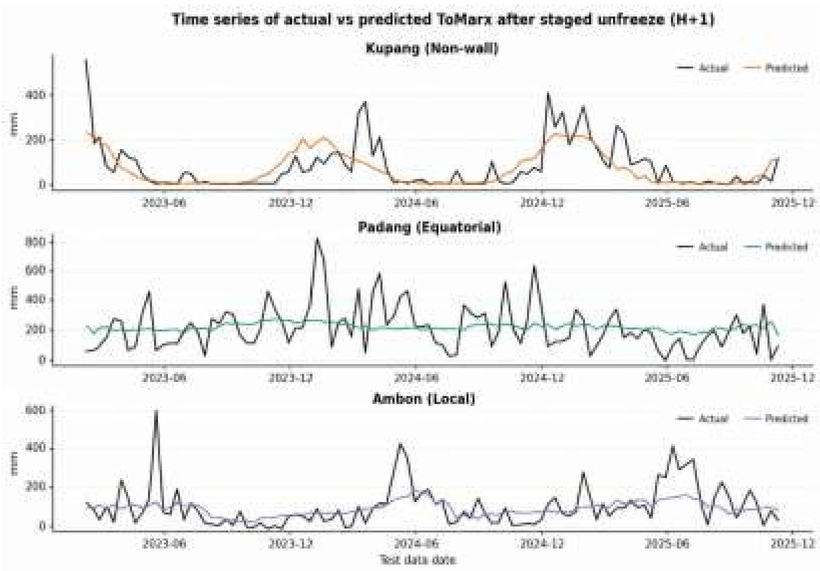

Figure 7. H1 time series of observed and predicted rainfall

DISCUSSION

Domain Shift Limits Zero-Shot Generalization

The negative zero-shot R^2 values on all targets show that the representation learned from West Java was not directly compatible with Kupang, Padang, and Ambon. Therefore, the main contribution of transfer learning in this experiment was not zero-shot generalization, but the use of the source model as a better initialization for target-domain fine-tuning.

Staged Unfreezing Provides the Most Consistent Adaptation

Staged unfreezing was the best transfer strategy across all targets. This pattern is plausible because head-only fine-tuning adjusts only the output layer, whereas staged unfreezing allows part of the internal temporal representation to adapt to the target domain. The improvement from negative to positive R^2 on all targets also indicates that fine-tuning changed the model from a failed out-of-domain predictor into a model with at least some target-domain predictive skill.

Target-Specific Interpretation

Kupang showed the largest improvement, with RMSE decreasing by 49.83% and staged R2 reaching 0.6190. Ambon showed a moderate improvement, with RMSE decreasing by 24.76% and staged R^2 reaching 0.2527. Padang was the most difficult target, with only a 9.13% RMSE reduction and staged R^2 of 0.0361. Based on these results, the equatorial Padang target should be treated as the strongest domain-shift case in this experiment.

Raw-Scale Results Prevent Overclaiming

The transformed-scale results show clear improvement after transfer learning, but the raw millimeter results show that the physical error magnitude remained large. In Padang, RMSE reached 168.3559 mm and CVaR95 reached 499.0328 mm. These results indicate that the transferred model should not be presented as a reliable extreme-rainfall prediction system. A more defensible interpretation is that transfer learning improved regional adaptation, while extreme rainfall prediction remains an open limitation.

Comparison with Climatology and Persistence

The baseline comparison provides an important boundary for interpretation. ToMarx outperformed persistence on all targets, indicating that it learned more useful temporal information than a previous-value assumption. However, climatology remained stronger in Kupang and Padang. This result shows that seasonal baselines can remain highly competitive in rainfall forecasting, especially when the target distribution is difficult to learn from limited target-domain adaptation. The advantage of ToMarx over climatology in Ambon indicates that transfer learning has potential, but this advantage was not consistent across targets.

Implications and Limitations

Methodologically, this study suggests that transfer learning for rainfall prediction should be evaluated at least through zero-shot testing, fine-tuning, raw-scale metrics, tail-risk metrics, and simple baselines. Without climatology, staged unfreezing would appear much stronger than it actually is; with climatology included, the results show a more realistic boundary of model usefulness. This study has several limitations. The target set contained only three sites, the transfer stage used a small fine-tuning sweep rather than full hyperparameter optimization, CSI@P95 remained 0.0000 on all targets, and the analysis did not explicitly incorporate additional physical predictors such as local topography or regional ocean-atmosphere drivers.

CONCLUSION

This study evaluated transfer learning for dekadal rainfall prediction across contrasting rainfall regimes in Indonesia. Zero-shot transfer produced negative R^2 in Kupang, Padang, and Ambon, showing that the source model was insufficient for direct cross-region generalization. Fine-tuning consistently improved performance, with staged unfreezing producing the best results and reducing RMSE by 49.83%, 9.13%, and 24.76% in Kupang, Padang, and Ambon, respectively. On the raw millimeter scale, the final model obtained RMSE values of 69.2977 mm, 168.3559 mm, and 92.9109 mm. The transferred model outperformed persistence on all targets but outperformed climatology only in Ambon. Therefore, transfer learning is a useful strategy for adapting rainfall prediction models to new regions, but claims of performance should be constrained by raw-scale evaluation, tail-risk analysis, and climatological baselines.

Data and Code Availability

ERA5-Land data are publicly available from the Copernicus Climate Data Store. The processed datasets, trained checkpoints, and analysis artifacts used to produce the reported results can be made available by the corresponding author upon reasonable request, subject to repository and journal policy.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the State College of Meteorology, Climatology, and Geophysics (STMKG), and the Copernicus Climate Change Service (C3S) Climate Data Store for providing the ERA5-Land data used in this study.

