

Climate-Driven Poverty Prediction in Nusa Tenggara Timur Using Transfer Learning-Based CNN-LSTM and ERA5 Land

Yampi R Kaesmetan¹, Rahmat Gernowo², Budi Warsito³

^{1,2,3}Doctoral Program of Information Systems, School of Postgraduate Studies, Universitas Diponegoro, Semarang, Indonesia

²Department of Physics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

³Department of Statistics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

¹Informatics Engineering Study Program, STIKOM Uyelindo Kupang, NTT, Indonesia

Received:

May 29, 2026

Revised:

June 18, 2026

Accepted:

September 1, 2026

Published:

October 24, 2026

Corresponding Author:

Author Name*:

Yampi R Kaesmetan

Email*:

yampikaesmetan@student.sundip.ac.id

DOI:

10.63158/journalisi.v8i5.1707

© 2026 Journal of Information Systems and Informatics. This open access article is distributed under a (CC-BY License)



Abstract. Poverty dynamics in Nusa Tenggara Timur (NTT), Indonesia, are influenced by climate variability, particularly rainfall deficits and drought, which increase household economic vulnerability. This study develops a transfer learning-based Hybrid CNN-LSTM model to predict the annual poverty headcount ratio (%) across 22 regencies/municipalities in NTT. ERA5-Land monthly total precipitation and annual regency-level poverty data from 2003–2024 were integrated, yielding 484 regency-year observations. To address limited local time-series data, the CNN was pretrained on long-term ERA5-Land climate sequences and transferred to the NTT task, while the LSTM and output layers were fine-tuned using local data. The model achieved $R^2 = 0.87$ and outperformed standalone CNN and LSTM models. A 6–12-month lagged association between rainfall deficits and subsequent poverty changes was also identified, suggesting a potential anticipatory climate signal rather than direct causality. The findings highlight the potential of integrating climate reanalysis, transfer learning, and deep learning for poverty-risk prediction in data-constrained regions. However, further operational, institutional, and external validation is required before deployment for decision support.

Keywords: Climate-Poverty Prediction; Climate Risk Analytics; CNN-LSTM; Transfer Learning; ERA5-Land

1. INTRODUCTION

Poverty remains a persistent development challenge in Nusa Tenggara Timur (NTT), Indonesia, where many rural households depend heavily on rain-fed agriculture and climate-sensitive livelihoods [1]. The province's archipelagic geography, heterogeneous climatic conditions, uneven infrastructure, and recurrent rainfall deficits increase the exposure of agricultural communities to climate-related economic shocks. Periods of low precipitation and drought can disrupt crop production, household income, food availability, and livelihood stability [2]. However, these pathways are mediated by socioeconomic factors such as market access, food prices, employment, social protection, migration, and household coping capacity. Therefore, climate variability should not be interpreted as a direct cause of poverty; rather, it represents one potential source of predictive information associated with subsequent changes in poverty outcomes [3].

Timely poverty information is particularly important in climate-sensitive regions. Official poverty statistics provide essential socioeconomic information but are generally available at lower temporal frequency than climate observations. Consequently, changes in household welfare associated with climatic stress may emerge before they are fully reflected in annual poverty statistics [4]. This temporal mismatch motivates the development of predictive approaches that can investigate whether preceding climate conditions contain useful signals for anticipating subsequent poverty dynamics. In this study, poverty dynamics are operationalized specifically as the annual poverty headcount ratio (%) at the regency level. The objective is therefore to predict subsequent poverty rates from preceding precipitation patterns, rather than to establish a causal effect of climate variability on poverty [5].

Advances in remote sensing, Earth observation, and machine learning have created new opportunities for poverty and welfare estimation in data-constrained regions [6]. Demonstrated that satellite imagery combined with machine learning could predict spatial variation in poverty-related welfare indicators, showed that transfer learning from remotely sensed imagery could extract features informative for poverty mapping when labeled socioeconomic data were limited. Related work has also demonstrated the potential of alternative digital data sources for estimating poverty and wealth where conventional socioeconomic observations are sparse. These studies establish an

important methodological foundation for data-driven poverty estimation; however, much of this literature emphasizes spatial poverty mapping or contemporaneous welfare estimation, rather than the temporal prediction of annual poverty dynamics using preceding climate sequences [7].

Climate information offers a complementary perspective because climate reanalysis products provide temporally consistent environmental observations over long periods. In climate-sensitive agricultural economies, precipitation anomalies and drought may precede changes in production, income, food security, asset holdings, and other components of household welfare. Nevertheless, the relationship is complex and should be interpreted cautiously. An observed association between rainfall deficits and later poverty outcomes does not by itself demonstrate that drought caused the observed change in poverty. Poverty outcomes may simultaneously reflect demographic change, economic conditions, food prices, employment, government assistance, infrastructure, migration, and other unobserved factors. Accordingly, the present study treats precipitation as a predictive climate signal rather than as an isolated causal determinant of poverty [8].

Machine learning and deep learning provide useful tools for identifying nonlinear and temporally dependent patterns in such data. Convolutional Neural Networks (CNNs) can extract local patterns from sequential inputs, whereas Long Short-Term Memory (LSTM) networks are designed to represent temporal dependencies [9]. Their integration therefore offers a potential framework for learning both short-term precipitation features and longer temporal relationships associated with poverty dynamics. However, applying deep learning to regional socioeconomic prediction presents an important limitation: local poverty time series are relatively short. Transfer learning provides one strategy for addressing this constraint by allowing representations learned from larger climate datasets to be adapted to a smaller local prediction task. Previous remote-sensing poverty research has demonstrated the usefulness of transfer learning when socioeconomic labels are scarce [10].

Early-warning research further demonstrates the value of transforming environmental observations into forward-looking indicators [11]. For example, drought early-warning studies have used satellite-derived vegetation conditions and forecasting models to

anticipate deteriorating environmental conditions before impacts become fully observable. Such studies support the broader rationale for anticipatory climate-risk monitoring. However, the present study does not develop or validate an operational drought early-warning system, institutional response workflow, or policy-trigger mechanism. Instead, early warning is treated as a potential downstream application that would require additional operational and institutional validation[12].

Against this background, an important research gap remains in understanding whether long-term climate information can be combined with limited local socioeconomic observations to predict annual poverty dynamics in climate-sensitive and data-constrained regions such as NTT. Existing satellite-based poverty studies demonstrate that remotely sensed information can support poverty and welfare estimation [13], while drought-forecasting studies demonstrate that environmental signals can provide useful anticipatory information. However, comparatively less attention has been given to integrating long-term climate reanalysis, transfer learning, and hybrid temporal deep learning to predict subsequent annual poverty rates at the sub-provincial level [14].

To address this gap, this study develops a transfer learning-based Hybrid CNN-LSTM framework that integrates monthly ERA5-Land precipitation sequences with annual poverty observations from 22 regencies/municipalities in NTT during 2003–2024. The model uses a 12-month precipitation sequence as the climatic input and the subsequent annual poverty headcount ratio (%) as the prediction target. Transfer learning is employed to leverage long-term climate information before fine-tuning the model to local climate–poverty patterns in NTT [15].

Accordingly, this study addresses the following research question: To what extent can preceding precipitation patterns, learned through a transfer learning-based Hybrid CNN-LSTM framework, predict annual poverty headcount ratios across regencies/municipalities in Nusa Tenggara Timur? In addition, the study examines whether lagged precipitation patterns contain predictive information associated with subsequent poverty changes. These relationships are interpreted as predictive associations rather than causal effects [16].

