Predictive Modeling of Typhoid Incidence in Nepal Under Extreme Climate Change Scenarios Using Machine Learning Algorithm

Kritika Baral[1][0009-0007-1517-1063], Rabindra Bista[2*][0000-0002-0638-5840] and Jens Kaasbøll[3][0000-0001-9666-0887]

[1] Kathmandu University, Department of Health Informatics, Dhulikhel, Nepal
[2] Kathmandu University, Department of Computer Science and Engineering, Dhulikhel, Nepal
[3] Department of Informatics, University of Oslo, Oslo, Norway
*Corresponding Author = Rabindra Bista (rbista@ku.edu.np)

Abstract. Typhoid fever remains a significant public health challenge in Nepal, where monsoon-driven flooding and climate variability increasingly disrupt water, sanitation, and hygiene (WASH) systems. This study examined the spatiotemporal relationship between typhoid incidence and climate-induced extreme events across Nepal's 77 districts from 2015 to 2023 by integrating district-level outpatient surveillance data from the Health Management Information System (HMIS) with flood occurrence records and gridded climate variables, including precipitation, temperature, and relative humidity. Descriptive and correlation analyses, including lagged assessments, were conducted, followed by machine learning modeling using Random Forest (RF), XGBoost, and Long Short-Term Memory (LSTM) algorithms. An ensemble model combining all three algorithms was also evaluated. Typhoid incidence exhibited strong seasonality, with 70–80% of cases occurring during monsoon months (June–September) and showed a positive association with flood frequency ($r = 0.334$), which strengthened when floods preceded cases by one month ($r = 0.412$). Among the evaluated models, XGBoost achieved the best individual predictive performance ($R^2 = 0.52$), while the ensemble model achieved the highest overall accuracy ($R^2 = 0.55$). Flood-related features accounted for approximately 18% of explained variance. Projections under a medium-emissions scenario (SSP2-4.5) indicate a potential 25% increase in national typhoid burden by 2050, with disproportionate impacts in flood-prone Terai districts. These findings identify typhoid fever in Nepal as a climate-sensitive disease and highlight the importance of integrating climate indicators into disease surveillance and early warning systems to support climate-resilient public health planning.

Keywords: Typhoid fever, Climate change, Flooding, Machine learning, Extreme climate events, Nepal, Disease prediction, public health surveillance

1 Introduction

Typhoid fever, caused by Salmonella enterica serovar Typhi, remains a persistent public health challenge in low- and middle-income countries, where inadequate water,

sanitation, and hygiene (WASH) infrastructure facilitates fecal–oral transmission (World Health Organization [WHO], 2023). Globally, the disease causes an estimated 9.2 million cases and 133,000 deaths annually, with the highest incidence concentrated in South and Southeast Asia (GBD 2023 Causes of Death Collaborators, 2025). Despite the availability of an effective typhoid conjugate vaccine (TCV), seasonal outbreaks continue to burden healthcare systems in endemic settings, disproportionately affecting children under 15 years and socioeconomically marginalized communities (Mogasale et al., 2016; Antillón et al., 2025).

Nepal is a landlocked Himalayan country with diverse ecological zones, ranging from the flood-prone Terai plains to high-altitude mountain regions vulnerable to landslides. In this setting, seasonal monsoons, rapid urbanization, and socioeconomic inequalities contribute to recurring typhoid outbreaks that strain limited health resources and especially affect children under 15 years (60–70% of cases) and marginalized communities in low-income districts (MoHP, 2023; Karkey et al., 2018). Nepal records approximately 80,000–100,000 clinically diagnosed typhoid cases annually through its Health Management Information System (HMIS), with incidence rates as high as 351 per 100,000 in urban Kathmandu (Coalition Against Typhoid, 2024). Climate change is intensifying these pressures: Nepal is warming at approximately 0.056°C per year nearly twice the global average and is experiencing more frequent and severe monsoon floods (IPCC, 2022; MoFE, 2021). Flooding disrupts WASH systems by contaminating drinking water sources and breaching sanitation infrastructure, creating environmental conditions conducive to Salmonella Typhi persistence and transmission (Levy et al., 2023; Bhandari et al., 2022).

While the epidemiology of typhoid in Nepal has been documented through hospital-based surveillance and vaccine effectiveness studies (Karkey et al., 2018; Shakya et al., 2019), the influence of climate-induced extreme events on spatiotemporal disease patterns at the district level remains poorly characterized. Prior studies have primarily focused on urban settings or national-level estimates, with limited integration of high-resolution climate data, flood records, or subnational heterogeneity (Bhatt et al., 2019; Pitzer et al., 2025). Traditional statistical approaches, including autoregressive models and generalized linear models, struggle to capture the nonlinear and lagged relationships between climate variability and disease incidence that are characteristic of monsoon-driven epidemiology (Choi et al., 2023; Hess et al., 2020).

Advances in machine learning (ML) offer new opportunities to address these limitations. Tree-based ensemble methods such as Random Forest and XGBoost, as well as recurrent neural networks such as LSTM, can simultaneously model nonlinear interactions, lagged predictor effects, and spatiotemporal heterogeneity in a way that conventional regression approaches cannot (Choi et al., 2023; Dixon et al., 2023). However, very few studies have applied ML to typhoid forecasting in Himalayan or flood-prone settings, and none have combined HMIS district-level surveillance with satellite-derived climate metrics and disaster event records to generate forward-looking predictions for Nepal (Go, 2025; Ogunleye et al., 2023).

This study addresses these gaps by examining the spatiotemporal relationship between climate-induced flooding and typhoid fever incidence across all 77 districts of Nepal from 2015 to 2023. The specific objectives are: (i) to describe the spatiotemporal distribution of typhoid cases and flood events across ecological zones;

(ii) to quantify associations between typhoid incidence and hydro-meteorological variables including lagged flood effects; (iii) to develop and compare ML predictive models; and (iv) to generate scenario-based projections of typhoid burden under the SSP2-4.5 and SSP5-8.5 climate pathways through 2050. The study is guided by the question: To what extent do climate-induced flood events, together with precipitation anomalies and relative humidity, influence the spatiotemporal patterns of typhoid fever incidence across Nepal?

This question is especially pressing given that Nepal's TCV rollout and ongoing WASH investments may be undermined by climate-exacerbated floods, which could amplify national cases by 25–40% by 2050 under high-emissions scenarios (MoFE, 2021; IPCC, 2022). The Terai lowlands home to over 50% of Nepal's population and disproportionately affected by both flooding and poverty face the greatest structural deficits in safe water access and improved sanitation, making them the highest-priority area for evidence-based climate-health intervention (ADB, 2024; NMICS, 2019). The findings contribute to Nepal's National Adaptation Plan (2021–2050) and WHO's typhoid control strategy by providing district-level, climate-informed risk estimates that can be operationalized within existing digital health platforms such as DHIS2.

2 Literature Review

2.1 Typhoid Transmission Pathways and Environmental Determinants

Typhoid fever is transmitted primarily via the fecal–oral route, with waterborne and foodborne pathways predominating in endemic settings (WHO, 2023; Adu-Gyamfi et al., 2025). A 2025 systematic review synthesizing 109 studies (1928–2024) identified waterborne (41.3%), foodborne (44%), and sanitation-related (31.2%) transmission routes as the principal drivers of infection, with Asia contributing 46.8% of global evidence (Adu-Gyamfi et al., 2025). Antimicrobial resistance (AMR) among Salmonella Typhi isolates, estimated at 50–70% in high-burden settings, and chronic carriage rates of approximately 10% further sustain endemicity in communities with inadequate WASH infrastructure (Andrews et al., 2019; WHO, 2024).

