Rozwój Regionalny i Polityka Regionalna 65: 13–35
https://doi.org/10.14746/rrpr.2023.65.03
2023

Daniel A. Griffith

University of Texas at Dallas

Does spatial autocorrelation matter for sustainable regional development planning and evaluation?

Abstract: Pursuing the various existing stainability dimensions obliges leaders of society to engage in more comprehensive monitoring of collective economic and other supplies and demands, particularly in a geographic context. In turn, the affected inputs, outputs, resources/goods/services stocks, and generated garbage/trash waste, which all exist and are tagged implicitly or explicitly in geographic space, are definite harborers of spatial autocorrelation. Harnessing this nearly ubiquitous georeferenced data property implants a capability of fostering efficient and effective sustainability ventures. Tessellation stratified random sampling to monitor environmental pollution alludes to one example of this assertion. This paper illustrates this exemplification with an examination of 2023 air quality data for Poland. In doing so, it translates a framework build upon idealized tessellations into one for the administrative districts of Poland. This methodological conversion enables governmental organizations to participate in and oversee any intended monitoring without additional jurisdictional complications. Serendipitous academic discoveries include an initial extension of the set of standard polygon shapes (e.g., square and hexagon) to the trapezoid for spatial sampling purposes, and the possibility that spatial autocorrelation impacts upon design-based statistics may far outweigh a violation of the conventional random sampling equiprobable commandment. Finally, the discerning conclusion reached through the analyses summarized in this exposé argues that spatial autocorrelation does matter for sustainable regional development planning and evaluation.

Key words: PM10, Poland, spatial autocorrelation, sustainable regional development, tessellation stratified random sampling

Introduction

Sustainable regional development commits to meeting economic, social, and environmental resource needs in a given geographic landscape for its present human generation without depleting them, and hence compromising the ability of future peoples to meet their own same resource needs (after Brundtland 1987). One consequence of this enactment is improved perpetual agricultural production, energy use, natural resource management, and industrial production (Rut-

kauskas et al. 2014, Dočekalová et al. 2015, Straková 2015) at regional, rather than solely local, national, and/or worldwide levels. One outcome is a transcending of the single imperative of alleviating regional disparities, converting it to a multiple obligation of inter-regional equity plus resources perpetuation.

Meanwhile, the necessary accompanying sustainability evaluation, the general topic of this paper, is the method of identifying, measuring, and appraising potential impacts of alternatives for this sustainability (Devuyst 2000). Accordingly, it requires both macro- and micro-level monitoring: the former involves gauging short- and long-run development of regional economies (e.g., analyses of intra- and inter-regional flows in addition to regional wealth stocks, such as natural resources); and, the latter involves supporting benchmarking that stimulates region-based learning behavior and innovations. Therefore, the principal challenge is to develop a science-base that promotes better understanding, standardization, and informed decision-making concerning resource usage, both of which entail more accurate quantification of and communication about the sustainability of agricultural and industrial products, green energy, and natural assets harvesting. Doing so compels proper environmental monitoring, particularly of the pollution by-products of these economic activities. Excess pollution diminishes or even prevents sustainability!

Within this context, the purpose of this paper is to discuss selected aspects of spatial sampling – with regard to statistical populations as well as superpopulations – for sustainable development pollution monitoring purposes that collects data for sustainability evaluation. The indirect focal element of these geospatial data traits is spatial autocorrelation (SA), whereas the direct focus is effective geographic sample size. In other words, the question this narrative helps answer asks whether or not SA matters for sustainable regional development planning (e.g., knowledge accumulation and comprehension), specifically through environmental monitoring ventures, and evaluation (e.g., resources renewal skills) by, among other contributions, enabling remarkably smaller sample sizes to support geographic landscape monitoring and surveillance. Quantitative model- and design-based inference become inescapable topics for macro-level global measures across a region (Griffith, Plant 2022). In contrast, qualitative inference becomes an unavoidable topic for micro-level individual human subject measurements within a region (Griffith 2013, Griffith et al. 2016, Brown et al. 2017). This troika of sampling strategies aligns with the following triplet of themes Graymore et al. (2008) promote: wellbeing (e.g., equity and essential sustenance; from the economic dimension), ecological footprint (e.g., waste handling; from the environmental dimension), and quality of life (e.g., individual time-specific perceptions vis-à-vis benchmark mensuration contrasts; from the social dimension). Their mutual underlying feature of interest is their potential effective sample size, say n*, usually materializing mostly from the presence of SA. A deviation of n* from the naive n of classical statistics impacts evidence-based inferences, their resulting decision-making judgments, and sustainability knowledge and cognition. Moreover, SA exists whether or not sustainability is engaged, rather than being a cause or effect of this sustainability. However, it offers potential welfare economies that

should be exploited (e.g., effective geographic sample size), and it flags a gap in the literature concerning state-of-the-art spatial sampling, namely ideal hexagonal geographic tessellations are awkward and difficult to implement (e.g., correcting for edge effects), more pliable square tessellations are less cumbersome but still remain difficult to implement, and little positive has been said analytically to date about utilizing existing administrative district irregular tessellations. In fact, these far more convenient latter surface partitionings seem to have been dismissed instead of embraced. This paper addresses this gap, and in doing so outlines spatial sampling methodology that contributes to improved techniques for evaluating sustainable development progress and continuation, highlighting its importance.

Effective geographic sample size

One interpretation of any correlation concept is the presence of replicate/redundant information. If correlation is zero, then any information overlap in data is completely absent, whereas if correlation is 1, then all information embedded in data is entirely shared, making this overlap an equivalency. In more traditional situations, this redundancy needs to materialize as a linear trend, whereas in contemporary situations, it also can manifest itself as a nonlinear trend. Non-zero correlation can contaminate probabilities (e.g., sampling without replacement, with finite population selection probabilities sequentially decreasing from 1/N, where N is the population size), multiple attributes (e.g., collinearity, with principal component/factor analysis covariance eigenvalues increasing from 1 to the number of attribute variables, P, accompanied by declining degrees of freedom), and/or observations (e.g., self-correlation, with sampling distribution behavior being that for sample sizes less than n, namely n*). This third possibility is the correlated data source of interest in this piece. In his treatment of this data property, Griffith (2020) presents n* calculations for a wide variety of self-correlation issues, not just those for SA, expanding upon Griffith (2005). The ensuing discussion restricts attention to only SA situations. It accentuates how SA does and does not change sampling situations: non-zero correlation means n misrepresents, whereas zero correlation means n properly measures, sample size, which also may need a degrees-of-freedom modification.

