OPEN ACCESS

Convergent Sciences REPORT

The Innovation Informatics

Scale-transition error: The unrecognized deterministic bias of data aggregation in evaluating climate-health impacts

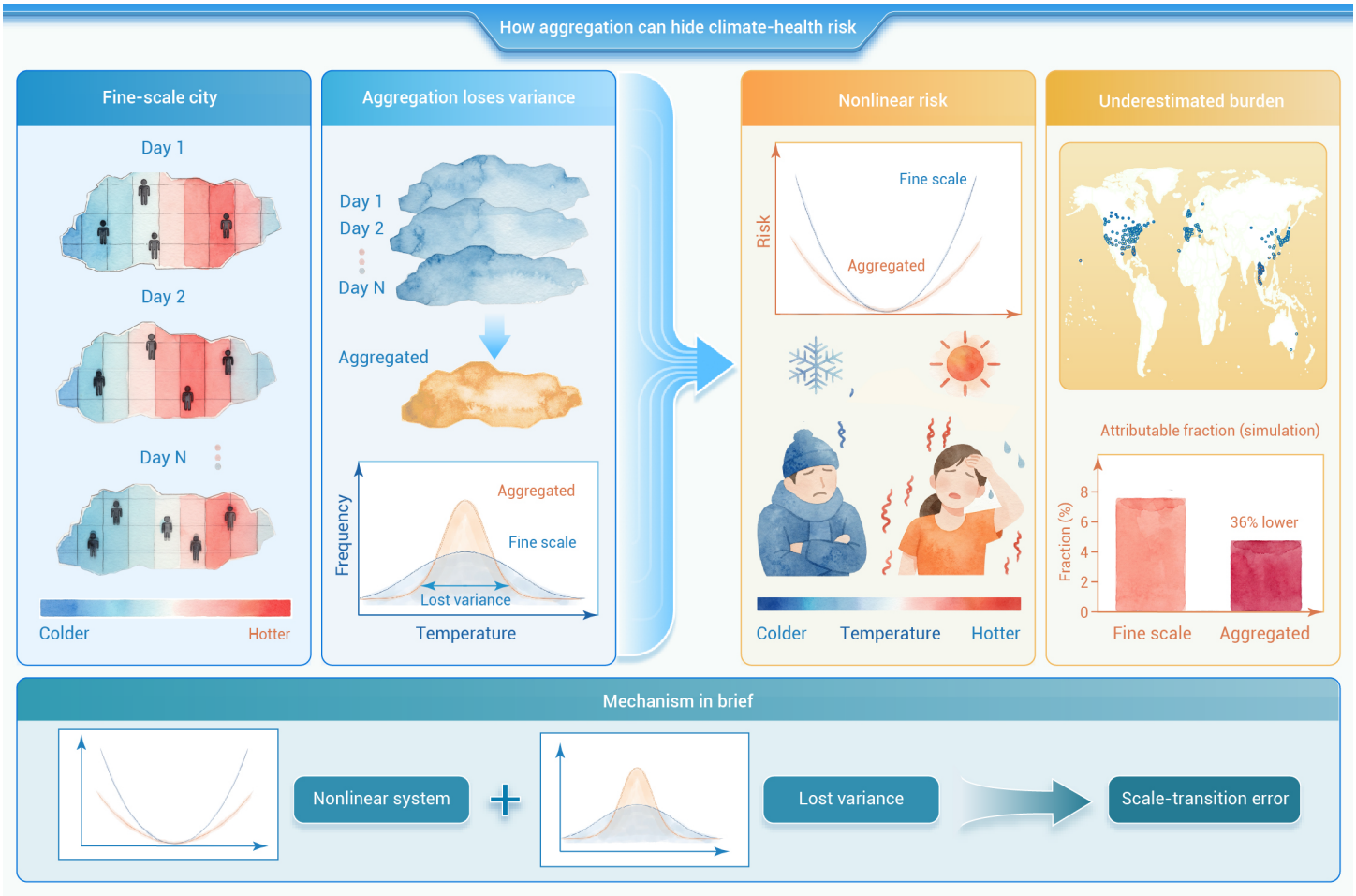
Hengyi Liu,[1,2] Jinting Guo,[1,2] Tao Xue,[1,2,3,4,5,6,*] and Tong Zhu[4,5,6,7]

*Correspondence: txue@hsc.pku.edu.cn

Received: April 2, 2026; Accepted: November 2, 2026; Published Online: November 2, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100074

© 2026 The Author(s). This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Aggregating climate-health data can systematically distort non-linear risk estimates.
- In simulations of 359 cities, coarser scales flattened exposure-response functions.
- The median attributable fraction fell from 7.10% at the micro-scale to 4.53% monthly.
- Lost spatiotemporal variance explained the bias, but real-world validation is still needed.

REPORT Convergent Sciences

OPEN ACCESS

www.the-innovation.org/the-innovation-informatics

Scale-transition error: The unrecognized deterministic bias of data aggregation in evaluating climate-health impacts

Hengyi Liu,[1,2] Jinting Guo,[1,2] Tao Xue,[1,2,3,4,5,6,*] and Tong Zhu[4,5,6,7]

[1]Institute of Reproductive and Child Health, National Health Commission Key Laboratory of Reproductive Health, Peking University, Beijing 100191, China
[2]Department of Epidemiology and Biostatistics, Ministry of Education Key Laboratory of Epidemiology of Major Diseases, Peking University, Beijing 100191, China
[3]National Institute of Health Data Science and Institute of Medical Technology, Peking University Health Science Center, Beijing 100191, China
[4]Ministry of Ecology and Environment Key Laboratory of Atmospheric Exposure and Health Risk Management, Peking University, Beijing 100871, China
[5]Institute of Tibetan Plateau, College of Environmental Sciences and Engineering, Peking University, Beijing 100871, China
[6]Center for Environment and Health, Peking University, Beijing 100871, China
[7]TUWAS-AEEEH, Xizang University, Lhasa, Xizang 850000, China
*Correspondence: txue@hsc.pku.edu.cn

Received: April 2, 2026; Accepted: November 2, 2026; Published Online: November 2, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100074

© 2026 The Author(s). This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Citation: Liu H., Guo J., Xue T., et al. (2026). Scale-transition error: The unrecognized deterministic bias of data aggregation in evaluating climate-health impacts. *The Innovation Informatics* **2**:100074.

Climate change is a fundamental threat to global health. Many climate-health analyses rely on aggregated data (e.g., city-scale time series), which may mask fine-scale spatiotemporal variation. For non-linear systems, this aggregation introduces a non-random bias, which we term scale-transition error (STE). We conducted a mathematical simulation to illustrate the mechanism of STE and characterize its potential consequences under specified data-generating assumptions. Estimates obtained under different aggregated observation scales were compared with those from a simulation-defined micro-scale benchmark. Coarser observation scales generally produced flatter estimated exposure–response functions and lower simulated attributable fraction estimates. At the city-monthly scale, the median simulated attributable fraction was 4.53%, compared with 7.10% under the micro-scale benchmark. Our findings indicate that, under the simulated settings, aggregation across increasingly coarse spatiotemporal scales can attenuate non-linear effect estimates and the corresponding estimates of temperature-attributable mortality. This study is intended to provide a mechanistic demonstration and internal validation of STE. Further empirical studies using observed health and exposure data at multiple resolutions are needed to evaluate the external validity and practical magnitude of STE.

