



OPEN ACCESS

EDITED BY

Mustapha Amodu,
University of Cape Coast, Ghana

REVIEWED BY

Jobson Santos,
Universidade de Pernambuco, Brazil
Niswatul Qona'Ah,
Sebelas Maret University, Indonesia

*CORRESPONDENCE

Shu Wei Chou-Chen
✉ shuwei.chou@ucr.ac.cr

RECEIVED 24 November 2025

REVISED 04 March 2026

ACCEPTED 11 March 2026

PUBLISHED 08 April 2026

CITATION

Chou-Chen SW and Parra-Rodríguez E (2026) Zero-inflated and distributed lag nonlinear models with random effects for assessing environmental impacts on respiratory health in peripheral regions of Costa Rica.

Front. Public Health 14:1753511.
doi: 10.3389/fpubh.2026.1753511

COPYRIGHT

© 2026 Chou-Chen and Parra-Rodríguez. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Zero-inflated and distributed lag nonlinear models with random effects for assessing environmental impacts on respiratory health in peripheral regions of Costa Rica

Shu Wei Chou-Chen^{1*} and Emanuelle Parra-Rodríguez^{2,3}

¹School of Statistics and Research Center in Pure and Applied Mathematics, University of Costa Rica, San José, Costa Rica, ²School of Mathematics, Costa Rica Institute of Technology, Cartago, Costa Rica, ³School of Mathematics, University of Costa Rica, San José, Costa Rica

Introduction: Climate change and air pollution are key determinants of public health, particularly in the onset and exacerbation of respiratory diseases. The main objective is to quantify the lagged and nonlinear effects of climate and air pollution on respiratory hospitalizations in peripheral regions of Costa Rica.

Methods: This study presents a methodological framework that combines Distributed Lag Nonlinear Models (DLNM) with Generalized Linear Mixed Models (GLMM), incorporating fixed and random effects, to assess the lagged and nonlinear effects of climatic variables and atmospheric pollutants on hospitalizations due to respiratory causes. The response specification was carried out using zero-inflated distributions, aiming to adequately capture the overdispersion and excess zeros present in the data. The analysis focused on peripheral climatological regions and subregions of Costa Rica-territories outside the Central Valley, including Caribbean and Pacific coasts and border areas, characterized by low population density. Weekly data (2000–2019) on temperature, precipitation, relative humidity, and aerosol optical depth (AOD) were combined with seasonal effects and a population offset to account for subregional differences.

Results: Northern and Central Pacific regions show similar climate–pollution impacts on respiratory health, while the South Pacific exhibits stronger and more persistent risks from moderate to high pollution, and Atlantic regions show consistently higher risks associated with intense rainfall and high humidity. Overall, precipitation extremes, high humidity, and AOD contribute more to respiratory hospitalizations than temperature.

Conclusion: This approach improves explanatory and predictive performance, yields robust relative risk estimates, and captures regional sensitivity to environmental conditions, supporting spatiotemporal health analysis and early warning systems in rural tropical settings.

KEYWORDS

distributed lag nonlinear model (DLNM), hospital discharges, mixed model, random effects, respiratory health, zero inflation

1 Introduction

Respiratory health is a crucial component of public health, encompassing a range of conditions such as asthma, chronic obstructive pulmonary disease (COPD), and acute respiratory infections. These disorders contribute substantially to global morbidity and mortality, placing significant pressure on healthcare systems (1).

Climate variability and air pollution are major environmental determinants of respiratory health (2). Climate change increases the frequency of extreme events such as floods and other natural disasters, while extremes of temperature, humidity, and seasonal fluctuations affect the transmission and severity of respiratory diseases (3). In addition, pollutants such as particulate matter, ozone, and nitrogen dioxide exacerbate symptoms and raise hospitalization rates. Collectively, these factors intensify healthcare burdens and pose escalating challenges for medical professional and policymakers worldwide (4–8).

Although several studies have documented these associations, methodological challenges remain in accurately capturing the complex nonlinear and lagged effects on respiratory health data. Temperature has been reported to have a nonlinear effect on respiratory outcomes, typically characterized by a U-shaped association. In other words, both extremely low and high temperatures are associated with increased risks of hospitalization (9–11). However, other studies have found that only extreme high temperatures significantly affect respiratory and cardiovascular diseases (12).

Other climatic exposures, such as precipitation and relative humidity, have also been linked to respiratory outcomes. For instance, Pedder et al. (10) reported that only low relative humidity (RH) was associated with increased pneumonia hospital admissions, whereas Tran et al. (11) observed that both extremely high and low RH were associated with respiratory hospitalizations. Similarly, higher precipitation levels have been found to correlate with an increase in respiratory and cardiovascular hospital admissions (13). Finally, air pollution—particularly particulate matter (PM_{10})—has been consistently reported to have a positive association with adverse respiratory health outcomes (13, 14). This variability in findings is further amplified by climatic fluctuations and geographically diverse contexts characterized by population and socioeconomic heterogeneity.

This study faces these challenges and seeks to address them through assessing the nonlinear and lagged associations between climatic variables (temperature, precipitation, and relative humidity), air pollution (measured by aerosol optical depth, AOD), and hospital discharges due to respiratory causes in Costa Rica, a Central American country with tropical and sub-tropical climates and a wide range of microclimates. Hospital discharge data are used because they provide accurate diagnoses of respiratory conditions within the country's health administrative system. A previous study applied several predictive machine learning methods to forecast respiratory hospitalizations based on climatic and pollution data (15).

Moreover, the analysis focuses on peripheral climatological regions of Costa Rica, including the Caribbean and Pacific coasts and border areas, which are characterized by lower population density and pronounced climatic variability (16). The Central

Valley is excluded due to its large population, high incidence of respiratory hospitalizations, and the need for a distinct modeling approach.

To characterize the dynamics between respiratory hospitalizations and their potential nonlinear and lagged climatic effects, we proposed the distributed lag nonlinear models (DLNMs) (17–19) with random effects to account subregional variability, focusing on weekly respiratory count data. Recent methodological developments have extended DLNMs to hierarchical and mixed-effects frameworks for multi-location environmental time-series analyses. These approaches allow simultaneous estimation of region-specific exposure-lag-response associations while accounting for between-area heterogeneity (20). Large-scale applications have combined DLNM structures with hierarchical meta-regression models to quantify spatial variability in temperature-related health risks (21). Contemporary respiratory health studies continue to apply DLNM-based specifications under diverse climatic regimes (22), supporting their integration with random-effects structures in geographically stratified epidemiological settings.

In peripheral tropical regions characterized by lower population density and pronounced climatic variability, weekly aggregated health counts frequently exhibit overdispersion and excess zeros due to sparse population structures or episodic disease occurrence (11, 23). To accommodate these low-count features, we incorporated zero-inflated negative binomial distributions, which allow the separation of structural absence of events from stochastic variability while addressing overdispersion (24, 25). Model estimation was implemented within flexible mixed-effects frameworks that support hierarchical structures and computational stability in complex regression settings (26–28). By jointly modeling nonlinear lagged exposure effects and inter-subregional heterogeneity within this framework, the proposed approach provides robust quantitative evidence of the influence of climatic and air-quality variables on respiratory hospital discharges.

From the selected models, Northern and Central Pacific regions exhibit broadly similar climate–pollution patterns in their impacts on respiratory health, reflecting shared climatic dynamics. In contrast, the South Pacific region demonstrates a stronger and more persistent increase in respiratory risk associated with moderate to high pollution levels. Atlantic regions, despite their heterogeneous microclimatic conditions, consistently show elevated respiratory risks linked to intense precipitation, high relative humidity, and cooler microclimates. In summary, extreme precipitation, elevated humidity, and increased aerosol loads play a more substantial role in driving respiratory hospitalizations than temperature, while satellite-derived aerosol optical depth likely underestimates surface-level pollution exposure in Pacific regions during dry and atmospherically stable conditions.

Overall, this study demonstrates a consistent methodological framework that provides quantitative evidence of the nonlinear lagged influence of climate variables on respiratory discharges, while incorporating random effects to account for inter-subregional heterogeneity in small-population settings. This evidence supports epidemiological surveillance and the development of early warning thresholds in tropical peripheral regions with differential risks, consistent with recent work in non-temperate climates (10, 11, 29) and previous applications of DLNM (30).

2 Materials and methods

2.1 Study area

This study focuses on Costa Rica, a country where diverse climatic conditions strongly influence the incidence of respiratory diseases (31). Illnesses such as asthma, bronchitis, and pneumonia are major causes of morbidity and hospitalization, particularly among children and the older population (32). The tropical and subtropical climate, characterized by high humidity and variable temperatures, plays a central role in both the prevalence and exacerbation of these conditions (6, 31, 33). According to the National Meteorological Institute (16), Costa Rica is divided into seven main climatic regions and several subregions, each defined by distinct patterns of temperature, precipitation, and dry season duration. Table 1 summarizes the principal characteristics of each climatic region and Figure 1a shows the map of the study area. VC is excluded from the study due to its high population density and large number of hospital discharges, which require a different modeling approach (see Figure 1b).