Hence, the relevant question asks about the ways to quantify sample size in the presence of non-zero SA. Griffith (2005) derives n* for model-based spatial statistical inferences, furnishing an equation to estimate it after calculating a SA parameter employing a spatial simultaneous autoregressive (SAR; aka spatial error in the spatial econometrics literature) model specification. This formulation is sufficiently general to capture both autoregressive response (AR; aka spatial lag in the spatial econometrics literature), and conditional autoregressive (CAR) models. A critical assumption here is the ceaselessly invoked independent and identically distributed (iid) one pervasive in classical statistics. The definition of n* becomes: the number of equivalent iid observations affiliated with all unduplicated areal unit information content latent in geotagged data. The independent

portion – the first i in iid – is the relevant part of this acronym. Griffith and Plant (2022) show that the id part of iid merely increases the effective geographic sample size n* in keeping with its counterpart for a finite mixture distribution subsample having the largest n* if it were the lone random variable. Griffith (2005) portrays a negative exponential declining trendline curve (i.e., concave upward, but shallowly) for this adjusted sample size, with n* = n when zero, and n* = 1 when perfect positive, correlation prevails (Fig. 1a displays a n = 100 simulated realization case). Especially remotely sensed pixel data for which is relatively close to 1 can have n* deflating to less than 1% of n (e.g., the Adirondack remotely sensed image in Griffith 2015). Socio-economic/demographic polygonal data for which a single parameter is closer to 0.5 frequently yield an n* approaching 20% or more of n (e.g., the Chicago insurance map adapted in Griffith 2020).

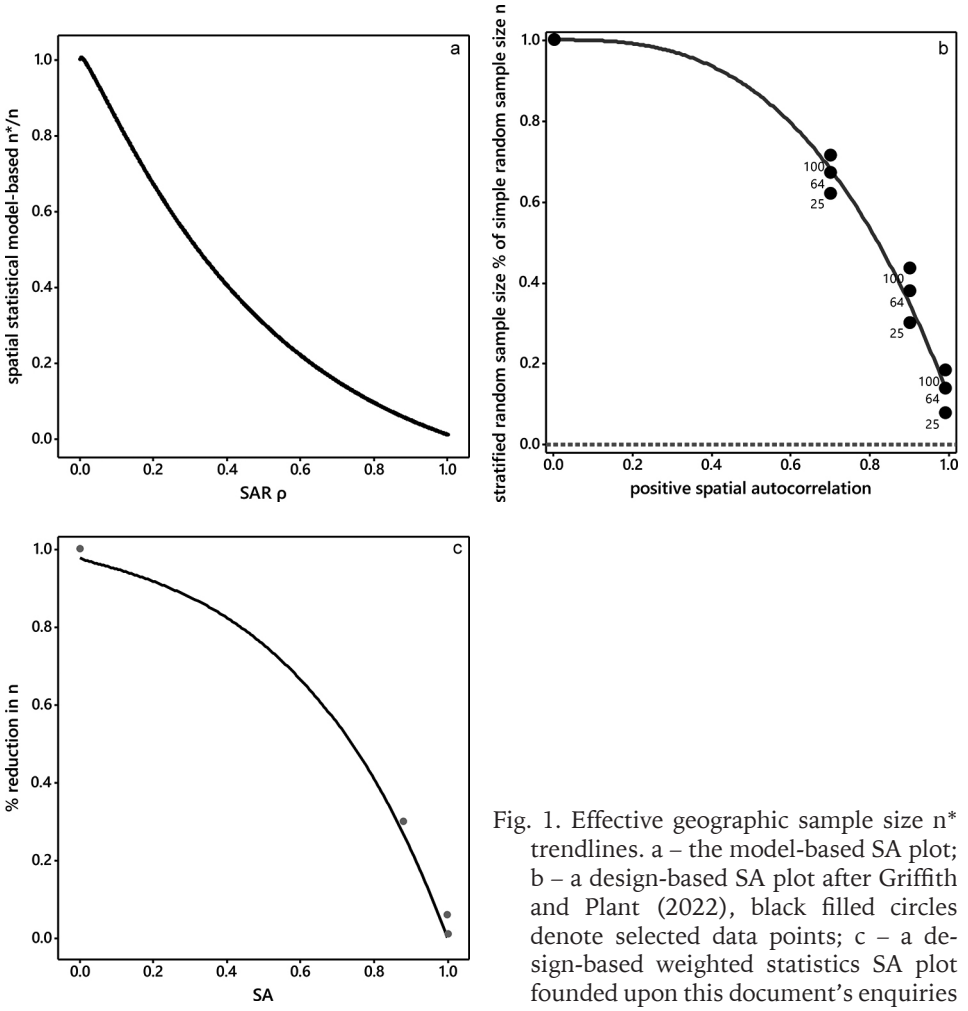

Fig. 1. Effective geographic sample size n* trendlines. a – the model-based SA plot; b – a design-based SA plot after Griffith and Plant (2022), black filled circles denote selected data points; c – a design-based weighted statistics SA plot founded upon this document's enquiries

Griffith and Plant (2022) conceptualize n* for design-based spatial statistical inferences, noting that the well-known and recognized conventional stratified random sampling (STRS) design effect – i.e., the variance ratio between an estimator calculated with a STRS divided by a simple random sampling (SRS) design – required to achieve a given statistical precision level and routinely labelled an effective sample size adjustment, implies that a tessellation STRS (TSTRS) design accounting for SA renders the following definition of n*: the number of equivalent SRS observations affiliated with all unduplicated information content latent in geotagged data isolated by a TSTRS design. Tessellation stratification tends to maximize SA within, while minimizing it between, grid polygons – the statistically most efficient geometric shape for this task being regular hexagons, although squares are nearly as statistically efficient (see the apropos subsequent discussion). As grid polygon size increases, SA effects tend to decrease, with resulting numerical map patterns becoming more random [i.e., the Moran Coefficient (MC) approaches $-1/(n-1)$, and the Geary Ratio (GR) approaches 1, their respective zero-SA null hypothesis values]. Figure 1b portrays an averaged (for n = 25, 64, and 100) bivariate plot reminiscent of typical standard multiple testing circumstances: a complex nonlinear negative exponential declining trendline curve[1] (i.e., concave downward) for this adjusted sample size, again with n* = n when zero, and n* = 1 when perfect positive, SA prevails.