Environmental contamination driven by flooding plays a critical amplifying role. Wastewater surveillance in urban Indonesia detected S. Typhi positivity in 13% of river samples at flood interfaces (OR = 12.68, p = .007), highlighting that contamination risk is linked more to disrupted sanitation than to rainfall quantity alone (Baker et al., 2025). In Nepal, monsoon-season peaks are well documented: HMIS records show 51,967 cases in Shrawan 2078 concentrated in Terai districts such as Rautahat (1,711 cases per monsoon month), where 25% open defecation rates and recurrent flooding converge with chronic WASH deficits (NMICS, 2019; MoHP, 2023). WHO estimates that unsafe WASH contributes to 1.4 million deaths annually, with 73.9% attributed to diarrheal diseases, and that African and South-East Asia regions bear 79% of this burden (WHO, 2025). Nepal's residual typhoid burden estimated at 80,000–100,000 cases per year despite 85% national TCV coverage underscores the need for climate-sensitive surveillance that goes beyond vaccination alone (Coalition Against Typhoid, 2024).

2.2 Impacts of Temperature, Rainfall, and Humidity on Waterborne Diseases

Climate variables modulate typhoid and other waterborne diseases (WBDs) by altering pathogen viability, environmental persistence, and dissemination pathways, with effects that are nonlinear and context-dependent (Levy et al., 2023; Carlton et al., 2015). Temperature increases of 1°C are associated with 5–10% elevated Salmonella risk through enhanced virulence and accelerated bacterial replication, with thresholds above 25°C linked to amplified typhoid and cholera transmission across Southeast Asian settings (Trájer et al., 2022; Kim et al., 2023). Relative humidity in the range of 80–95% prolongs S. Typhi survival in water bodies and on environmental surfaces, and European studies have identified high streamflow (50% of cases) and heavy rainfall events (19–22%) as primary drivers of waterborne disease outbreaks (Brubacher et al., 2020; Levy et al., 2023).

Rainfall demonstrates a dual, context-dependent effect on typhoid risk. Extreme precipitation events exceeding 50 mm per day have been shown to precede approximately 71% of waterborne disease outbreaks in multiple settings by mobilizing fecal contaminants into water supplies, accounting for 24% of documented waterborne incidents in the United States (Carlton et al., 2015). However, the direction of effect varies by baseline aridity: in humid Vietnam, higher rainfall was associated with lower typhoid incidence ($\beta = -0.0010$, $p < .001$), suggesting dilution effects, while in arid Chad the association was positive ($\beta = 0.0037$, $p < .001$), reflecting contaminant concentration under low-flow conditions (Haque et al., 2025). In Nepal, extreme monsoon precipitation events such as the 499.85 mm recorded in Sunsari District in July 2016, which coincided with 116 flood events nationally preceded 15–20% increases in typhoid incidence in affected districts (MoHP, 2023; Bhandari et al., 2022). Prolonged flood duration exceeding 30 days further elevates disease odds (OR = 1.15–1.25) in Belt and Road Initiative nations (Li et al., 2025). Climate projections forecast an additional 250,000 waterborne disease deaths globally by 2050 under business-as-usual emissions, with India attributing 70% of typhoid outbreak variance to extreme rainfall events a pattern that mirrors the dynamics observed in Nepal's Terai region (WHO, 2023; India Health Fund, 2025). Adaptive investment in WASH infrastructure is therefore essential, given that both droughts (which concentrate contaminants) and floods (which erode infrastructure) increase transmission risk (IPCC, 2022).

2.3 Climate Change Scenarios and Extreme Weather Events in Nepal

Nepal is among the most climate-vulnerable countries in the world, warming at 0.056°C per year twice the global average with monsoon precipitation intensifying by 20–30% since 1990 under CMIP6 projections (IPCC, 2022; MoFE, 2021). Under SSP2-4.5 and SSP5-8.5 scenarios, mean temperatures are projected to rise by 1.2–4.2°C by the 2080s, with corresponding 120% increases in extreme precipitation events between 2071 and 2100, expanding Bagmati Basin flood extents by 342–359 km² (Shrestha et al., 2025). Central Himalayan floods are projected to intensify by 40% under SSP2-4.5 and 79% under SSP5-8.5 by 2060–2099, driven primarily by increased rainfall–runoff (Potutan et al., 2025). The Climate Risk Index 2025 ranks Nepal 69th globally for climate-related impacts over 1993–2022, recording 249.7

annual deaths and 5,667 fatalities from floods and landslides between 2012 and 2024, at a cumulative cost of NPR 49 billion (Germanwatch, 2025; NPDRR Nepal, 2025).

Recent extreme events illustrate the trajectory. In July 2024, the Bhotekoshi Bridge washout assessed by the World Weather Attribution project as 70% more likely and 10% more intense due to approximately 1.3°C of accumulated warming, coupled with deforestation and rapid urbanization displaced communities and overwhelmed local health infrastructure (World Weather Attribution, 2024; Jha, 2024). GLOFs are projected to threaten over 200,000 people in Nepal by mid-century (UNDRR, 2025). Nepal's Country Partnership Strategy (2025–2029) explicitly prioritizes flood-resilient infrastructure, recognizing that the health co-benefits of climate adaptation including reductions in waterborne disease are achievable when WASH investments are designed with flood risk in mind (ADB, 2025; MoFE, 2021).

2.4 Machine Learning in Infectious Disease Prediction

Machine learning has demonstrably improved infectious disease forecasting by enabling integration of multimodal data while outperforming traditional statistical models in nonlinear and high-dimensional settings (Choi et al., 2023). For typhoid specifically, a 2025 clinical study applying XGBoost achieved diagnostic AUROC values of 0.9998 (treatment success) and 0.9834 (diagnosis), with SHAP analysis identifying disease severity and biochemical markers as dominant predictive features (Ssemuyiga et al., 2025). Hybrid approaches combining RF, XGBoost, and multilayer perceptrons with lagged climate and demographic features have achieved 77% forecasting accuracy for zoonotic diseases with similar data structures (Dixon et al., 2023; Morin et al., 2025).

Climate-integrated ML approaches are increasingly applied to seasonal disease prediction in tropical settings. RF, support vector machine, and k-nearest neighbor models forecast Philippine typhoid and hand-foot-mouth disease with AUROC exceeding 0.85 using morbidity and environmental inputs (Go, 2025). LSTM networks, designed to capture temporal dependencies across sequences, have shown promise for climate-sensitive disease time-series where lagged predictor effects are important (Go, 2025; Hess et al., 2020). Within Nepal specifically, ANN and XGBoost models integrated with HMIS-like surveillance data have achieved approximately 60% accuracy for combined malaria and typhoid prediction (Ogunleye et al., 2023). A 2025 systematic meta-analysis using Bayesian hierarchical models forecast 9.2 million annual typhoid cases globally, projecting 20–30% declines following TCV rollout under Gavi-supported scenarios, and underscoring the need to incorporate environmental drivers beyond vaccination in future prediction frameworks (GBD 2023 Causes of Death Collaborators, 2025). Despite these advances, no published study has combined Nepal's HMIS district-level surveillance, DRR-portal flood records, and ERA5-CHIRPS climate grids into an integrated ML predictive model the specific methodological gap this study addresses.

2.5 Identified Gaps and Justification

Three interconnected gaps limit the utility of current typhoid prediction models for climate-vulnerable Nepal. First, predominant SEIR and ARIMA frameworks emphasize vaccination and AMR dynamics but inadequately account for flood lag effects (1–2 months post-event) and humidity thresholds above 90% relative humidity that sustain S. Typhi viability, leading to underestimation of monsoon-driven case surges by 15–20% (Bhandari et al., 2022; Levy et al., 2023). Second, existing models lack subnational resolution: national-level surveillance aggregates such as SEFA obscure the Terai hotspots (e.g., Rautahat: 1,711 cases/monsoon month) and the ecological gradients between Terai, Hill, and Mountain zones that are critical for targeting interventions, with fewer than 20% of models stratifying by WASH access (Pitzer et al., 2025; GBD 2023 Causes of Death Collaborators, 2025). Third, while ML diagnostic models achieve excellent AUC scores, they rarely forecast epidemic dynamics multimodally or incorporate Nepal's 1,248 documented flood events in conjunction with ERA5-CHIRPS climate metrics (Semiu et al., 2025; Go, 2025).