2.2 Data

Daily hospital discharge records due to respiratory causes [ICD-10 codes J00-J99 (34)], provided by the Costa Rican Social Security Fund (CCSS), were aggregated by epidemiological week to ensure consistency with the temporal resolution of environmental exposures. The dependent variable therefore corresponds to the weekly count of respiratory hospital discharges per climatic subregion from 2000 to 2019 across 27 peripheral climatological subregions of Costa Rica (16). The Central Valley is excluded because of its large population, high rates of respiratory hospitalizations, and a separate modeling strategy is required (Figures 1b, c). See also Supplementary Figure S1 for time series plots.

Climatic exposures were obtained using satellite-derived data on mean temperature (T) (35), precipitation (P) (36), relative humidity (RH) (37), and aerosol optical depth (AOD) (38), as a proxy for air pollution (39, 40). Both AOD and relative humidity data contained missing values across several subregions. To address missing values in aerosol optical depth (AOD) and relative humidity (RH), we applied an inverse distance weighting (IDW) interpolation method (41, 42), using Euclidean distances between each satellite grid cell and the centroid of each climatic subregion. For a given week and subregion, exposure values were aggregated as a weighted average of surrounding grid cells according to

$$\hat{X}_{it} = \frac{\sum_{j=1}^m w_{ij} X_{jt}}{\sum_{j=1}^m w_{ij}}, \quad w_{ij} = \frac{1}{d_{ij}^p},$$

where \hat{X}_{it} denotes the interpolated exposure (AOD or RH) for subregion i at week t , X_{jt} is the observed exposure at grid cell j , d_{ij} is the Euclidean distance between the centroid of subregion i and grid cell j , and p is the distance decay parameter. In this study, we set $p = 1$, assigning greater influence to nearby grid cells while preserving smooth spatial transitions.

Following an exploratory analysis, a marked spatial heterogeneity was observed both in the magnitude and variability of respiratory hospital discharges, as well as in climatic conditions. Although the overall proportion of zero counts in the full dataset was 5.44%, these zeros were primarily concentrated in subregions with low population density, supporting the use of count models capable of handling zero inflation (For further details, see the selected summary statistics in Table 2.

2.3 Model

Each region exhibits unique climatic conditions with only slight variation across its subregions; therefore, our aim is to capture differences in nonlinear and lagged associations across climatic regions. For this end, we propose a distributed lag nonlinear model (DLNM) with a random intercept for subregions to model weekly hospital discharges for each region, using Poisson (PO), negative binomial (NegBin) distributions, and their zero-inflated counterparts (ZIP and ZINBI, respectively) to account for the excess zeros.

For a given climate region, the hospital discharges y_{it} in subregion i at week t are modeled using a count distribution. Specifically, $Y_{it} \sim \text{Poisson}(\mu_{it})$ or $Y_{it} \sim \text{NegBin}(\mu_{it}, \kappa)$ with mean μ_{it} :

$$\log(\mu_{it}) = \beta_0 + \sum_{v=1}^4 s_v(X_{vit}, l_v) + \gamma_s' D_t + b_i + \log(N_{it}),$$

where β_0 is the intercept, $b_i \sim \mathcal{N}(0, \sigma_b^2)$ is the random intercept capturing subregional variability, D_t is a vector of 51 dummy variables that account for the seasonal effect with 52 levels, and γ denotes the corresponding coefficients, and $\log(N_{it})$ is the offset accounting the population at risk. Additionally, $s_v(\cdot)$ is cross-basis functions representing nonlinear exposure-response and lag associations for each exposure predictor v , of the form

$$s_v(x_t; \beta) = \sum_{l=0}^L \sum_{j=1}^{v_x} \sum_{k=1}^{v_l} b_j(x_{t-l}) h_k(l) \beta_{jk},$$

with x_t denoting exposure at time t ; L the maximum lag (with $l = 0, 1, \dots, L$); v_x and v_l the degrees of freedom for spline basis functions in the exposure and lag spaces, respectively; $b_j(\cdot)$ the j th basis in exposure evaluated at x_{t-l} ; and $h_k(\cdot)$ the k th basis function evaluated at l (17, 18, 43, 44).

Overdispersion was assessed in the Poisson models by examining the ratio of residual deviance to degrees of freedom. In all climatic regions, this ratio exceeded 1, indicating extra-Poisson variability (45, 46). Consequently, negative binomial models were considered, as they include an additional dispersion parameter allowing the variance to exceed the mean. Model selection was further supported by information criteria (AIC, BIC) and predictive accuracy measures (MAE, RMSE), ensuring both goodness-of-fit and parsimony.

In the presence of excess zeros, we extend the count family to zero-inflated specifications with density function:

$$f_Y = \pi_{it} \delta_0 + (1 - \pi_{it}) f_Y(\mu_{it}),$$

TABLE 1 Description of climatic and subclimatic regions of Costa Rica.

Climate region	Description	Subregions
North Pacific Region (PN)	Temperate and tropical climates with a dry season, influenced by geographic factors.	Western Nicoya Peninsula (PN1), Central North Pacific (PN2), foothills of the Guanacaste and Tilarán ranges (PN3), and the lower basins of the Barranca and Grande de Tárcoles rivers (PN4).
Central Pacific Region (PC)	Tropical climate with a marked rainy season and a short, moderate dry season.	Parrita Valley (PC1), Naranjo River Basin, Quepos (PC2), and Barú River Basin, Dominical (PC3).
South Pacific Region (PS)	Unique rainfall regime with a very short dry period and intense rainfall throughout the year.	Valle del General and Coto Brus (PS1), Diquís Valley (PS2), Coto Colorado Valley (PS3), Osa Peninsula (PS4), and the Pacific slopes of the Talamanca Range (PS5).
Southern Mountain Region (RMS)	Highland valleys that result in warm temperatures in lower areas and cold conditions in higher zones.	Upper Basins of the Turrubares and Tulin rivers, and the Candelaria Valley (RMS1), and Upper Basin of the Pirris or Parrita River (RMS2).
Central Valley Region (VC)	This region is influenced by both Caribbean and Pacific foothill climates, with areas where the dry season is reduced to one month and pockets of temperate climate.	Western Central Valley (VC1), Eastern Central Valley (VC2), and the slopes of the Central Volcanic Range (Poás, Barva, and Irazú) (VC3). This region has the highest population density and urbanization levels and is also among the most exposed to environmental pollutants.
Northern Region (RN)	It belongs to the Caribbean precipitation regime, which is rainy year-round, with relatively drier months in February, March, and October.	Eastern slopes of the Guanacaste and Tilarán ranges (RN1), Northern slopes of the Central Volcanic Range (RN2), Guatusos Plains (RN3), San Carlos Plains (RN4), and Sarapiquí Plains (RN5).
Atlantic Region (RA)	This humid tropical region experiences abundant rainfall, especially in the mountainous areas where it rains year-round.	Basins of the Macho, Grande de Orosi, and Pejibaye rivers (RA1), Atlantic slopes of the Irazú-Turrialba Massif and the Talamanca Range – Reventazón Valley (RA2), Tortuguero Plains (RA3), Santa Clara and Matina Plains and the Banano River Basin (RA4), and the southern area of the Banano River Basin (RA5).

where π_{it} (logit link) models structural zeros and $f_{Y_{it}}(\cdot)$ is either Poisson or negative binomial (27, 28, 47). In the zero-inflated specifications, the inflation probability is modeled using a logit link with an intercept-only structure, while all climatic exposures, cross-basis terms, seasonal indicators, offset, and subregional random effects are included exclusively in the count component. Thus, zero inflation captures structural excess zeros not explained by the covariates, without introducing additional predictors in the inflation equation. We refer them as DLNM-ZIP_{rf} and DLNM-ZINBI_{rf}, respectively. This decouples zero generation from count dynamics—critical in sparsely populated subregions, especially peripheral or coastal areas of Costa Rica with low population density, limited hospital access, and higher frequency of incomplete records.

The specification of the DLNM requires determining the relevant lag structure and degree of nonlinearity for each climatic exposure. To this end, an exploratory dependence analysis between each exposure and the response variable was conducted for lags up to 14 weeks in each region and subregion. A maximum lag of 14 weeks was selected as an upper bound, consistent with epidemiological evidence indicating that delayed respiratory responses to climatic exposures may extend over multiple weeks, while avoiding unnecessarily long lag windows that could induce overfitting (15, 17, 18).

Pearson correlation (capturing linear dependence) and distance correlation and mutual information (capturing both linear and nonlinear dependence) were computed at each lag. Statistical significance of all measures was assessed using

a non-parametric permutation test (44, 48–52), allowing explicit comparison between purely linear associations and more general dependence structures. The maximum lag was defined as the largest lag at which at least two of the three dependence measures remained statistically significant ($p < 0.05$), ensuring robustness under both linear and nonlinear dependence.

To capture nonlinear exposure-lag relationships, cubic spline bases were specified for both the predictor and lag spaces. Alternative specifications with 2 and 4 degrees of freedom were evaluated; however, 3 degrees of freedom provided a parsimonious balance between flexibility and stability, as indicated by improved predictive accuracy metrics and information criteria. This choice is also consistent with standard DLNM applications in environmental epidemiology (17, 18). Knot placement in both the exposure and lag dimensions followed the default quantile-based specification of the natural cubic spline function implemented in the `dlnm` package (i.e., internal knots placed at equally spaced quantiles of the observed distribution), without manual tuning.