Griffith (2013), Griffith et al. (2016), and Brown et al. (2017) establish that SA is an important factor to account for when determining qualitative – which rarely involve any form of SRS – as well as, by default, mixed methods, geographic sampling designs, crucial for sustainable quality of life features. Latent SA in georeferenced socio-economic/demographic and political household attributes tends to foster moderate-to-strong geographic clustering (i.e., SA) of individual time-specific viewpoints (e.g. Martin, Webster 2020, Fages, Cerda 2022), which, in turn, can promote perception discovery failure and misleading saturation levels with the relatively small sample sizes often employed in qualitative research (Mocănașu 2020, Sebele-Mpofu 2021, Subedi 2021, Mthuli et al. 2022). Exacerbating this situation is a lack of convincing numerical rules or techniques guiding a qualitative researcher pertaining to what constitutes a proper sample size. On the one hand, a quality biography can have a sample size of n = 1 (e.g., the person of interest) On the other hand, larger sample traditions such as grounded theory rarely exceed the desirable quantitative minimum sample size of 30 (Mthuli et al. 2022, Nasheeda 2022), let alone 100 or more. The preceding TSTRS procedure furnishes insights for this setting, too: spacing respondents relatively far apart in geographic space enables social, political, and other preferences to be more independent by diminishing the degree of operational SA, thus remediating discovery failure and misleading saturation levels primarily attributable to SA. In other words, a sample size increases from n* toward n. This more unambiguous context suggests that snowball/chain-referral sampling, with its concomitant geographic dimension (Griffith et al. 2016), warrants special consideration during

[1] $n^* = n\{1/n + [(n-1)/n](1-\rho)^{1/3}\exp(-0.6\rho^{8.8})$, where exp denotes the anti-natural logarithm.

the sample size determination stage of a qualitative research design, even one employed in sustainability monitoring.

Although all three geographic sampling scenarios can pertain to sustainability confirming exercises, this examination explores the second of them while addressing environmental pollution appraisal so critically important to sustainability efforts. In summary, non-zero SA means that a properly designed spatial sample can have a much smaller n than would be required by a classical SRS, to achieve the same level of precision, tremendously reducing pollution and other types of monitoring costs for sustainable development progress and continuation surveillance. This focus on SA and the physical environment automatically alludes to spatial statistical techniques, such as krigging and Moran eigenvector spatial filtering (MESF).

Selected empirical demonstrations

This section summarizes specimen regional development sustainability illustrations that embrace SA while spanning three different geographic resolutions. Poland supplies the empirical geographic landscape. The sustainability data exemplar concerns air pollution by voivodeships, poviats, and gminas across Poland, primarily because air pollution is harmful when it accumulates in the lower ground-level atmosphere beyond certain threshold concentrations. High intensities can cause human and wildlife health problems, as well as agricultural crop (e.g., increase plant susceptibility to diseases, pests, and other environmental stresses) and tree (e.g., reduced tree seedling growth and survivability) damage, among other sustainability compromisers. Therefore, achieving and maintaining sustainability requires environmental monitoring of air pollution to determine the need for and activation of remediation action The subsequent spatial statistical analysis dedicated to conceptualizing this task engages geostatistics (ordinary co-kriging to create a fine resolution geographic landscape), spatial autoregression, and MESF to explore observed geotagged instrument sensed 2023 PM10 (i.e., inhalable aerosol particulate matter no more than 10 micrometers in diameter) measurements, overlaid with administrative district TSTRS to probe n* in its articulation. In addition, because air pollution is a continuous, fluid phenomenon, it naturally encompasses spatial spillovers, and hence SA. This geospatial feature and situation are so obvious that, for example, the United States (US) Clean Air Act's "good neighbor" provision [section 110(a)(2)(D)(i)(I)], dating back decades, requires states to implement controls prohibiting any interstate transport of emissions interfering with downwind states' ability to attain and maintain national ambient air quality standards. A sustainable world needs clean air. The simplest way to achieve this end is to devise a spatial sampling design that builds upon existing administrative surface partitionings to take advantage of governmental jurisdictions. This goal is the reason for presenting foundational work in this section that underpins the ensuing illustrative analysis.

The administrative spatial hierarchy of Poland[2]

In 1999, pursuant to a series of acts passed by its parliament in 1998, and replacing its 49 smaller voivodeships existing from 1975 until then, Poland was reorganized into the following nested administrative areal unit sets: 16 voivodeships [Nomenclature of Territorial Units for Statistics (NUTS) 2]; 373 poviats (similar to counties), of which 65 were cities—these have been consolidated into 369 areal units for purposes of this paper; and, 2,489 gminas (similar to municipalities), 11 of which were in the capital city, Warsaw, with another 302 being urban (including 107 cities), 638 being mixed rural-urban, 1,537 being rural, and one retained for the purposes of this article that had been abolished since the utilized shapefile's creation (Fig. 2). This scheme was revised in January of 2022 to the following: 16 voivodeships; 314 poviats, of which 66 are cities; and, 2,477 gminas, of which, respectively, 302, 662, and 1,513 are urban, mixed rural-urban,