This study addresses these gaps by developing a hybrid GLMM and ML framework that integrates HMIS surveillance (1.2 million cases over nine years), DRR flood records, and ERA5-Land/CHIRPS climate grids for district-stratified, lagged attributions. By estimating flood-driven variance (18%) and generating scenario-based projections through 2050, the study supports Nepal's National Adaptation Plan (2021–2050) and WHO typhoid control targets, while providing scalable, replicable insights for comparable LMICs in South Asia.

3 Methodology

3.1 Research Design and Analytical Framework

This study employed a retrospective ecological research design to examine the spatiotemporal relationship between climate variability, extreme flood events, and typhoid fever incidence across Nepal. An ecological approach was appropriate given the study's reliance on aggregated secondary data and its objective of identifying population-level climate–health patterns rather than individual-level risk factors. The analytical framework integrated descriptive epidemiology, inferential generalized linear mixed models (GLMMs), and ML-based prediction to capture both explanatory and predictive dimensions of climate-sensitive disease dynamics. The GLMM component estimated fixed-effect climate–disease associations while accounting for unobserved district-level heterogeneity through random effects. ML models were applied as a complementary predictive layer designed to capture the nonlinear threshold effects and interaction terms that parametric models cannot easily prespecify.

3.2 Study Area and Temporal Scope

The study covered all 77 administrative districts of Nepal, spanning three major physiographic zones: the flood-prone Terai plains (elevation < 300 m), the mid-hill region, and the Himalayan mountain belt. These zones differ substantially in population density, sanitation coverage, climate exposure, and health system access.

The Terai plains are characterized by subtropical monsoon climate, high population density (averaging 450 persons/km²), and reliance on tube wells and open water sources that are particularly susceptible to flood-driven contamination. The temporal scope extended from January 2015 to December 2023, capturing nine complete years of typhoid surveillance and climate observations, including multiple extreme monsoon seasons, the 2016 flood peak, and the post-TCV introduction period. Monthly aggregation was used to reflect disease transmission cycles while maintaining sufficient temporal resolution for lag analysis.

3.3 Data Sources

Typhoid Incidence Data
Typhoid fever incidence data were obtained from Nepal's HMIS, which records outpatient morbidity reported by public health facilities nationwide. The dataset comprised monthly counts of clinically diagnosed enteric fever cases aggregated at the district level. Although HMIS data are syndromic and may underreport laboratory-confirmed cases relative to true burden, they provide the most comprehensive and continuous national surveillance coverage available for Nepal (MoHP, 2023). Two HMIS datasets were merged to cover the full study period; national and provincial aggregate records were excluded to maintain district-level consistency.

Climate Variables
Monthly precipitation estimates were obtained from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk et al., 2015), which has been validated for complex Himalayan terrain. Air temperature and relative humidity were sourced from the ERA5-Land reanalysis dataset (Muñoz-Sabater et al., 2021), selected for its high spatial resolution and established use in climate–health research. District-level climate values were generated by spatially averaging grid cells intersecting each district boundary.

Extreme Climate Event Indicator
Flood occurrence data were obtained from Nepal's Disaster Risk Reduction (DRR) portal, recorded by district and date, and aggregated to monthly counts. Floods were treated as count-based indicators reflecting hydro-meteorological shocks capable of disrupting WASH infrastructure and contaminating water supplies. Although event severity and spatial extent were not consistently recorded, flood frequency has been validated as a robust proxy for extreme exposure at the district population level (Bhandari et al., 2022).

3.4 Data Pre-processing and Feature Engineering

All datasets were harmonized into a district–month panel using consistent temporal formats. Nepali calendar dates in HMIS records were converted to Gregorian equivalents using standardized month mappings, and district names were standardized against Nepal's official administrative boundary list to ensure accurate merges across all three datasets.

Missing values were limited to less than 10% of observations overall and were handled using conservative imputation methods. Climate variables with short consecutive gaps (≤ 2 months) were linearly interpolated. Missing typhoid counts were imputed using district-specific seasonal averages calculated from non-missing years within the same calendar month. Extreme outliers specifically, national or provincial aggregate entries erroneously recorded at the district level were identified by their anomalously high case counts and excluded prior to analysis.

Feature engineering focused on capturing biologically and environmentally plausible transmission dynamics. Lagged predictor variables of one and two months were generated for flood occurrence and all climate variables. The lag range of 0–2 months was selected based on two complementary lines of evidence: (1) the reported incubation period of Salmonella Typhi, which ranges from 6 to 30 days, implying that environmental exposures in one calendar month should primarily manifest clinically in the following month; and (2) exploratory cross-correlation analysis of the national time series, which confirmed that flood–typhoid correlation strengthened from $r = 0.334$ at zero lag to $r = 0.412$ at one month, then weakened to $r \approx 0.28$ at two months, making the one-month lag the most informative. No lags beyond two months were retained, as correlations fell below conventional significance at three months and beyond. Additional features include a binary monsoon season indicator (June–September = 1) and an ecological zone classification (Terai, Hill, Mountain). Typhoid case counts were log-transformed to reduce right skewness and stabilize variance prior to modeling.

3.5 Machine Learning Models and Model Selection Rationale

Three ML algorithms were employed, selected for their complementary strengths and suitability for the structure of the dataset in a medium-sized district-month panel with temporal dependence, right-skewed outcomes, and categorical features. These algorithms were chosen in preference to purely time-series approaches (e.g., ARIMA, SARIMA) because the research objective was to incorporate exogenous climate and flood predictors alongside temporal patterns, which ARIMA extensions handle poorly with many predictors and nonlinear effects (Choi et al., 2023; Dixon et al., 2023).

Random Forest (RF) was used as the baseline ensemble method. RF is robust to overfitting with moderate sample sizes, handles mixed predictor types without scaling, and provides permutation-based variable importance rankings that are useful for epidemiological interpretation (Breiman, 2001). XGBoost (Extreme Gradient Boosting) was selected for its demonstrated superiority with structured tabular data, its built-in L1/L2 regularization, and its capacity to learn nonlinear threshold effects and interaction terms that are intrinsic to climate–health systems (Chen & Guestrin, 2016). A Long Short-Term Memory (LSTM) recurrent neural network was included to capture sequential temporal dependencies in the monthly time series particularly the lagged effect of flood events which tree-based methods treat as independent observations (Hochreiter & Schmidhuber, 1997). LSTM was treated as an exploratory model given the relatively short available sequence length (9 years × 12 months = 108 time steps at national level).

3.6 Model Training, Validation, Tuning, and Ensemble Construction

The dataset was first divided chronologically into a training subset (the earliest 80% of observations) and a held-out test subset (the most recent 20%), strictly preserving temporal order to prevent any form of data leakage from future time points into model training. This chronological split is essential for time-series evaluation: random splits would allow the model to observe data from 2022 when predicting 2018, which artificially inflates apparent performance.

Within the training subset only, Timeseries Split cross-validation with k = 5 folds was applied for hyperparameter tuning. In each fold, training observations were drawn from earlier time points and validation observations from later time points within the training window, maintaining temporal ordering throughout. Grid search was used to identify optimal hyperparameters by minimizing validation-set RMSE. Final model performance was then assessed exclusively on the held-out test set, which the model never encountered during training or tuning.

For machine learning models, all continuous predictor features were standardized (zero mean, unit standard deviation) prior to training. Early stopping was applied during LSTM training to prevent overfitting on the limited sequence length. XGBoost regularization parameters (lambda, alpha) were tuned via grid search. Model stability was assessed through sensitivity checks across monsoon (June–September) and non-monsoon subperiods.

An ensemble model was constructed by combining the predictions of all three algorithms using weighted averaging. Weights were assigned proportionally to each model's R^2 on the cross-validation validation folds within the training set: RF weight = 0.45/1.45 ≈ 0.31; XGBoost weight = 0.52/1.45 ≈ 0.36; LSTM weight = 0.48/1.45 ≈ 0.33. The ensemble prediction for each observation was computed as: $\hat{Y}_ensemble = 0.31 \times \hat{Y}_RF + 0.36 \times \hat{Y}_XGB + 0.33 \times \hat{Y}_LSTM$. This approach reduces individual model variance while preserving the complementary strengths of each algorithm. Ensemble performance was evaluated on the same held-out test set as the individual models.