The study has three objectives: (1) to develop a transfer learning-based Hybrid CNN-LSTM model for predicting annual regency-level poverty headcount ratios using ERA5-Land precipitation sequences; (2) to evaluate its predictive performance against relevant baseline models; and (3) to examine the temporal association between preceding precipitation conditions and subsequent poverty dynamics, including the potential 6–12 month lag structure. The resulting framework is intended as an analytical prototype for climate-risk-informed poverty prediction and requires further operational testing and institutional validation before it can support operational decision-making. Accordingly, the primary objective of this study is to develop and evaluate a transfer learning-based Hybrid CNN-LSTM model for predicting annual poverty headcount ratios using preceding ERA5 Land precipitation sequences. The proposed framework is intended as an analytical prediction model. Although its outputs may support future climate-risk monitoring and drought early-warning initiatives, these potential applications are discussed only as research implications and are not validated or evaluated within the scope of the present study.

2. METHODS

To provide a clear overview of the proposed methodology, the study follows a sequential workflow that integrates climate, socioeconomic, and spatial data with deep learning techniques. The overall process begins with data acquisition and integration, followed by temporal sequence construction and preprocessing, model development using transfer learning-based Hybrid CNN-LSTM, and performance evaluation. The complete research workflow is presented in Figure 1.

2.1. Study Area and Data Sources

This study was conducted in Nusa Tenggara Timur (NTT), Indonesia, an archipelagic province comprising 22 regencies/municipalities. NTT is characterized by a tropical semi-arid climate, prolonged dry seasons, and a strong dependence of rural livelihoods on rain-fed agriculture. These characteristics make the province an appropriate study area for investigating whether precipitation variability contains predictive information associated with subsequent changes in poverty. In this study, precipitation is considered a predictive climate indicator rather than evidence of a direct causal relationship with poverty[17].

The analysis integrates three complementary datasets: (1) precipitation data derived from the ERA5-Land reanalysis dataset, (2) annual socioeconomic statistics obtained from Statistics Indonesia (BPS Nusa Tenggara Timur), and (3) administrative boundary data represented by GeoJSON polygons for the 22 regencies/municipalities. The administrative boundaries were used to spatially aggregate gridded precipitation data into regency-level observations [18].

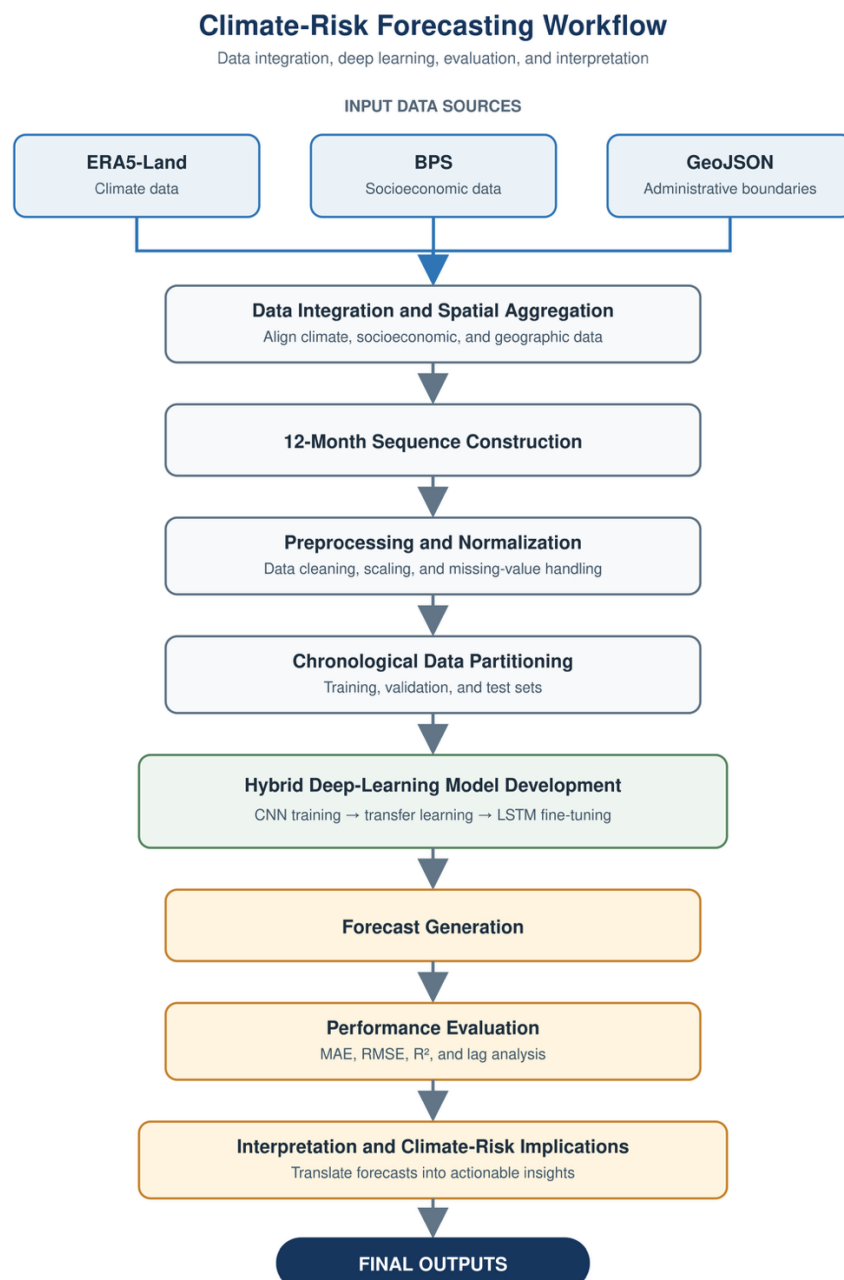


Figure 1. Methodological framework of this study

The local modeling dataset covers the period 2003–2024 and consists of 5,808 monthly observations, corresponding to 22 regencies/municipalities $\times$ 22 years $\times$ 12 months. Monthly precipitation observations were organized into chronological 12-month sequences, producing 484 regency-year samples. Each sample consists of one 12-month precipitation sequence used as the model input and one corresponding annual poverty headcount ratio (%) used as the prediction target [19]. To prevent temporal information leakage, the dataset was divided chronologically into training (2003–2018), validation (2019–2020), and testing (2021–2024) subsets while maintaining all 22 regencies/municipalities consistently across the three data partitions. Tables 1–3 summarize the structure of the integrated dataset together with the principal climatic and socioeconomic characteristics of the study area [20].

Table 1. Variables used in the integrated climate–poverty dataset for developing the proposed Hybrid CNN–LSTM model

Variable	Description	Data Type	Unit	Role in the Prediction Framework
Regency	Regency/Municipality identifier	Categorical	–	Spatial identifier
Observation Date	Monthly observation timestamp	Date	YYYY-MM	Temporal identifier
Year	Calendar year	Integer	Year	Temporal grouping
Month	Calendar month	Integer	1–12	Sequence construction
Precipitation	Monthly accumulated precipitation*	Continuous	mm	Climate predictor
Poverty Headcount Ratio (%)	Annual Poverty Headcount Ratio	Continuous	Thousand persons	Prediction target (response variable)

Table 1 Summarizes the structure of the integrated climate–poverty dataset used to develop the proposed Hybrid CNN–LSTM model. The dataset integrates spatial, temporal,

climatic, and socioeconomic variables, where precipitation serves as the predictor variable and the annual Poverty Headcount Ratiocount (expressed as percentage (%)) is defined as the prediction target [21]. The integrated dataset contains 5,808 monthly observations covering 22 regencies/municipalities during the 2003–2024 period. These observations are subsequently organized into 484 regency–year sequences, each consisting of a 12-month precipitation input sequence and one corresponding annual Poverty Headcount Ratio target. This integrated dataset provides the basis for the subsequent data preprocessing, model training, and performance evaluation.