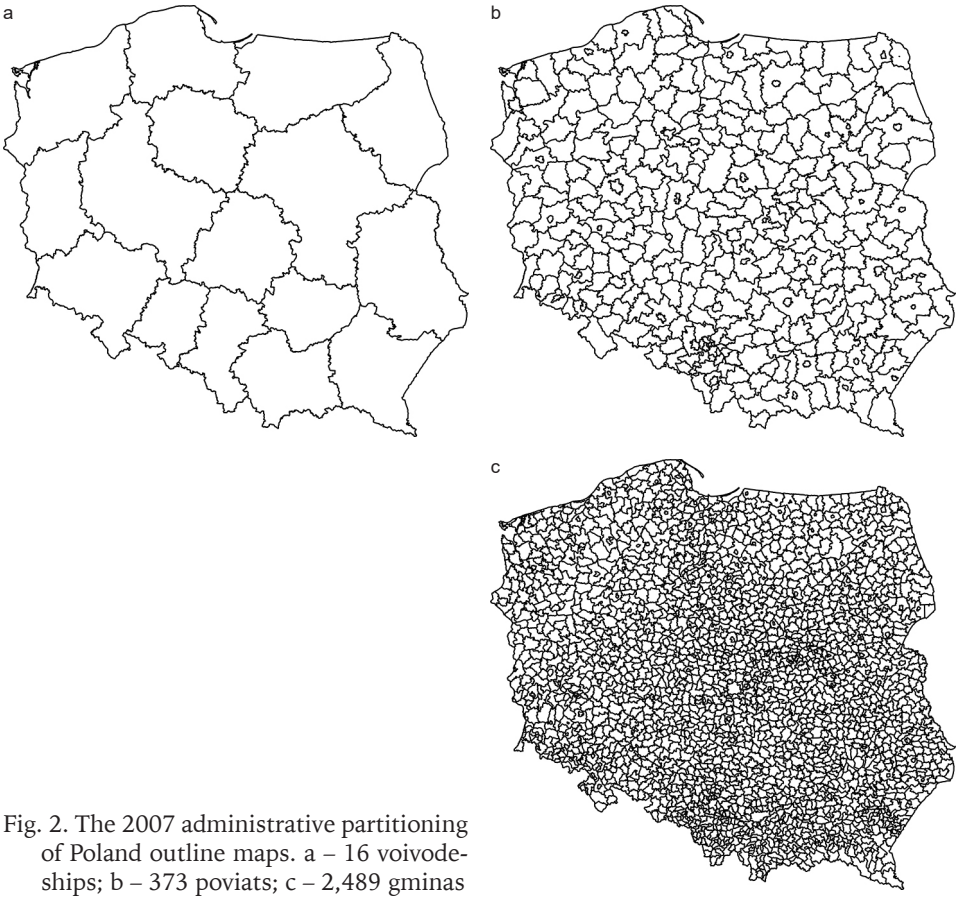

Fig. 2. The 2007 administrative partitioning of Poland outline maps. a – 16 voivodeships; b – 373 poviats; c – 2,489 gminas

[2] See Bradley and Zaucha (2017), in which the word municipalities rather than gminas appears in conjunction with poviats and voivodeships terminology.

and rural. Research finalized for this narrative did not employ this newer delineation to avoid any possible beta-testing nuisances.

Figure 2 portrays the three geographic surface partitionings employed to implement TSTRS simulation experiments investigating n* within the context of design-based inference. It parallels the varying size hexagons Griffith (2005) uses to study n* within the context of model-based inference. Furthermore, because of its coarseness, the voivodeship surface partitioning (Fig. 2a) serves as the testbed geography for special sustainability air pollution monitoring probes.

The high-resolution elevation covariate for co-kriging

Hart et al. (2013), among others, note that common covariates used to analyze ambient air pollution include elevation, and population density. Because a high resolution version (i.e., 311,448 geographic pixels covering Poland's 312,608 km^2 of surface area) of elevation via a digital elevation model (DEM) is readily accessible (via https://www.eea.europa.eu/data-and-maps/data/digital-elevation-model-of-europe), the numerical analysis for this effort adopts it as a co-kriging covariate (Fig. 3a). This DEM under-represents the official physical area of Poland by 11,127 km^2. This discrepancy of roughly 3.5% relegates the analyses summarized here to an insightful exploratory role. Meanwhile, the co-krigging interpolation surface constitutes 309,632 raster points, likewise slightly too few by nearly 3,000 for a one-to-one correspondence matching with an invisible 1 km^2 quadrat mesh overlaying the country.

PM10 data described in the next section were retrieved from a publicly available archive (i.e., https://discomap.eea.europa.eu/App/AirQualityStatistics/index.html). The number of monitor stations networking Poland is rather meager,

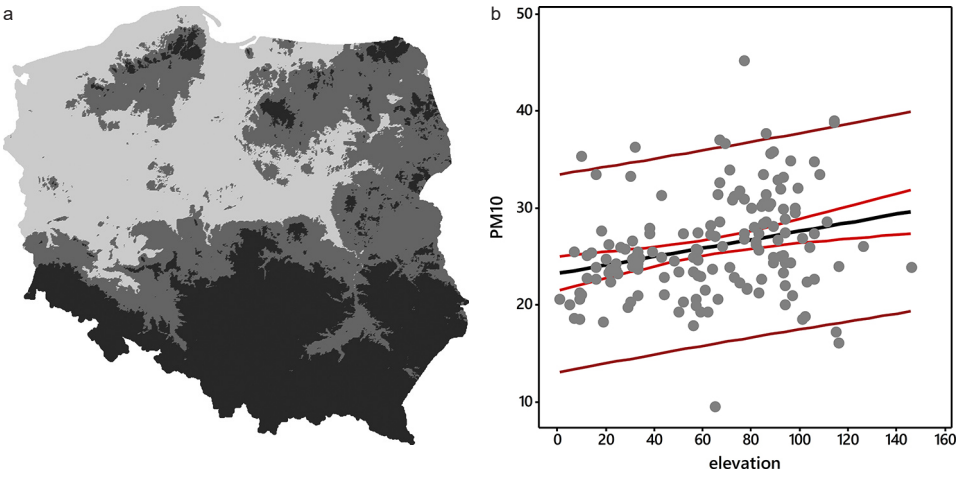

Fig. 3. Poland DEM. a – 1km-by-1km tertile choropleth DEM map; 0–183 values range; elevation directly proportional to grayscale; b – scatterplot portraying the association between elevation and PM10 (gray filled circles); black, red, and burgundy lines respectively denote the linear trend, 95% confidence intervals, and 95% prediction intervals

numbering only 154. This DEM furnishes a covariate (see Fig. 3b; elevation accounts for roughly 7% of the country's PM10 geographic variation) for co-kriging to couple with SA in order to interpolate PM10 to a fine pixel grid, bolstering the analysis capabilities undertaken for this current work.