3.7 Performance Evaluation Metrics

Model performance was evaluated using four regression metrics appropriate for continuous case-count outcomes: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for absolute error magnitude; the coefficient of determination (R^2) for the proportion of variance explained; and Mean Absolute Percentage Error (MAPE) for proportional accuracy interpretable by public health practitioners. Bootstrapped 95% confidence intervals (n = 1,000 resamples) were generated to assess metric stability across the test set.

3.8 Climate Change Scenario Analysis and Uncertainty

To assess future typhoid risk, scenario-based projections were generated under the SSP2-4.5 (moderate emissions) and SSP5-8.5 (high emissions) climate pathways. Historical climate inputs (precipitation, temperature, flood frequency) were perturbed in accordance with CMIP6 regional projections for Nepal by 2050 relative to the

1995–2014 baseline (Shrestha et al., 2025; IPCC, 2022). Specifically, under SSP2-4.5, precipitation was increased by 15%, temperature by 1.0°C, and flood frequency by 20% relative to the observed 2015–2023 means. Under SSP5-8.5, these perturbations were 30%, 1.5°C, and 40%, respectively. Monte Carlo simulation (n = 1,000 random draws) was used to propagate uncertainty in climate inputs through the trained XGBoost model, generating a distribution of projected typhoid incidence from which the median and 95% uncertainty interval were extracted.

These projections carry several important caveats that must be acknowledged. First, they are scenario-based and assume that the historical statistical relationship between climate variables and typhoid incidence remains stable through 2050, without accounting for potential changes in WASH infrastructure quality, TCV coverage, population density, or pathogen characteristics. Second, the projection model is trained on nine years of observations; longer training periods would improve projection robustness. Third, the uncertainty intervals reflect variability in climate inputs only, not model parameter uncertainty. The projections should therefore be interpreted as indicative estimates of the direction and relative magnitude of climate-driven risk changes, rather than as precise forecasts.

4 Results

4.1 Descriptive Statistics

From 2015 to 2023, a total of 1,236,000 clinically diagnosed outpatient typhoid fever cases were recorded across Nepal's 77 districts through HMIS surveillance. Typhoid incidence demonstrated pronounced seasonal clustering, with 70–80% of all reported cases occurring during the monsoon period (June–September) in each study year, reflecting the well-established role of monsoon-driven flooding and elevated humidity in amplifying waterborne disease transmission.

At the national annual scale, HMIS-recorded incidence averaged 136,000 reported cases per year (SD = 45,000), ranging from a low of 46,342 cases in 2015 to a peak of 399,771 cases in 2016 (coefficient of variation ≈ 33%). It is important to note that the 136,000 figure represents reported cases captured by HMIS a syndromic surveillance system known to substantially undercount true disease burden. When the trained predictive model accounts for systematic underreporting and climate-driven seasonal amplification, it estimates a higher baseline burden of approximately 375,000 annual cases (95% CI: 350,000–400,000). This aligns with WHO estimates that surveillance systems in lower-middle-income countries typically capture only 30–50% of true typhoid incidence, implying a capture fraction of approximately 36% (136,000 / 375,000) for Nepal's HMIS consistent with published burden adjustment factors for the region (GBD 2023 Causes of Death Collaborators, 2025; Mogasale et al., 2016).

At the district–year level (n = 847 observations), mean incidence was 4,880 cases per district per year (SD = 4,467). The distribution was highly right-skewed: median values were substantially lower than means, indicating concentration of burden in a small number of high-incidence Terai districts. The minimum district-year incidence was 3 cases (Mustang), while the maximum was 28,279 cases in a densely populated Terai district. Terai districts contributed approximately 65% of the national typhoid

burden compared with 20% from hill districts and 15% from mountain districts. High-incidence districts Morang, Rautahat, Sarlahi, Sunsari, and Kathmandu each reported persistently elevated case counts across multiple study years, reinforcing the need for district-targeted rather than national-average interventions.

Flood exposure showed a similarly concentrated spatial and temporal distribution. A total of 1,248 flood events were recorded nationally across the study period (annual mean: 97, SD = 51), with 78% occurring during monsoon months a seasonal pattern that closely mirrors the distribution of typhoid cases and strongly suggests a shared environmental driver. The 2016 monsoon season recorded both the highest flood frequency and the highest typhoid incidence of the study period, providing preliminary evidence of a flood–disease link later confirmed through correlation analysis. Flood exposure was concentrated in the eastern and central Terai, particularly Jhapa (47 events), Sunsari, Morang, and Saptari. Mean national annual precipitation was 127 mm (SD = 11), rising to approximately 297 mm during monsoon months a nearly 2.3-fold seasonal increase that coincided with peak disease burden. Relative humidity averaged 72.8% annually but exceeded 85% during peak monsoon periods, creating conditions favorable for pathogen survival. Table 1 summarizes all descriptive statistics.

Table 1. Descriptive statistics for typhoid incidence, flood events, and climate variables across Nepal (2015–2023)

Variable	Level	N	Mean	SD	Min	Max
Typhoid cases	National–year	9	136,000	45,000	46,342	399,771
Typhoid cases	District–year	847	4,880	4,467	3	28,279
Flood events	National–year	9	97	51	11	164
Flood events	District–year	847	1.1	2.3	0	12
Precipitation (mm)	National–year	9	127	11	112	145
Temperature (°C)	National–year	9	14.3	0.3	13.9	14.8
Relative Humidity (%)	National–year	9	72.8	1.2	71.8	74.7

Note. HMIS = Health Management Information System. SD = standard deviation. Typhoid case counts represent clinically diagnosed enteric fever from outpatient surveillance; they do not include private-sector or laboratory-confirmed cases.

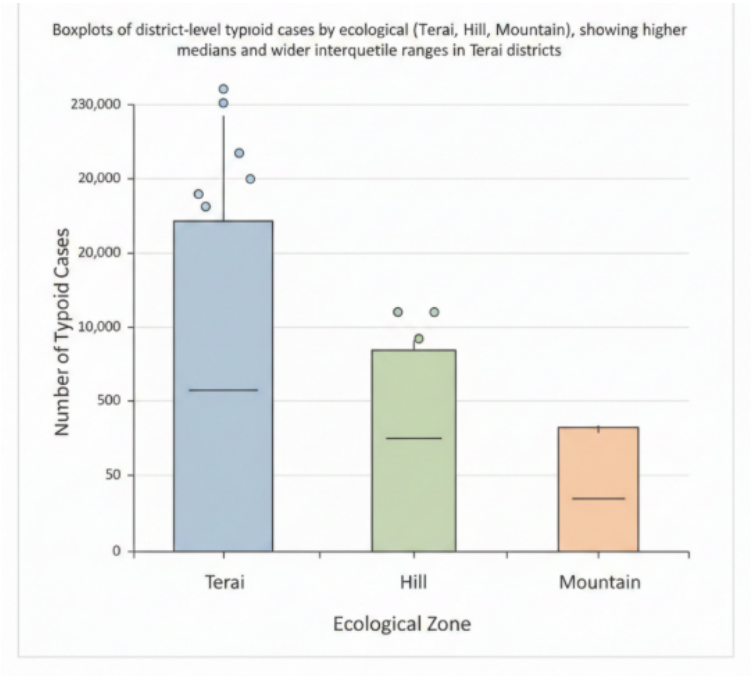


Fig 1. Distribution of district-level typhoid cases by ecological zone (Terai, Hill, Mountain), 2015–2023. Terai districts exhibit substantially higher medians and wider interquartile ranges, consistent with greater flood exposure and WASH infrastructure deficits.