INTRODUCTION

Climate change has emerged as a paramount risk factor for global health. Yet, the precise quantification of these climate-health effects remains highly susceptible to a pervasive, systematic, and largely unrecognized methodological pitfall: the Scale-Transition Error. Indeed, across diverse data-driven disciplines, ranging from econometrics and macroecology to epidemiology, accurately characterizing complex, non-linear relationships is fundamental to decoding underlying system dynamics.[1–2] However, constrained by the resolution limits of observational data, researchers are frequently forced to rely on macro-level aggregated metrics (e.g., regional means or monthly sums) to infer micro-level mechanistic processes.[3–5] While this practice is mathematically permissible for linear systems, it introduces systematic distortion when applied to non-linear contexts. We term this bias as Scale-Transition Error (STE), which is rooted in Jensen's Inequality:[6] the expectation of a non-linear function does not equal the function of the expectation ($E[f(x)] \neq f(E[x])$). Consequently, evaluating responses based on average exposure fails to capture the true average of individual responses, triggering STE that undermines macro-scale inference.[7]

This discrepancy manifests across various scientific domains under different nomenclatures, each capturing a specific facet of this underlying scale-transition phenomenon. In geography and spatial statistics, it is fundamentally recognized as the Modifiable Areal Unit Problem (MAUP),[8] which demonstrates how aggregating high-resolution data into coarser macro-units inherently smooths out local variance and distorts spatial patterns. In epidemiology and sociology, this loss of variance constitutes a mathematically deterministic driver of the Ecological Fallacy, specifically occurring when non-linear, individual-level health risks are inferred from group-level exposure averages.[9–10] Concurrently, theoretical ecology provides a rigorous mechanistic framework through Scale Transition Theory (STT). It explicitly models how this retained or lost fine-scale variance interacts with the non-linear properties of biological responses.[11–12] Despite these disciplinary boundaries, the main point is clear: the artificial smoothing of fine-scale spatiotemporal variance during data aggregation fundamentally distorts the quantification of non-linear processes.

The field of environmental epidemiology, specifically the assessment of temperature-related mortality, provides an ideal context for investigating the STE. This is primarily because a robust foundation of high-quality, multi-country observational studies has already established the biological relationship between thermal stress and health outcomes as profoundly non-linear, universally characterized by steep "U- or J-shaped" Exposure-Response Functions (ERFs).[13–17] Concurrently, ambient temperature inherently possesses immense spatiotemporal variance.[18–19] However, a critical bottleneck in global climate-health assessments is data accessibility. In many developing nations and vast geographical regions, health surveillance and mortality registries are only reported at coarse aggregated scales, such as weekly or monthly intervals.[20–22] While some methodological studies have increasingly recognized the STE-induced by data aggregation and have recently developed unbiased estimation techniques to circumvent this issue,[23–24] the field still lacks a comprehensive, quantitative evaluation of the STE and its underlying deterministic mechanisms across diverse global settings.

To systematically quantify the impact of STE on climate-health effects evaluation, and decode its underlying mechanics, we designed a statistical simulation encompassing 359 cities. First, by systematically varying spatial and temporal aggregation pathways, we empirically demonstrate how the smoothing of spatiotemporal variance flattens the extreme tails of non-linear ERFs, resulting in an asymmetric underestimation of attributable fractions (AF). Second, through a simulation experiment that standardizes the exposure distribution, we isolate the mathematical mechanism driving this bias. Drawing upon STT and a second-order Taylor expansion, we derive a mathematical approximation for this systematic deviation and evaluate it against simulation-derived bias. We show that the magnitude of this underestimation is not a fluke, but rather a predictable outcome driven by the lost spatiotemporal variance and the non-linearity of the health effect.

MATERIALS AND METHODS

Data acquisition and baseline establishment

To construct an empirically parameterized simulation framework with broad global coverage, we extracted city-specific, non-linear temperature-mortality ERFs for 359 cities from previous Multi-Country Multi-City network,[13] designating these as the data-generating reference risk function. For the exposure data, we utilized high-resolution daily temperatures from the ERA5-Land reanalysis grid (0.1° × 0.1° resolution) from 1 January 2020 to 31 December 2024. Corresponding spatial population data were acquired from LandScan and resampled to align with the temperature grid. Crucially, the 0.1° × 0.1° grid cell served as our fundamental spatial unit for generating

simulated mortality counts. For each grid cell-day, we generated a simulated mortality count based on a unified baseline crude mortality rate, the grid-specific population, and the relative risk specified by the corresponding reference ERF at the grid-day temperature. This data generating process yielded a simulation-defined micro-scale dataset with a spatial resolution of 0.1°and a temporal resolution of one day, against which estimates obtained after spatial and temporal aggregation were compared.

Simulation design and spatiotemporal aggregation

A high-resolution, bottom-up simulation framework was constructed with two parallel scenarios. In the Real-world Scenario (i.e., heterogeneous ERFs), we generated simulated reference mortality counts at the micro-grid level using grid-specific temperatures and the city-specific ERFs. In parallel, the Simplified Scenario (i.e., a standardized ERF) applied a universal, percentile-based ERF across all cities to isolate the aggregation-related mechanism. The simulated grid-day data were subsequently aggregated into four macroscopic scales: 1. City-Daily: Spatial aggregation, where daily grid-level temperatures were averaged and mortality counts were summed across the city; 2. Station-Daily: Spatial proxy, where the daily temperature of the most populated grid cell represented the entire city's exposure, and mortality counts were summed across the city; 3. City-Weekly: Spatiotemporal aggregation, where City-Daily temperatures were averaged and mortality counts were summed over each week; 4. City-Monthly: Spatiotemporal aggregation, where City-Daily temperatures were averaged and mortality counts were summed over each month. At each macroscopic aggregation scale, conditional quasi-Poisson regression models were fitted separately to estimate the apparent ERFs, Minimum Mortality Temperatures (MMTs), and heat/cold-related AFs. More details of the simulation framework are available in Supplementary Information S1. The standardized ERF is shown in Figure S1.

To ensure the observed aggregation bias was not an artifact of our specific simulation parameters, we conducted targeted sensitivity analyses (e.g., varying crude death rates, introducing negative binomial overdispersion, and incorporating a simplified lag structure). These checks confirmed that the fundamental signature of STE—the systematic flattening of the risk curve—remains highly robust across varying statistical and epidemiological assumptions (Supplementary Information S4). In addition, we conducted a mathematical stress test by artificially amplifying the within-group variance to extreme levels to evaluate the boundary conditions and stability of our second-order Taylor approximation (Supplementary Information S6). Sensitivity-analysis results and the estimand–estimator decomposition are shown in Figures S5–S6 and Table S1.

Theoretical derivation and validation

To provide a mechanistic explanation for the observed biases, we derived a mathematical formula based on Poisson regression. For clarity, Equations (1)-(3) present the centered equal-baseline special case of the derivation. The simulation additionally allows baseline expected mortality to vary across grid cells through population size. The corresponding weighted form, and its relationship to the arithmetic exposure means used in the aggregated models, are given explicitly in Supplementary S2.

Let y_{it} denote the health outcome, where the subscript i indicates the spatial unit and t the temporal unit:

$$y_{it} \sim Poisson\left(\exp\left(f\left(x_{it}\right)\right)\right) \tag{1}$$

In the imposed micro-level benchmark, i represents a 0.1°×0.1° grid cell and t represents a single day. When data are aggregated to a macroscopic scale (e.g., i becomes a city, t becomes a week or month), the aggregated outcome $Y = \sum_{i,t} y_{it}$ has a true expectation of $\Lambda = \sum_{i,t} \exp(f(x_{it}))$. Conversely, a naive macroscopic model relies on the aggregated mean exposure $\bar{x}$, yielding an expected value of $N\exp(f(\bar{x}))$, where N is the total number of spatiotemporal micro-units within the aggregated macro-unit.