The geographic distribution of PM10 across Poland

Breathable air quality is a sustainable regional development concern: as air pollution worsens, public health risks tend to increase. Locational indices quantifying it often incorporate concentration values for: both PM2.5 and PM10 particulate

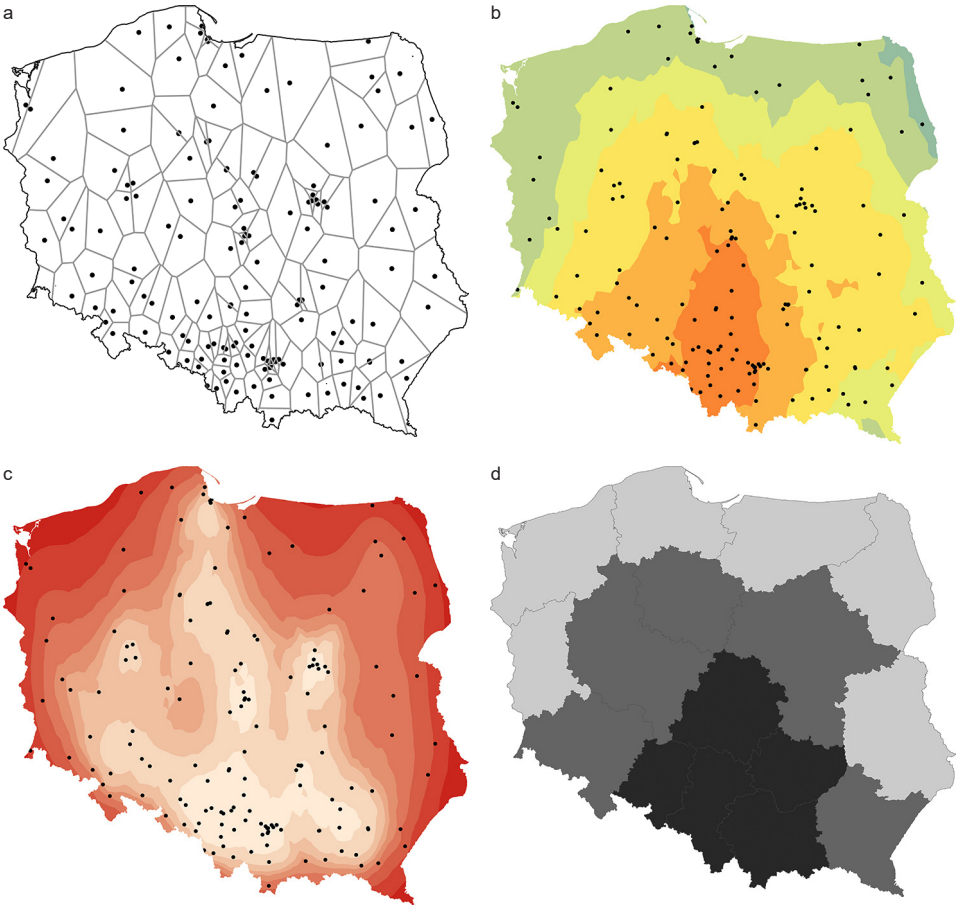

Fig. 4. Monitoring station sites, and tertile choropleth maps depicting the geographic distribution of PM10. a – Thiessen polygon surface partitioning based upon monitoring station sites; b – a K-Bessel function co-krigged surface; black dots denote sample locations; c – the krigging prediction error surface; black dots denote sample locations; increasing prediction error is directly proportional to the darkness of red; d – voivodeships resolution local PM10 maxima