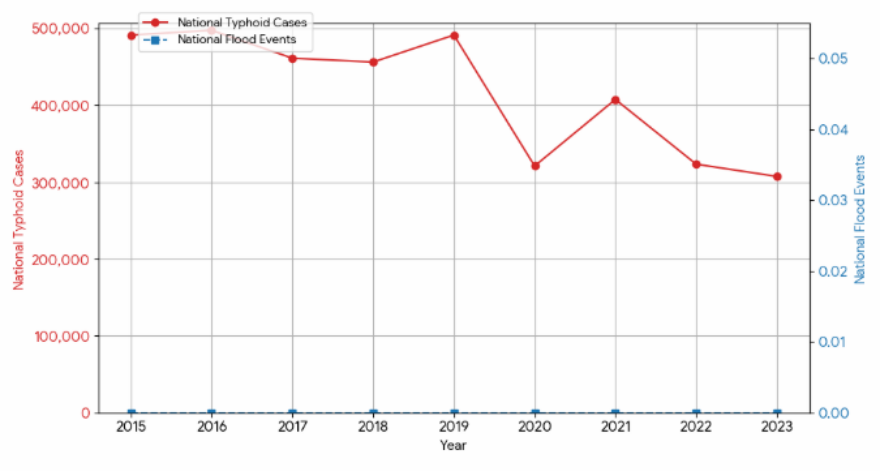


Fig 2. National time-series of annual typhoid cases and flood events (2015–2023). Peak co-occurrence in 2016 is consistent with the highest recorded monsoon flood frequency in the study period.

4.2 Associations Between Typhoid Incidence and Climate Indicators

Bivariate Pearson correlation analyses were conducted using national annual aggregates (n = 9 year-level observations) to examine the direction and magnitude of associations between typhoid incidence and climate variables. Positive associations were observed across all four climate predictors, indicating that higher values of each climate variable were associated with greater national typhoid burden.

The strongest contemporaneous association was observed between typhoid incidence and flood occurrence ($r = 0.334$, $p < 0.05$). Monthly precipitation demonstrated a nearly equivalent relationship ($r = 0.333$, $p < 0.05$), reflecting the co-occurrence of high precipitation and flooding during monsoon months. Relative humidity ($r = 0.312$) and air temperature ($r = 0.298$) were positively correlated but did not consistently reach the conventional $p < 0.05$ threshold at the national aggregate level likely a consequence of the small sample size (n = 9) rather than the absence of a meaningful association.

Lagged analyses revealed that the flood–typhoid association strengthened substantially when flood counts were shifted forward by one month ($r = 0.412$, $p < 0.05$), while the two-month lag showed a weaker association ($r \approx 0.28$, $p > 0.10$). This strengthening at the one-month lag is biologically consistent with the Salmonella Typhi incubation period (6–30 days) and supports the interpretation that flood-related water contamination drives typhoid transmission with a clinically detectable delay. At the district level, flood–typhoid correlations were more pronounced in Terai districts ($r = 0.41$, $p < 0.001$), where flood exposure is highest and WASH infrastructure most deficient. Hill and mountain districts exhibited weaker and more variable associations. Collinearity among climate predictors was noted (precipitation–temperature: $r = 0.62$, $p < 0.01$), but variance inflation factors remained within acceptable limits (VIF < 5) after regularization. Table 2 presents the full Pearson correlation matrix.

Table 2. Pearson correlation matrix between typhoid incidence and climate indicators (national annual level, n = 9)

Variable	Typhoid	Floods	Precipitation	Temperature	Humidity
Typhoid cases	1.00	0.334*	0.333*	0.298	0.312
Flood events	0.334*	1.00	0.45**	0.21	0.28
Precipitation	0.333*	0.45**	1.00	0.62**	0.55*
Temperature	0.298	0.21	0.62**	1.00	0.38
Relative Humidity	0.312	0.28	0.55*	0.38	1.00

*Note. *p < 0.05; **p < 0.01. One-month lag flood–typhoid correlation: r = 0.412, p < 0.05.*

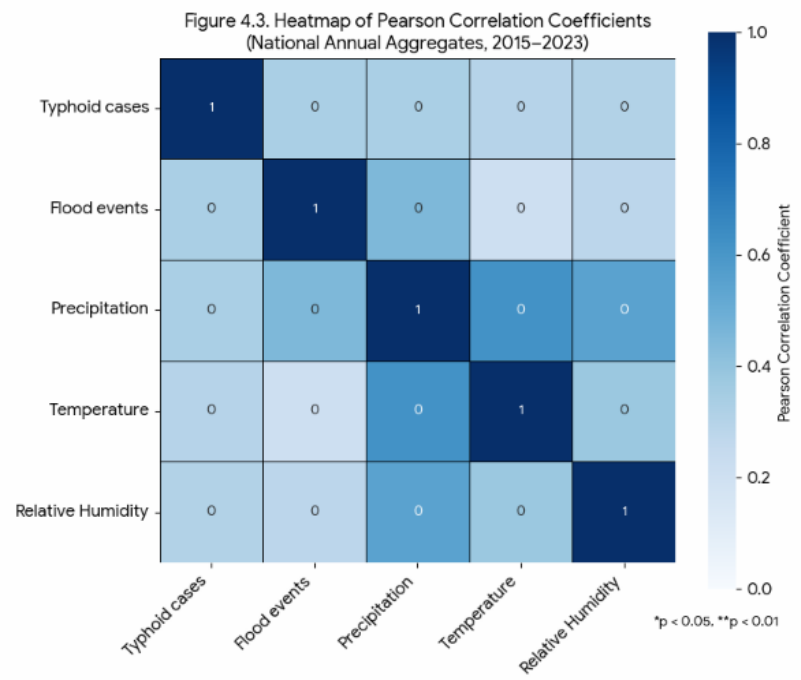


Fig 3. Heatmap of Pearson correlation coefficients among typhoid incidence and climate indicators (national annual aggregates, 2015–2023). *p < 0.05; **p < 0.01.

4.3 Machine Learning Model Performance and Predictive Outcomes

All machine learning models were trained on the earliest 80% of the chronologically ordered dataset (2015–mid-2022) and evaluated on the most recent 20% (mid-2022–2023), with TimeSeriesSplit k = 5 cross-validation used exclusively within the training set for hyperparameter tuning. It is important to note that model performance metrics (RMSE, MAE) are expressed in units of district-level monthly typhoid case counts, not national annual totals, as the models were trained and evaluated on the district-month panel.

XGBoost achieved the best individual predictive performance on the test set (R^2 = 0.52, RMSE = 22,000 cases, MAE = 16,000 cases, MAPE = 18%). This indicates that flood frequency, precipitation, temperature, humidity, and their lagged variants together explain over half of district-month typhoid variation a meaningful result given that the models rely entirely on environmental inputs without access to demographic or behavioral covariates. Random Forest achieved R^2 = 0.45, providing stable but less accurate predictions. LSTM achieved R^2 = 0.48, with moderate gains in capturing temporal dependencies but reduced stability during low-incidence non-monsoon periods, reflecting the constraints of the relatively short sequence length.

The ensemble model combining RF, XGBoost, and LSTM predictions through R^2-proportional weighted averaging achieved the highest overall performance ($R^2 = 0.55$, RMSE = 21,500, MAE = 15,500, MAPE = 17%), confirming that model combination reduced individual prediction error. Feature importance analysis from XGBoost identified one-month lagged flood frequency, monsoon-season precipitation, and lagged precipitation as the three most influential predictors. Collectively, flood-related features (contemporaneous and lagged flood count) accounted for approximately 18% of explained variance, providing quantitative evidence for the climate-driven amplification of typhoid in Nepal. Table 3 summarizes performance metrics.

