2026 Collaborative Technologies and Data Science for Smart Cities Applications (CODASSCA) | 979-8-3315-5221-3/26/$31.00 ©2026 IEEE | DOI: 10.1109/CODASSCA69992.2026.00035

# Tracking Distributional Shifts in Air-Pollution Exposure Across Time and Space

Thomas A. Weber
*Chair of Operations, Economics and Strategy*
*École Polytechnique Fédérale de Lausanne*
Lausanne, Switzerland
thomas.weber@epfl.ch

***Abstract*—Air pollution is usually observed through sparse station networks and model-supported concentration fields, whereas public policy is concerned with population exposure. We develop a methodological approach for monitoring pollutant-specific distributional shifts when the available data are rich in time in one setting and rich in space in another. The central statistic is a rank-based regime-shift index (RSI), defined relative to a baseline distribution and reported as a relative log-RSI anchored at a specified date. We apply the method in two complementary designs. In Zurich, daily exposure signals combine station measurements with populations across 216 statistical zones, and distributional shifts are tracked using 365-day rolling windows. In Switzerland, annual national pollutant maps are combined with gridded population data to compare population-weighted spatial distributions of mapped pollutant values with unweighted distributions over populated cells. The Zurich monitor shows long-run downward movements for $NO_2$, $PM_{10}$, and $PM_{2.5}$, while $O_3$ follows a nonmonotone path; population weighting matters most for $NO_2$. In the Swiss design, population weighting materially changes both exposure levels and measured distributional shifts, especially for historical $NO_2$ and $PM_{10}$ regimes. Since 2020, the population-weighted Swiss spatial relative log-RSI has declined more strongly than the unweighted populated-cell counterpart for all four pollutants, a pattern consistent with relatively greater improvement where population is concentrated. The method provides a common rank-based basis for monitoring population-relevant exposure across complementary temporal and spatial data designs without having to depend on a unified parametric model.**

***Keywords*—Air pollution, population exposure, population weighting, regime-shift index, relative log-RSI, spatial monitoring, temporal monitoring, Switzerland, Zurich**

## I. Introduction

Air pollution represents the largest environmental risk to public health worldwide; specifically, the World Health Organization estimates that exposure to ambient and household $PM_{2.5}$ air pollution was associated with close to 7 million premature deaths in 2019 [1]. Air-quality regulation and health assessment focus on the exposure experienced by populations. This issue becomes important when pollution is observed on a spatial support that is imperfectly aligned with residential location. A sparse station network can be informative about environmental conditions and still provide an incomplete measure of exposure unless the population distribution enters the construction of the monitoring signal. This concern is consistent with current air-quality guidelines and with the European and Swiss literature on long-term exposure [2]–[7].

A second issue is statistical, as environmental improvement or deterioration need not be well described by a change in the mean alone. The upper tail, the dispersion, and the general displacement of the exposure distribution may all fluctuate, and these changes need not move together. A useful monitor should therefore capture distributional movement without imposing more parametric structure than the data warrant. Spatial misalignment and metric choice can materially affect epidemiologic and policy interpretation [8], [9]. Neighborhood-scale health assessments also show substantial intra-urban heterogeneity in population-relevant air-pollution burdens when concentration, disease-rate, and population data are resolved at fine spatial scales [10].

A third issue concerns data architecture, since data availability may differ significantly across settings. Zurich provides long, high-frequency station series with sparse spatial support. Switzerland produces annual national pollutant maps with rich spatial detail and low temporal frequency. The mapped values are pollutant-specific: the $NO_2$, $PM_{10}$, and $PM_{2.5}$ fields report annual mean concentrations, whereas the $O_3$ fields give an annual ozone peak statistic. The two settings form complementary empirical designs under a common monitoring principle. This treatment accords with the WHO overview of exposure-assessment methods and with the complementary spatial and temporal capabilities documented for current urban monitoring systems [1], [11]. It is also consistent with air-pollution mapping practice ranging from city-scale high-resolution modeling to broad spatial exposure models [12], [13]. The same structure supports multi-rate urban monitoring, with station observations, spatial fields, and population data entering at their native update frequencies.

Proceeding accordingly, for each pollutant we construct a population-relevant ambient-exposure signal from a pollution field and a system of population weights. The RSI is developed theoretically in [14] and applied to climate monitoring in [15]. The present results extend its empirical use to population-weighted air-pollution exposure in complementary temporal and spatial designs. For display, we use the relative log-RSI, centered at an anchor date or year, so that proportional change appears in additive form. The objective is continuous distributional monitoring rather than the estimation of abrupt structural breaks.

979-8-3315-5221-3/26/$31.00 ©2026 IEEE
DOI 10.1109/CODASSCA69992.2026.00035

There are three main contributions. First, we adapt the RSI to pollutant-specific population-relevant exposure and develop a relative log-RSI for baseline-centered monitoring. Second, we give a common distributional formulation for two data architectures: temporally dense station records with sparse spatial support and spatially dense annual pollutant maps with coarse temporal support. Third, we isolate the effect of population weighting on exposure levels and distributional movement, including differences between population-weighted and unweighted spatial monitors.

The two empirical designs yield different forms of evidence. In Zurich, the monitored object is the temporal distribution of a daily citywide exposure signal within a rolling window. Short windows transmit the annual cycle into the monitoring statistic, whereas a 365-day window allows the monitored distribution to span that cycle and produces a substantially more stable series. In Switzerland, the monitored object is the spatial cross-section of mapped pollutant values within each year. Each annual national map provides a large cross-section, and the corresponding spatial distribution becomes the natural monitoring object. The same RSI functional is applied in both settings, but its numerical levels are design-specific and are not compared directly across the two applications.

By design, the results are pollutant-specific. In Zurich, population weighting matters most clearly for $NO_2$, less for $O_3$, and only modestly for $PM_{10}$ and $PM_{2.5}$. In Switzerland, the differences between population-weighted and unweighted spatial monitors are historically most pronounced for $NO_2$ and $PM_{10}$. Across the post-2020 Swiss observations, each population-weighted monitor falls more than its unweighted populated-cell counterpart, conditional on the available native raster supports. For $NO_2$ and $PM_{10}$, the Swiss mapped-value and rank-shift contrasts show that many of the largest recent absolute reductions occur where mapped values were initially highest, a pattern consistent with compression of the spatial concentration distribution.