To quantify the discrepancy between the mean log-risk implied by micro-scale benchmark and the naive model, a second-order Taylor expansion is applied to $f(x_{it})$ around the mean $\bar{x}$:

$$f(x_{it}) \approx f(\bar{x}) + f'(\bar{x})\,\delta_{it} + \frac{1}{2} f''(\bar{x})\,\delta_{it}^2 \tag{2}$$

where $\delta_{it} = x_{it} - \bar{x}$. By incorporating standard numerical approximations for the exponential link and logarithmic functions, we mathematically deduce that the true macroscopic risk systematically deviates from the naive model by an additive bias term:

$$\Delta_{bias}(\bar{x}) = \log\left(\frac{\Lambda}{N}\right) - f(\bar{x}) \approx \frac{1}{2}\sigma^2\left[f''(\bar{x}) + \left(f'(\bar{x})\right)^2\right] \tag{3}$$

where σ^2 is defined as the intra-group spatiotemporal variance of the exposure which is not represented by the aggregated mean. Equation (3) provides a second-order local approximation to the aggregation-induced discrepancy on the log-risk scale. The magnitude and direction of this discrepancy depend jointly on the within-group exposure variance and the local slope and curvature of the ERF. The comprehensive, step-by-step mathematical proof of this derivation is detailed in Supplementary Materials S2.

RESULTS

The global landscape of spatiotemporal variance: drivers of STE

The 359 cities included in our simulation are primarily concentrated in North America, Europe, and East/Southeast Asia (Figure 1A). As detailed in Table 1, these cities cover diverse climatic zones, including cold (< 10°C, n = 40), temperate (10-20°C, n = 225), and hot (>20°C, n = 94) regions. In the Real-World Scenario, city-specific ERFs were used as the reference exposure–response relationships for the simulation. The corresponding MMTs shifted across climate zones, with mean values of 18.8°C in cold zones to 27.7°C in hot zones.

Across these diverse geographical and climatic settings, the cities exhibited highly heterogeneous patterns of spatiotemporal temperature variance (Figure 1B and Table 1). Overall, the magnitude of temporal variation ($\sigma_{temporal}$, overall mean 7.0°C) was generally larger than spatial variation ($\sigma_{spatial}$ overall mean 0.9°C). Furthermore, $\sigma_{temporal}$ demonstrated a clear climate dependency, with cold regions experiencing much wider temporal fluctuations (10.1°C) compared to hot regions (3.3°C). Conversely, spatial variation was highly sensitive to city size and morphology. While the overall mean $\sigma_{spatial}$ was 0.9°C, geographically expansive megacities frequently recorded intra-city spatial variances exceeding 2.0°C to 5.0°C, indicating that individual cities possess distinctly different combinations of spatial and temporal exposure variability. Finally, the station mismatch, representing the absolute difference between the most populated grid cell (defined as "station") and the city-wide mean, averaged 0.6°C (SD = 1.1°C) across all cities (Table 1).

Scale-transition error systematically attenuates estimated ERFs and attributable fractions

To quantitatively evaluate the magnitude of STE, we compared the health impact estimates derived from different macroscopic scales against those from the simulation-defined micro-scale benchmark under the Real-World Scenario. Figure 2 illustrates how STE distorts the estimated ERF curvature in four representative cities. At the city-daily scale, the estimated ERFs generally exhibited relatively small deviations from the reference ERFs. Notably, when daily exposures were represented by a single proxy station (station-daily), some of the ERFs exhibited noticeable horizontal shifts, as observed in Beijing and Tokyo. This horizontal displacement is primarily driven by spatial mismatch (i.e., systematic differences between the proxy station's temperature and the true city-wide mean), rather than variance-driven flattening (detailed further in Figure S2). However, as the observational scale coarsened to weekly and monthly intervals, the estimated ERFs became progressively flatter, particularly at the tails, and fine-scale temperature extremes were less distinctly represented.

This dual manifestation of STE directly translated into a systematic underestimation of the simulated AF estimates (Figure 3A). Across the 359 cities, the median AF decreased sequentially from 7.10% (interquartile range [IQR]: 5.26, 9.62) under the micro-scale benchmark to 4.53% [IQR: 2.71, 6.76] at the city-monthly scale. This STE-induced underestimation exhibited a pronounced asymmetry. The cold-related AF showed a substantial reduction (median AF dropping from 6.20% [IQR: 4.36, 8.61] to 3.88% [IQR: 1.98, 6.03]), while the heat shrank from 0.34% [IQR: 0.16, 0.61] to 0.25% [IQR: 0.06, 0.54] at the city-monthly scale. Figure 3B confirms that this STE-induced underestimation is structurally linked to the lost spatiotemporal variance. Across all

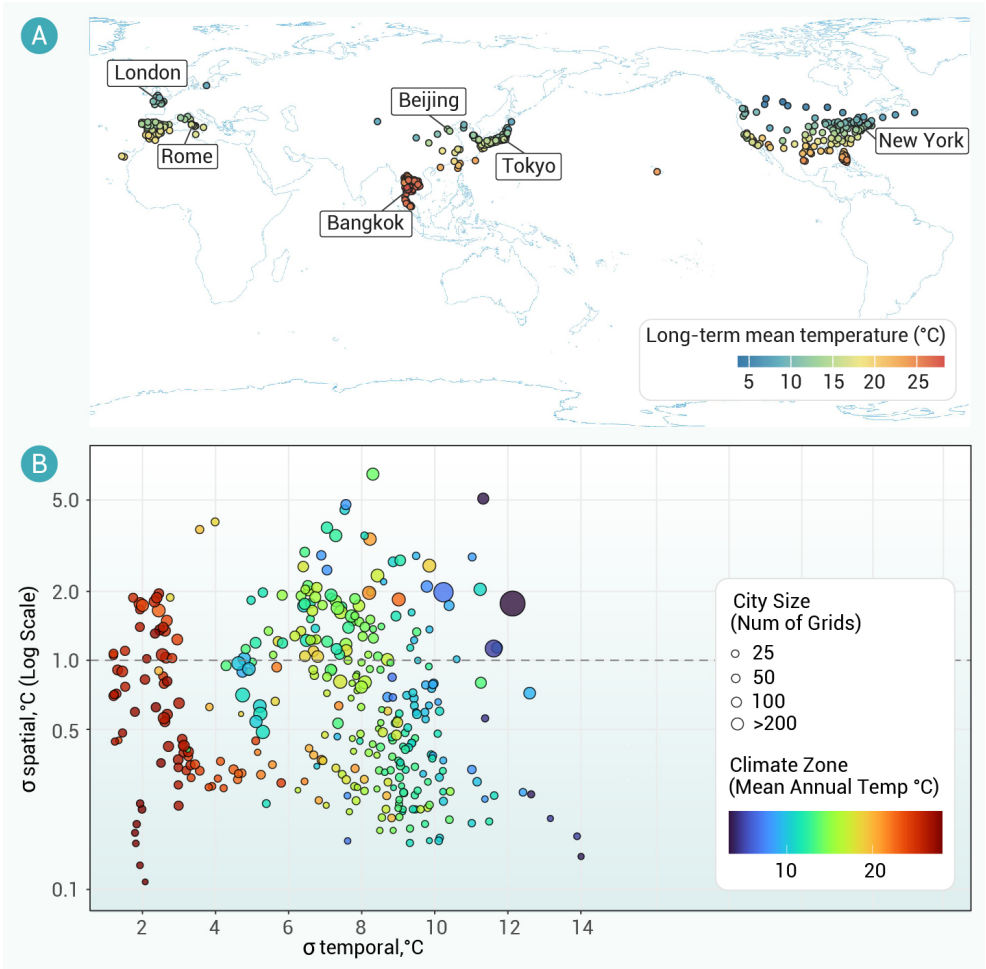


Figure 1. Global distribution and spatiotemporal temperature variance of the studied 359 cities (A) Geographical locations of the 359 cities. (B) The relationship between temporal standard deviation ($\sigma_{temporal}$, x-axis) and spatial standard deviation ($\sigma_{spatial}$, y-axis, shown on a logarithmic scale) for each city. The base map was adapted from the officially approved standard world map (review number GS(2016)2950), Ministry of Natural Resources of China; city markers and labels were added by the authors.