We also fitted these models without the lagged and nonlinear specifications, that is, as generalized linear models with random effects for all specified distributions (GLM-Poisson_{rf}, GLM-NegBin_{rf} and GLM-ZIP_{rf}), in order to evaluate the improvement of incorporating the DLNM framework.

Model selection was based on information criteria (AIC, BIC) (53–57), and accuracy measures, including mean absolute error (MAE), and root mean square error (RMSE) (58, 59, 59–63). The

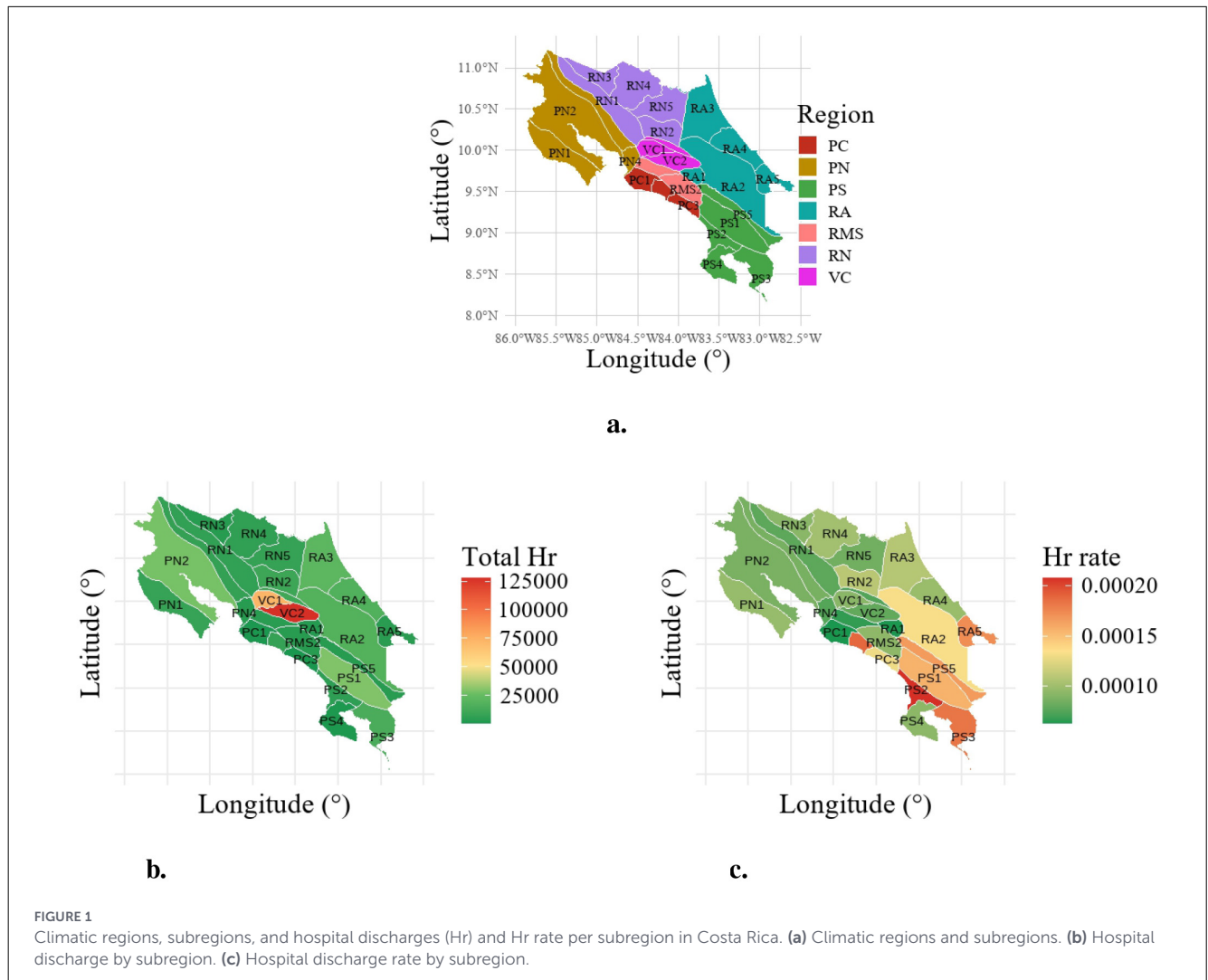


FIGURE 1 Climatic regions, subregions, and hospital discharges (Hr) and Hr rate per subregion in Costa Rica. (a) Climatic regions and subregions. (b) Hospital discharge by subregion. (c) Hospital discharge rate by subregion.

RMSE defined as follows

$$RMSE = \sqrt{\frac{1}{Tn} \sum_{i,t} (Y_{i,t} - \hat{Y}_{i,t})^2},$$

where $t = 1, \dots, T$ denotes the number of weeks for $i = 1, \dots, n$ subregion in each climatic region; $Y_{i,t}$ and $\hat{Y}_{i,t}$ are the observed and estimated hospital discharges, respectively, for subregion i at week t . The MAE is computed as follows:

$$MAE = \frac{1}{Tn} \sum_{i,t} |Y_{i,t} - \hat{Y}_{i,t}|.$$

We also incorporate a composite scoring methodology following the principles of rganisation for Economic Co-operation and Development (OECD) (64), where normalization is essential for ensuring comparability among heterogeneous indicators. Denote by $m_{i,j}$ the value of the j -th performance metric (MAE, RMSE, AIC, BIC) for model i . Two normalization schemes were adopted: (i) the *min-max* normalization, defined as

$$z_{i,j}^{(1)} = 1 - \frac{m_{i,j} - \min(m_j)}{\max(m_j) - \min(m_j)},$$

which rescales all metrics into the interval $[0, 1]$ with inverse orientation (smaller values implying better performance); and (ii) the *z-score* standardization,

$$z_{i,j}^{(2)} = -\frac{m_{i,j} - \bar{m}_j}{s_j},$$

where \bar{m}_j and s_j are the mean and standard deviation of metric j , respectively. For each model i , the overall composite score was obtained as the arithmetic mean across all normalized indicators:

$$S_i^{(k)} = \frac{1}{p} \sum_{j=1}^p z_{i,j}^{(k)},$$

where $k = 1, 2$, and p denotes the number of metrics considered. We refer to $S_i^{(1)}$ and $S_i^{(2)}$ as the *Min-max score* and the *Z-score*, respectively. This normalization-based aggregation enhances interpretability and reduces implicit weighting biases among indicators (65). Sensitivity analysis was performed to compare the two composite scores, revealing a strong linear association ($r > 0.9$), which confirms the robustness of the ranking and the internal consistency of the derived indicators. In addition, raw residuals, Pearson residuals, and deviance residuals were inspected;

TABLE 2 Descriptive statistics (mean, first-third quantiles, maximum and percentage of zeros) for weekly hospital discharges (Y) and weekly mean climatic covariates—temperature (T), precipitation (P), relative humidity (RH), and aerosol optical depth (AOD)—by climatic subregion during 2000–2019.

Subregion	Hospital discharge (Y)						P	T	RH	AOD
	Mean	Q1	Median	Q3	Max	Zeros (%)				
PC1	1.06	0	1	2	5	35.96	9.61	29.3	0.372	1.94 × 10 ⁻⁴
PC2	7.67	5	7	10	22	0.10	12.7	29.8	0.370	1.92 × 10 ⁻⁴
PC3	0.763	0	1	1	5	48.56	14.4	27.6	0.369	1.91 × 10 ⁻⁴
PN1	8.02	6	8	10	23	0.19	7.92	29.8	0.375	1.97 × 10 ⁻⁴
PN2	27.4	21	27	32.5	62	0.00	6.96	30.0	0.376	1.96 × 10 ⁻⁴
PN3	7.03	5	6	9	20	0.19	8.26	26.9	0.375	1.94 × 10 ⁻⁴
PN4	3.79	2	4	5	14	3.65	7.75	29.0	0.373	1.93 × 10 ⁻⁴
PS1	28.7	22	27	34	63	0.00	12.0	23.5	0.368	1.88 × 10 ⁻⁴
PS2	5.38	3	5	7	20	0.87	11.3	27.9	0.367	1.90 × 10 ⁻⁴
PS3	14.7	11	14	18	45	0.00	14.9	28.4	0.367	1.89 × 10 ⁻⁴
PS4	1.25	0	1	2	7	32.31	14.3	30.1	0.366	1.91 × 10 ⁻⁴
PS5	3.89	2	4	5	17	3.56	12.6	20.7	0.368	1.87 × 10 ⁻⁴
RA1	3.28	2	3	4	13	5.87	12.4	18.7	0.370	1.89 × 10 ⁻⁴
RA2	18.1	14	17	22	49	0.00	15.2	21.6	0.370	1.86 × 10 ⁻⁴
RA3	21.4	17	21	25	48	0.00	18.2	28.4	0.372	1.86 × 10 ⁻⁴
RA4	17.0	13	16	21	40	0.10	18.7	27.6	0.371	1.84 × 10 ⁻⁴
RA5	4.33	2	4	6	20	3.65	15.0	28.8	0.369	1.84 × 10 ⁻⁴
RMS1	4.33	3	4	6	12	1.83	11.1	24.9	0.372	1.92 × 10 ⁻⁴
RMS2	3.88	2	3	5	18	3.75	12.8	21.5	0.370	1.90 × 10 ⁻⁴
RN1	6.92	5	7	9	20	0.19	11.7	27.2	0.375	1.93 × 10 ⁻⁴
RN2	10.1	7	10	12	30	0.00	13.6	23.4	0.373	1.90 × 10 ⁻⁴
RN3	3.78	2	3	5	15	5.00	9.81	30.9	0.375	1.92 × 10 ⁻⁴
RN4	6.14	4	6	8	21	1.06	13.0	29.5	0.375	1.91 × 10 ⁻⁴
RN5	7.33	5	7	9	23	0.38	16.1	28.6	0.374	1.89 × 10 ⁻⁴

among competing model specifications, those minimizing residual autocorrelation were prioritized to ensure the model captures underlying systematic temporal patterns (28, 66–68).