## II. Monitoring Framework

### A. Regime-Shift Index

Following the general construction in [14], let $X$ be a scalar exposure variable with current distribution $F$, and let $G$ be a baseline cumulative distribution function (CDF). Suppose that $0 < G(X) < 1$ almost surely under $F$. For $r > 0$, define

$$G^r(x) = \big(G(x)\big)^r. \tag{1}$$

The family $\mathcal{G} = \{G^r : r > 0\}$ is ordered monotonically in $r$ and therefore provides a convenient one-parameter comparison class around the baseline distribution $G$.

**Definition 1 (Regime-Shift Index).** *Provided that* $-\mathbb{E}[\log G(X)] \in (0, \infty)$*, the quantity*

$$R(F|G) = (-\mathbb{E}[\log G(X)])^{-1} \tag{2}$$

*is called the* regime-shift index *of* $F$ *relative to* $G$*.*

Further ordering, mixture, and divergence properties of the functional are developed in [14]. When $G$ is continuous, $R(G|G) = 1$. If the current regime $F$ places relatively more mass at high baseline ranks than $G$, then $R(F|G) > 1$; if it places relatively more mass at low baseline ranks, then $R(F|G) < 1$. Here, a regime shift denotes distributional displacement relative to a baseline and need not be abrupt. The logarithmic form has an information-theoretic interpretation.[1]

**Remark 1 (Rank Invariance).** If $\phi : \mathbb{R} \to \mathbb{R}$ is strictly increasing and $Y = \phi(X)$, then $G_Y(Y) = G(X)$ almost surely, where $G_Y$ is the image of $G$ under $\phi$. Hence, $R(F_Y|G_Y) = R(F|G)$. The RSI is invariant under strictly increasing transformations, depends only on baseline ranks, and is independent of measurement units [14]. As a scalar summary, it does not recover the full distribution, so distributional displays and spatial contrasts remain complementary.

### B. Empirical Estimation

Let $X_1^0, \ldots, X_{n_0}^0$ be baseline observations with normalized weights $b_j \geq 0$, $\sum_{j=1}^{n_0} b_j = 1$, and define

$$\widehat{G}_{0,b}(x) = \sum_{j=1}^{n_0} b_j \mathbf{1}_{\{X_j^0 \leq x\}}. \tag{3}$$

For $\varepsilon_-, \varepsilon_+ > 0$ with $\varepsilon_- + \varepsilon_+ < 1$, define the regularized baseline rank function

$$G_{\mathrm{ref},b}(x) = \max\Big\{\min\{\widehat{G}_{0,b}(x), 1 - \varepsilon_+\}, \varepsilon_-\Big\}. \tag{4}$$

In all reported calculations, $\varepsilon_- = \varepsilon_+ = 0.005$, so that empirical baseline ranks are restricted to $[0.005, 0.995]$. The lower bound ensures a finite log-score below the empirical baseline support, while the upper bound regularizes observations at or above its maximum.

For current observations $X_1, \ldots, X_n$ with normalized weights $v_i \geq 0$, $\sum_{i=1}^n v_i = 1$, the weighted empirical RSI is

$$\widehat{R}_v = \left(-\sum_{i=1}^n v_i \log G_{\mathrm{ref},b}(X_i)\right)^{-1}. \tag{5}$$

The baseline weights determine the empirical rank scale, and the current weights determine the average log-rank. With equal weights, $b_j = 1/n_0$ and $v_i = 1/n$, this reduces to

$$\widehat{R}_n = \left(-\frac{1}{n}\sum_{i=1}^n \log G_{\mathrm{ref}}(X_i)\right)^{-1}, \tag{6}$$

where $G_{\mathrm{ref}}$ is based on the ordinary empirical baseline CDF. Neither estimator requires density estimation.

[1]The orientation of the RSI is relevant for monitoring. By contrast, the Wasserstein distance is a metric and is therefore symmetric in its arguments; it quantifies separation between distributions without assigning a direction to a shift [16]. Meanwhile, the Kullback–Leibler divergence is asymmetric and nonnegative [17]. In the RSI construction, Kullback–Leibler divergence supplies the projection criterion on the ordered family $\{G^r : r > 0\}$, while the fitted exponent $r$ captures direction through $r > 1$ or $r < 1$ [14].

### C. Relative Log-RSI

For monitoring purposes, it is convenient to work with the logarithm of the RSI and to anchor it at a selected date or year. Let $\widehat{R}_\tau$ denote an RSI computed for a time block or year indexed by $\tau$, and let $\tau_0$ denote the anchor. Consider

$$A_{\tau;\tau_0} = \log \widehat{R}_\tau - \log \widehat{R}_{\tau_0}. \tag{7}$$

Then $A_{\tau_0;\tau_0} = 0$ and

$$\exp(A_{\tau;\tau_0}) = \frac{\widehat{R}_\tau}{\widehat{R}_{\tau_0}}, \tag{8}$$

so the relative log-RSI represents multiplicative change relative to the anchor in additive form.

### D. Baseline and Anchor

The baseline period determines the reference distribution $G$ and hence the rank scale against which every monitored regime is evaluated. Changing the baseline can therefore alter the magnitude and time profile of the measured shift. The anchor $\tau_0$ enters only after the RSI has been computed and fixes the zero of the displayed series through (7). Replacing it changes all displayed values by the same additive constant, leaving pairwise changes, temporal orderings, and turning points unchanged.

An operational baseline should be substantively meaningful, precede the main monitoring interval, and contain enough observations to estimate the relevant rank distribution. For high-frequency variables with a seasonal cycle, it should span complete annual cycles. A common baseline facilitates comparison when supports permit it, while alternative plausible periods or leave-one-year-out calculations provide sensitivity checks. Choosing the end of the baseline as the anchor gives a direct transition to the post-baseline monitoring interval.