evaluated macroscopic scales, a consistent negative correlation was observed: cities with greater variance loss tended to exhibit larger downward deviations from the micro-scale benchmark.

Simulation-based evaluation of the scale-transition error

In the Simplified Scenario with a standardized ERF, Figure 4A isolates the structural impact of STE: a systematic flattening of the pooled exposure-response curves across increasingly coarser observational scales. At the extreme heat end (e.g., 99th percentile), the true RR of 1.43 was attenuated to 1.31 under the city-weekly scale and further reduced to 1.21 under the city-monthly scale. Similar STE-induced attenuation occurred at the cold extreme (1st percentile), with the true RR dropping from 1.73 to 1.51 at the city-monthly scale. The absolute AF error increased progressively across the City-Daily, City-Weekly, and City-Monthly aggregation scales (Figure 4B).

To examine the role of exposure variance loss, cities were grouped according to the magnitude of lost spatiotemporal variance (σ^2) under the city-monthly scale (Figure 4C). Groups with greater variance loss generally exhibited more pronounced attenuation of the estimated ERFs. In the group with the highest variance loss, the estimated RR at the extreme heat percentile decreased to 1.21. Conversely, groups with the smaller variance loss retained ERF shapes that were closer to the reference curve.

Figure 4D provides an oracle-based evaluation of how the second-order approximation propagates to AF underestimation under the Simplified Scenario. For each city, Equation (3) was evaluated using the known reference-ERF derivatives and lost variance, and the resulting attenuated curve was passed through the same AF calculation used in the simulation. The x- and y-axes show the group means of the predicted and simulation-observed percentage underestimation, respectively, across 10 lost-variance groups. The two quantities were strongly correlated ($R^2 = 0.91$), indicating that the approximation captured the variance-related ordering, although the points lay below the 1:1 line (linear fit: $y = 1.41x - 15.2$ percentage points), indicating imperfect absolute calibration. This does not conflict with the underprediction at extreme variances in Figure S8: Figure 4D evaluates a downstream AF measure that additionally incorporates the log-risk-to-RR transformation, mortality weighting, truncation of protective contributions, and monthly spline estimation, whereas Figure S8 directly compares the exact and second-order discrepancies on the local log-risk scale. Additional agreement diagnostics are shown in Figure S7.

DISCUSSION

Using a high-resolution, bottom-up simulation parameterized framework, we examined how spatial and temporal aggregation affected the estimation of non-linear temperature–mortality relationships and attributable fractions. Relative to the simulation-defined micro-scale benchmark, increasingly coarse observation scales generally produced flatter estimated ERFs and lower attributable fraction estimates. The second-order approximation derived from Scale Transition Theory and the Poisson model reproduced the overall pattern of the simulation-observed discrepancies, supporting the internal consistency of the proposed variance-related mechanism. Consequently, the resulting underestimation of health risks is a deterministic mathematical penalty rather than a random statistical artifact.

While diverse disciplines have long recognized the risk of data aggregation through concepts like MAUP and the Ecological Fallacy, these traditional frameworks have largely operated in methodological silos.[8–9,25–26] MAUP primarily emphasizes the empirical instability of statistical patterns driven by arbitrary spatial zoning, whereas the Ecological Fallacy focuses on the inferential hazard of deducing micro-level risks from macro-level group averages.[27] Despite capturing different phenomenological facets of data aggregation, both concepts merely serve as qualitative cautions; they highlight the potential for bias but lack a unified mathematical mechanism to predict its exact magnitude. By formalizing this multidimensional challenge as STE, we bridge these fragmented observations using the mathematical mechanics of STT.[11–12] The STE concept advances the scientific discourse by redefining a vague methodological vulnerability into a quantifiable variance penalty. It explicitly demonstrates that the distortion across any spatial or temporal scale is deterministically governed by the interaction between the unobserved lost variance and the non-linearity of the exposure-response function.[23,28–29]Consequently, adopting the STE framework not only provides a universal diagnostic metric for evaluating macroscopic inferences but also establishes the exact mathematical foundation required for prospective bias correction across all data-driven sciences.

When applied to the highly non-linear temperature-mortality ERF, STE manifests as profound, multidimensional epidemiological distortions. A primary consequence of this progressive variance loss is a severe flattening of the ERF tails. Our theoretical framework precisely decodes the asymmetry in this risk underestimation. Biological responses to extreme heat are typically characterized by substantially steeper exponential trajectories (i.e., significantly larger curvature f'' and squared slope $(f')^2$ than those for cold stress.[30–32] While heat effects often manifest as acute mortality spikes over a few days, cold effects typically exhibit a prolonged "long tail" covering a

Table 1. Descriptive characteristics of the sampled cities stratified by climate zones. Values are presented as median [interquartile range] for city size and mean (standard deviation, SD) for the rest symmetrically-distributed variables.

	Cold (< 10°C)	Temperate (10-20°C)	Hot (> 20°C)	Overall
Baseline				
Number of cities	40	225	94	359
City size (num. of grids)	28.0 [14.5, 95.5]	35.0 [19.0, 98.0]	62.5 [37.5, 94.8]	42.0 [21.0, 96.5]
Minimum Mortality Temp. (°C)	18.8 (3.2)	23.4 (2.7)	27.7 (1.5)	24.0 (3.6)
Temp. by day and grid ($T_{grid,day}$; °C)				
Original scale	8.2 (1.9)	14.3 (2.6)	25.2 (2.3)	16.5 (6.0)
Spatial variation [$\sigma_{spatial}$ = SD($T_{\cdot,day}$)]	0.9 (0.9)	1.0 (0.9)	0.8 (0.7)	0.9 (0.9)
Temporal variation [$\sigma_{temporal}$ = SD($T_{grid,\cdot}$)]	10.1 (2.2)	8.0 (1.6)	3.3 (1.9)	7.0 (2.9)
Station mismatch*	0.6 (1.2)	0.6 (1.2)	0.6 (0.9)	0.6 (1.1)

*Station mismatch: for each day and city, we calculate the difference between within-city average of grid-day scale temperature and the in-situ observation at most-populated grid within a city.