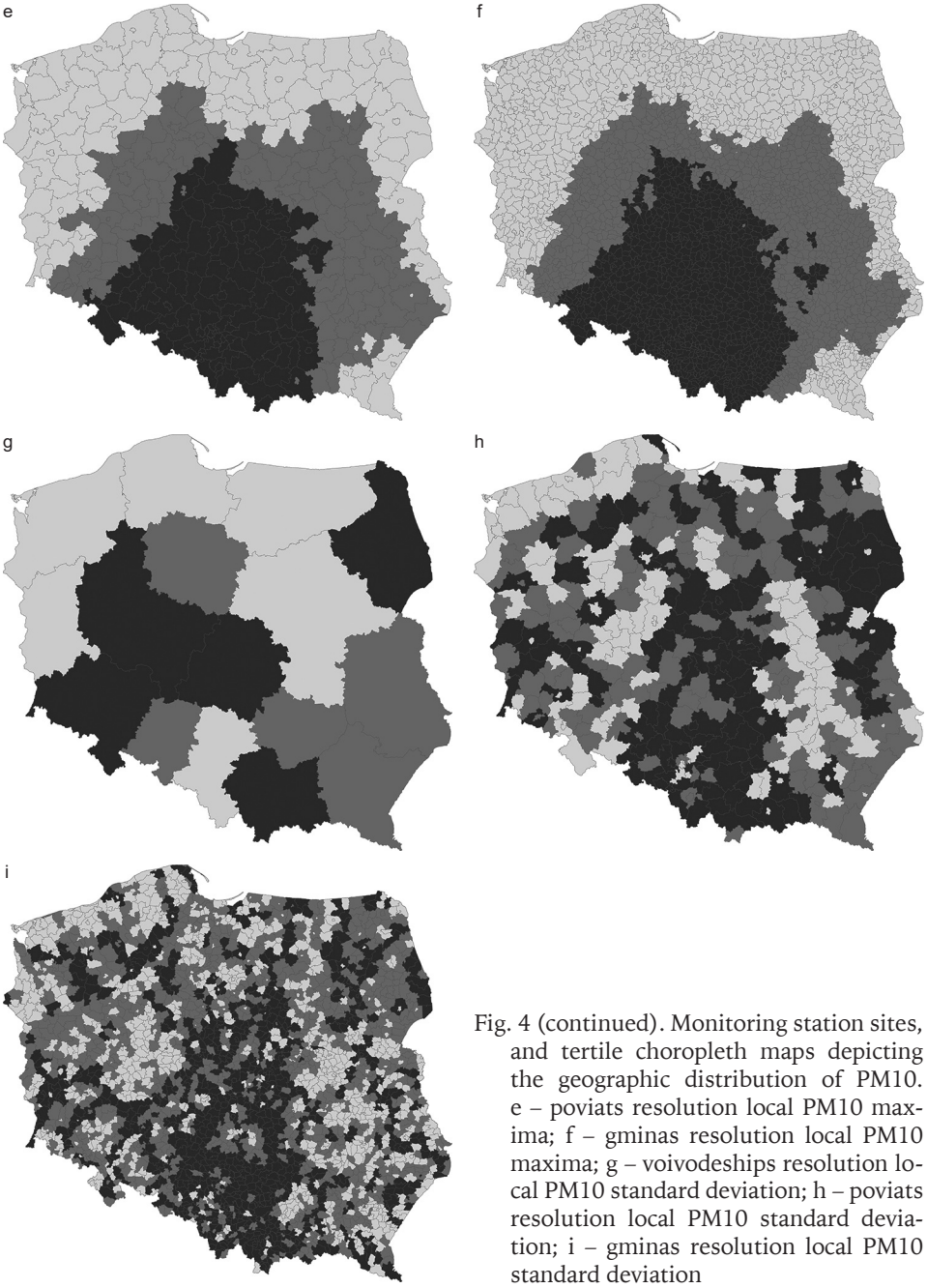

Fig. 4 (continued). Monitoring station sites, and tertile choropleth maps depicting the geographic distribution of PM10. e – poviats resolution local PM10 maxima; f – gminas resolution local PM10 maxima; g – voivodeships resolution local PM10 standard deviation; h – poviats resolution local PM10 standard deviation; i – gminas resolution local PM10 standard deviation

matter, ozone (O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and carbon monoxide (CO), among other pollutants. Meanwhile, factors impacting constructed effluence scales, sometimes even inducing small-area geographic varia-

tion, include: wind speed and direction; terrain/topography/elevation; precipitation, smoke plumes; vehicular and other traffic; and, sundry fine dust emitters. Government agencies attempt to sustain clean air by interventions curtailing or scrubbing airborne pollution, suppressing it in at least some places, although its fluidity allows it to backfill from nearby unhampered spots. Wind currents circulate aerial contamination, relocating it from one locality to another. These competitive types of displacement mechanisms imply that SA for these phenomena should contain a mixture of positive and negative components. PM10 is the contaminant examined in this subsection. Three-year-average PM10 data were obtained[3] for 154 monitoring stations (Fig. 4a), co-krigged (using ordinary kriging and DEM elevation as a covariate) with a K-Bessel function semivariogram model (it links to an SAR model; its effective range estimate is 8.8km) to interpolate the data at an additional 309,632 sites across the country (Fig. 4b, c shows the accompanying prediction error map). Following this data generation was an aggregation of interpolated PM10 values as small geographic area averages for Poland's three different polygon resolutions; Figures 4d–f portray the geographic distribution of local PM10 maxima (this often is a target when checking for clean air) by administrative areal unit aggregations, whereas Figures 4g–i visualize within areal unit PM10 variability.

Preliminary spatial statistical analysis of the aggregated average PM10 geographic distribution across Poland

Because the sample of monitoring places procured for this paper is not random, spatial statistical analyses require model-based inference, whose soundness requires a minimal of model assumption violations. A critical feature is that the data conform to a bell-shaped (i.e., normal or Gaussian) frequency distribution. Accordingly, Figure 5 supplies diagnostic normal quantile plots: Figure 5a represents the raw data, whose Shapiro-Wilk test statistic H_0 probability is 0.0032; Figure 5b represents the best Box-Cox transformation, with a power exponent of zero coupled with a translation parameter of roughly 26, which essentially over-shrinks the magnitude of the upper outlier while marginally exacerbating the extremeness of the lower outlier, increasing the preceding Shapiro-Wilk statistic probability to 0.0405; and, Figure 5c characterizes the best Manley transformation, which effectively over-shrinks the magnitude of the lower outlier while slightly exacerbating the extremeness of the upper outlier, increasing the aforementioned Shapiro-Wilk statistic probability to 0.0595. Complications introduced (e.g., back-transformation hurdles) into an analysis by employing either of these transformations seems to far outweigh their meager benefits (see Fig. 5a). Thus, the raw data remain the target of analytical treatment. Furthermore, because these PM10 data may well harbor a positive-negative SA mixture, the enhanced specification of choice here is a composite SAR and spatial moving

[3] Source: https://discomap.eea.europa.eu/App/AirQualityStatistics/index.html.

average (MA) model (i.e., SAR-MA), complemented by an MESF specification as an affirmatory tool.

Griffith et al. (2022a,b) and Griffith (2023) demonstrate the presence of positive-negative SA mixtures in a wide variety of real-world georeferenced data. These publications imply that it conceivably is a universal sensation. A Thiessen polygon partitioning of Poland based upon monitoring station locations reveals the same categorization for PM10 (Table 1; recalling that the MA coefficient sign is the opposite of its SA nature). Although the SAR-MA residuals fail to perfectly, or nearly so, conform to a bell-shaped curve, their Shapiro-Wilk diagnostic statistic supplementing their reported K-S statistic is 0.96 (re the Shapiro-Wilk diagnostic statistic for Fig. 5a is 0.97), which is not very far from 1. Hence, their normal approximation seems to lack a dramatic substantive deviation from a bell-shaped curve (e.g., its skewness is well within a six-sigma confidence interval).