### E. Population-Weighted Exposure Signals

The relevant exposure signal depends on the empirical design. With high-frequency concentrations observed over spatial units carrying population weights, the natural signal is a population-weighted arithmetic mean. Fine spatial grids available at lower frequency instead provide a cross-sectional distribution of mapped ambient pollution at residential locations; see, e.g., [1], [8]–[10].

Let $c_{mst}$ denote the concentration of pollutant $m$ in spatial unit $s$ at time $t$, and let $\omega_s \geq 0$ satisfy

$$\sum_s \omega_s = 1. \tag{9}$$

The arithmetic exposure signal

$$E_{mt} = \sum_s \omega_s c_{mst} \tag{10}$$

is a population-weighted mean concentration suitable for high-frequency temporal monitoring.

For a spatial annual-map design, let $c_{msy}$ denote the mapped pollutant value in spatial unit $s$ and year $y$. The distribution at the residential location of a randomly drawn resident is

$$F_{my}(x) = \sum_s \omega_s \mathbf{1}_{\{c_{msy} \leq x\}}, \tag{11}$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function. This population-weighted distribution is suitable for tracking shifts in annual model-based pollutant maps.

## III. Empirical Design

To illustrate the distributional monitoring signal in (7), we use two data environments. The Zurich design combines hourly station measurements, statistical-zone geometry, and annual zone populations [18]–[20]. The Swiss design combines annual model-supported pollutant maps published by the Federal Office for the Environment (FOEN) with gridded federal population data [21]–[25].

### A. City-Level Temporal Design

Hourly Zurich station measurements are first aggregated to daily means. The city is partitioned into 216 statistical zones, including four structural zero-population lake zones retained for geometric completeness. The annual zone-population panel measures the economic resident population and is available for 1993–2025; for later dates, the 2025 weights are carried forward.

The design uses the four long-running monitoring stations Heubeeribüel, Rosengartenstrasse, Schimmelstrasse, and Stampfenbachstrasse. Pollutant-specific records begin on 01-Jan-1983 for $NO_2$, 01-Jan-1987 for $O_3$, 03-Jan-2001 for $PM_{10}$, and 01-Jan-2017 for $PM_{2.5}$, and extend through 06-Apr-2026. Station availability varies by pollutant and date. All four stations appear in the full record for each pollutant, but the 2014–2019 baseline contains observations from all four stations for $NO_2$ and $O_3$ and from three stations for $PM_{10}$ and $PM_{2.5}$. It covers all six calendar years for $NO_2$, $O_3$, and $PM_{10}$; because the $PM_{2.5}$ record begins in 2017, its effective baseline covers 2017–2019. The post-baseline monitoring period begins on 01-Jan-2020, and the anchor date $\tau_0$ is 31-Dec-2019. Earlier observations are evaluated retrospectively on the same rank scale for historical context.

Each statistical zone is assigned by proximity to one of the four stations, so that its concentration inherits the corresponding station series. Let $\mathcal{K}^{\mathrm{obs}}_{m,t}$ denote the set of stations reporting pollutant $m$ on day $t$, and let $\mathcal{Z}^{\mathrm{obs}}_{m,t}$ denote the zones assigned to those stations. When one or more station observations are unavailable, the corresponding zones are excluded for that day and the remaining population weights are renormalized. We construct

$$E^{\mathrm{fix}}_{mt} = \frac{\sum_{z \in \mathcal{Z}^{\mathrm{obs}}_{m,t}} \omega^0_z c_{mzt}}{\sum_{z \in \mathcal{Z}^{\mathrm{obs}}_{m,t}} \omega^0_z}, \tag{12}$$

$$E^{\mathrm{real}}_{mt} = \frac{\sum_{z \in \mathcal{Z}^{\mathrm{obs}}_{m,t}} \omega^{\mathrm{real}}_{z,t} c_{mzt}}{\sum_{z \in \mathcal{Z}^{\mathrm{obs}}_{m,t}} \omega^{\mathrm{real}}_{z,t}}, \tag{13}$$

and

$$E_{mt}^{\text{naive}} = \frac{1}{K_{m,t}^{\text{obs}}} \sum_{k \in \mathcal{K}_{m,t}^{\text{obs}}} y_{mkt}. \tag{14}$$

Here $\omega_z^0$ is the fixed 2019 population weight for zone $z$, $\omega_{z,t}^{\text{real}}$ is the annual population weight applicable at date $t$, $c_{mzt}$ is the station-assigned concentration in zone $z$, $y_{mkt}$ is the observed concentration at station $k$, and $K_{m,t}^{\text{obs}} = |\mathcal{K}_{m,t}^{\text{obs}}|$. The naive signal is a contemporaneous average over reporting stations and is included only as a benchmark.

The Zurich RSI is constructed from 365-day rolling windows requiring at least 330 observed days. Shorter windows transmit seasonality to the monitoring signal, whereas an annual window allows the monitored distribution to span the seasonal cycle.

### B. National-Level Spatial Design

The Swiss mapped values are annual mean concentrations for $NO_2$, $PM_{10}$, and $PM_{2.5}$ and the maximum monthly 98th-percentile concentration (MP98) for $O_3$. The available supports are summarized in Table I. Fixed population weights are constructed from the 2024 STATPOP grid. The retained support contains 347,736 positive-population cells on a 100 m lattice, with a total recorded population of 9,123,704.

To preserve each pollutant map on its native grid, the fixed 2024 population weights are aggregated from the 100 m population support to each source raster by assigning every population cell to the source-grid cell whose spatial footprint contains it. Thus, each mapped field remains on its native support and only the population weights are moved.[2]

The Swiss spatial RSI uses baseline years 2015–2019 and anchor year 2019. For each pollutant and design, the valid populated-cell values from the five baseline fields are pooled to construct a separate baseline rank distribution. The weighted design uses the aligned fixed 2024 population weights, whereas the unweighted design gives equal weight to the populated cells. Within each pollutant-year, both designs use the same valid positive-population cells and differ only in their weights. Because the native raster grid can change across years, however, this cell support is not fixed through time. Comparable Swiss health-risk assessments are known to depend materially on such exposure-construction choices [26].