REFERENCES

[1] I. Ramli, S. Rusdiana, A. Achmad, Azizah, and M. E. Yolanda, 'Forecasting of Rainfall Using Seasonal Autoregreressive Integrated Moving Average (SARIMA) Aceh, Indonesia', Math. Model. Eng. Probl., 2023, doi: 10.18280/mmep.100216.
[2] C. Liyew and H. A. Melese, 'Machine learning techniques to predict daily rainfall amount', J. Big Data, vol. 8, 2021, doi: 10.1186/s40537-021-00545-4.
[3] P. Rejekiningrum et al., 'Optimising Water Management in Drylands to Increase Crop Productivity and Anticipate Climate Change in Indonesia', Sustainability, 2022, doi: 10.3390/su141811672.
[4] D. Lee et al., 'Maize yield forecasts for Sub-Saharan Africa using Earth Observation data and machine learning', Glob. Food Sec., 2022, doi: 10.1016/j.gfs.2022.100643.
[5] W. Atiah, E. Bendito, and F. Muthoni, 'Evaluating Seasonal Rainfall Forecast Gridded Models over Sub-Saharan Africa', Hydrology, 2025, doi: 10.3390/hydrology12100251.
[6] M. Ariska, Suhadi, Supari, M. Irfan, and I. Iskandar, 'Detection of Dominant Rainfall Patterns in Indonesian Regions Using Empirical Orthogonal Function (EOF) and Its Relation with ENSO and IOD Events', Sci. Technol. Indones., 2024, doi: 10.26554/sti.2024.9.4.1009-1023.
[7] G. Alsepan and S. Minobe, 'Relations between Interannual Variability of Regional-Scale Indonesian Precipitation and Large-Scale Climate Modes during 1960–2007', J. Clim., 2020, doi: 10.1175/jcli-d-19-0811.1.
[8] M. Yamanaka et al., 'Maritime continent coastlines controlling Earth's climate', Prog. Earth Planet. Sci., vol. 5, pp. 1–28, 2018, doi: 10.1186/s40645-018-0174-9.
[9] M. Marzuki et al., 'Long-Term Spatial–Temporal Variability, Trends and Extreme Rainfall Events Over Indonesia Based on 43 Years of CHIRPS Data', Int. J. Climatol., vol. 45, 2025, doi: 10.1002/joc.70107.
[10] Q. Gu and Q. Dai, 'A novel active multi-source transfer learning algorithm for time series forecasting', Appl. Intell., vol. 51, pp. 1326–1350, 2020, doi: 10.1007/s10489-020-01871-5.
[11] P. Harder et al., 'Benchmarking the geographic generalization of deep learning models for precipitation downscaling', Sci. Rep., vol. 16, 2026, doi: 10.1038/s41598-025-34557-4.
[12] S. Siswanto, G. v. d. Schrier, and B. v. d. Hurk, 'Observed Increase of Urban Extreme Rainfall as Surface Temperature Rise: The Jakarta Case', J. Meteorol. Soc. Japan, vol. 100, no. 3, pp. 475–492, 2022, doi: 10.2151/jmsj.2022-023.
[13] M. Yu, Q. Huang, and Z. Li, 'Deep learning for spatiotemporal forecasting in Earth system science: a review', Int. J. Digit. Earth, vol. 17, 2024, doi: 10.1080/17538947.2024.2391952.
[14] A. Vaswani et al., 'Attention Is All You Need', 2017, doi: 10.48550/arxiv.1706.03762.
[15] M. Sit, B. Z. Demiray, Z. Xiang, G. J. Ewing, Y. Sermet, and İ. Demir, 'A Comprehensive Review of Deep Learning Applications in Hydrology and Water Resources', 2020, doi: 10.31223/osf.io/xs36g.
[16] A. Maghridlo et al., 'Implementation of LSTM for Weather Prediction in Kupang City, East Nusa Tenggara', in 2025 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS), 2025, pp. 1–6. doi: 10.1109/AIMS66189.2025.11227176.
[17] Y. Xu et al., 'Deep Transfer Learning based on Transformer for Flood Forecasting in Data-sparse Basins', J. Hydrol., 2023, doi: 10.1016/j.jhydrol.2023.129956.
[18] F. Ghobadi, Z. Yaseen, and D. Kang, 'Long-term streamflow forecasting in data-scarce

regions: Insightful investigation for leveraging satellite-derived data, Informer architecture, and concurrent fine-tuning transfer learning', J. Hydrol., 2024, doi: 10.1016/j.jhydrol.2024.130772.

[19] M. He, X. Xu, S. Wu, C.-X. Kang, and B. Huang, 'Multi-step ahead forecasting of daily streamflow based on the transform-based deep learning model under different scenarios', Sci. Rep., vol. 15, 2025, doi: 10.1038/s41598-025-89837-w.

[20] S. Golian et al., 'Dynamical–statistical seasonal forecasts of winter and summer precipitation for the Island of Ireland', Int. J. Climatol., vol. 42, pp. 5714–5731, 2022, doi: 10.1002/joc.7557.

[21] M. Krouma, D. Specq, L. Magnusson, C. Ardilouze, L. Batté, and P. Yiou, 'Improving subseasonal forecast of precipitation in Europe by combining a stochastic weather generator with dynamical models', Q. J. R. Meteorol. Soc., vol. 150, pp. 2744–2764, 2024, doi: 10.1002/qj.4733.

[22] R. Ashrit, K. Sharma, A. Dube, G. Iyengar, A. Mitra, and E. Rajagopal, 'Verification of short range forecasts of extreme rainfall during monsoon', MAUSAM, 2021, doi: 10.54302/mausam.v66i3.548.