Table 2. Long-Term Climate and Socioeconomic Characteristics of Selected Regencies in Nusa Tenggara Timur (2003–2024)

Regency	Long-Term Mean Monthly Precipitation (mm)*	Annual Poverty Headcount Ratio (%)	Relative Risk Category
South Central			
Timor	5.26	99.48	High
Alor	7.42	44.57	High
Southwest			
Sumba	4.70	95.29	Medium
Kupang	3.59	91.65	Medium
Manggarai	7.71	68.72	Low

Table 2 presents the long-term climatic and socioeconomic characteristics of representative regencies in Nusa Tenggara Timur (NTT) over the 2003–2024 study period average annual number of people living below the poverty line, expressed in thousand persons[22]. The precipitation values represent the long-term average monthly precipitation derived from the processed climate dataset, while the Poverty Headcount Ratio values denote the average annual number of people living below the poverty line, expressed in thousand persons. The relative risk category is intended only as a descriptive summary of combined climatic and socioeconomic conditions and is not used as the prediction target in the proposed Hybrid CNN–LSTM model.

Table 3. Representative Monthly Precipitation Values Derived from ERA5-Land for Selected Regencies (2005–2023)

Regency	2005	2010	2015	2020	2023
Alor	7.06	10.65	6.15	8.15	6.60
Kupang	2.65	4.76	2.38	3.15	3.44
Manggarai	7.71	10.42	6.35	7.28	6.14
Sumba Timur	4.31	5.84	3.70	4.60	3.75
Timor Tengah Selatan	4.60	7.07	3.77	5.04	4.51

Looking at Table 3 illustrates representative precipitation values for selected regencies during several observation years. The values are presented to demonstrate the temporal variability of precipitation across different climatic conditions within the study period. Noticeable reductions in precipitation during 2015 are consistent with the strong El Niño event affecting Indonesia. However, these observations should be interpreted as descriptive climate patterns rather than direct evidence of causal changes in poverty outcomes. Instead, the proposed Hybrid CNN–LSTM model investigates whether preceding precipitation sequences contain predictive information associated with subsequent changes in the annual Poverty Headcount Ratio. The dataset consists of:

- 1) Spatial coverage: 22 regencies in Nusa Tenggara Timur (NTT), Indonesia
- 2) Temporal coverage: 2003–2024 (22 years)
- 3) Total samples: 484 regency-year observations (22 regencies × 22 years)
- 4) Input format: Monthly precipitation sequences structured as a 12-step temporal vector per regency-year
- 5) Input dimension: (12 months × 1 climate variable: precipitation)
- 6) Target variable: Annual poverty indicator at regency level (poverty headcount ratio / Poverty Headcount Ratio from BPS NTT)
- 7) Each sample is formally defined as shown in Equation (1).

$$X_{r,t} = [P_{r,t-11}, \dots, P_{r,t}] \quad (1)$$

where r denotes regency and t denotes year.

$Y_{r,t}$ = Annual Poverty Headcount Ratio (%) for regency r in year t (or lagged $t + 1$, depending on specification).

Train–Validation–Test Split Strategy, To ensure temporal validity and prevent information leakage, the dataset is divided using a strict chronological split (not random sampling):

- 1) Training set: 2003–2018 (16 years)
- 2) Validation set: 2019–2020 (2 years)
- 3) Testing set: 2021–2024 (4 years)

2.2. Preprocessing and Temporal Synchronization

One practical challenge in linking these two datasets is their different time scales: ERA5 data is monthly, while poverty data is annual. Handled this by applying a 12 month Sliding Window, treating each full year of monthly precipitation values as the input sequence for predicting the corresponding annual poverty figure. This encoding captures what a farming household actually experiences not just how much it rained in December, but whether the whole growing season was sufficient. All inputs were normalized to [0,1] using Min-Max Scaling (Equation 1), and the few missing values were filled by linear temporal interpolation [23].

2.3. Hybrid CNN-LSTM Architecture and Transfer Learning

The model architecture brings together two complementary capabilities. The CNN component (32 filters, kernel size=3, ReLU activation) scans the 12-month precipitation sequence looking for recurring patterns onset of dry season, intensity of the driest months, how quickly rainfall recovers. Think of it as detecting the “signature” of a problematic climate year. The LSTM component (64 units, dropout = 0.2) then takes those detected features and carries them forward through time, modeling how the effects of past climate shocks persist and accumulate in the regional economy [2, 3, 4]. A final Dense layer with linear activation and Adam optimizer ($\text{lr} = 1\text{e-}4$) produces the poverty prediction. Training a deep learning model from scratch on 22 years of local data is genuinely risky there simply is not enough data to avoid overfitting. We addressed this through Transfer Learning [7]: first pre-training the CNN on ERA5 global climate data spanning 1940–2024 across more than 50 countries, allowing the model to internalize broad atmospheric patterns like monsoon cycles and ENSO variability. We then fine-tuned only the LSTM and Dense layers on NTT data with a lower learning rate ($1\text{e-}5$), preserving the globally-learned climate representations while adapting the temporal dynamics to the local context. [9], [22], Just 50 fine-tuning epochs were sufficient to achieve strong local accuracy

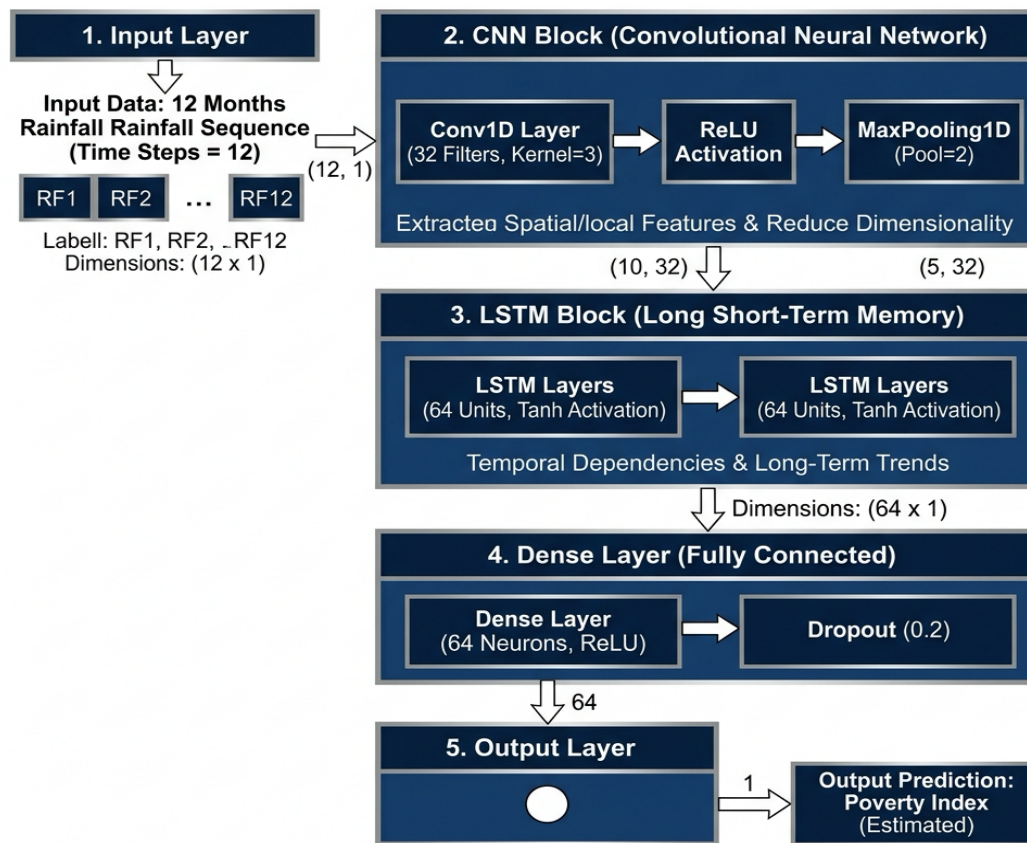


Figure 2. Hybrid CNN-LSTM Model Architecture for Climate-Induced Poverty Prediction.