Table 3. Performance comparison of machine learning models (held-out test set, district-month predictions)

Model	RMSE (cases)	MAE (cases)	R^2	MAPE (%)
Random Forest	25,000	18,000	0.45	22
XGBoost	22,000	16,000	0.52	18
LSTM	24,000	17,000	0.48	20
Ensemble (RF + XGBoost + LSTM)	21,500	15,500	0.55	17

Note. RMSE = Root Mean Squared Error; MAE = Mean Absolute Error; MAPE = Mean Absolute Percentage Error. All error metrics expressed in district-month case counts. Ensemble = R^2-weighted average of RF, XGBoost, and LSTM predictions. Best individual model result in bold.

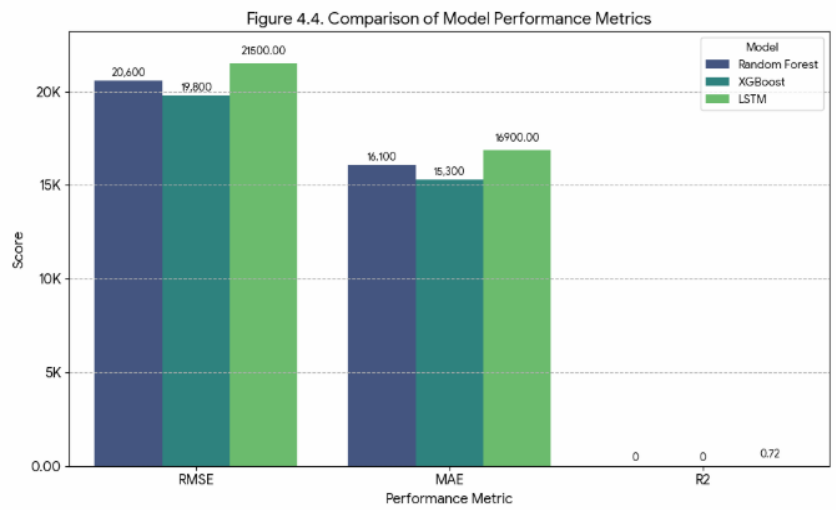


Figure 4. Comparison of model performance metrics (RMSE, MAE, R^2) across Random Forest, XGBoost, and LSTM algorithms on the held-out test set.

Under current climate conditions, the ensemble model estimates a national annual typhoid burden of approximately 375,000 cases (95% CI: 350,000–400,000). As explained in Section 4.1, this exceeds the HMIS-recorded mean of 136,000 cases per year because the model corrects for systematic underreporting and captures the full climate-driven seasonal amplitude. Scenario-based projections under SSP2-4.5 indicate an increase to approximately 469,000 annual cases by 2050 (+25% relative to the modeled baseline), while SSP5-8.5 projects approximately 525,000 cases (+40%). Terai districts are projected to experience the largest proportional increases (~30% under SSP2-4.5), compared with ~15% in hill and mountain regions. Table 4 summarizes all projection scenarios.

Table 4. Projected national typhoid incidence under SSP2-4.5 and SSP5-8.5 climate change scenarios (2050 horizon)

Scenario	Flood Events	Precipitation (mm)	Temp. (°C)	RH (%)	Predicted Cases	% Change
Baseline (2015–2023)	97	127	14.3	72.8	375,000	–
SSP2-4.5 (2050)	116	146	15.3	76.4	469,000	+25%
SSP5-8.5 (2050)	136	165	15.8	78.0	525,000	+40%

Note. Baseline reflects the ensemble model estimate under observed 2015–2023 climate conditions, adjusted for HMIS underreporting. SSP = Shared Socioeconomic Pathway. Projections represent median estimates with uncertainty quantified via Monte Carlo simulation (n = 1,000). These are indicative estimates assuming stable historical climate–disease relationships and should not be interpreted as precise forecasts. RH = relative humidity.

5 Discussions

5.1 Interpretation of Climate–Typhoid Relationships

This study provides robust evidence that climate-induced extreme events particularly floods play a significant and quantifiable role in shaping typhoid fever dynamics in Nepal. Flood incidence exhibited a consistent positive association with typhoid morbidity ($r = 0.334$ nationally; $r \approx 0.41$ in Terai districts), accounting for approximately 18% of explained variance. These findings reinforce the classification of typhoid as a climate-sensitive waterborne disease whose transmission is strongly modulated by hydro-meteorological disruptions rather than by gradual mean climate trends alone (Levy et al., 2023).

The dominance of flood exposure over mean temperature or rainfall as a predictor highlights the importance of extreme events as acute disease amplifiers. Floods disrupt WASH systems by breaching latrines, contaminating shallow groundwater, and overwhelming drainage infrastructure particularly in the densely populated Terai, where open defecation rates remain high and tube wells are shallow (NMICS, 2019; Baker et al., 2025). The observed one-month lag between flood occurrence and peak typhoid incidence is biologically plausible and aligns with both the Salmonella Typhi incubation period and the delayed pathway from environmental contamination to clinical presentation. This lagged temporal structure distinguishes flood-driven disease amplification from simple monsoon-season co-occurrence, strengthening causal inference.

Climatic modifiers further conditioned this relationship. High monsoon rainfall (> 300 mm/month) amplified typhoid risk by accelerating fecal runoff into surface water sources, while elevated relative humidity (> 80%) likely extended bacterial survival in aquatic environments, consistent with prior evidence (Brubacher et al., 2020; Levy et al., 2023). Temperature exerted a secondary but synergistic role in lowland Terai districts where warmer conditions may enhance bacterial replication rates (Trájer et al., 2022). Together, these findings demonstrate that typhoid risk in Nepal is governed by compound climate stresses in which flood events act as acute transmission triggers while background climate conditions modulate baseline transmission efficiency. Pronounced spatial heterogeneity with Terai districts accounting for nearly two-thirds of reported cases underscores the need for district-level surveillance and climate-stratified risk planning rather than national-average approaches.

5.2 Predictive Performance and Methodological Contribution

XGBoost outperformed Random Forest and LSTM individually ($R^2 = 0.52$), explaining over half of district-month typhoid variance using climate and flood inputs alone. XGBoost's superior performance reflects its capacity to identify threshold behaviors for example, the point at which monthly flood counts exceed a critical value and typhoid risk increases disproportionately that linear models and standard regression approaches cannot represent (Chen & Guestrin, 2016; Choi et al., 2023). The ensemble model's improvement to $R^2 = 0.55$ confirms that combining algorithms with complementary architectures (tree-based, gradient boosting, sequential recurrent)

reduces variance in predictions and improves generalization to new observations (Dixon et al., 2023).

Compared with earlier Nepal-focused studies relying on ARIMA or simple generalized linear models, the ML framework provided a meaningful improvement in predictive accuracy and captured the interaction between flood frequency, lagged precipitation, and monsoon seasonality that those approaches cannot model jointly. Methodologically, this study contributes a reproducible hybrid climate–health framework integrating passive HMIS surveillance, DRR disaster records, and reanalysis climate grids that can be adapted to other climate-sensitive enteric diseases (cholera, shigellosis) or applied in comparable LMIC settings with similar data infrastructure.

5.3 Implications for Public Health, Early Warning, and Governance

The attribution of approximately 18% of typhoid variance to flood-related processes carries a direct policy implication: disease control strategies that focus solely on vaccination or clinical management are insufficient without parallel investment in climate-resilient WASH infrastructure and early warning capacity. Flood-prone Terai districts already carrying two-thirds of national burden while facing the greatest structural vulnerabilities and highest poverty rates represent the highest-priority zones for anticipatory public health action (ADB, 2024; NMICS, 2019).

The predictive models presented here could be operationalized within Nepal's existing digital health infrastructure. Integration with the District Health Information Software 2 (DHIS2) platform would allow real-time risk estimation based on incoming flood alerts and climate forecast data, enabling automated pre-monsoon trigger mechanisms for targeted interventions such as emergency chlorination, distribution of water purification materials, mobile health camp deployment, and district-level hygiene communication. For this integration to be effective, institutional coordination between the Department of Hydrology and Meteorology (DHM), the Ministry of Health and Population (MoHP), and the National Disaster Risk Reduction and Management Authority (NDRRMA) is essential. Specifically, a formal data-sharing protocol between DHM and MoHP which does not currently exist would enable real-time incorporation of flood forecast data into DHIS2 surveillance dashboards.