The Zurich RSI summarizes the temporal distribution of daily citywide signals, whereas the Swiss RSI summarizes an annual spatial cross-section of mapped values. Their numerical levels are therefore not compared directly.

## IV. Zurich Results: Intertemporal Monitoring

Combining the Zurich station data with the annual population counts across the 216 statistical zones yields the daily exposure signals in (12)–(14). Since the fixed-weight and realized-weight signals are almost identical over the displayed horizons, Fig. 1 reports only the fixed-weight signal,[3] together with the naive station mean. The underlying signals are daily, and the figure shows monthly means for compact presentation. The shaded block marks the baseline period: 2014–2019 for $NO_2$, $O_3$, and $PM_{10}$, and 2017–2019 for $PM_{2.5}$.

TABLE I
Empirical designs, supports, and roles

| Entity | Support (space) | Support (time) | Role |
|---|---|---|---|
| Zurich stations | 4 stations | daily | temporal monitoring |
| Zurich zones | 216 zones | fixed | exposure attribution |
| Zurich population | 216 zones | 1993–2025 | population weights |
| Swiss population | 347,736 cells | 2024 | spatial weights |
| $NO_2$ maps | national raster | 1990–2024 | spatial monitoring |
| $O_3$ MP98 maps | national raster | 1990–2024 | spatial monitoring |
| $PM_{10}$ maps | national raster | 1998–2024 | spatial monitoring |
| $PM_{2.5}$ maps | national raster | 2015–2024 | spatial monitoring |

For $NO_2$, the population-weighted exposure signal lies above the unweighted station mean over a substantial part of the sample, consistent with higher $NO_2$ concentrations in more densely inhabited parts of the city and in line with the known fine-scale spatial heterogeneity of $NO_2$ within Zurich [12]. For $O_3$, the two signals are much closer, consistent with weaker within-city heterogeneity. For $PM_{10}$ and $PM_{2.5}$, slight level differences are present. Thus, in Zurich, population weighting matters most for $NO_2$ levels, much less for the other pollutants.

Figure 2 shows the corresponding relative log-RSI monitors. The RSI is computed from daily data by means of a 365-day rolling window and is displayed as monthly means. Shorter windows inherit a substantial seasonal component because the monitored distribution samples only part of the year. The 365-day monitoring distribution spans the full annual cycle and yields substantially smoother trajectories.

The dominant empirical feature in Fig. 2 is the long-run downward movement for $NO_2$, $PM_{10}$, and $PM_{2.5}$, with shorter-run fluctuations in the particulate monitors. $O_3$ follows a distinct, nonmonotone trajectory. The difference between the population-weighted and unweighted monitors is clearest for $NO_2$, smaller for $O_3$, and modest for the particulates. The Zurich evidence bears primarily on temporal trends, with population weighting acting as a secondary correction.

**Remark 2 (Rank-Based Shift Interpretation).** The magnitude of the observed decline is also interpretable on the RSI scale. Since $\exp(A_{\tau;\tau_0}) = \widehat{R}_\tau / \widehat{R}_{\tau_0}$, a fall of 2 in relative log-RSI means that the RSI is multiplied by $\exp(-2) \approx 0.135$ relative to its anchor value. Equivalently, the average baseline log-rank depth $-\mathbb{E}[\log G(X)]$ is multiplied by $\exp(2) \approx 7.4$. Within the comparison family $\mathcal{G}$ in (1), the fitted exponent is therefore reduced to approximately $13.5\%$ of its anchor value. A negative relative log-RSI indicates movement toward lower baseline ranks relative to the anchor and hence, for $NO_2$, $PM_{10}$, and $PM_{2.5}$, an improvement according to the rank-based monitor.

[2]In the analyzed raster inventory, some older $NO_2$ rasters are on a $1105 \times 1745$ grid, many $O_3$, $PM_{10}$, and $PM_{2.5}$ rasters are on a $2210 \times 3490$ grid, and the 2024 $O_3$ raster is on a $2210 \times 3501$ grid. FOEN documents broader changes in the construction and resolution of the annual map series [23].

[3]The fixed-weight signal holds zone population weights at their 2019 values and therefore isolates changes in the concentration field from annual population redistribution, conditional on the stations observed on each date.

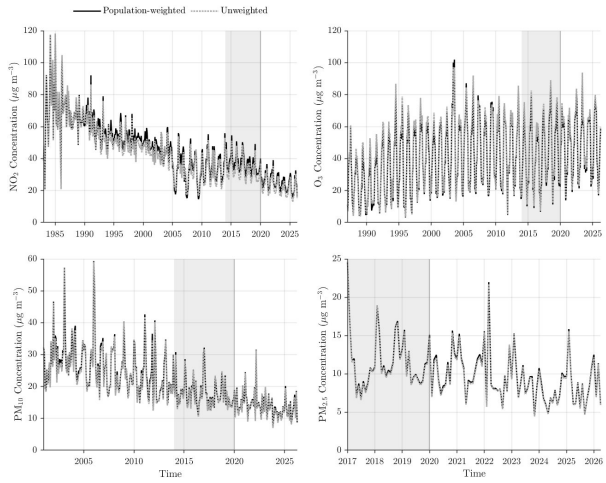



Fig. 1. Zurich exposure signals for $NO_2$, $O_3$, $PM_{10}$, and $PM_{2.5}$. The daily signals are shown as monthly means. The solid line is based on fixed 2019 population weights, and the dotted line is the naive unweighted station mean. The realized-weight signal is omitted because it closely tracks the fixed-weight signal. The shaded block marks the baseline period: 2014–2019 for $NO_2$, $O_3$, and $PM_{10}$, and 2017–2019 for $PM_{2.5}$.