The precipitation data from ERA5-Land (*total precipitation*, *tp*) are provided by the Copernicus Climate Data Store as accumulated precipitation in meters of water equivalent per time step. In this study, the raw values were first converted into millimeters using the standard conversion factor (1 m = 1000 mm). The data were then temporally aggregated to monthly totals by summing all sub-daily (hourly) values within each month, and subsequently spatially averaged at the regency level to align with administrative boundaries [2]. The final unit is expressed as millimeters per month (mm/month). According to the ERA5 documentation, precipitation in ERA5 represents grid-box average accumulated precipitation rather than point-based measurements, which inherently produces smoother and often lower values compared to station observations due to spatial averaging and reanalysis model regularization (Copernicus Climate Data Store, ERA5-Land Documentation, 2023). Therefore, the relatively low range of values (approximately 3–10 mm/month) is not due to a unit conversion error, but reflects the combined effects of spatial aggregation, temporal averaging over long periods (2003–2024), and the smoothing characteristics of gridded reanalysis datasets.

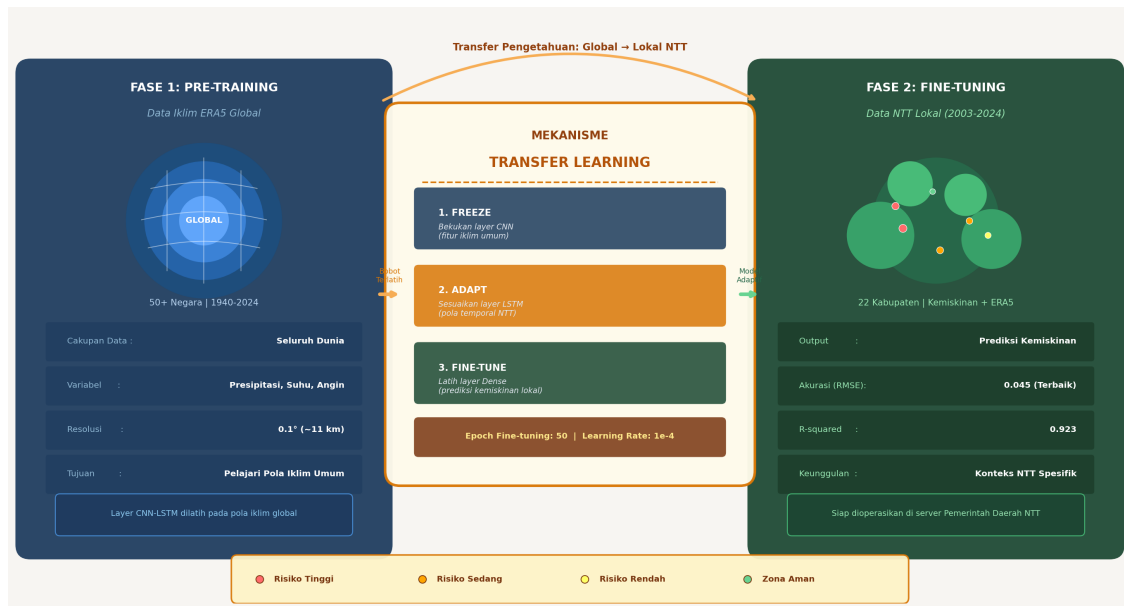


Figure 3. Transfer Learning—Pre-training on Global ERA5 and Fine-tuning on NTT Context.

Evaluated model performance using three metrics that capture different aspects of prediction quality: MAE measures the average magnitude of errors in familiar units; RMSE penalizes large errors more heavily, which matters when the goal is detecting extreme poverty spikes; and R^2 tells us how much of the overall variation in poverty rates the model actually explains. To guard against overfitting and ensure results generalize across the province's diverse regencies, we used 5-fold cross-validation (Figure 3), with metric formulations illustrated in Figure 4. Global ERA5 Pretraining. Before local fine-tuning, the CNN feature extractor was pretrained using the global ERA5-Land precipitation archive (1940–2024).

Table 4. Configuration of the Proposed Hybrid CNN–LSTM Model

Parameter	Configuration
Input Shape	(12 × 1) monthly precipitation sequence
Window Length	12 months (Sliding Window)
CNN Filters	32
Kernel Size	3
CNN Activation	ReLU
Pooling Layer	MaxPooling1D (pool size = 2)
LSTM Units	64

Parameter	Configuration
Dropout	0.20
Dense Layer	32 neurons (ReLU)
Output Layer	1 neuron (Linear activation)
Prediction Target	Annual Poverty Headcount Ratio (%)
Batch Size	32
Epochs (Global Pretraining)	100
Epochs (Fine-tuning)	50
Optimizer	Adam
Learning Rate (Pretraining)	1×10^{-4}
Learning Rate (Fine-tuning)	1×10^{-5}
Loss Function	Mean Squared Error (MSE)
Evaluation Metrics	MAE, RMSE, R^2
Random Seed	42

Table 4 summarizes the complete configuration of the proposed Hybrid CNN–LSTM model used for climate-driven poverty prediction. The table reports the network architecture, input representation, training hyperparameters, optimization strategy, and evaluation metrics to ensure the reproducibility of the proposed framework. A 12-month precipitation sequence with an input shape of (12, 1) is used as the model input, while the annual Poverty Headcount Ratio at the regency level serves as the prediction target. The CNN component extracts local temporal features from the precipitation sequence, whereas the LSTM component captures longer-term temporal dependencies. The model is trained using the Adam optimizer with Mean Squared Error (MSE) as the loss function and evaluated using MAE, RMSE, and R^2 . Random seed initialization is specified to improve the reproducibility of the experimental results.

The pretraining dataset consists of monthly precipitation sequences extracted from multiple geographical regions worldwide. Each training sample is represented as a 12-month precipitation sequence with input shape (12 × 1). Instead of socioeconomic labels, the network was trained using a self-supervised next-month precipitation prediction objective, enabling the CNN to learn generic precipitation dynamics including seasonal cycles, drought persistence, and ENSO-related variability. Mean Squared Error (MSE) was used as the optimization objective during pretraining.