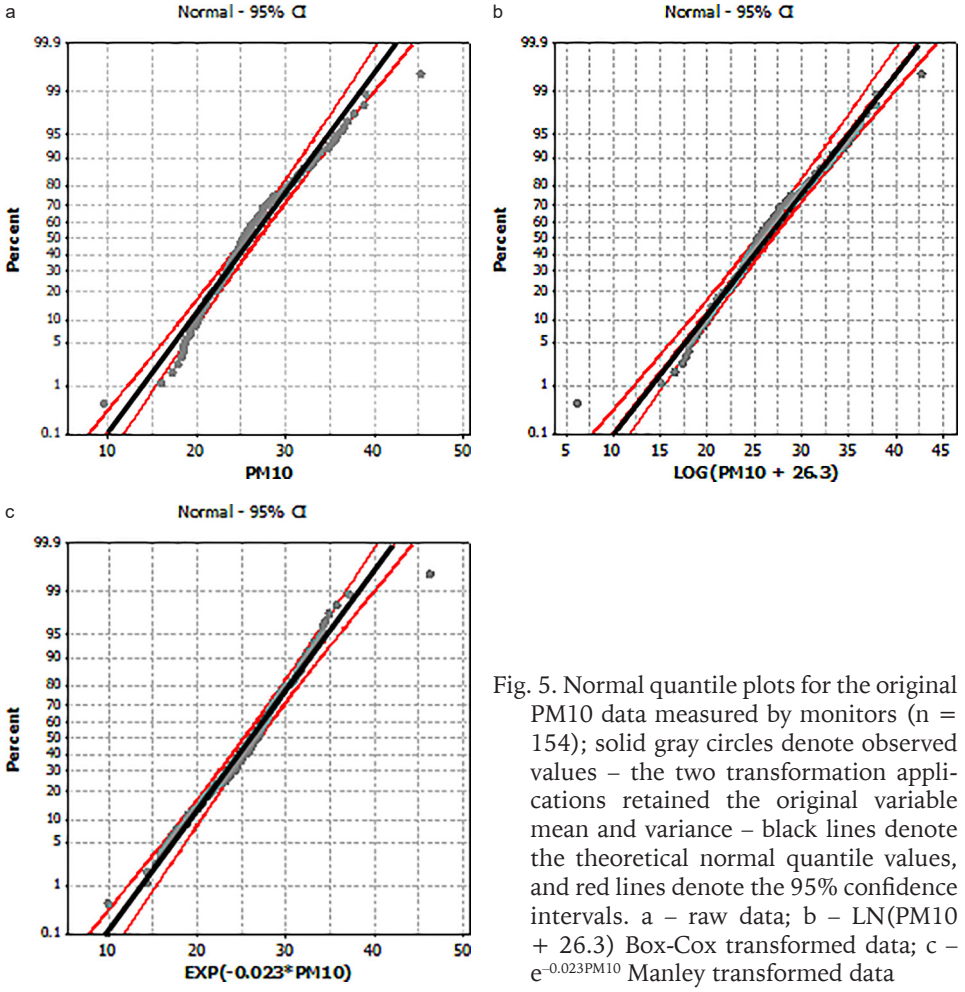

Fig. 5. Normal quantile plots for the original PM10 data measured by monitors (n = 154); solid gray circles denote observed values – the two transformation applications retained the original variable mean and variance – black lines denote the theoretical normal quantile values, and red lines denote the 95% confidence intervals. a – raw data; b – LN(PM10 + 26.3) Box-Cox transformed data; c – $e^{-0.023PM10}$ Manley transformed data

Meanwhile, K-Bessel function krigging markedly smooths interpolated PM10 values, not surprisingly removing their latent negative SA component. Regardless, positive SA constituents across Poland's geographic resolutions are consistent, resembling those for most remotely sensed images (e.g., is in the 0.9+ range), a consequence also conspicuously disclosed by a visual inspection of Figure 4b.

Table 1. Selected summary spatial autoregressive estimation outcomes for 2023 PM10

Geographic resolution	n; points (points/polygon‡)	SAR	SAR-MA				SA index	Residual SA p†
		$\hat{\rho}$	$\hat{\rho}$	$r_{\hat{\rho}\hat{\theta}}$	$\hat{\theta}$	K-S p†		
Thiessen polygon	154; 154 (1; 1; 1)	0.475 (0.092)	0.976 (0.028)	0.86	0.866 (0.081)	< 0.01 (0.10)	MC	0.32
							GR	0.43
Gmina	2,489; 309,507‡ (2; 124; 668)	0.999 (0.001)	0.997 (0.001)	0.29	–0.167 (0.040)	< 0.01 (0.06)	MC	0.39
							GR	0.25
Poviat	369; 309,632 (14; 839; 3,024)	0.998 (0.002)	0.995 (0.004)	0.29	–0.239 (0.72)	< 0.01 (0.06)	MC	0.14
							GR	0.31
Voivodeship	16; 309,632 (9,030; 19,352; 35,514)	0.878 (0.113)	0.651 (0.320)	0.68	–0.841 (0.364)	> 0.15 (0.14)	MC	0.20
							GR	0.26

‡semicolons separate the minimum, arithmetic mean, and maximum point #s in each sequence
†p denotes Type I error probability; n > 2000 dictates using K-S (Kolmogorov-Smirnov)
‡unaligned peripheral concomitant gmina, poviats, and voivodeship borders omit 125 points (see Appendix A).
NOTE: parenthetical values denote parameter estimate standard errors or K-S Type I p