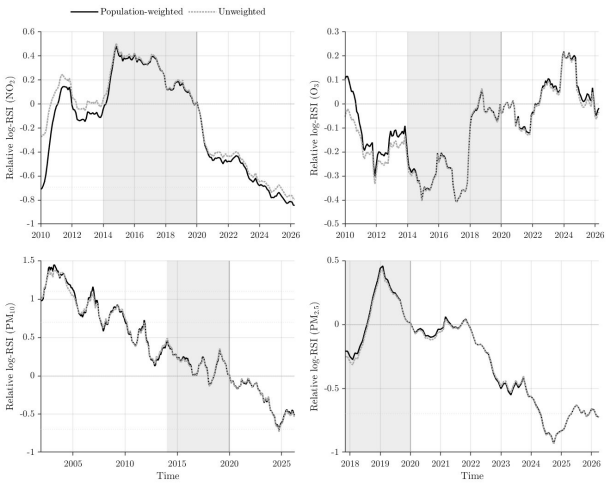



Fig. 2. Zurich 365-day rolling relative log-RSI for $NO_2$, $O_3$, $PM_{10}$, and $PM_{2.5}$, displayed as monthly means. The solid line is based on fixed 2019 population weights, and the dotted line on the naive unweighted station mean. The shaded block marks the baseline period: 2014–2019 for $NO_2$, $O_3$, and $PM_{10}$, and 2017–2019 for $PM_{2.5}$; the anchor date is 31-Dec-2019.

## V. Swiss Results: Spatial Monitoring

The Swiss design is both annual and cross-sectional. Figure 3 shows the 2019 and 2024 mapped pollutant fields, the corresponding 2024-minus-2019 mapped-value contrasts, and the associated baseline-rank contrasts. For $NO_2$ and $PM_{10}$, the highest values lie along the main urban and transport corridor of the Swiss Plateau, consistent with broad spatial patterns found in Swiss and Western European modeling studies [13], [22]. The largest absolute reductions occur mainly in those same locations. The rank-shift contrasts show a similar pattern, consistent with compression of the spatial concentration distribution. $PM_{2.5}$ shows a less pronounced pattern, while $O_3$ MP98 has a distinct spatial configuration.

Table II provides a level comparison. Over 2020–2024, the population-weighted mean mapped value exceeds the unweighted populated-cell mean by about $20\%$ for $NO_2$, $11\%$ for $PM_{10}$, 9–$10\%$ for $PM_{2.5}$, and only 1–$2\%$ for $O_3$. Over each pollutant's full available panel, the average weighted-to-unweighted ratios are $1.268$ for $NO_2$, $1.017$ for $O_3$, $1.100$ for $PM_{10}$, and $1.098$ for $PM_{2.5}$. Population weighting therefore matters most for $NO_2$, by about $10\%$ for both particulate measures, and little for $O_3$.

A distributional comparison is shown in Fig. 4. Let $A^{\mathrm{w}}_{my;y_0}$ and $A^{\mathrm{uw}}_{my;y_0}$ denote the weighted and unweighted populated-cell relative log-RSI for pollutant $m$ in year $y$, anchored at $y_0 = 2019$. By setting $D_{my} = A^{\mathrm{w}}_{my;y_0} - A^{\mathrm{uw}}_{my;y_0}$, one obtains

$$\exp(D_{my}) = \frac{\widehat{R}^{\mathrm{w}}_{my}/\widehat{R}^{\mathrm{w}}_{my_0}}{\widehat{R}^{\mathrm{uw}}_{my}/\widehat{R}^{\mathrm{uw}}_{my_0}}. \tag{15}$$

This normalized weighting factor compares the population-weighted RSI change with the unweighted populated-cell RSI change, each relative to 2019. Values above 1 mean that the normalized population-weighted RSI is larger; values below 1 mean that it is smaller.[4]

Historically, this factor is well above 1 for $NO_2$ and $PM_{10}$. For $NO_2$, it reaches about $1.55$ in the early 1990s; for $PM_{10}$, it exceeds $2.18$ at its maximum. Relative to the fixed 2024 population grid, historically high mapped values for these pollutants were therefore disproportionately concentrated in more densely populated locations. $PM_{2.5}$ is much closer to neutrality over its shorter sample. $O_3$ follows a different pattern and has a full-panel mean below 1.

Over 2020–2024, the mean normalized weighting factor is below 1 for all four pollutants: $0.738$ for $NO_2$, $0.783$ for $O_3$, $0.922$ for $PM_{10}$, and $0.896$ for $PM_{2.5}$. Within each pollutant-year, the weighted and unweighted calculations use the same populated raster cells and differ only in their weights. The population-weighted spatial relative log-RSI therefore declines more strongly in every case examined. Because some native raster grids change over time, this comparison remains conditional on the available raster supports.

## VI. Interpretation and Limitations

The proposed method is intended for monitoring environments with heterogeneous data availability, across time and space. In our application, Zurich provides high-frequency temporal information under sparse spatial support, whereas Switzerland provides broad spatial information at an annual frequency. In both settings, the objective is a population-relevant distributional signal that can be updated as new observations

[4]Equivalently, values above 1 indicate a larger increase or a smaller decline under population weighting, whereas values below 1 indicate a larger decline or a smaller increase.