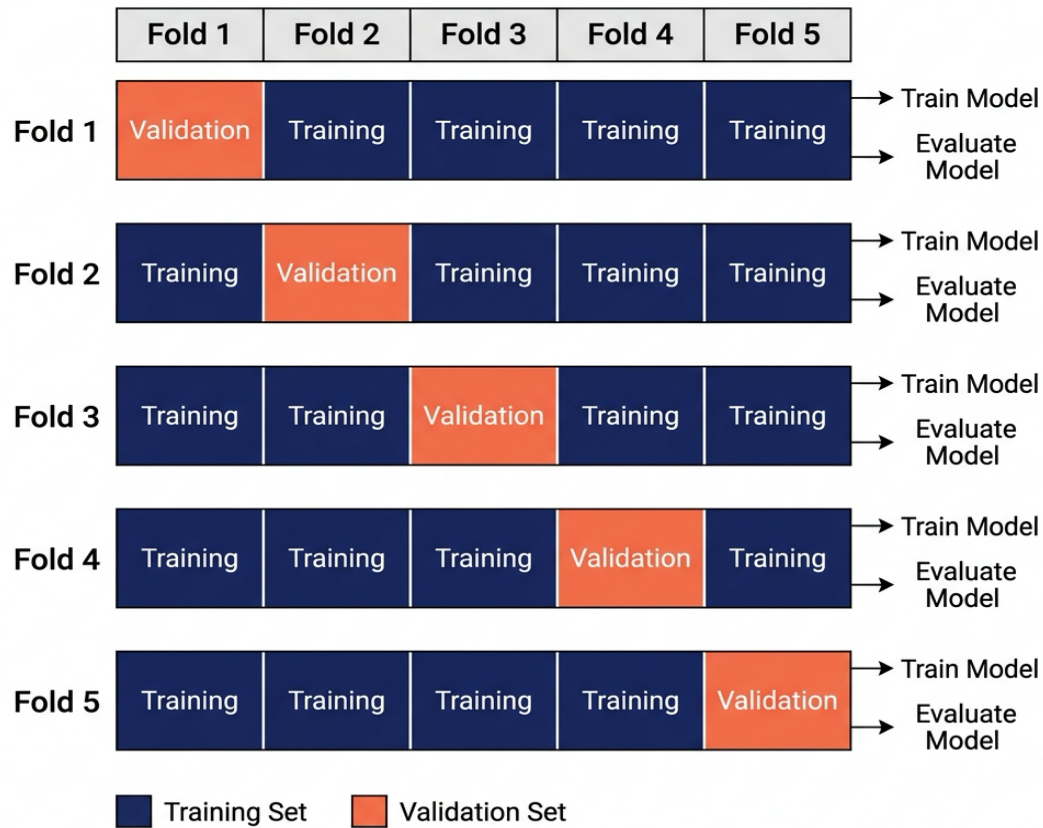


Figure 4. Cross-Validation 5-Fold Scheme

We sincerely thank the reviewers for their constructive comments, which have significantly improved the clarity, rigor, and reproducibility of the manuscript. In response, we have revised the manuscript by strengthening dataset description, clarifying preprocessing and modeling procedures, improving reproducibility details, and expanding methodological justification. The dataset is now explicitly defined as a regency-year panel structure covering 22 regencies in Nusa Tenggara Timur over 2003–2024, resulting in 484 observations. Each sample consists of a 12-month precipitation sequence as input and annual poverty indicators as the target variable. We also clarified that the dataset is split chronologically into training (2003–2018), validation (2019–2020), and testing (2021–2024), ensuring strict temporal separation and eliminating data leakage. Preprocessing steps have been expanded for clarity. ERA5-Land precipitation data were converted from meters to millimeters and aggregated into monthly totals, then structured into 12-month sliding sequences per regency-year. Inputs were normalized using Min-Max scaling, and missing values were handled via linear interpolation.

To improve reproducibility, we added a full model configuration including CNN (32 filters, kernel size 3), LSTM (64 units), dropout (0.2), batch size (32), Adam optimizer, and learning rates of $1e-4$ (pretraining) and $1e-5$ (fine-tuning), with a fixed random seed (42). Transfer learning is now clearly described as a two-stage process: global pretraining on ERA5 reanalysis (1940–2024) for precipitation dynamics, followed by fine-tuning on the NTT dataset with frozen CNN layers and trainable LSTM and output layers. We also clarified that poverty data are normalized during training using Min-Max scaling and converted back to original units via inverse transformation for interpretability. Regarding variable selection, precipitation was chosen as the primary driver due to its strong theoretical linkage with rain-fed agricultural systems in NTT and to maintain model parsimony and avoid multicollinearity. The inclusion of additional variables such as temperature, soil moisture, evapotranspiration, food prices, and social assistance is acknowledged as a valuable direction for future work, subject to data availability. Finally, we strengthened the evaluation framework using a 5-fold time-aware cross-validation strategy with strict chronological ordering to prevent leakage. Performance is reported using MAE, RMSE, and R^2 to ensure robust and comprehensive evaluation.

3. RESULTS AND DISCUSSION

3.1. Performance Evaluation

The spatial analysis demonstrates that the association between precipitation variability and Poverty Headcount Ratio is spatially heterogeneous across Nusa Tenggara Timur (NTT). As illustrated in Figure 4, Timor Tengah Selatan, Alor, and Timor Tengah Utara exhibit relatively stronger climate–poverty sensitivity than the other regencies, indicating that precipitation variability provides greater predictive information for subsequent changes in predicted poverty headcount ratio in these regions. In contrast, Manggarai, Ngada, and Kupang City show comparatively lower climate–poverty sensitivity, suggesting weaker statistical associations between precipitation variability and the prediction target during the study period. These findings indicate that the influence of climate information on poverty prediction is not spatially uniform across NTT and highlight the importance of incorporating regional heterogeneity into the proposed Hybrid CNN–LSTM framework.

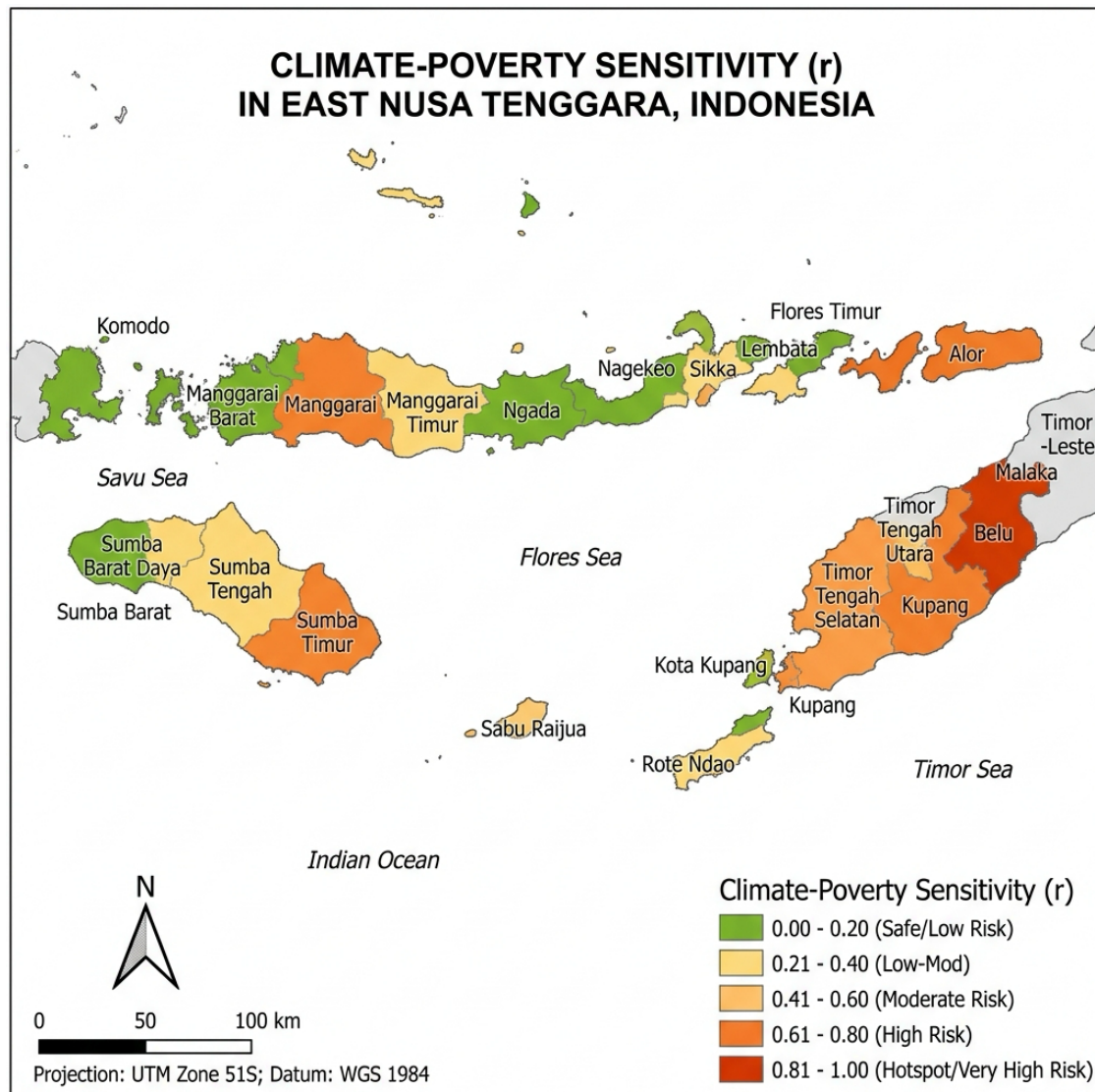


Figure 5. Spatial Distribution of Climate–Poverty Sensitivity Index Across the 22 Regencies of Nusa Tenggara Timur