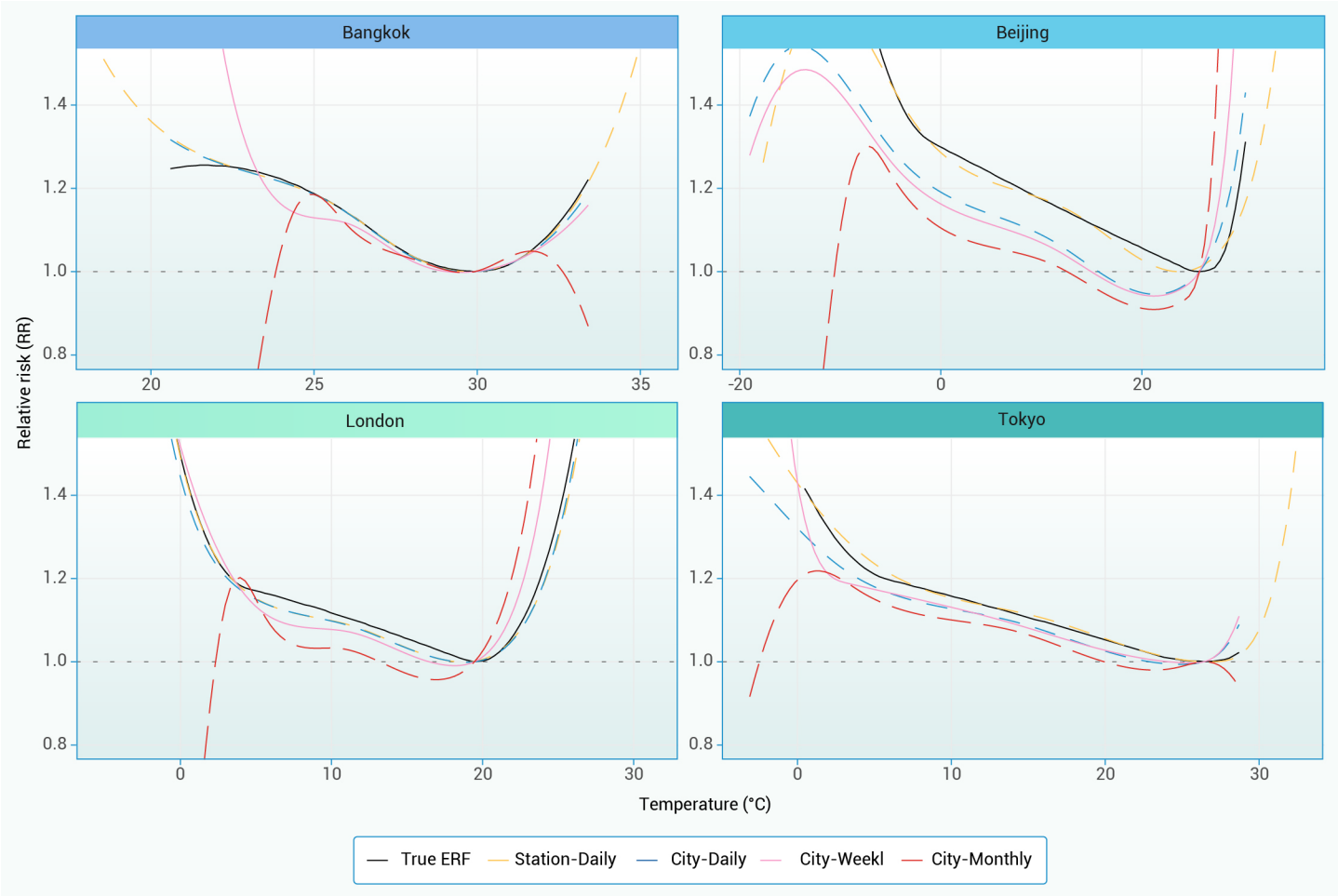


Figure 2. Estimated exposure-response functions across different observation scales in four representative cities The aggregated models represent ERFs estimated using exposure data at various macroscopic observation scales: a single proxy station at the daily level (Station-Daily), and city-wide averages at daily (City-Daily), weekly (City-Weekly), and monthly (City-Monthly) intervals. All exposure-response functions are centered at their respective reference Minimum Mortality Temperature.

substantial portion of the year. Therefore, the generalization that cold-related burdens suffer a larger absolute underestimation relies on specific conditions: the underlying ERF must be U- or J-shaped, and the population must experience a substantially higher frequency of non-optimal cold days relative to non-optimal heat days. Consequently, the continuous smoothing of prolonged cold temperatures accumulates into a significantly larger absolute underestimation of the cold AF. It should nevertheless be emphasized that the finest resolution considered in this study—a 0.1° grid-day—still aggregates exposure across multiple individuals. Further disaggregation could reveal additional within-grid exposure variation. In particular, the baseline AF of 7.10% represents the median estimate obtained under the simulation-defined micro-scale benchmark and should not be interpreted as either an estimate or a conservative lower bound of the true temperature-attributable mortality burden in real-world populations. Beyond the flattening effect of STE, the simulation also highlights a spatial mismatch error when employing a single proxy station. Larger differences between the proxy temperature and the city-