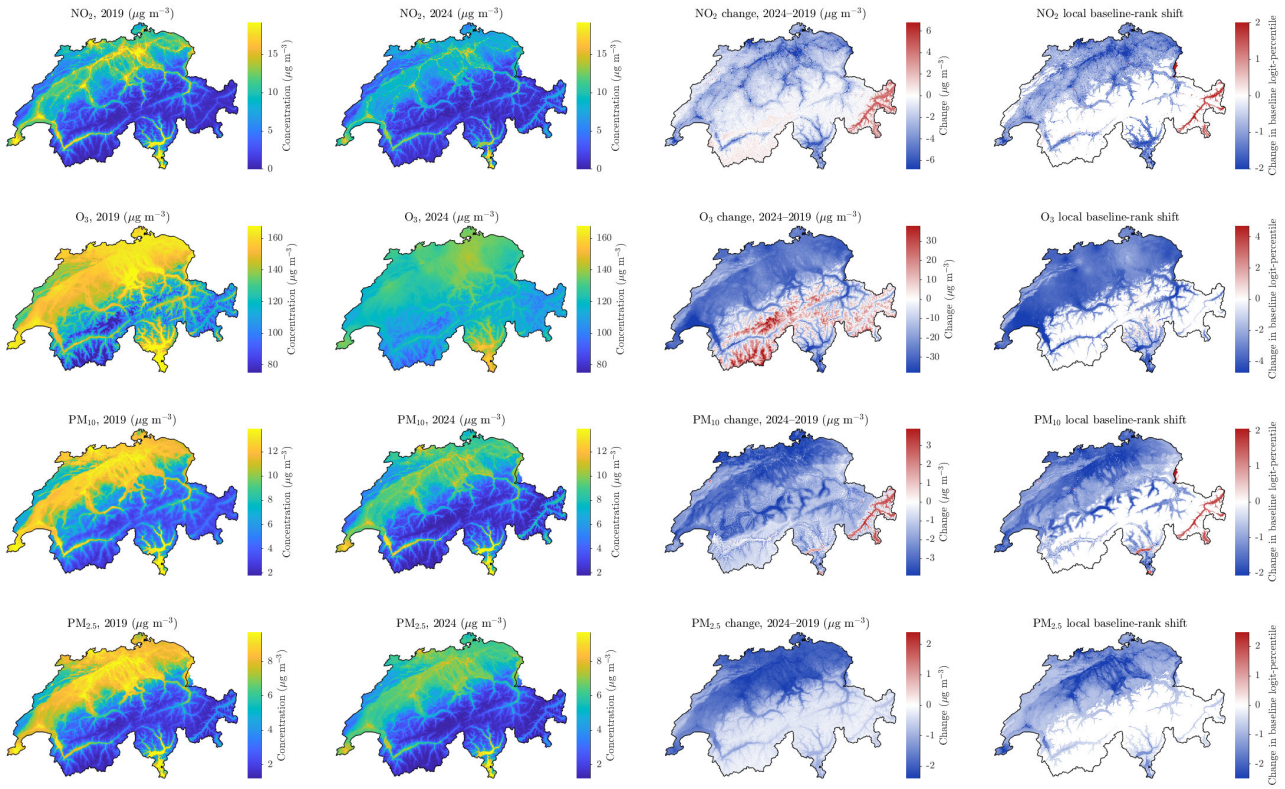



Fig. 3. Swiss annual mapped pollutant fields and spatial contrasts. For each pollutant, the first two columns show the 2019 and 2024 fields, the third shows the 2024-minus-2019 mapped-value contrast, and the fourth shows the corresponding baseline-rank contrast. $NO_2$, $PM_{10}$, and $PM_{2.5}$ are annual means; $O_3$ is MP98. For $NO_2$ and $PM_{10}$, the pattern is consistent with compression of the spatial concentration distribution.

TABLE II
Swiss population weighting in levels and spatial monitoring

| Pollutant | Mean level ratio weighted / unweighted | Full-panel mean $\exp(D_{my})$ | 2020–2024 mean $\exp(D_{my})$ |
|---|---|---|---|
| $NO_2$ | 1.268 | 1.204 | 0.738 |
| $O_3$ | 1.017 | 0.892 | 0.783 |
| $PM_{10}$ | 1.100 | 1.621 | 0.922 |
| $PM_{2.5}$ | 1.098 | 1.005 | 0.896 |

become available. A "regime shift" here denotes displacement relative to a baseline and need not imply an abrupt structural break. Statistical estimation of break dates is treated in the change-point literature [27]. The method admits a multi-rate operational implementation: station observations can update the urban temporal monitor, revised pollutant maps can update the national spatial monitor, and population weights can enter at their own frequency. Once the baseline CDF has been fixed, these data streams enter the same rank-based construction without repeated parametric estimation. Denser sensor networks and sensor–model fusion can improve the underlying concentration field, subject to calibration and uncertainty control, without changing the monitoring statistic [11], [28]. Alternative baselines can represent different comparison regimes.

Several design constraints persist. In Zurich, the zone-to-station assignment limits spatial resolution, and changes in station availability alter the represented zones, although weights are renormalized over the stations observed on each date. Fixed 2019 weights isolate concentration-field changes from annual population redistribution, conditional on station availability; realized weights also incorporate annual residential redistribution. In Switzerland, the fixed 2024 population grid holds residential location constant, whereas time-varying weights would measure the joint evolution of mapped pollution and residential location. Because native raster supports can change, this interpretation is conditional on the available pollutant grids. Both designs provide residential ambient-exposure proxies rather than personal exposure measures: they do not incorporate indoor infiltration, mobility, commuting, or individual time-activity patterns [29]. The RSI is also a scalar summary, so distinct distributional changes may offset on the average log-rank scale; maps and distributional plots remain complementary.

The RSI is descriptive and does not by itself identify health effects or policy impacts. Health assessment requires an outcome model incorporating concentration–response relationships, exposure lags, population susceptibility, and relevant confounders. Policy evaluation requires an identification

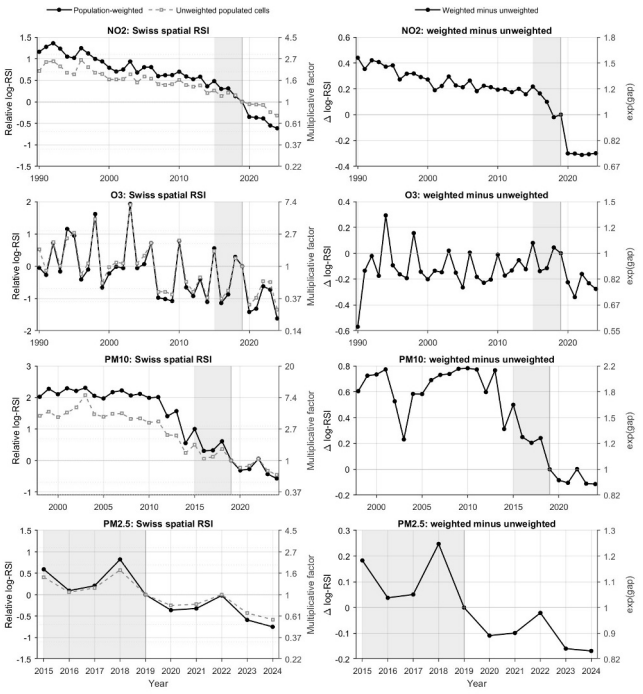



Fig. 4. Swiss annual relative log-RSI by pollutant and design. In the left panels, solid lines are population-weighted and dashed lines are unweighted over populated cells; the left axis reports relative log-RSI and the secondary scale gives the corresponding RSI ratio relative to 2019. The right panels show $D_{my}$ on the left axis and $\exp(D_{my})$ on the secondary axis. Baseline: 2015–2019; anchor: 2019. $O_3$ denotes MP98.