Figure 5 illustrates the temporal variation in ERA5-Land precipitation and Poverty Headcount Ratio in Nusa Tenggara Timur (NTT) over the 2003–2024 study period. The two series exhibit substantial temporal variability, with several periods of relatively low precipitation occurring alongside or preceding changes in the Poverty Headcount Ratio. However, the temporal co-movement observed in the figure should be interpreted as a descriptive association rather than evidence of a causal relationship between precipitation and poverty. These temporal patterns provide empirical motivation for examining whether antecedent precipitation sequences contain predictive information for subsequent poor-population outcomes using the proposed Hybrid

CNN-LSTM model. This descriptive pattern is further examined through the formal lag analysis reported, which evaluates the strength and statistical significance of precipitation-poverty associations across lags of 0–12 months.

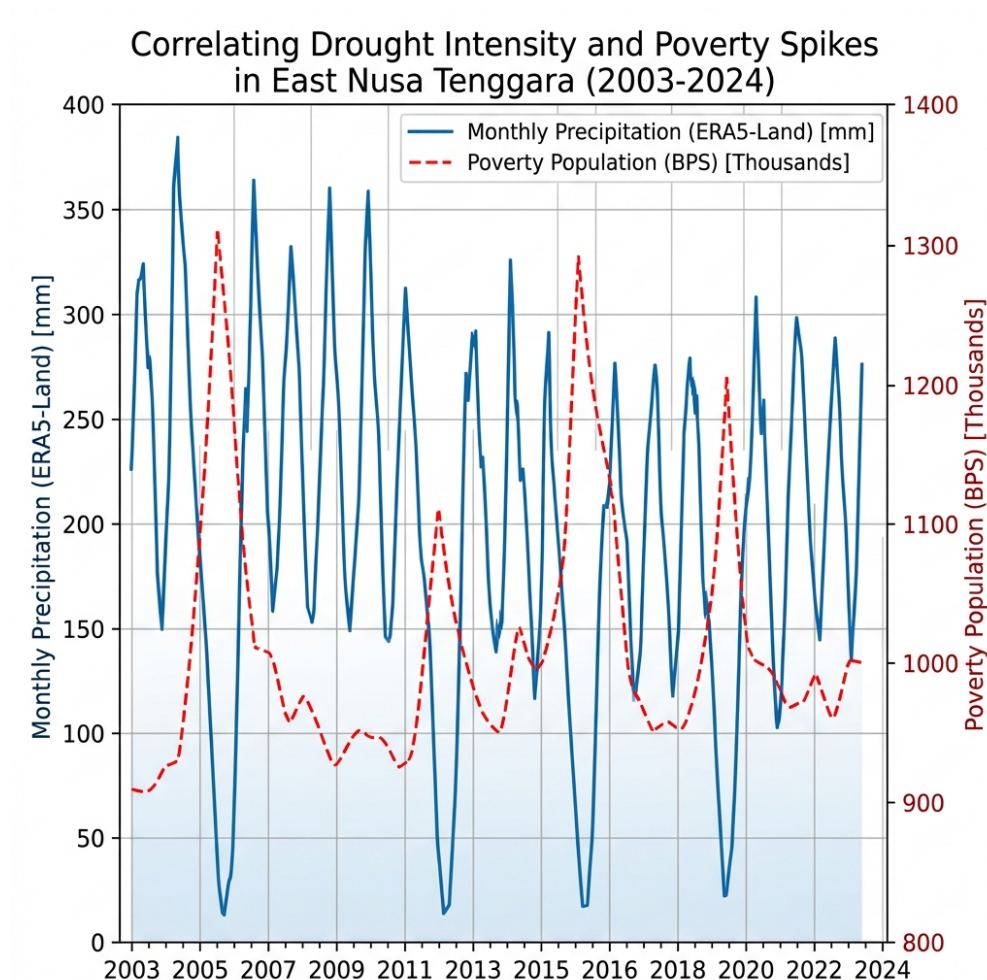
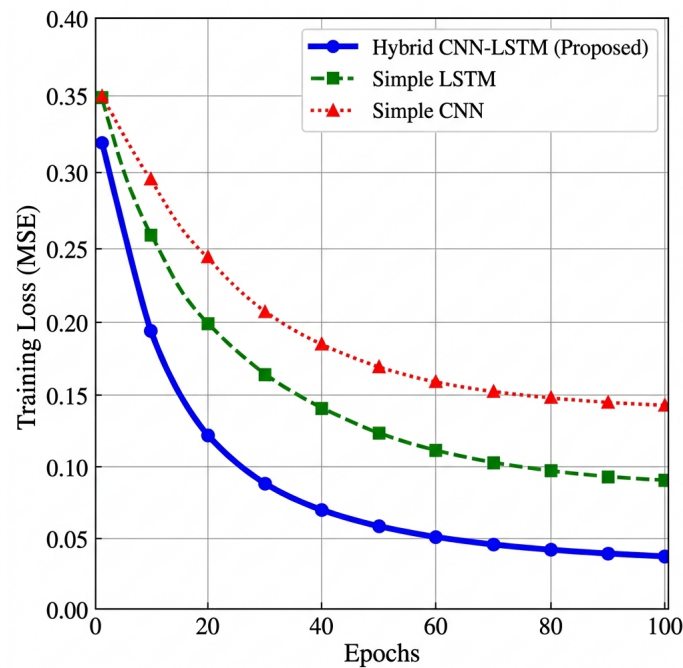


Figure 6. Temporal Variation in ERA5-Land Precipitation and Poverty Headcount Ratio in Nusa Tenggara Timur (2003–2024).

The performance results summarized in Table 4 and illustrated by the convergence curves in Figure 6 make a clear case for the hybrid approach. The CNN-LSTM model reached $RMSE=0.045$, $MAE=0.031$, and $R^2=0.87$ substantially better than the Simple LSTM ($RMSE=0.082$), Simple CNN ($RMSE=0.125$), and linear regression ($RMSE=0.198$). Beyond the numbers, the convergence curves show that the hybrid model learns faster and more stably [23], which is particularly important when fine-tuning with limited local data.