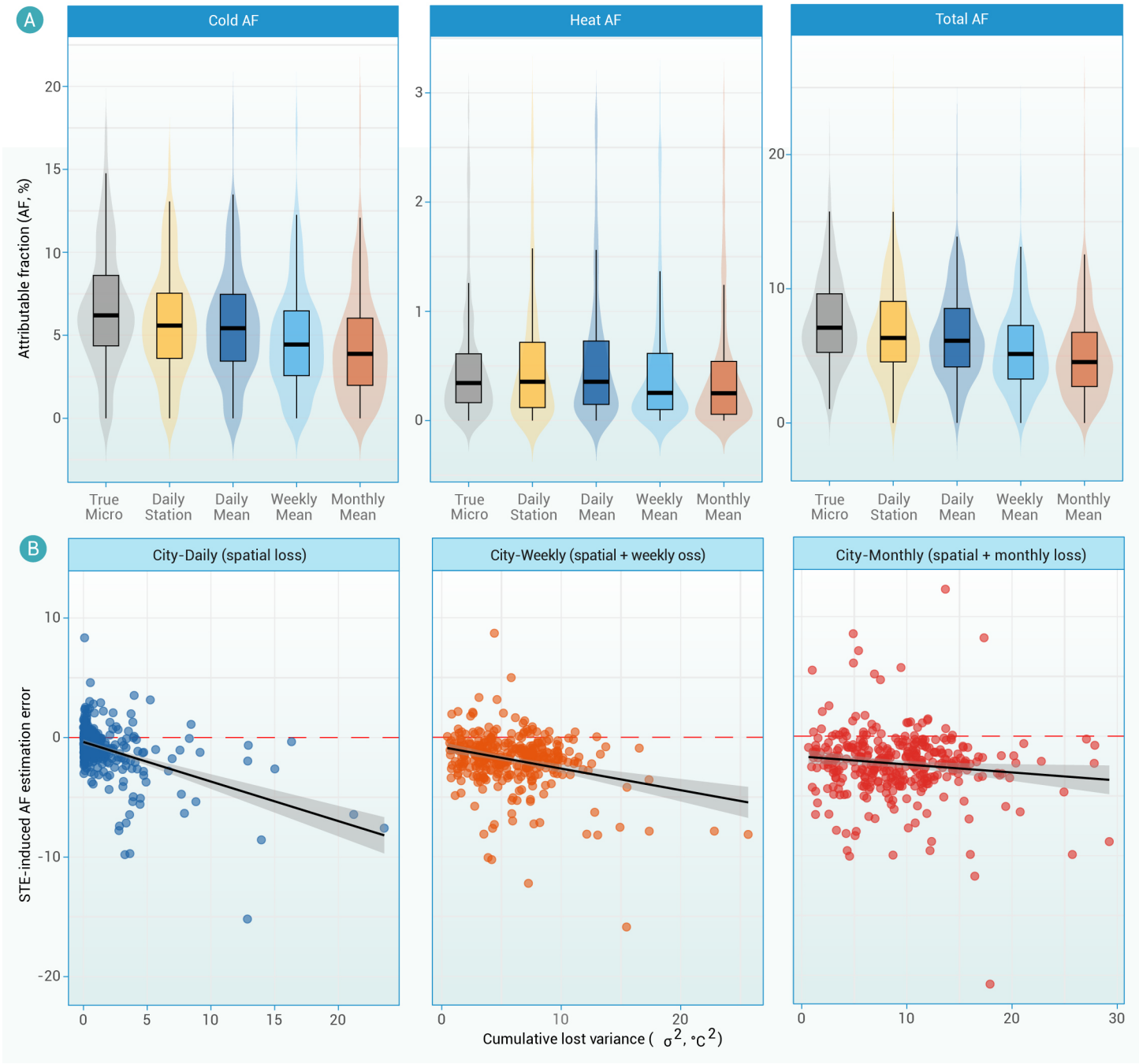


Figure 3. Estimated Attributable Fractions (AF) and estimation errors across different observation scales (A) Distributions of heat-related, cold-related, and total attributable fractions estimated across n = 359 global cities under Real-World Scenario. The boxplots represent the median (solid line) and the interquartile range (IQR, 25th to 75th percentiles). Numbers above the plots denote the exact median [IQR] values. (B) Scatter plots illustrating the relationship between the cumulative lost spatiotemporal variance (σ^2, x-axis) and the STE-induced AF estimation error. Each point represents a single city. Solid black lines indicate linear fits with 95% confidence intervals.

wide mean were associated with more pronounced horizontal shifts in the estimated ERF and corresponding changes in the MMT (Figure S2).[33–34]

Importantly, our results demonstrate that mainstream city-daily level studies (the standard design in primary environmental epidemiology) remain largely unbiased. However, the implications of STE reveal a major methodological weakness in a growing body of macroscopic climate-health research. Recently, there has been an increase in macroscopic assessment—frequently secondary analyses utilizing Global Burden of Disease (GBD) estimates—that link national-level, annual mortality statistics against highly aggregated climate metrics.[35–38] Our theoretical framework dictates that the magnitude of STE is fundamentally driven by the unobserved loss of spatiotemporal variance. At a national-annual scale, the amount of smoothed temperature fluctuation is enormous. Consequently, applying naive non-linear models to such heavily aggregated datasets inevitably leads to massive STE. Yet, these macroscopic studies rarely examine whether this level of aggregation is acceptable, risking profound mischaracterizations of true biological exposure-response relationships. Beyond these flawed methodological choices, STE also intersects with data accessibility constraints to global health inequity. Health surveillance registries in many Low- and Middle-Income Countries frequently provide only coarse weekly or monthly mortality statistics, even though these regions are often disproportionately vulnerable to extreme heat.[19,39–40] If researchers continue to apply naive non-linear models to these smoothed datasets, our findings indicate that the impacts of climate extremes in these regions will be systematically underestimated. This reliance on aggregated mean values effectively exacerbates global health inequity by masking the true severity of thermal stress in

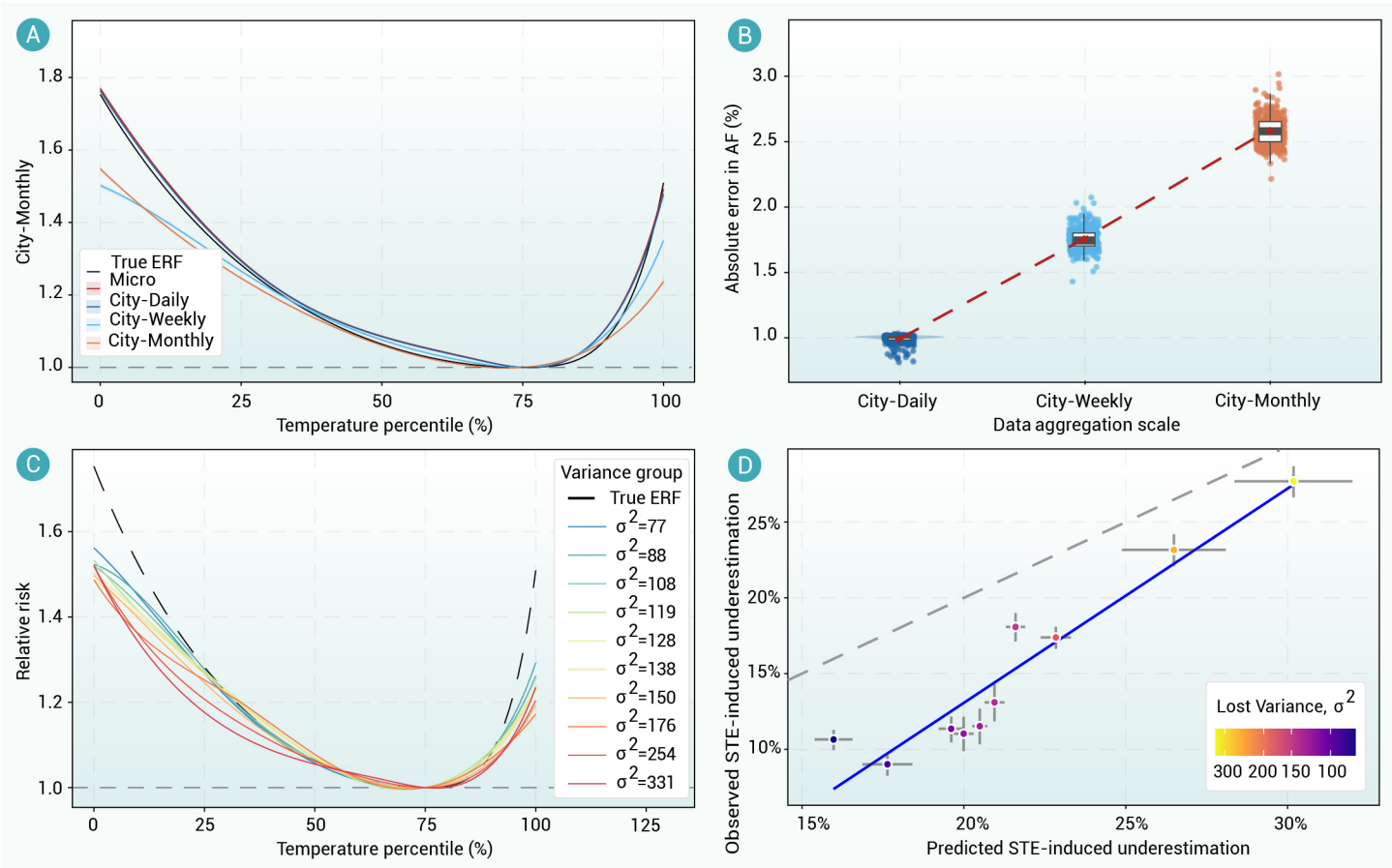


Figure 4. Mechanistic validation of aggregation-induced attenuation under the standardized exposure scenario (Simplified Scenario) (A) Pooled ERFs across all 359 cities. (B) Absolute error in AF across the City-Daily, City-Weekly, and City-Monthly aggregation scales. (C) Stratified ERFs for City-Monthly models; cities were divided into 10 groups according to the magnitude of spatiotemporal variance loss (σ^2). (D) Association between approximation-predicted and simulation-observed AF underestimation; each point represents the mean within one of the 10 lost-variance groups.