Finally, exposure-lag surfaces and relative risk curves with 95% confidence intervals were reported for interpretation (43, 47). In addition, cumulative incidence rate ratios (IRRs) were derived from the fitted DLNM by aggregating lag-specific contributions across the entire lag period and exponentiating the corresponding linear predictor, following the standard DLNM framework (17, 43). To account for potential nonlinear and asymmetric exposure–response relationships, percentile-based contrasts were defined by comparing the 95th and 5th percentiles of each exposure to the regional median (50th percentile), as commonly applied in environmental epidemiology studies using distributed lag models (69, 70). Confidence intervals were obtained from the model-based variance–covariance matrix.

All analyses were conducted using the open-source software R (version 4.4.3) (71). The DLNM with random effects was fitted using the `dlnm` package (version 2.4.10) (43), while the zero-inflated versions were implemented within the Generalized Additive Model for Location, Scale, and Shape

(GAMLSS) framework using the `gamlss` package (version 5.5-0) (27, 28).

3 Results

Table 3 reports the maximum lag for each covariate in across regions, based on an exploratory correlation analysis using Pearson correlation, distance correlation, and mutual information to capture both linear and nonlinear relationships. A higher lag value indicates more persistent lagged associations; therefore, it is important to include variables up to those lags in the model. Similar patterns are observed across subregions within each region (see Supplementary Tables S1 and Supplementary Figures S2, S3 for results by subregion).

Across all exposures, nonlinear dependence predominated at the most pronounced lags, where significant Pearson correlations were comparatively smaller in magnitude than distance correlation and mutual information values, further supporting the need for a nonlinear lag specification.

TABLE 3 Maximum lag (L) selected for each climatic exposure (temperature, precipitation, relative humidity, and AOD) in the DLNM across climatic regions, based on the correlation analysis (Pearson, distance correlation, and mutual information) between each exposure and the hospitalization.

Exposure	PC	PN	PS	RN	RA	RMS
Temperature	9	14	11	7	5	14
Precipitation	9	6	7	4	12	11
Relative humidity	9	8	7	4	7	13
AOD	11	13	13	11	11	11

By using the specification of the relevant lag structure and degree of nonlinearity for each climatic exposure, we fit all described models. Table 4 compares the goodness of fit of all fitted models for each region. The composite scores indicate that the DLNM with random effects and a negative binomial distribution (DLNM-NegBin_{rf}) achieved the highest normalized performance in most peripheral climatic regions (PC, PN, PS, and RN), reflecting a superior balance between model fit and parsimony.

In contrast, the zero-inflated variant (DLNM-ZINBI_{rf}) exhibited slightly higher scores in regions RA and RMS, where overdispersion and structural zeros are more pronounced. These results suggest that negative binomial mixed specifications provide the most robust and stable modeling framework overall, while zero-inflated extensions offer localized improvements under data sparsity or excess-zero conditions.

The inclusion of nonlinear lagged effects of climatic covariates improved model fit. Moreover, subregional random effects adequately captured unexplained spatial heterogeneity, reducing residual variance and further enhancing overall model performance. This advantage was consistent compared to alternatives without random effects, which exhibited overdispersion in the response variable. The performance of the best-fitting model for each region was also evaluated through visual inspection of the Pearson residuals (for more details, see Supplementary Figures S10–S15). Although the residuals of the fitted models in some subregions still exhibited autocorrelation, we consider them appropriate for the subsequent analysis.

The main advantage of the DLNM framework lies in its ability to represent the lagged and nonlinear effects of each covariate by estimating the exposure-response-lag relationship. These effects were obtained from the best-fitting model for each region by exponentiating the covariate effects derived from the product of the cross-basis functions. These values were compared with a reference value (their average) to quantify the relative risk (RR) associated with increases in hospital discharges.

For each climatic region, we extract the relevant lags and exposure ranges for each covariate at 5% significance level, and the summaries are visualized in Figure 2 (For precise ranges for each covariate across regions, see Supplementary Tables S3, S4, and the three-dimensional surfaces, contour maps, and lag-response curves of the selected model across the six regions can be found in Supplementary Figures S4–S9). Colored rectangles denote exposure-lag combinations with relative risks (RR) significantly different from 1 at the 5% level. Green indicates decreased risk ($RR < 1$) and red increased risk ($RR > 1$). The

horizontal and vertical extents represent exposure ranges and lag (weeks), respectively. The 5th (blue dashed line) and 95th (orange dashed line) percentiles of each climatic exposure for each region are included to facilitate meaningful interpretation (For more details on the descriptive statistics of the climatic exposures, see Supplementary Table S2).

Cumulative effects were derived from the DLNM–GLMM by integrating lag-specific contributions across the entire lag period, yielding cumulative incidence rate ratios (IRRs) as measures of the total relative risk associated with a given exposure level. Because climatic exposure–response relationships are typically non-linear and potentially asymmetric, percentile-based contrasts were used instead of linear increments. Specifically, cumulative IRRs were computed comparing the 95th percentile to the median (high exposure contrast) and the 5th percentile to the median (low exposure contrast). This bidirectional percentile approach enables a balanced assessment of risk associated with both elevated and reduced climatic conditions and is consistent with recent environmental epidemiology studies summarizing non-linear climate-related health impacts using distributed lag models (69, 70). The resulting regional cumulative IRRs, along with their corresponding confidence intervals and statistical significance, are presented in Table 5.

Across climatic regions, we summarize these findings as follows:

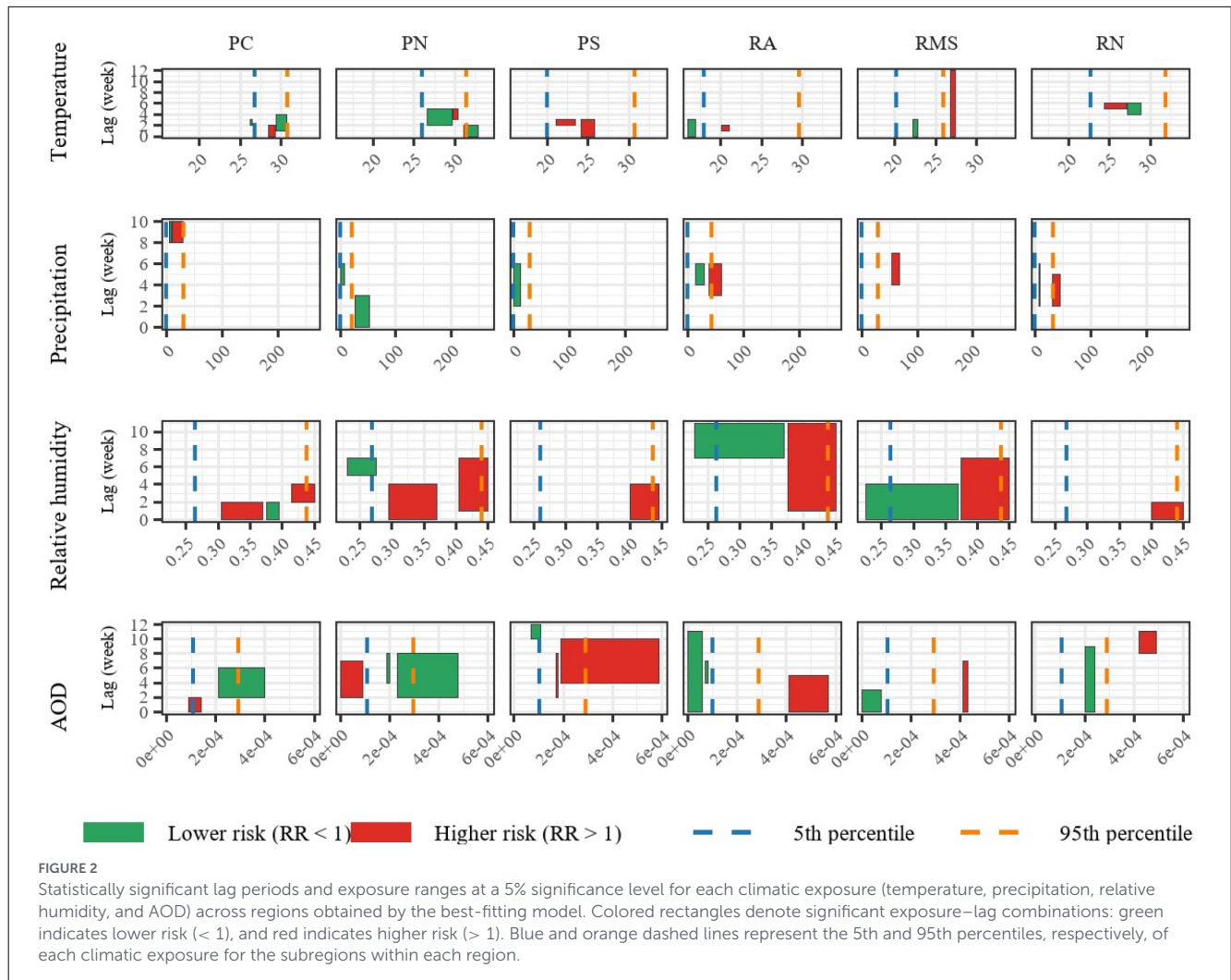
- PC and PN share similar patterns in almost all variables since both belong to the North and Central Pacific, with similar climate dynamics.
- PS presents a more pronounced relative risk for medium or high pollution levels, which persist for 5–10 weeks, a considerable difference from the other Pacific regions.
- The RA, RMS, and RN regions show heterogeneous responses in climate and pollution variables. They exhibit higher precipitation levels than the other peripheral regions, associated with an increased risk of hospital discharges for respiratory illnesses. This is consistent with the climatic characteristics of these three regions, which are characterized by more intense rainfall and cooler microclimates.