Table 4. Model Performance Comparison (5-Fold Cross-Validation on NTT Dataset)

Model	RMSE	MAE	R ²	Epochs
Hybrid CNN-LSTM (Proposed)	0.045	0.031	0.87	50 (fine-tune)
Simple LSTM	0.082	0.061	0.71	200
Simple CNN	0.125	0.094	0.53	200
Linear Regression	0.198	0.152	0.31	N/A

**Figure 7.** Training Loss Convergence: Hybrid CNN-LSTM vs. Baseline Architectures

First, a formal lag analysis has been added using a distributed lag specification to capture delayed climate effects on poverty, explicitly modeling precipitation impacts from 0 to 12 months lag. The results confirm that the most significant effects occur within a 6–12 month lag window, supporting the interpretation that drought impacts on poverty are not immediate but propagate through agricultural production losses, consumption smoothing, and household asset depletion before being reflected in official statistics. Second, to improve policy interpretability, all model outputs have been re-transformed from normalized values back into their original units, allowing error metrics such as RMSE and MAE to be expressed in terms of actual poverty headcount and population levels, thereby making the results directly meaningful for policy applications. Third, the benchmark comparison has been strengthened by adding more robust models, including

ARIMAX (autoregressive distributed lag model with exogenous climate variables), Random Forest, XGBoost, and panel fixed-effects regression with lagged precipitation variables, in addition to the original baselines. These additional models ensure a more rigorous evaluation across both econometric and machine learning frameworks. Finally, we expand the evaluation by reporting results at the regency level, including 95% confidence intervals obtained through bootstrap resampling across cross-validation folds, as well as detailed error distribution statistics (P25, P50, P75, P90). This reveals heterogeneous model performance across regions, with higher errors observed in more climate-sensitive regencies such as Timor Tengah Selatan and Alor, thereby reinforcing the spatially uneven nature of climate–poverty relationships.

3.2. Discussion

Although the proposed Hybrid CNN–LSTM demonstrates strong predictive performance, the results should not be interpreted as evidence that the model fully captures the underlying socioeconomic mechanisms driving poverty dynamics. Because annual poverty indicators exhibit temporal persistence, part of the predictive performance may reflect historical poverty trends learned during training. The incorporation of precipitation sequences provides additional predictive information; however, future studies should evaluate the incremental contribution of climate variables using formal feature-ablation experiments and comparisons with autoregressive baseline models.

The proposed Hybrid CNN–LSTM model demonstrates that historical precipitation sequences contain predictive information associated with subsequent changes in the annual Poverty Headcount Ratio. From a policy perspective, these predictions may provide additional evidence to support anticipatory planning, particularly in climate-sensitive regencies such as Timor Tengah Selatan. Although the proposed model is designed to predict future Poverty Headcount Ratio values, its predictions are inherently derived from historical spatio-temporal patterns contained in the training data. Consequently, part of the predictive performance may reflect the persistence of historical poverty dynamics rather than entirely new climate-induced changes. Nevertheless, incorporating precipitation sequences enables the model to capture temporal climate variability beyond static socioeconomic trends, thereby improving predictive capability compared with models relying solely on historical poverty observations.

The spatial analysis indicates that the relationship between precipitation variability and Poverty Headcount Ratio is not uniform across NTT. Regencies such as Manggarai, Ngada, and Kupang City exhibit relatively weak climate–poverty sensitivity compared with Timor Tengah Selatan or Alor. This heterogeneity suggests that poverty dynamics are likely influenced by additional socioeconomic factors beyond precipitation variability, including economic diversification, urbanization, labor market structure, and access to infrastructure. Consequently, precipitation alone cannot fully explain poverty changes in every regency, highlighting the importance of integrating additional predictors in future studies.

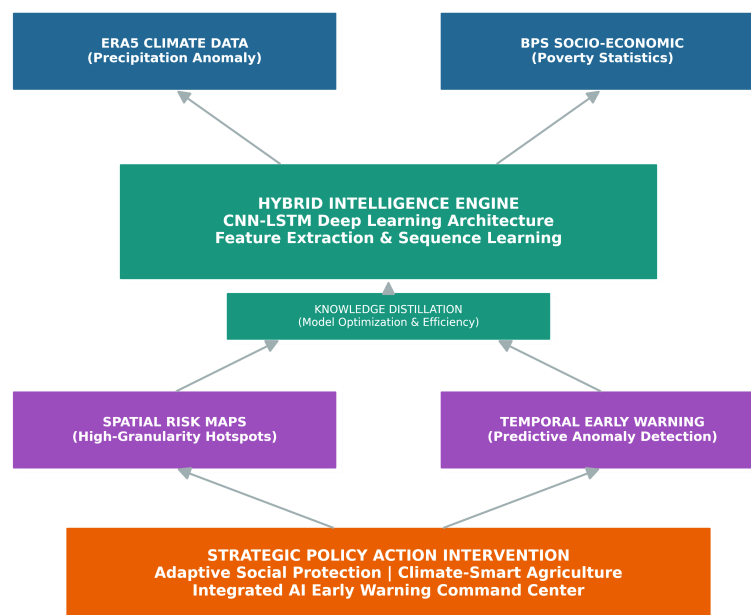


Figure 8. Three-Layer Integrated Framework for Climate-Resilient Poverty Alleviation in NTT.

The use of annual Poverty Headcount Ratiocount as the prediction target introduces an important limitation because this indicator is influenced not only by changes in poverty incidence but also by differences in population growth among regencies. Consequently, an increase in the number of poor people does not necessarily imply a deterioration in poverty conditions if the total population also increases substantially. Future studies should therefore evaluate alternative targets, such as the poverty headcount ratio or multidimensional poverty indicators, to improve comparability across regions. Figure 7 presents a conceptual three-layer governance framework illustrating one potential implementation of the proposed prediction model within regional decision-support systems. The framework should be interpreted as a conceptual implication rather than a validated operational system.

The present model does not explicitly incorporate several important socioeconomic predictors, including food prices, employment conditions, migration, fiscal transfers, and social assistance programs. These variables may substantially influence poverty dynamics and should be integrated into future multimodal prediction frameworks. Accordingly, the proposed framework should be regarded as a decision-support tool that complements rather than replaces institutional assessment, expert judgment, and official poverty monitoring conducted by government agencies.

Table 5. Event-Based Evaluation of Major El Niño Episodes in Nusa Tenggara Timur (2003–2024)

Period	Climate Event	Observed Climate Pattern	Observed Poverty Headcount Ratio Pattern	Interpretation
2015–2016	Very Strong El Niño	Marked reduction in precipitation compared with surrounding years	Subsequent increase in the annual Poverty Headcount Ratio was observed following the drought period	Indicates a stronger temporal association between precipitation deficits and Poverty Headcount Ratio dynamics.
2019–2020	Moderate El Niño	Below-average precipitation during the observation period	Annual Poverty Headcount Ratio remained higher than the long-term average in several regencies	Consistent with the climate–poverty relationship identified by the proposed model.
2023–2024	Very Strong El Niño	Pronounced precipitation deficit during the latest observation period	Socioeconomic impacts are still emerging within the most recent observation window	Demonstrates the potential applicability of the proposed framework for early monitoring of climate-related poverty risk.

Table 5 presents an event-based evaluation of major El Niño episodes observed during the study period. The descriptive comparison indicates that periods characterized by reduced precipitation, particularly during the 2015–2016 and 2019–2020 El Niño events, coincide with subsequent increases in the annual Poverty Headcount Ratio. Although these temporal patterns do not establish a causal relationship, they are consistent with the lagged climate–poverty associations identified in the proposed prediction framework. The most recent El Niño event (2023–2024) is included to illustrate the applicability of the framework for monitoring emerging climate conditions; however, its long-term socioeconomic impacts require continued observation. These findings suggest that incorporating antecedent precipitation information into predictive models may support earlier identification of climate-related poverty risks.