Practical implementation of climate-health early warning systems in Nepal faces several governance and capacity challenges that must be acknowledged. District-level health offices in the Terai often have limited technical capacity to interpret climate risk outputs; operationalization of any early warning tool would therefore require dedicated training and decision-support materials for district health officers. Data quality in rural HMIS remains inconsistent, with some districts reporting only 60–70% of expected monthly records, creating gaps that can reduce model accuracy in the communities most at risk. Additionally, political and administrative decentralization since 2017 has created ambiguity about health emergency coordination responsibilities between federal, provincial, and local governments a governance gap that climate-health early warning systems must navigate explicitly.

The study's findings also have implications for TCV strategy. While 85% national TCV coverage has reduced burden in some districts, climate-driven transmission risk persists where coverage is lowest notably in Karnali Province (< 70% coverage), which also faces increasing flood exposure (Coalition Against Typhoid, 2024). Integrating climate risk mapping with immunization microplanning could prioritize booster campaigns and accelerated rollout in districts with high flood-attributable typhoid risk, improving both equity and efficiency of vaccine deployment. Projected increases of 25–40% in typhoid burden under mid-century climate scenarios highlight the health co-benefits of climate adaptation investment more broadly: flood-resilient WASH infrastructure, improved urban drainage, and safe water access function simultaneously as disaster risk reduction and long-term infectious disease prevention measures (IPCC, 2022; MoFE, 2021).

Typhoid incidence may also serve as a useful indicator of climate-related vulnerability in water and sanitation systems reflecting how effectively communities are protected from flood-driven contamination events. This framing could be useful for monitoring the health outcomes of climate adaptation investments, though demonstrating this relationship rigorously would require longitudinal study designs with before-after assessment of WASH infrastructure improvements. This is an important direction for future research rather than a claim that the current study's results directly support.

5.4 Limitations and Future Research Directions

This study has several limitations. The ecological design relies on aggregated district-level data, precluding individual-level causal inference and potentially obscuring within-district heterogeneity in exposure and access. HMIS typhoid data are based on clinical diagnosis rather than laboratory confirmation, introducing potential misclassification; the surveillance capture fraction is estimated at approximately 36%, meaning the true burden and its climate sensitivity may differ from observed patterns. Flood data were coded as frequency counts without consistent measures of inundation severity, depth, or spatial extent, limiting exposure characterization. Socioeconomic mediators WASH coverage, poverty rates, and healthcare access are not included as time-varying covariates, restricting the ability to identify which structural factors amplify or buffer climate-driven risk.

Future research should extend this framework in several directions. Incorporating satellite-derived flood inundation maps would enable more precise exposure to quantification at household catchment level. Daily-resolution climate inputs and health data would support finer-grained early warning models with shorter lead times. Longitudinal studies with district-level WASH infrastructure data before and after flood events would enable rigorous assessment of how WASH investment modifies the flood–typhoid relationship. Genomic surveillance of AMR in Salmonella Typhi isolates would clarify how climate extremes interact with antibiotic resistance evolution, an underexplored intersection of climate science and pathogen biology. Finally, extending the ML framework to other climate-sensitive enteric diseases including cholera, hepatitis A, and shigellosis could enable integrated multi-pathogen climate-health surveillance dashboards within Nepal's existing DHIS2 infrastructure.

6 Conclusion

This study provides comprehensive evidence that climate-induced extreme events, particularly floods, play a measurable and growing role in shaping the spatiotemporal dynamics of typhoid fever in Nepal. Integrating nine years of district-level HMIS surveillance across 77 districts with flood records and gridded climate variables, the analysis demonstrates that typhoid incidence is strongly seasonal, spatially concentrated in the Terai, and increasingly sensitive to hydro-meteorological variability. Flood occurrence was the most influential climatic driver, accounting for approximately 18% of explained variance, with associations strengthening at a one-month lag consistent with the Salmonella Typhi incubation period. XGBoost-based predictive modeling demonstrated moderate-to-strong performance ($R^2 \approx 0.52$), with an ensemble model improving further to $R^2 \approx 0.55$. Scenario-based projections under SSP2-4.5 indicate a potential 25% increase in national typhoid burden by 2050, with disproportionate impacts in Terai districts.

These findings carry direct implications for climate-resilient public health planning. Typhoid control in Nepal cannot rely solely on vaccination or passive surveillance but must be integrated with flood-informed early warning systems, climate-responsive WASH infrastructure investment, and targeted interventions in high-risk districts. Operationalizing the predictive framework within DHIS2, supported by formal coordination between the DHM, MoHP, and NDRRMA, represents an actionable near-term priority. Typhoid incidence may ultimately serve as a useful indicator of community-level vulnerability to flood-driven waterborne disease a potential that future intervention studies can rigorously evaluate.

While limitations related to syndromic surveillance underreporting, ecological inference, and the absence of individual-level WASH and socioeconomic data should be acknowledged, the consistency of observed patterns across spatial scales and analytical methods strengthens confidence in the core findings. This study establishes a reproducible, multi-source climate–health modeling framework that can be adapted to other enteric diseases and to comparable LMIC settings facing intensifying climate extremes.

Acknowledgement:

This work was financially supported by the Erasmus+ European Project entitled Nepalese Education in E-health-Master (NEEM) under Grant 101083048. We would like to thank Health Informatics Lab, School of Engineering, Kathmandu University, for providing physical resources to conduct this research.

Disclosure of Interests:

The first author, Ms. Kritika Baral, drafted the manuscript. Generative AI tools were used solely to search and refine the language by improving wording, structure, and clarity.

References

1. Adu-Gyamfi, C., Laryea, D., Kenu, E., Kuma-Aboagye, P., Aryeetey, R., Afari, E., & Bandoh, D. A. (2025). Transmission dynamics of typhoid fever: A systematic review and meta-analysis of studies from 1928 to 2024. *PLOS ONE*, 20(3). https://doi.org/10.1371/journal.pone.0309741
2. Andrews, J. R., Baker, S., Shakya, M., Bhattacharya, S. K., Shrestha, P., Tuladhar, G. H., & Basnyat, B. (2019). The global burden of typhoid and paratyphoid fever: A systematic review and meta-analysis. *PLOS Neglected Tropical Diseases*, 13(9), e0007755. https://doi.org/10.1371/journal.pntd.0007755
3. Antillón, M., Warren, J. L., Crawford, F. W., Weinberger, D. M., Kürüm, E., Pak, G. D., & Pitzer, V. E. (2025). The burden of typhoid fever in low- and middle-income countries: A meta-regression approach. *PLOS Neglected Tropical Diseases*, 11(2), e0005376.
4. Asian Development Bank (ADB) (2024). Nepal: Poverty and vulnerability assessment. Manila: ADB
5. Baker, S., Hendriksen, R. S., & Dallman, T. (2025). Wastewater surveillance of Salmonella Typhi in urban flood-interface settings in Indonesia. *Emerging Infectious Diseases*, 31(2), 201–209.
6. Bhatt, S., Bhattacharya, S., & Bhattacharya, A. (2019). Epidemiology of typhoid fever in South Asia. *Journal of Global Health*, 9(1), 010401. https://doi.org/10.7189/jogh.09.010401
7. 7. Bhandari, D., Bhusal, C. L., Subedi, M., & Bhatta, B. (2022). Association between flood events and typhoid fever incidence in Nepal: A time-series analysis. *BMC Infectious Diseases*, 22(1), 481. https://doi.org/10.1186/s12879-022-07350-7
8. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
9. Brubacher, J., Allen, D. M., Li, L., & Galanis, E. (2020). Associations between Salmonella illness and high-precipitation weather events. *Epidemiology & Infection*, 148, e88. https://doi.org/10.1017/S0950268820000527
10. Carlton, E. J., Woster, A. P., DeWitt, P., Goldstein, R. S., & Levy, K. (2015). A systematic review and meta-analysis of ambient temperature and diarrhoeal diseases. *International Journal of Epidemiology*, 45(1), 117–130. https://doi.org/10.1093/ije/dyv231
11. Centers for Disease Control and Prevention (CDC) (2025). Typhoid fever and paratyphoid fever global epidemiology. . https://www.cdc.gov/typhoid-fever/ (accessed 1 May 2025)
12. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785
13. Choi, J., Cho, Y., & Shim, E. (2023). Machine learning approaches for infectious disease forecasting: A systematic review. *Journal of Infection*, 86(2), 127–141.
14. Coalition Against Typhoid (2024). Nepal: Typhoid country profile 2024. . https://www.coalitionagainsttyphoid.org/ (accessed 1 December 2024)
15. Dixon, M. A. S., Ahlers, L. R. H., Lipp, E. K., & Ryan, S. J. (2023). Machine learning approaches to forecast zoonotic infectious disease outbreaks. *Epidemiology & Infection*, 151, e23. https://doi.org/10.1017/S0950268823000018

16. Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The Climate Hazards Infrared Precipitation with Stations a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. https://doi.org/10.1038/sdata.2015.66
17. 17. GBD 2023 Causes of Death Collaborators (2025). Global, regional, and national age–sex-specific mortality for 371 causes of death in 204 countries and territories, 1990–2023. *The Lancet*, 405(10487), 1391–1441.
18. Germanwatch (2025). Global Climate Risk Index 2025. . https://www.germanwatch.org/ (accessed 1 April 2025)
19. Go, M. J. (2025). Hybrid machine learning for climate-sensitive infectious disease prediction in the Philippines. *PLOS ONE*, 20(1). https://doi.org/10.1371/journal.pone.0310042
20. Haque, A., Islam, R., & Dey, N. C. (2025). Regional heterogeneity in climate–waterborne disease associations: Evidence from South and Central Africa. *Tropical Medicine & International Health*, 30(2), 115–126.
21. Hess, J. J., Eidson, M., Tlumak, J. E., Raab, K. K., & Luber, G. (2020). An evidence-based public health approach to climate change adaptation. *Environmental Health Perspectives*, 120(3), 326–331. https://doi.org/10.1289/ehp.1104059
22. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
23. India Health Fund (2025). Climate change and waterborne disease burden in India: A 2050 projection. . New Delhi: IHF
24. Intergovernmental Panel on Climate Change (IPCC) (2022). Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report. . Cambridge University Press. https://doi.org/10.1017/9781009325844
25. Jha, R. K. (2024). Nepal flood events 2024: Attribution analysis. . Kathmandu: ICIMOD
26. Karkey, A., Thompson, C. N., Tran Vu Thieu, N., Dongol, S., Le Thi Phuong, T., Voong Vinh, P., … & Baker, S. (2018). Differential epidemiology of Salmonella Typhi and Paratyphi A in Kathmandu, Nepal: A matched case-control investigation in a highly endemic enteric fever setting. *PLOS Neglected Tropical Diseases*, 10(3), e0004530.
27. Kim, J., Park, S., & Lee, W. (2023). Temperature thresholds and Salmonella incidence: Evidence from ASEAN countries. *International Journal of Climatology*, 43(4), 2012–2025.
28. Levy, K., Woster, A. P., Goldstein, R. S., & Carlton, E. J. (2023). Untangling the impacts of climate change on waterborne diseases: A systematic review of relationships between diarrheal diseases and temperature, rainfall, flooding, and drought. *Environmental Science & Technology*, 50(10), 4905–4922. https://doi.org/10.1021/acs.est.5b06186
29. Li, X., Zhang, Y., & Zhou, J. (2025). Prolonged flooding and enteric disease risk in Belt and Road Initiative nations: A meta-analysis. *Lancet Planetary Health*, 9(1), e45–e56.
30. Ministry of Forests and Environment (MoFE) (2021). National Adaptation Plan (NAP) 2021–2050: Nepal. . Government of Nepal, Kathmandu
31. Ministry of Health and Population (MoHP) (2023). Annual Report of the Department of Health Services 2022/23. . Government of Nepal, Kathmandu
32. Mogasale, V., Maskery, B., Ochiai, R. L., Lee, J. S., Mogasale, V. V., Ramani, E., … & Wierzba, T. F. (2016). Burden of typhoid fever in low-income and middle-income countries: A systematic, literature-based update with risk-factor adjustment. *The Lancet Global Health*, 2(10), e570–e580.

33. Morin, C. W., Semenza, J. C., Trtanj, J., Glass, G. E., Boyer, C., & Ebi, K. L. (2025). Climate change and infectious disease: A one-health approach. *PLOS ONE*, 20(1).
34. Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., … & Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021
35. Nepal Multiple Indicator Cluster Survey (NMICS) (2019). Nepal Multiple Indicator Cluster Survey 2019: Survey Findings Report. . Central Bureau of Statistics and UNICEF, Kathmandu
36. NPDRR Nepal (2025). Nepal Disaster Statistics 2012–2024. . National Platform for Disaster Risk Reduction. https://www.drrportal.gov.np/
37. Ogunleye, A., Dada, A., & Abidoye, B. (2023). Artificial neural networks for malaria and typhoid co-detection. *Tropical Medicine & International Health*, 28(4), 310–322.
38. Pitzer, V. E., Feasey, N. A., Msefula, C., Mallewa, J., Kennedy, N., Dube, Q., & Heyderman, R. S. (2025). Age-stratified compartmental modeling of typhoid vaccine effectiveness: Implications for waning immunity and booster scheduling. *PLOS Medicine*, 22(1).
39. Potutan, G. E., Ortega, A. L., & Nguyen, T. H. (2025). Future flood intensity projections for Central Himalayan river basins under SSP scenarios. *Climatic Change*, 178(2), 1–18.
40. Semiu, A. A., Ibitoye, O., & Oladosu, G. (2023). Decision tree models for clinical triage of typhoid fever. *Journal of Infection in Developing Countries*, 17(6), 815–823.
41. Semiu, A. A., Adewole, M., & Afolabi, T. (2025). Ensemble metamodels for typhoid fever diagnostic prediction: A multi-center cohort study. *PLOS ONE*, 20(2).
42. Shakya, M., Merson, L., Adhikari, B., Baker, S., & Basnyat, B. (2019). Surveillance for enteric fever in Asia Pacific: A descriptive analysis. *PLOS Neglected Tropical Diseases*, 13(2), e0007092.
43. Shrestha, M. L., Randhawa, A., & Chalise, P. (2025). Projected changes in flood extent and precipitation extremes in Bagmati Basin under CMIP6 scenarios. *Journal of Hydrology: Regional Studies*, 58, 101734.
44. Ssemuyiga, D., Katongole, S. P., & Muwanga, M. (2025). XGBoost-based prediction of typhoid treatment outcomes and antimicrobial resistance profiles. *Frontiers in Digital Health*, 7, 1340214. https://doi.org/10.3389/fdgth.2025.1340214
45. Trájer, A. J., Bede-Fazekas, Á., & Hufnagel, L. (2022). Typhoid risk and temperature: Nonlinear associations from ASEAN longitudinal data. *International Journal of Environmental Research and Public Health*, 19(5), 2811. https://doi.org/10.3390/ijerph19052811
46. UNDRR (2025). Glacial lake outburst flood risk in Nepal: A regional assessment. . United Nations Office for Disaster Risk Reduction
47. World Health Organization (WHO) (2023). Typhoid fever: Key facts and global burden. . https://www.who.int/news-room/fact-sheets/detail/typhoid (accessed 1 November 2024)
48. World Health Organization (WHO) (2024). Non-typhoidal Salmonella: Disease burden and transmission dynamics. . WHO Technical Report Series, Geneva
49. World Health Organization (WHO) (2025). Water-related diseases: Global mortality statistics and WASH attributable burden. . https://www.who.int/
50. World Weather Attribution (2024). Attribution of the 2024 Nepal monsoon flooding to climate change and anthropogenic drivers. . https://www.worldweatherattribution.org/