populations least equipped to adapt.[41–42]

Several limitations of the theoretical derivation and simulation framework should be considered when interpreting the findings. First, the present study is a simulation-based mechanistic demonstration rather than an empirical validation using observed health data. The findings should therefore be interpreted as an aid to understanding the concept, potential mechanisms, and possible consequences of STE under the specified assumptions, rather than as epidemiological evidence of the aggregation bias in real-world populations. Further validation using observed health and exposure data at multiple resolutions is needed to establish the real-world relevance of STE. Second, mathematical derivation depends on both a second-order Taylor approximation and assumptions associated with the Poisson framework. The Taylor expansion omits higher-order terms and therefore provides only a local approximation whose accuracy may decline when the ERF is highly nonlinear or within-group exposure variance is large. In addition, the derivation relies on the Poisson mean structure and aggregation property, and its applicability to outcomes with substantial overdispersion or complex spatiotemporal dependence requires further evaluation. Third, the simulation simplified several features of real-world epidemiological data. In particular, it does not fully represent complex spatial and temporal dependence in mortality and exposure, and it adopts a simplified demographic structure without explicitly accounting for age-specific baseline mortality, population composition, or vulnerability across population subgroups. The data-generating process also assumes time-invariant city-specific ERFs and omits complex distributed lag structures. Although the sensitivity analyses showed that the general attenuation pattern remained qualitatively consistent across the alternative assumptions evaluated, these analyses cannot reproduce the full complexity of real-world epidemiological processes. Finally, a fully developed and systematic correction framework for macroscopic risk assessment is still lacking. Although a proof-of-principle post-hoc adjustment was explored in Supplementary Materials S3 to examine its preliminary feasibility, this analysis represents a conceptual demonstration rather than a fully developed correction framework. The procedure has not yet been validated using observed multiresolution health data, and its implementation in real-world aggregated datasets may be constrained by the availability and uncertainty of the within-group exposure variance and ERF shape required for correction. We consider the derivation in this study as the mathematical backbone, setting the stage for future methodological work to build comprehensive correction tools. The corresponding ERF and AF comparisons are shown in Figures S3–S4.

Moving forward, the epidemiological and climate science communities must transition from merely acknowledging the STE to actively correcting it. Fortunately, pioneering methodological innovations have begun to tackle this challenge with remarkable success. For example, recently a study developed a strategy to preserve daily non-linearities by applying polynomial transformations to high-resolution exposures prior to temporal summation.[43] Concurrently, other outstanding frameworks utilize the superposition property of the Poisson distribution to construct a joint likelihood function, enabling unbiased inferences directly from aggregated outcomes.[24] Though these highly sophisticated solutions represent major leaps forward in environmental epidemiology, they naturally face inherent operational constraints. Pre-transforming exposures often require rigid parametric assumptions about the functional form, while inference-level likelihood optimizations impose substantial computational burdens. The latter approach frequently encounters convergence difficulties and the risk of identifying local optima in non-concave functions. By mathematically adjusting naive macroscopic estimates using the lost spatiotemporal variance, researchers could achieve robust bias correction with significantly lower computational thresholds and broader model flexibility. Beyond methodological advancements, addressing STE has profound practical policy implications. Currently, national and international health policies, including climate adaptation funding and healthcare

resource allocation, heavily rely on macroscopic burden estimates. By underestimating mortality burdens, especially in regions relying on coarse data, STE may lead policymakers to under-prioritize climate-health interventions. Developing practical methods to address variance-driven bias is therefore important for producing more accurate and equitable policy-relevant estimates.

CONCLUSION

The aggregation of spatiotemporal data in non-linear modeling does not merely introduce random statistical noise; it can generate a systematic Scale Transition Error. Scale Transition Error is an inherent mathematical consequence, whereas the simulation in this study provides a controlled setting in which this mechanism can be isolated and illustrated. This Scale Transition Error is fundamentally driven by the unobserved intra-group variance lost during data aggregation and the intrinsic curvature of the true exposure-response function. To ensure the reliability of scientific inference and safeguard equity in risk assessment, future analytical frameworks across data-driven disciplines must transition beyond mean-centric models. Further empirical studies using observed health and exposure data at multiple resolutions are needed to validate the practical magnitude of STE in real-world settings.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used OpenAI Codex to assist with language editing, document formatting, and vector reconstruction of author-provided figures. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