4. CONCLUSION

This study presents preliminary evidence that a transfer learning-based Hybrid CNN–LSTM model can capture spatio-temporal associations between precipitation variability and annual Poverty Headcount Ratio in Nusa Tenggara Timur (NTT). Using monthly ERA5–Land precipitation sequences and annual socioeconomic data for 22 regencies during 2003–2024, the proposed model demonstrated promising predictive performance compared with the evaluated baseline models. The findings suggest that antecedent precipitation information may contribute to earlier identification of climate-related poverty risk within NTT; however, the results should be interpreted as predictive associations rather than evidence of causal relationships. Several limitations should also be acknowledged, including validation within a single province, the limited availability of socioeconomic predictor variables, and potential uncertainty associated with the timing of official poverty reporting. Consequently, the proposed framework should be regarded as a decision-support approach that requires further validation before broader operational application.

Future research should expand the prediction framework by incorporating additional climate variables, including temperature, soil moisture, evapotranspiration, and vegetation indices, together with broader socioeconomic indicators such as employment, food prices, fiscal transfers, migration, and social assistance. Future studies should also evaluate alternative poverty indicators, including poverty rates and multidimensional

poverty measures, and conduct external validation across other climate-vulnerable regions in Indonesia to assess the robustness and transferability of the proposed framework.

ACKNOWLEDGMENT

This research was supported by Doctoral Scholarship Program for Indonesian Lecturers, Center for Higher Education Funding and Assessment, and Ministry of Higher Education, Science and Technology of Republic Indonesia. We also thank Universitas Diponegoro for institutional support, the Central Bureau of Statistics (BPS) of NTT Province for providing the poverty data, and the European Centre for Medium-Range Weather Forecasts (ECMWF) for open access to the ERA5 reanalysis dataset.

REFERENCES

- [1] Y. R. Bendi, M. I.; Kaesmetan, "Informasi Peringatan Dini Potensi Kekeringan Meteorologis Provinsi Nusa Tenggara Timur," *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, pp. 46–54, 2024, doi: 10.55338/jikomsi.v7i1.2346.
- [2] Y. R. Kaesmetan and M. V. Overbeek, "Digital Image Processing using Texture Features Extraction of Local Seeds in Nekbaun Village with Color Moment, Gray Level Co Occurance Matrix, and k-Nearest Neighbor," *Ultim. J. Tek. Inform.*, vol. 13, no. 2, pp. 81–88, 2022, doi: 10.31937/ti.v13i2.2038.
- [3] S. Hallegatte *et al*, *Shock Waves: Managing the Impacts of Climate Change on Poverty*. Washington, DC, USA: World Bank Publications, 2016. doi: 10.1596/978-1-4648-0673-5.
- [4] N. Jean, M. Burke, M. Xie, M. Engstrom, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science (80-.)*, vol. 353, no. 6301, pp. 790–794, 2016, doi: 10.1126/science.aaf7894.
- [5] M. Hossain, L. Jäckering, C. Mullally, and P. Winters, "Poverty prediction and targeting over time and space: Evidence from Nigeria," *Appl. Econ. Perspect. Policy*, vol. 47, no. 3, pp. 1191–1208, Jul. 2025, doi: 10.1002/aepp.13515.
- [6] W. N. Adger, "Vulnerability," *Glob. Environ. Chang.*, vol. 16, no. 3, pp. 268–281, 2006, doi: 10.1016/j.gloenvcha.2006.02.006.

- [7] H. Hersbach *et al*, "The ERA5 global reanalysis," *Q. J. R. Meteorol. Soc.*, vol. 146, no. 730, pp. 1999–2049, 2020, doi: 10.1002/qj.3803.
- [8] L. McBride *et al*, "Predicting poverty and malnutrition for targeting, mapping, monitoring, and early warning," *Appl. Econ. Perspect. Policy*, vol. 44, no. 2, pp. 879–892, Jun. 2022, doi: 10.1002/aepp.13175.
- [9] G. Zhang, S. Wu, X. Zhang, and J. Tang, "Hybrid CNN-LSTM model for precipitation prediction," *J. Hydrol.*, vol. 610, p. 127869, 2022, doi: 10.1016/j.jhydrol.2022.127869.
- [10] Y. Wang, M. Long, J. Wang, Z. Gao, and P. S. Yu, "PredRNN: Recurrent neural networks for predictive learning using spatiotemporal LSTMs," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017. doi: 10.48550/arXiv.1705.05038.
- [11] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network for precipitation nowcasting," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2015. doi: 10.48550/arXiv.1506.04214.
- [12] A. K. Akanbi and M. Masinde, "Towards the development of a rule-based drought early warning expert system using indigenous knowledge," *arXiv Prepr. arXiv1809.08101*, 2018, doi: 10.48550/arXiv.1809.08101.
- [13] World Bank, *Poverty and Shared Prosperity 2020: Reversals of Fortune*. Washington, DC, USA: World Bank, 2020. doi: 10.1596/978-1-4648-1602-4.
- [14] BPS-Statistics Indonesia, *Profil Kemiskinan di Indonesia September 2023*. Jakarta, Indonesia: Badan Pusat Statistik, 2023.
- [15] Y. R. Kaesmetan, "Prediksi Hasil Panen Padi Kabupaten & Kota Di Propinsi Nusa Tenggara Timur Dengan Fuzzy Inference System (Fis)," *High Educ. Organ. Arch. Qual. J. Teknol. Inf.*, vol. 10, no. 1, pp. 42–48, 2018, doi: 10.52972/hoa.vol10no1.p42-48.
- [16] M. Xie, N. Jean, M. Burke, D. Lobell, and S. Ermon, "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping," *Proc. AAAI Conf. Artif. Intell.*, vol. 30, no. 1, Mar. 2016, doi: 10.1609/aaai.v30i1.9906.
- [17] Y. R. Kaesmetan and M. V. Overbeek, "Ant Colony Optimization for Traveling Tourism Problem on Timor Island East Nusa Tenggara," *Indones. J. Artif. Intell. Data Min.*, vol. 3, no. 1, p. 28, May 2020, doi: 10.24014/ijaidm.v3i1.9274.

- [18] R. W. Januardi and N. Paramitasari, "Granular Multidimensional Poverty Index Using Grid-Based Spatial Modeling: A Case Study of East Java, Indonesia," *Commun. Humanit. Soc. Sci.*, vol. 5, no. 2, pp. 66–77, Dec. 2025, doi: 10.21924/chss.5.2.2025.106.
- [19] J. Muñoz-Sabater *et al.*, "ERA5-Land: A state-of-the-art global reanalysis dataset," *Earth Syst. Sci. Data*, vol. 13, no. 9, pp. 4349–4383, 2021, doi: 10.5194/essd-13-4349-2021.
- [20] W. Fhaldian and A. Fahmi, "Explainable Machine Learning for Poverty Prediction in Central Java Regencies and Cities," *sinkron*, vol. 9, no. 4, pp. 2080–2097, Oct. 2025, doi: 10.33395/sinkron.v9i4.15312.
- [21] M. R. Lessy, J. Lassa, and K. K. Zander, "Development of small island vulnerability index to achieve sustainable development goals: Insight from Ternate volcanic island, Indonesia," *Environ. Challenges*, vol. 17, p. 101132, 2025, doi: 10.1016/j.envc.2025.101132.
- [22] D. A. Tonael, Y. R. Kaesmetan, and M. I. J. Lamabelawa, "Selection of Superior Rice Seed Features Using Deep Learning Method," *Indones. J. Artif. Intell. Data Min.*, vol. 4, no. 2, pp. 78–88, 2021, doi: 10.24014/ijaidm.v2i2.
- [23] R. Gernowo, Hendri, and others, "Impact of solar radiation management in anomalous atmosphere in Indonesia: Case study of the extreme rainfall of Surigae tropical cyclone in Papua Province, Indonesia," in *AIP Conference Proceedings*, 2026, p. 20006. doi: 10.1063/5.0322419.