strategy separating an intervention from concurrent changes in emissions, economic activity, technology, and meteorological conditions. Statistical uncertainty must likewise be treated differently across designs. The Zurich RSI is a functional of serially dependent daily observations, so moving-block bootstrap or long-run variance methods can provide confidence intervals when needed. The Swiss statistics are computed from model-based national rasters and a population grid; resampling cells would ignore spatial dependence and omit uncertainty in the underlying pollution model. Confidence intervals would therefore require multiple plausible realizations of each pollutant field, or equivalent information about the spatial uncertainty in the underlying raster model.

Our approach uses pollutant concentration as the main unit of analysis. This preserves the observed heterogeneity across pollutants and avoids imposing additional aggregation assumptions. A joint air-quality index would require explicit choices concerning normalization, substitution across pollutants, health-risk weights, and exposure dependence [30]. Pollutant-specific monitoring leaves these choices transparent and permits them to be introduced for a particular application rather than embedded in the basic statistic. It also provides a basis for subsequent source-attribution, health-risk, intervention, or urban-policy analysis.

## VII. Conclusion

Air-pollution monitoring commonly begins with concentration measurements from a sparse network of stations or from spatial fields available only at relatively low frequency. For distributional exposure monitoring, however, concentrations must be related to the locations of the population. Population weighting is therefore part of the measurement problem rather than a subsequent adjustment. The monitoring design must also accommodate the resolution of the available data: station observations provide temporal detail under limited spatial support, whereas national pollutant maps provide broad spatial coverage at coarse temporal frequency.

The relative log-RSI provides a common rank-based construction for these different environments. Each monitored distribution is evaluated on the rank scale defined by a baseline distribution and expressed relative to an anchor date. The resulting statistic is nonparametric, invariant under strictly increasing transformations, and applicable without imposing a common parametric model on the underlying pollutant fields. It can be constructed from station-based exposure signals or directly from spatial distributions on their native raster supports. Fixed population weights hold the residential distribution constant and focus the comparison on changes in the concentration field, whereas time-varying weights incorporate the joint evolution of pollution and residential location.

The Zurich application illustrates the temporal use of the method. A 365-day rolling window spans the seasonal cycle and produces smoother monitoring trajectories than shorter windows. The fixed- and realized-weight exposure signals track each other closely over the observed period, while the difference between population-weighted exposure and the naive station mean is most pronounced for $NO_2$. The corresponding relative log-RSI monitors show long-run downward movements for $NO_2$, $PM_{10}$, and $PM_{2.5}$, with shorter-run fluctuations, whereas $O_3$ follows a distinct, nonmonotone path. In this setting, the principal empirical signal is longitudinal change, with population weighting providing an important but generally secondary correction.

The Swiss application shows that population weighting can affect both mapped-value levels and measured spatial distributional shifts. Relative to the fixed 2024 population grid, historically high $NO_2$ and $PM_{10}$ values were disproportionately concentrated in more densely populated locations. Over 2020–2024, the population-weighted relative log-RSI declined more strongly than the unweighted populated-cell counterpart for all four pollutants within the native-support design. The national monitor is therefore consistent with relatively stronger recent improvement where population is concentrated. The spatial contrasts support this interpretation for $NO_2$ and $PM_{10}$: the largest absolute reductions occur mainly where mapped values were initially highest, a pattern consistent with compression of the spatial concentration distribution.

The method thus converts spatially and temporally dispersed pollution data into pollutant-specific distributional monitoring signals with a common rank-based interpretation. Population

weighting can change the measured direction and magnitude of distributional movement relative to a common anchor, and is therefore beyond a marginal correction to average levels. Naturally the RSI cannot by itself identify health effects, structural breaks, or causal policy impacts. Those questions require additional outcome models, identification strategies, and uncertainty information. Its role is to provide the exposure-side monitoring signal while keeping the baseline, anchor, spatial support, and population-weighting assumptions explicit.

## REFERENCES


[1] World Health Organization, *Overview of Methods to Assess Population Exposure to Ambient Air Pollution*. Geneva, Switzerland: WHO, 2023.

[2] World Health Organization, *WHO Global Air Quality Guidelines: Particulate Matter ($PM_{2.5}$ and $PM_{10}$), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. Geneva, Switzerland: WHO, 2021.

[3] European Environment Agency, *Harm to Human Health From Air Pollution in Europe: Burden of Disease Status, 2024*. Copenhagen, Denmark: EEA, 2024.

[4] U. Ackermann-Liebrich, B. Leuenberger, J. Schwartz, C. Schindler, C. Monn, G. Bolognini, et al., "Lung function and long term exposure to air pollutants in Switzerland," *American Journal of Respiratory and Critical Care Medicine*, vol. 155, no. 1, pp. 122–129, Jan. 1997.

[5] R. Beelen, O. Raaschou-Nielsen, M. Stafoggia, Z. J. Andersen, G. Weinmayr, B. Hoffmann, K. Wolf, et al., "Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project," *Lancet*, vol. 383, no. 9919, pp. 785–795, Mar. 2014.

[6] M. Stafoggia, B. Oftedal, J. Chen, et al., "Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: results from seven large European cohorts within the ELAPSE project," *Lancet Planetary Health*, vol. 6, no. 1, pp. e9–e18, Jan. 2022.