Regarding to the temperature, the intervals are small and dispersed; temperature effects do not appear to dominate the weekly variability of discharges. However, the RMS region, which generally has lower temperatures than the other regions, shows a high relative risk near 27 °C for a prolonged period of 0 to 12 weeks. This suggests that averaged weekly temperature alone does not accurately capture the relevant physiological exposure (minimum temperature or temperature range would probably be more informative).

In terms of precipitation, in RA, RMS, and RN, high values (above the 95th percentile) are associated with higher risk, reflecting the impact of intense rainfall events, which are typically linked to increased ambient humidity. At low precipitation levels, the protective effect may be due to greater atmospheric stability and fewer fluctuations in relative humidity. PC region presents a high risk at medium values (10–30 mm), which are medium to high precipitation in this region and a drier climate is common. This could reflect transitional episodes (beginning or end of the rainy

TABLE 4 Goodness-of-fit metrics (predictive accuracy, information criteria, and composite scores) of different fitted models for hospital discharge in six peripheral regions.

Region	Model	MAE	RMSE	AIC	BIC	Min–Max score	Z-score
PC	GLM-Poisson _{rf}	1.785	2.558	12,398.460	12,736.250	0.029	-1.676
	GLM-NegBin _{rf}	1.780	2.608	11,732.010	12,075.840	0.222	-1.145
	DLNM-ZIP	1.469	2.210	10,992.310	11,529.160	0.898	0.444
	DLNM-ZINBI	1.476	2.237	10,825.780	11,368.670	0.931	0.542
	DLNM-Poisson _{rf}	1.450	2.182	10,915.300	11,452.150	0.954	0.580
	DLNM-NegBin _{rf}	1.455	2.202	10,771.750	11,314.630	0.985	0.670
	DLNM-ZIP _{rf}	1.450	2.183	10,898.830	11,453.730	0.955	0.585
PN	GLM-Poisson _{rf}	3.499	5.129	23,541.300	23,901.530	0.028	-1.693
	GLM-NegBin _{rf}	3.513	5.157	22,520.000	22,886.540	0.416	-0.605
	DLNM-ZIP	3.267	4.721	22,843.380	23,405.830	0.706	0.030
	DLNM-ZINBI	3.286	4.774	22,204.150	22,772.920	0.919	0.640
	DLNM-Poisson _{rf}	3.247	4.693	22,821.550	23,384.000	0.749	0.136
	DLNM-NegBin _{rf}	3.270	4.748	22,204.880	22,773.660	0.948	0.709
	DLNM-ZIP _{rf}	3.247	4.692	22,802.780	23,390.420	0.751	0.144
	DLNM-ZINBI _{rf}	3.287	4.774	22,204.160	22,772.990	0.918	0.638
PS	GLM-Poisson _{rf}	3.187	4.673	27,386.570	27,753.030	0.070	-1.666
	GLM-NegBin _{rf}	3.210	4.726	26,487.680	26,860.680	0.277	-0.852
	DLNM-ZIP	3.047	4.528	26,583.890	27,166.290	0.556	-0.206
	DLNM-ZINBI	3.050	4.500	26,399.690	26,988.630	0.633	0.050
	DLNM-Poisson _{rf}	2.974	4.436	26,245.750	26,828.150	0.818	0.571
	DLNM-NegBin _{rf}	2.981	4.459	25,654.000	26,242.950	0.973	1.154
	DLNM-ZIP _{rf}	2.974	4.437	26,218.870	26,833.930	0.820	0.579
	DLNM-ZINBI _{rf}	3.015	4.462	26,253.590	26,874.640	0.743	0.371
RA	GLM-Poisson _{rf}	4.292	5.770	33,765.130	34,131.530	0.135	-1.705
	GLM-NegBin _{rf}	4.467	6.106	30,928.610	31,301.550	0.302	-1.129
	DLNM-ZIP	3.770	5.173	30,577.350	31,159.660	0.770	0.095
	DLNM-ZINBI	3.808	5.228	29,244.000	29,832.860	0.887	0.466
	DLNM-Poisson _{rf}	3.663	5.070	30,130.690	30,713.000	0.876	0.391
	DLNM-NegBin _{rf}	3.678	5.096	28,976.190	29,565.050	0.988	0.740
	DLNM-ZIP _{rf}	3.664	5.071	30,083.320	30,698.320	0.879	0.400
	DLNM-ZINBI _{rf}	3.678	5.096	28,950.060	29,571.590	0.989	0.743
RMS	GLM-Poisson _{rf}	1.800	2.292	9,157.490	9,478.260	0.166	-1.067
	GLM-NegBin _{rf}	1.801	2.294	9,095.490	9,421.890	0.354	-0.480
	DLNM-ZIP	1.737	2.227	9,073.770	9,574.620	0.527	-0.263
	DLNM-ZINBI	1.736	2.228	9,041.210	9,547.690	0.626	0.043
	DLNM-Poisson _{rf}	1.721	2.204	9,042.990	9,543.840	0.743	0.318
	DLNM-NegBin _{rf}	1.722	2.206	9,015.580	9,522.060	0.816	0.549
	DLNM-ZIP _{rf}	1.722	2.205	9,033.980	9,545.990	0.749	0.339
	DLNM-ZINBI _{rf}	1.722	2.206	9,008.770	9,526.370	0.820	0.562
RN	GLM-Poisson _{rf}	2.598	3.389	26,935.300	27,301.700	0.023	-1.673
	GLM-NegBin _{rf}	2.606	3.394	26,231.330	26,604.280	0.361	-0.669
	DLNM-ZIP	2.511	3.266	26,466.560	27,048.570	0.550	-0.320
	DLNM-ZINBI	2.515	3.272	26,000.700	26,589.550	0.771	0.338
	DLNM-Poisson _{rf}	2.476	3.224	26,282.280	26,864.590	0.774	0.262
	DLNM-NegBin _{rf}	2.478	3.227	25,846.510	26,435.370	0.988	0.895
	DLNM-ZIP _{rf}	2.477	3.224	26,254.970	26,869.820	0.777	0.272
	DLNM-ZINBI _{rf}	2.478	3.227	25,830.140	26,451.240	0.987	0.894



season), where there is high variability and particles resuspended by the first rain (72).

On the other hand, high RH values (>0.43) coincide with higher risks, reflecting saturated environments in which respiratory illnesses tend to increase. In PC and PN, average relative humidity values (0.36 – 0.40) are also associated with elevated risk, possibly due to the influence of Pacific maritime conditions, exposure to which has been shown to have detrimental effects on human health (73, 74). Low RH values (<0.27) are consistently protective, as drier conditions reduce the survival of biological agents and hygroscopic particles (75).

In terms of statistically significant cumulative effects, exposure to low temperatures (p5 vs. p50) was associated with a 15.4% increased risk in RN (IRR = 1.154, *p* = 0.0141), while significant protective effects were observed in PC, PS, PN, and RMS. Exposure to high temperatures (p95 vs. p50) showed significant protective effects in PC and PS, but an increased risk in RMS. Relative humidity exhibited the most consistent and strongest positive associations across regions, particularly in RA, PC, PS, PN, and RMS, where high exposure contrasts produced IRRs substantially greater than 1. Aerosol optical depth (AOD) showed significant inverse associations at high percentiles in most regions, whereas low AOD levels were associated with increased risk in RN and

PN. In contrast, precipitation effects were generally weaker and less consistent, although significant associations were detected in selected regions for low exposure contrasts.

Finally, high AOD values imply greater risk of respiratory hospitalization in most regions, reflecting the fine aerosol load and its relationship with respiratory illnesses. In PC and PN, however, the opposite pattern is observed. This could be explained by two factors related to thermal inversion and atmospheric stability during the dry season: (1) low AOD measured by satellite despite high local surface concentrations due to limited vertical mixing, and (2) the presence of local non-optically active pollutants, such as ozone or gases, that do not increase AOD but do increase risk (76). This suggests that in the Pacific, satellite-detected aerosols may underestimate surface concentrations during dry and stable periods.

4 Discussion

We proposed a methodological framework based on the DLNM that integrates subregional random effects, zero-inflated distributions, and nonlinear and delayed effects of climatic and

TABLE 5 Regional cumulative incidence rate ratios (IRR) comparing extreme percentiles against the median.