1. Goldberger A.L. (1996). Non-linear dynamics for clinicians: Chaos theory, fractals, and complexity at the bedside. *Lancet* **347**:1312–1314. DOI:10.1016/S0140-6736(96)90948-4
2. Froese R., Andrino A., Giudice R., et al. (2023). Describing complex interactions of social-ecological systems for tipping point assessments: An analytical framework. *Front. Clim.* **5**:1145942. DOI:10.3389/fclim.2023.1145942
3. Wakefield J. (2008). Ecologic studies revisited. *Annu. Rev. Public Health* **29**:75–90. DOI:10.1146/annurev.publhealth.29.020907.090821
4. Bhaskaran K., Gasparrini A., Hajat S., et al. (2013). Time series regression studies in environmental epidemiology. *Int. J. Epidemiol.* **42**:1187–1195. DOI:10.1093/ije/dyt092
5. Stehle S. (2022). Temporal aggregation bias and gerrymandering urban time series. *Geoinformatica* **26**:233–252. DOI:10.1007/s10707-021-00452-z
6. Denny M. (2017). The fallacy of the average: On the ubiquity, utility and continuing novelty of jensen's inequality. *J. Exp. Biol.* **220**:139–146. DOI:10.1242/jeb.140368
7. Ruel J.J. and Ayres M.P. (1999). Jensen's inequality predicts effects of environmental variation. *Trends Ecol. Evol.* **14**:361–366. DOI:10.1016/S0169-5347(99)01664-X
8. Jelinski D.E. and Wu J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecol.* **11**:129–140. DOI:10.1007/BF02447512
9. Greenland S. and Morgenstern H. (1989). Ecological bias, confounding, and effect modification. *Int. J. Epidemiol.* **18**:269–274. DOI:10.1093/ije/18.1.269
10. Greenland S. (2001). Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int. J. Epidemiol.* **30**:1343–1350. DOI:10.1093/ije/30.6.1343
11. Chesson P. (2012). Scale transition theory: Its aims, motivations and predictions. *Ecol. Complexity* **10**:52–68. DOI:10.1016/j.ecocom.2011.11.002
12. Melbourne B.A. and Chesson P. (2006). The scale transition: Scaling up population dynamics with field data. *Ecology* **87**:1478–1488. DOI:10.1890/0012-9658(2006)87[1478:TSTSUP]2.0.CO;2
13. Gasparrini A., Guo Y., Hashizume M., et al. (2015). Mortality risk attributable to high and low ambient temperature: A multicountry observational study. *Lancet* **386**:369–375. DOI:10.1016/S0140-6736(14)62114-0
14. Chen R., Yin P., Wang L., et al. (2018). Association between ambient temperature and mortality risk and burden: Time series study in 272 main Chinese cities. *BMJ* **363**:k4306. DOI:10.1136/bmj.k4306
15. Kephart J.L., Sánchez B.N., Moore J., et al. (2022). City-level impact of extreme temperatures and mortality in Latin America. *Nat. Med.* **28**:1700–1705. DOI:10.1038/s41591-022-01872-6
16. Burkart K.G., Brauer M., Aravkin A.Y., et al. (2021). Estimating the cause-specific relative risks of non-optimal temperature on daily mortality: A two-part modelling approach applied to the Global Burden of Disease Study. *Lancet* **398**:685–697. DOI:10.1016/S0140-6736(21)01700-1
17. Ribeiro-Silva R.D.C.S., Luz L.P., Sacramento L.S.S., et al. (2026). Ambient temperature and child nutritional status of more than 6 million children in Brazil: A cohort study. *Lancet Planet. Health* **10**:101428. DOI:10.1016/j.lanplh.2026.101428
18. Heaviside C., Macintyre H. and Vardoulakis S. (2017). The urban heat island: Implications for health in a changing environment. *Curr. Environ. Health Rep.* **4**:296–305. DOI:10.1007/s40572-017-0150-3
19. Bathiany S., Dakos V., Scheffer M., et al. (2018). Climate models predict increasing temperature variability in poor countries. *Sci. Adv.* **4**:eaar5809. DOI:10.1126/sciadv.aar5809
20. Li P., Wu J., Xue T., et al. (2025). Associations between monthly temperature anomalies and mortality among children younger than five years in low- and middle-income countries. *Commun. Med.* **5**:365. DOI:10.1038/s43856-025-01101-4
21. Ai S., Lu H., Liu H., et al. (2024). All-cause mortality attributable to long-term changes in mean temperature and diurnal temperature variation in China: A nationwide quasi-experimental study. *Environ. Res. Lett.* **19**:014002. DOI:10.1088/1748-9326/ad0d3d
22. Liu H., Tong M., Guo F., et al. (2022). Deaths attributable to anomalous temperature: A generalizable metric for the health impact of global warming. *Environ. Int.* **169**:107520. DOI:10.1016/j.envint.2022.107520
23. Rhomberg L.R., Chandalia J.K., Long C.M., et al. (2011). Measurement error in environmental epidemiology and the shape of exposure-response curves. *Crit. Rev. Toxicol.* **41**:651–671. DOI:10.3109/10408444.2011.563420
24. Basagaña X. and Ballester J. (2024). Unbiased temperature-related mortality estimates using weekly and monthly health data: A new method for environmental epidemiology and climate impact studies. *Lancet Planet. Health* **8**:e766–e777. DOI:10.1016/S2542-5196(24)00212-2
25. Wakefield J. (2007). Disease mapping and spatial regression with count data. *Biostatistics* **8**:158–183. DOI:10.1093/biostatistics/kxl008
26. Morgenstern H. (1995). Ecologic studies in epidemiology: Concepts, principles, and methods. *Annu. Rev. Public Health* **16**:61–81. DOI:10.1146/annurev.pu.16.050195.000425
27. Greenland S. (1992). Divergent biases in ecologic and individual - level studies. *Stat. Med.* **11**:1209–1223. DOI:10.1002/sim.4780110907
28. Goldman G.T., Mulholland J.A., Russell A.G., et al. (2011). Impact of exposure measurement error in air pollution epidemiology: Effect of error type in time-series studies. *Environ. Health* **10**:61. DOI:10.1186/1476-069X-10-61
29. Sarnat S.E., Sarnat J.A., Mulholland J., et al. (2013). Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. *J. Expo. Sci. Environ. Epidemiol.* **23**:593–605. DOI:10.1038/jes.2013.41
30. Analitis A., Katsouyanni K., Biggeri A., et al. (2008). Effects of cold weather on mortality: Results from 15 european cities within the PHEWE project. *Am. J. Epidemiol.* **168**:1397–1408. DOI:10.1093/aje/kwn266
31. Lakhoo D.P., Blake H.A., Chersich M.F., et al. (2022). The effect of high and low ambient temperature on infant health: A systematic review. *Int. J. Environ. Res. Public Health* **19**:9109. DOI:10.3390/ijerph19159109
32. Anderson B.G. and Bell M.L. (2009). Weather-related mortality: How heat, cold, and heat waves affect mortality in the United States. *Epidemiology* **20**:205–213. DOI:10.1097/EDE.0b013e318190ee08
33. Brauer M., Hoek G., Van Vliet P., et al. (2003). Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. *Epidemiology* **14**:228–239. DOI:10.1097/01.EDE.0000041910.49046.9B
34. Gryparis A., Paciorek C.J., Zeka A., et al. (2009). Measurement error caused by spatial misalignment in environmental epidemiology. *Biostatistics* **10**:258–274. DOI:10.1093/biostatistics/kxn033
35. Zhang P., Zhou C., Zhao K., et al. (2023). Associations of air pollution and greenness with global burden of breast cancer: An ecological study. *Environ. Sci. Pollut. Res.* **30**:103921–103931. DOI:10.1007/s11356-023-29579-2
36. Qu C., Chen Y., Liu C., et al. (2024). Burden of stroke attributable to nonoptimal temperature in 204 countries and territories: A population-based study, 1990–2019. *Neurology* **102**:e209299. DOI:10.1212/WNL.0000000000209299
37. Zheng J., Yao L., Lei K., et al. (2025). Economic burden attributable to high BMI-caused cancers: A global level analysis between 2002 and 2021. *BMC Med.* **23**:297. DOI:10.1186/s12916-025-04109-8
38. Xu Q., Zhou Q., Chen J., et al. (2023). The incidence of asthma attributable to temperature variability: An ecological study based on 1990–2019 GBD data. *Sci. Total Environ.* **904**:166726. DOI:10.1016/j.scitotenv.2023.166726
39. Berry P., Enright P.M., Shumake-Guillemot J., et al. (2018). Assessing health vulnerabilities and adaptation to climate change: A review of international progress. *Int. J. Environ. Res. Public Health* **15**:2626. DOI:10.3390/ijerph15122626
40. Leal Filho W., Ternova L., Fayyaz M.M., et al. (2022). An analysis of climate change and health hazards: Results from an international study. *Int. J. Clim. Change Strateg. Manag.* **14**:375–398. DOI:10.1108/IJCCSM-08-2021-0090
41. N S.V.S.C., Gunther S.H., Kjellstrom T., et al. (2025). Advancing heat wave definitions: A policy review towards prioritizing health impacts of extreme heat. *Environ. Res. Lett.*

20:033004. DOI:10.1088/1748-9326/adb5a0
42. Hoffman J.S., Shandas V. and Pendleton N. (2020). The effects of historical housing policies on resident exposure to intra-urban heat: A study of 108 US urban areas. *Climate* 8:12. DOI:10.3390/cli8010012
43. Callahan C.W., Trok J., Wilson A.J., et al. (2026). Increasing risk of mass human heat mortality if historical weather patterns recur. *Nat. Clim. Chang.* 16:26–32. DOI:10.1038/s41558-025-02480-1

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant 42422507). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Hengyi Liu: Writing – original draft; Formal analysis/Data analysis; **Jinting Guo**: Data curation/Data support; **Tao Xue**: Writing – review & editing; Supervision; **Tong Zhu**: Supervision. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ETHICAL STATEMENT AND PATIENT CONSENT

This simulation study used no individual-level human participant data; ethical approval and informed consent were not applicable.

DATA AND CODE AVAILABILITY

The simulation code and datasets are available at https://github.com/darcklaw/STE_code.

SUPPLEMENTAL INFORMATION

It can be found online at https://doi.org/10.59717/j.xinn-inform.2026.100074