[7] D. Vienneau, M. Stafoggia, et al., "Association between exposure to multiple air pollutants, transportation noise and cause-specific mortality in adults in Switzerland," *Environmental Health*, vol. 22, no. 1, art. 29, Mar. 2023.

[8] F. Gilliland, E. Avol, P. Kinney, M. Jerrett, T. Dvonch, F. Lurmann, et al., "Air pollution exposure assessment for epidemiological studies of pregnant women and children," *Environmental Health Perspectives*, vol. 113, no. 10, pp. 1447–1454, Oct. 2005.

[9] G. Hoek, R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer, and D. Briggs, "A review of land-use regression models to assess spatial variation of outdoor air pollution," *Atmospheric Environment*, vol. 42, no. 33, pp. 7561–7578, Oct. 2008.

[10] V. A. Southerland, S. C. Anenberg, M. Harris, J. S. Apte, P. Hystad, A. van Donkelaar, R. V. Martin, M. Beyers, and A. Roy, "Assessing the distribution of air pollution health risks within cities: A neighborhood-scale analysis leveraging high-resolution data sets in the Bay Area, California," *Environmental Health Perspectives*, vol. 129, no. 3, art. 037006, Mar. 2021.

[11] J. S. Apte and C. Manchanda, "High-resolution urban air pollution mapping," *Science*, vol. 385, pp. 380–385, Jul. 2024.

[12] M. D. Mueller, M. Wagner, I. Barmpadimos, and C. Hueglin, "Two-week $NO_2$ maps for the City of Zurich, Switzerland, derived by statistical modelling utilizing data from a routine passive diffusion sampler network," *Atmospheric Environment*, vol. 106, pp. 1–10, Apr. 2015.

[13] K. de Hoogh, J. Chen, J. Gulliver, B. Hoffmann, O. Hertel, M. Ketzel, M. Bauwelinck, et al., "Spatial $PM_{2.5}$, $NO_2$, $O_3$ and BC models for Western Europe: evaluation of spatiotemporal stability," *Environment International*, vol. 120, pp. 81–92, Nov. 2018.

[14] T. A. Weber, "A rank-based measure of distributional shifts," Working Paper, Chair of Operations, Economics and Strategy, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2025.

[15] T. A. Weber, "A regime-shift index for climate monitoring," *2026 International Conference on Sustainable Engineering and Digital Innovation (ICSEDI)*, Muscat, Oman, Feb. 2026, IEEE, pp. 1–6.

[16] C. Villani, *Optimal Transport: Old and New*, Grundlehren der Mathematischen Wissenschaften, vol. 338. Berlin, Germany: Springer, 2009.

[17] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[18] City of Zurich, "Stündlich aktualisierte Luftqualitätsmessungen, seit 1983," Open Data Zurich, updated Jul. 18, 2026. [Online]. Available: https://data.stadt-zuerich.ch/dataset/ugz_luftschadstoffmessung_stundenwerte. [Accessed: Jul. 20, 2026].

[19] City of Zurich, "Statistische Zonen," Open Data Zurich, updated Jul. 10, 2026. [Online]. Available: https://data.stadt-zuerich.ch/dataset/geo_statistische_zonen. [Accessed: Jul. 20, 2026].

[20] City of Zurich, "Bevölkerung nach statistischer Zone," Open Data Zurich, updated Feb. 10, 2026. [Online]. Available: https://data.stadt-zuerich.ch/dataset/bev_bestand_jahr_statzone_od3241. [Accessed: Jul. 20, 2026].

[21] Federal Office for the Environment, "Air quality: data," [Online]. Available: https://www.bafu.admin.ch/en/air-quality-data. [Accessed: Jul. 20, 2026].

[22] Federal Office for the Environment, "Maps of annual values," Aug. 14, 2024. [Online]. Available: https://www.bafu.admin.ch/en/maps-of-annual-values. [Accessed: Jul. 20, 2026].

[23] T. Künzle, *Karten der Luftbelastung in der Schweiz: Dokumentation zu Datengrundlagen, Berechnungsmethoden und Resultaten der Jahreswerte 2024*. Bern, Switzerland: Meteotest AG, final report commissioned by the Federal Office for the Environment, Jul. 31, 2025. [Online]. Available: https://www.bafu.admin.ch/en/maps-of-annual-values. [Accessed: Jul. 20, 2026].

[24] Federal Office for the Environment, "National Air Pollution Monitoring Network (NABEL)," Dec. 5, 2025. [Online]. Available: https://www.bafu.admin.ch/en/national-air-pollution-monitoring-network-nabel. [Accessed: Jul. 20, 2026].

[25] Federal Statistical Office, *Statistik der Bevölkerung und Haushalte (STATPOP), Geodaten 2024*. Neuchâtel, Switzerland: FSO, 2025. [Online]. Available: https://www.bfs.admin.ch/asset/en/36079999. [Accessed: Jul. 20, 2026].

[26] A. Castro, M. Röösli, K. de Hoogh, R. Kappeler, M. Kutlar Joss, D. Vienneau, and N. Künzli, "Methods matter: a comparative review of health risk assessments for ambient air pollution in Switzerland," *Public Health Reviews*, vol. 43, art. 1604431, Apr. 2022.

[27] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowledge and Information Systems*, vol. 51, no. 2, pp. 339–367, May 2017.

[28] P. Schneider, N. Castell, M. Vogt, F. R. Dauge, W. A. Lahoz, and A. Bartonova, "Mapping urban air quality in near real-time using observations from low-cost sensors and model information," *Environment International*, vol. 106, pp. 234–247, Sep. 2017.

[29] U.S. Environmental Protection Agency, *Guidelines for Human Exposure Assessment*, EPA/100/B-19/001. Washington, DC, USA: Risk Assessment Forum, U.S. EPA, Oct. 2019. [Online]. Available: https://www.epa.gov/sites/default/files/2020-01/documents/guidelines_for_human_exposure_assessment_final2019.pdf. [Accessed: Jul. 20, 2026].

[30] M. Oakes, L. Baxter, and T. C. Long, "Evaluating the application of multipollutant exposure metrics in air pollution health studies," *Environment International*, vol. 69, pp. 90–99, Aug. 2014.