Region	Variable	Contrast	IRR	CI _{low}	CI _{high}	p
RN	T	p95 vs. p50	1.022	0.902	1.158	0.7338
		p5 vs. p50	1.154	1.029	1.294	0.0141
	P	p95 vs. p50	0.984	0.883	1.097	0.7712
		p5 vs. p50	1.035	0.932	1.149	0.5236
	RH	p95 vs. p50	1.153	1.092	1.218	0.0000
		p5 vs. p50	1.049	0.962	1.143	0.2771
	AOD	p95 vs. p50	0.768	0.726	0.813	0.0000
p5 vs. p50		1.157	1.099	1.219	0.0000	
RA	T	p95 vs. p50	1.172	0.775	1.773	0.4524
		p5 vs. p50	1.467	0.901	2.389	0.1237
	P	p95 vs. p50	1.155	0.984	1.357	0.0788
		p5 vs. p50	1.097	1.049	1.147	0.0001
	RH	p95 vs. p50	1.556	1.411	1.717	0.0000
		p5 vs. p50	0.832	0.754	0.918	0.0003
	AOD	p95 vs. p50	0.802	0.765	0.840	0.0000
p5 vs. p50		0.922	0.881	0.965	0.0005	
PC	T	p95 vs. p50	0.855	0.749	0.977	0.0215
		p5 vs. p50	0.714	0.534	0.955	0.0232
	P	p95 vs. p50	1.009	0.736	1.383	0.9575
		p5 vs. p50	0.968	0.628	1.492	0.8820
	RH	p95 vs. p50	1.357	1.193	1.543	0.0000
		p5 vs. p50	1.075	0.889	1.301	0.4557
	AOD	p95 vs. p50	0.847	0.718	1.000	0.0495
p5 vs. p50		1.031	0.840	1.266	0.7685	
PS	T	p95 vs. p50	0.694	0.605	0.796	0.0000
		p5 vs. p50	0.574	0.447	0.737	0.0000
	P	p95 vs. p50	0.970	0.878	1.072	0.5551
		p5 vs. p50	0.750	0.654	0.860	0.0000
	RH	p95 vs. p50	1.174	1.095	1.259	0.0000
		p5 vs. p50	1.056	0.931	1.198	0.3958
	AOD	p95 vs. p50	0.908	0.845	0.976	0.0089
p5 vs. p50		0.971	0.896	1.053	0.4774	
PN	T	p95 vs. p50	0.993	0.963	1.023	0.6294
		p5 vs. p50	0.824	0.726	0.936	0.0029
	P	p95 vs. p50	1.018	0.917	1.130	0.7416
		p5 vs. p50	0.825	0.713	0.954	0.0094
	RH	p95 vs. p50	1.739	1.579	1.916	0.0000
		p5 vs. p50	0.906	0.785	1.047	0.1803
	AOD	p95 vs. p50	0.824	0.774	0.876	0.0000
p5 vs. p50		1.127	1.047	1.214	0.0016	
RMS	T	p95 vs. p50	1.199	1.015	1.416	0.0327
		p5 vs. p50	0.688	0.534	0.888	0.0040
	P	p95 vs. p50	1.105	0.911	1.340	0.3118
		p5 vs. p50	0.931	0.808	1.074	0.3256
	RH	p95 vs. p50	2.048	1.620	2.590	0.0000
		p5 vs. p50	0.574	0.354	0.932	0.0248
	AOD	p95 vs. p50	0.768	0.663	0.890	0.0004
p5 vs. p50		1.039	0.879	1.228	0.6525	

IRR denotes cumulative incidence rate ratio. p95 and p5 correspond to the 95th and 5th percentiles of each exposure relative to the regional median (p50). CI low and CI high correspond to the lower and upper bounds of the 95% confidence interval.

pollution variables to model weekly hospital discharges due to respiratory causes in Costa Rica. The results confirm that DLNM with count distribution with subregional random effects capture nonlinear and lagged relationships between climate, pollution, and respiratory morbidity/mortality in high-variability contexts. In particular, we identified critical risk-increase windows associated with environmental variables, with heterogeneity within regions and by exposure level.

The regional heterogeneity observed in the exposure-response relationships is likely driven by climatic, environmental, and demographic differences across Costa Rican subregions. Previous studies have shown that the health impacts of temperature and precipitation depend strongly on local baseline climate conditions, population adaptation, and infrastructure characteristics (7, 69). For instance, warmer regions may exhibit attenuated heat effects due to long-term acclimatization, whereas cooler or highland regions may show stronger relative risks for equivalent temperature deviations. Similar regional variability has been documented in respiratory morbidity studies using distributed lag models (9, 12, 13).

In Costa Rica, climatic gradients between coastal and inland regions are well established (16), and previous national studies have reported spatially heterogeneous associations between climate variables and health outcomes (15, 77). Differences in humidity patterns, precipitation regimes, and aerosol composition (72, 78) may partly explain why certain regions exhibited stronger cumulative IRRs at extreme percentiles. These findings are consistent with the broader environmental epidemiology literature, which emphasizes that exposure-lag-response associations are context-specific and influenced by regional climatic baselines and adaptive capacity (19, 79).

The diversity of results reported in the literature may be attributed to differences in climatic conditions, data quality, study designs, and analytical methods. Nevertheless, the findings of our study are consistent with those of some previous works. An additional consideration is that in peripheral regions of Costa Rica, as in several tropical countries, cold weather is generally perceived as comfortable compared to temperate regions. Conversely, during periods of extreme heat, people tend to spend more time indoors to maintain comfort, often relying on artificial cooling systems.

These patterns are consistent with evidence from tropical/subtropical climates reporting nonlinear and delayed environmental effects on respiratory outcomes, strong spatial heterogeneity of associations, and the usefulness of hierarchical models with zero-inflated distributions when structural zeros and temporal dependence are present. The superiority of models with random effects over non-hierarchical alternatives has also been documented for improving estimate stability and the identification of regional thresholds.

Although the results cannot be interpreted as causal, the nonlinear lagged associations detected in this study are consistent with the existing literature. The exposure-lag thresholds enable the translation of model findings into surveillance and operational alerts; however, their direct applicability to policy decisions requires careful consideration. We recommend (1) defining activation thresholds based on extreme percentiles (e.g., P90/P95) of climatic exposures by subregion; (2) prioritizing peripheral regions with limited access to health services; (3) using hospital discharges as

a sensitive indicator for detecting weekly peaks and standardizing monitoring; and (4) integrating these models into epidemiological surveillance systems to trigger preventive actions and guide resource allocation.

Our findings are consistent with previous studies conducted in Latin America examining climate-health associations. In Colombia, Rúa et al. (80) reported significant lagged associations between local climatic variability and dengue transmission using time-series models. Similarly, Vásquez et al. (81) identified nonlinear relationships between temperature and precipitation with dengue incidence in Costa Rica. Although these investigations focused on vector-borne diseases rather than respiratory outcomes, they consistently demonstrated that climatic exposures exert region-dependent and temporally lagged effects on health indicators, reinforcing the relevance of modeling delayed and nonlinear exposure-response relationships in tropical contexts.

Beyond cumulative incidence rate ratios, an important contribution of the present study lies in the identification of statistically significant exposure thresholds associated with relative risks above and below unity. By jointly examining exposure level, lag structure, and changes in hospital discharges, we were able to characterize regions of increased and decreased risk within the exposure-lag surface. This approach allows for a more nuanced interpretation than single-point estimates, as it captures the interaction between intensity of exposure and temporal displacement effects, consistent with distributed lag non-linear modeling frameworks (19, 79).

More broadly, studies conducted in tropical and subtropical settings have emphasized the importance of accounting for local climatic regimes when evaluating climate-sensitive health outcomes (10, 11). The presence of region-specific thresholds and heterogeneous lag structures in our results supports the notion that climatic baselines and adaptive conditions modulate health responses. In heterogeneous environments such as Costa Rica, percentile-based contrasts provide a scale-invariant framework for comparing effects across subregions while preserving the interpretation of relative risk in epidemiological terms.

The cumulative results highlight substantial regional heterogeneity in climate-health associations. Relative humidity emerged as the most consistent risk factor, suggesting that sustained high humidity conditions may play a critical role in respiratory morbidity in tropical peripheral regions. Temperature effects showed asymmetry, with cold-related impacts predominating in some regions (e.g., RN, PN), while heat-related increases were evident in RMS. The inverse association observed for high AOD percentiles in several regions may reflect complex exposure dynamics or residual confounding related to atmospheric conditions. Overall, these findings reinforce the importance of modeling non-linear and lagged effects, as both elevated and reduced climatic conditions contribute differently to respiratory risk across regions.

Among the limitations encountered in this study: (1) the series built from administrative records and satellite sources may contain measurement error, omissions, or inconsistencies; although data cleaning and imputation procedures were applied, potential biases remain. (2) DLNM interpretation may be affected by multicollinearity and by parametrization decisions for the basis functions; these were guided by correlation criteria, information metrics, and predictive performance, without

guaranteeing global optimality. (3) Extending the model to include random slopes to account for subregional climatic effects increases computational cost and the risk of non-convergence. (4) Although the DLNM design reduces residual autocorrelation, it does not eliminate it completely given the complexity of dependencies; thus, inferences should focus primarily on exposure-lag associations.

Taking these limitations into account, future work may include:

(1) exploring more flexible spatio-temporal models for low-count hospitalizations; (2) integrating additional pollutants and socio-environmental variables; and (3) using data with higher spatio-temporal resolution to enhance predictive capacity and improve early detection of extremes.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://github.com/Emanuelle-Parra/DLNM-GLMM-health>.

Author contributions

SC-C: Conceptualization, Validation, Writing – review & editing, Supervision, Data curation, Project administration, Methodology, Investigation, Writing – original draft, Funding acquisition, Visualization. EP: Conceptualization, Data curation, Investigation, Methodology, Writing – review & editing, Visualization, Formal analysis, Writing – original draft.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Shu Wei Chou Chen received funding from the Vice-Rectorate for Research of the University of Costa Rica (C4196 and C3175).

Acknowledgments

This work was conducted as part of the thesis titled “Modelo de la influencia de los factores climáticos y la contaminación

del aire en la incidencia de enfermedades respiratorias en las regiones y subregiones climáticas de Costa Rica,” carried out by the second author under the supervision of the first author, within the framework of project C4196. The authors express their gratitude to Luis A. Barboza and Fabio Sanchez for serving as thesis readers. Finally, the authors gratefully acknowledge the Health Statistics Area of the CCSS for providing the data utilized in this manuscript.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. We used ChatGPT to assist with language editing and proofreading of the manuscript. The author(s) take full responsibility for the content and interpretation.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2026.1753511/full#supplementary-material>

References

- Marciniuk D, Ferkol T, Nana A, de Oca MM, Rabe K, Billo N, Zar H. Respiratory diseases in the world. realities of today-opportunities for tomorrow. *Afr J Respir Med.* (2014) 9.
- Chang JH, Lee YL, Chang LT, Chang TY, Hsiao TC, Chung KF, et al. Climate change, air quality, and respiratory health: a focus on particle deposition in the lungs. *Ann Med.* (2023) 55:2264881. doi: 10.1080/07853890.2023.2264881
- D’Amato G. Climate change and respiratory diseases. *Eur Respir Rev.* (2014) 23:161–9. doi: 10.1183/09059180.00001714.
- Saleri N, Ryan ET. Respiratory infections. In: Keystone JS, Kozarsky PE, Connor BA, Nothdurft HD, Mendelson M, Leder K, editors. *Travel Medicine, 4th Edn.* Amsterdam: Elsevier (2019). p. 527–37. doi: 10.1016/B978-0-323-54696-6.00059-8

5. Levine S, Marciniuk D. Global impact of respiratory disease: what can we do, together, to make a difference? *Chest*. (2022) 161:1153–4. doi: 10.1016/j.chest.2022.01.014
6. Zhao Y, Huang Z, Wang S, Hu J, Xiao J, Li X, et al. Morbidity burden of respiratory diseases attributable to ambient temperature: a case study in a subtropical city in China. *Environ Health*. (2019) 18:89. doi: 10.1186/s12940-019-0529-8
7. Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*. (2009) 20:205–13. doi: 10.1097/EDE.0b013e318190ee08
8. Hajat S, Armstrong B, Gouveia A, Wilkinson P. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology*. (2005) 16:613–20. doi: 10.1097/01.ede.0000164559.41092.2a
9. Chen Y, Kong D, Fu J, Zhang Y, Zhao Y, Liu Y, et al. Associations between ambient temperature and adult asthma hospitalizations in Beijing, China: a time-stratified case-crossover study. *Respir Res*. (2022) 23:38. doi: 10.1186/s12931-022-01960-8
10. Pedder H, Kapwata T, Howard G, Naidoo RN, Kunene Z, Morris RW, et al. Lagged association between climate variables and hospital admissions for pneumonia in South Africa. *Int J Environ Res Public Health*. (2021) 18:6191. doi: 10.3390/ijerph18126191
11. Tran QA, Le VTH, Ngo VT, Le TH, Phung DT, Berman JD, et al. The association between ambient temperatures and hospital admissions due to respiratory diseases in the capital city of Vietnam. *Front Public Health*. (2022) 10:903623. doi: 10.3389/fpubh.2022.903623
12. Lin S, Luo M, Walker RJ, Liu X, Hwang SA, Chinery R. Extreme high temperatures and hospital admissions for respiratory and cardiovascular diseases. *Epidemiology*. (2009) 20:738–46. doi: 10.1097/EDE.0b013e3181ad5522
13. Tsangari H, Paschalidou AK, Kassomenos AP, Vardoulakis S, Heaviside C, Georgiou KE, et al. Extreme weather and air pollution effects on cardiovascular and respiratory hospital admissions in Cyprus. *Sci Total Environ*. (2016) 542:247–53. doi: 10.1016/j.scitotenv.2015.10.106
14. Peng W, Li H, Peng L, Wang Y, Wang W. Effects of particulate matter on hospital admissions for respiratory diseases: an ecological study based on 125 years of time series data in Shanghai. *Environ Health*. (2022) 21:12. doi: 10.1186/s12940-021-00828-6
15. Chou-Chen SW, Barboza LA. Forecasting hospital discharges for respiratory conditions in Costa Rica using climate and pollution data. *Math Biosci Eng*. (2024) 21:6539–58. doi: 10.3934/mbe.2024285
16. Solano J, Villalobos R. *Regiones y subregiones climáticas de Costa Rica*. Instituto Meteorológico Nacional (2020) Available online at: <https://www.imn.ac.cr/documents/10179/20909/Regionalizaci%C3%B3n+clim%C3%A1tica+de+Costa+Rica> (Accessed March 10, 2024).
17. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med*. (2010) 29:2224–34. doi: 10.1002/sim.3940
18. Gasparrini A. Distributed lag linear and non-linear models in R: the dlnm package. *Stat Med*. (2013) 33:5287–99. doi: 10.18637/jss.v043.i08
19. Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. *Stat Med*. (2014) 33:881–99. doi: 10.1002/sim.5963
20. Sera F, Armstrong B, Blangiardo M, Gasparrini A. An extended mixed-effects framework for multi-location time-series analyses. *Stat Med*. (2022) 41:2325–43. doi: 10.1002/sim.8362
21. Gasparrini A, Guo Y, Sera F, Vicedo-Cabrera AM, et al. Global, regional, and national burden of mortality associated with non-optimal ambient temperatures. *Lancet Planet Health*. (2022) 6:e360–9. doi: 10.1016/S2542-5196(21)00081-4
22. Wu J, Zha X, Ren X, Feng B, Zhang Y, Wang L, et al. The short-term effects of different cold spell definitions on asthma outpatient visits in Lanzhou, China. *BMC Pulm Med*. (2025) 25:145. doi: 10.1186/s12890-025-03605-0
23. Motlogeloa T, Wright CY, Kapwata T. Assessing the impact of climatic variability on acute respiratory diseases across diverse climatic zones in South Africa. *Sci Total Environ*. (2024) 912:170661. doi: 10.1016/j.scitotenv.2024.170661
24. Min Y, Agresti A. Random effect models for repeated measures of zero-inflated count data. *Stat Model*. (2005) 5:1–19. doi: 10.1191/1471082X05st084oa
25. Feng CX. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *J Stat Distrib Appl*. (2021) 8:1–19. doi: 10.1186/s40488-021-00121-4
26. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J*. (2017) 9:378–400. doi: 10.32614/RJ-2017-066
27. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape (with discussion). *Appl Stat*. (2005) 54:507–54. doi: 10.1111/j.1467-9876.2005.00510.x
28. Stasinopoulos MD, Rigby RA, Heller GZ, Voudouris V, De Bastiani F. *Flexible Regression and Smoothing: Using GAMLSS in R*. Boca Raton, FL: CRC Press. (2017). doi: 10.1201/b21973
29. Motlogeloa O, Fitchett JM. Assessing the impact of climatic variability on acute respiratory diseases across diverse climatic zones in South Africa. *Sci Total Environ*. (2024) 918:170661. doi: 10.1016/j.scitotenv.2024.170661
30. Zhang Y, Li C, Feng R, Zhu Y, Wu K, Tan X, et al. The short-term effect of ambient temperature on mortality in Wuhan, China: a time-series study using a distributed lag non-linear model. *Int J Environ Res Public Health*. (2016) 13:722. doi: 10.3390/ijerph13070722
31. Pan American Health Organization. *Costa Rica Country Profile*. (2024). Available online at: <https://hia.paho.org/en/countries-22/costa-rica-country-profile> (Accessed May 20, 2024).
32. World Health Organization. *Costa Rica: Country Mortality Details*. (2024). Available online at: <https://platform.who.int/mortality/countries/country-details/MDB/costa-rica> (Accessed May 20, 2024).
33. Zhang Y, Liu N, Zhang S, Li C, Ren Z, Wang Y. *The Lag-Effects of Meteorological Factors and Air Pollutants on Child Respiratory Diseases in Fuzhou, China*. (2024). Available online at: https://www.researchgate.net/publication/362747671_The_lag-effects_of_meteorological_factors_and_air_pollutants_on_child_respiratory_diseases_in_Fuzhou_China
34. Centers for Medicare and Medicaid Services, National Center for Health Statistics. *ICD-10-CM Chapter 10: Diseases of the Respiratory System (J00–J99)*. (2024). Available online at: <https://www.icd10data.com/ICD10CM/Codes/J00-J99> (Accessed April 27, 2024).
35. Verdin A, Funk C, Peterson P, Landsfeld M, Tuholske C, Grace K. Development and validation of the CHIRTS-daily quasi-global high-resolution daily temperature data set. *Sci Data*. (2020) 7:1–18. doi: 10.1038/s41597-020-00643-7
36. Funk c, Peterson M, Landsfeld D, Pedreros J, Verdin S, Shukla G, et al. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci Data*. (2015) 2:150066. doi: 10.1038/sdata.2015.66
37. Global Modeling and Assimilation Office (GMAO). *MERRA-2 tavg3_3d_cld_Nx: 3d,3-Hourly,Time-Averaged,Cloud Diagnostics on a 0.5x0.625 Degree Grid V5.12.4*. (2020). Available online at: https://disc.gsfc.nasa.gov/datasets/M2T3NVCLD_5.12.4/summary (Accessed April 26, 2024).
38. MODIS Atmosphere Science Team. *MODIS/Terra Aerosol Cloud Water Vapor Ozone Daily L3 Global 1Deg CMG. NASA Level 1 and Atmosphere Archive and Distribution System Distributed Active Archive Center*. (2017). Available online at: <https://www.earthdata.nasa.gov/es/data/catalog/laads-mod08-d3-6.1#toc-copy-citation> (Accessed August 20, 2024).
39. Paciorek CJ, Liu Y. Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. *Environ Health Perspect*. (2009) 117:904–9. doi: 10.1289/ehp.0800360
40. He Q, Wang M, Yim SHL. The spatiotemporal relationship between PM 25 and aerosol optical depth in China: influencing factors and implications for satellite PM 25 estimations using MAIAC aerosol optical depth. *Atmos Chem Phys*. (2021) 21:18375–91. doi: 10.5194/acp-21-18375-2021
41. Shepard D. A two-dimensional interpolation function for irregularly spaced data. In: *Proceedings of the 1968 ACM National Conference*. New York, NY (1968). p. 517–24. doi: 10.1145/800186.810616
42. Lu GY, Wong DW. An adaptive inverse-distance weighting spatial interpolation technique. *Comput Geosci*. (2008) 34:1044–55. doi: 10.1016/j.cageo.2007.07.010
43. Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *J Stat Softw*. (2011) 43:1–20. doi: 10.18637/jss.v043.i08
44. Gasparrini A. Modelling lagged associations in environmental time series data: a simulation study. *Epidemiology*. (2016) 27:835–42. doi: 10.1097/EDE.0000000000000533
45. Cameron AC, Trivedi PK. *Regression Analysis of Count Data, 2nd Edn*. Cambridge: Cambridge University Press (2013). doi: 10.1017/CBO9781139013567
46. Hilbe JM. *Negative Binomial Regression, 2nd Edn*. Cambridge: Cambridge University Press (2011). doi: 10.1017/CBO9780511973420
47. Young DS, Roemmele ES, Shi X. Zero-inflated modeling part II: zero-inflated models for complex data structures. *WIREs Comput Stat*. (2022) 14:e1540. doi: 10.1002/wics.1540
48. Freund JE, Perles BC. *Modern Elementary Statistics*. London: Pearson Education (2003).
49. Cover TM, Thomas JA. *Elements of Information Theory, 2nd Edn*. Hoboken, NJ: John Wiley & Sons (2006).
50. Székely GJ, Rizzo ML, Bakirov NK. Measuring and testing dependence by correlation of distances. *Ann Stat*. (2007) 35:2769–94. doi: 10.1214/009053607000000505
51. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E*. (2004) 69:066138. doi: 10.1103/PhysRevE.69.066138
52. Good PI. *Permutation, Parametric, and Bootstrap Tests of Hypotheses, 3rd Edn*. Cham: Springer (2005).
53. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Cham: Springer (2002).
54. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. (1974) 19:716–23. doi: 10.1109/TAC.1974.1100705

55. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. *Psychon Bull Rev.* (2004) 11:192–6. doi: 10.3758/BF03206482
56. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* (1995) 90:773–95. doi: 10.1080/01621459.1995.10476572
57. Schwarz G. Estimating the dimension of a model. *Ann Stat.* (1978) 6:461–4. doi: 10.1214/aos/1176344136
58. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast.* (2006) 22:679–88. doi: 10.1016/j.ijforecast.2006.03.001
59. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim Res.* (2005) 30:79–82. doi: 10.3354/cr030079
60. Armstrong JS. *Long-Range Forecasting: From Crystal Ball to Computer.* Hoboken, NJ: John Wiley & Sons (1985).
61. De Myttenaere A, Golden B, Le Grand B, Rossi F. Mean absolute percentage error for regression models. *Neurocomputing.* (2016) 192:38–48. doi: 10.1016/j.neucom.2015.12.114
62. Montgomery DC, Peck EA, Vining GG. *Introduction to Linear Regression Analysis.* Hoboken, NJ: John Wiley & Sons (2012).
63. Makridakis S. Measures of accuracy: theoretical and practical concerns. *Int J Forecast.* (1995) 9:527–9. doi: 10.1016/0169-2070(93)90079-3
64. Organisation for Economic Co-operation and Development (OECD). *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD Publishing (2008). Available online at: <https://www.oecd.org/sdd/42495745.pdf>
65. Mazziotta M, Pareto A. Everything you always wanted to know about normalization (but were afraid to ask). *Riv Ital Econ Demogr Stat.* (2021) LXXV:67–78. Available online at: <https://www.rieds-journal.org/rieds/article/view/29> (Accessed October 24, 2025).
66. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast.* (2006) 22:679–88. doi: 10.1016/j.ijforecast.2006.03.001
67. Wood SN. *Generalized Additive Models: An Introduction with R, 2nd Edn.* Boca Raton, FL/London: Chapman and Hall/CRC (2017). doi: 10.1201/9781315370279
68. Dobson AJ, Barnett A. *An Introduction to Generalized Linear Models, 3rd Edn.* Boca Raton, FL/London: CRC Press/Chapman & Hall (2008). doi: 10.1201/9780367807849
69. Gasparrini A, Guo Y, Hashizume M, Lavigne E, Zanobetti A, Schwartz J, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet.* (2015) 386:369–75. doi: 10.1016/S0140-6736(14)62114-0
70. Vicedo-Cabrera AM, Scovronick N, Sera F, Royé D, Schneider R, Tobias A, et al. The burden of heat-related mortality attributable to recent human-induced climate change. *Nat Clim Chang.* (2021) 11:492–500. doi: 10.1038/s41558-021-01058-x
71. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna (2025). Available online at: <https://www.R-project.org/>
72. Murillo JH, Rojas Marin JF, Roman SR, Beita Guerrero VH, Arias DS, Ramos AC, et al. Temporal and spatial variations in organic and elemental carbon concentrations in PM10/PM25 in the metropolitan area of Costa Rica, Central America. *Atmos Pollut Res.* (2013) 4:53–63. doi: 10.5094/APR.2013.006
73. Bowen KJ, Ebi KL, Woodward A, McIver L, Tukuitonga C, Schwerdtle PN. Human health and climate change in the Pacific: a review of current knowledge. *Clim Dev.* (2024) 16:119–33. doi: 10.1080/17565529.2023.2185479
74. Tremblay LA, Chariton AA, Li MS, Zhang Y, Horiguchi T, Ellis JI. Monitoring the health of coastal environments in the Pacific region—a review. *Toxics.* (2023) 11:277. doi: 10.3390/toxics11030277
75. Kudo E, Song E, Yockey LJ, Rakib T, Wong PW, Homer RJ, et al. Low ambient humidity impairs barrier function and innate resistance against influenza infection. *Proc Natl Acad Sci.* (2019) 116:10905–10. doi: 10.1073/pnas.1902840116
76. Li H, Liu B, Ma X, Ma Y, Jin S, Fan R, et al. The influence of temperature inversion on the vertical distribution of aerosols. *Remote Sens.* (2022) 14:4428. doi: 10.3390/rs14184428
77. David Alberto GA. Situación actual del dengue como enfermedad reemergente en Costa Rica. *Rev Costarric Salud Pública.* (2018) 27:35–41. Available online at: http://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1409-14292018000100035&lng=en (Accessed March 22, 2026).
78. Herrera J, Rodríguez S, Rojas JF, Cárdenas B. Source apportionment of PM25 in the metropolitan area of Costa Rica using receptor models. *Atmos Clim Sci.* (2013) 3:562–75. doi: 10.4236/acs.2013.34059
79. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L. Time series regression studies in environmental epidemiology. *Int J Epidemiol.* (2013) 42:1187–95. doi: 10.1093/ije/dyt092
80. Rúa G, Suárez C, Chauca P, José adn Ventosilla, Almanza R. Modelado del efecto de la variabilidad climática local sobre la transmisión de dengue en Medellín (Colombia) mediante análisis de series temporales. *Rev Bioméd.* (2013) 33:142–52. doi: 10.7705/biomedica.v33i0.1444
81. Vásquez P, Loria A, Sánchez F, Barboza LA. Variables climáticas como predictores de la incidencia de dengue en Costa Rica: un enfoque de modelo aditivo generalizado y bosques aleatorios. *Rev Mat Teor Apl.* (2019) 27:1–21. doi: 10.15517/rmta.v27i1.39931