Table 2 tabulations are in the spirit of confirmatory spatial statistical data analysis, utilizing generalized linear model (GLM) theory coupled with a beta random variable – DEM values are integers ranging from 0 to 255, with division by 255 converting them to proportions in the unit interval. The statistical significance subset eigenvector selection criterion was a Bonferroni-adjusted $\alpha = 0.10/k$, where k is the number of candidate eigenvectors, sometimes changing by increasing it – at most to the common stepwise default value 0.20/k – as a trade-off to minimize residual SA as measured jointly by the MC and GR indices. Generation of the interpolated PM10 values for all three Polish administrative district resolutions involved only this phenomenon's positive SA. However, the co-kriging DEM covariate is a potential negative SA pathway that allows a mixture to emerge. Table 2 entries endorse an anticipated bias toward positive SA for these three surface partitionings. The coarsest resolution (i.e., voivodeships) retains the unmodified Bonferroni-adjusted α. MESF constructs a strictly positive SA eigenvector spatial filter (ESF) that accounts for roughly 90% of the geographic variation in average PM10 values in this case; its model-based inference diagnostics authenticate a sound inferential basis for it. The intermediate degree of coarseness (i.e., poviats) does not fare as well: although its purely positive SA ESF accounts for nearly as much geographic variation, its inferential basis is not very sound. In this second case, estimation involved weighting by areal unit PM10 variance to avoid selection of all candidate eigenvectors (i.e., 250) The use

of smoothed krigged values necessitates this weighting because a lack of residual stochastic error ascribable to smoothing biases coefficient standard errors downward, unveiling false positives. In addition, the next selection by a modified Bonferroni adjustment is a negative SA eigenvector, counter to the prevailing K-Bessel function spawned expectation. Pragmatically the same description holds for the gminas resolution, which has a candidate set of 1,738 eigenvectors. One difference is that its ESF contains negative SA eigenvectors before any alterations are made to its Bonferroni adjustment. Liberalizing the selection criterion here incrementally adds three more positive SA eigenvectors without substantially improving its diagnostic statistics.

Table 2. Selected PM10 MESF estimation and diagnostic results across geographic resolutions

Feature	Thiessen polygon	Gmina	Poviat	Voivodeship
stepwise selection α	0.12/133	0.10/1738	0.10/250	0.10/12
# PSA† vectors	7	76	32	3
# NSA‡ vectors	2	3	0	0
linear regression R^2	0.3277	0.8584	0.8621	0.9252
ESF_{PSA} R^2	0.2664	0.0109	0	0
ESF_{NSA} R^2	0.0613	0.8693	0.8621	0.9252
$MC_{residual}$ (H_0 p value)	0.19	< 0.0001	< 0.0001	0.33
$GR_{residual}$ (H_0 p value)	0.08	< 0.0001	< 0.0001	0.40
$S\text{-}W_{residual}$ (H_0 p value)	0.01	< 0.0001	< 0.0001	0.05

†PSA denotes positive SA; ‡NSA denotes negative SA.

A geospatial simulation sampling experiment: toward a better design-based n* understanding

The preceding section surveys pertinent spatial statistical methodology for monitoring air pollution in a geographic landscape for sustainability progress indexing and/or maintenance. The simulation experiments in this section build upon an adaptation of the TSTRS protocol outlined in Griffith and Plant (2022), who emulate coauthored work by Overton and Stehman (1993, 1996), and reviewed in the preceding section. Using interpolated rather than observed values magnifies the similarity of nearby PM10 quantities (e.g., see Griffith and Liau 2021): an aftermath attributable to an artificial absence of stochastic noise (a la their being conditional expectations) coupled with exceptionally high positive SA (a la their being predictions; see Tables 1 and 2). This situation reinforces the argument championing an exploitation of redundant information in geospatial data (i.e., SA), while offering an insightful testbed for augmenting Figures 1a and 1b.

Particularly because of compactness (i.e., maximum interpoint distance minimization partiality) and pervasive SA, the ideal tessellation is a regular polygon mesh, the two most popularly implemented ones being the square and the

hexagon (e.g., Birch et al. 2007). This latter plane geometry form is the preferred tessellation for spatial sampling purposes (e.g., Stough et al. 2020), in part because it reduces sampling bias arising from grid shape edge effects. For kriging of geotagged phenomena, hexagons yield the lowest average and maximum standard errors (Olea 1984), favored and desirable statistical properties. Such spatial sampling advantages prompted the US Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP) scientists to devise a 40-square-kilometer hexagonal tessellation covering the entire coterminous US (Carr et al. 1992, p. 235, White 1992) for spatial sampling purposes. The US Geological Survey has released this regular hexagonal grid for public use. Interestingly, this regular hexagonal grid predilection continues today with, among others, forestry inventories (e.g., Frank, Monleon 2021).

Although Webster and Oliver (2007, p. 188) graphically demonstrate that the most precise TSTRS design is the one conducted using a regular hexagonal grid, which geographically spreads out nearer sample points to the greatest extent possible, their sketch also shows that a square grid has almost the same precision. This comparison is fortunate because sampling plan designers generally tend to approve harnessing the more application-friendly square grid mesh. Furthermore, if a georeferenced random variable has a high degree of positive SA, then two nearby locations exhibit a strong tendency to deliver exceedingly similar, if not identical, values, an important spatial resampling and sampling distribution construction property (re the best TSTRS plan selects only one draw from each polygon, which SA helps guarantee is locally representative), implying that attribute sampling variance directly relates to the average randomly chosen pairwise interpoint distance within polygons. Bäsel (2021) summarizes already known, as well as derives new, statistical moments for the distribution of this distance genre within regular polygons. For such squares and hexagons enclosing the same area, namely $\frac{3\sqrt{3}}{2} r^2$, for a hexagon with radius $r > 0$ and a square with sides $r\sqrt{\frac{3\sqrt{3}}{2}}$, the random straight line segment length variance for this general square is given by

$$\sigma^2_{SQ} = r^2\{25[LN(1 + \sqrt{2})]^2 + 10(2 + \sqrt{2})LN(1 + \sqrt{2}) + 4\sqrt{2} - 69\} / (50\sqrt{3}) \quad (1),$$

whereas the parallel variance for this general hexagon is given by

$$\sigma^2_{HEX} = r^2\{-36[LN(2 + \sqrt{3})]^2 + (2052LN[3] + 1008\sqrt{3} - 336)LN(2 + \sqrt{3}) - 29241[LN(3)]^2 + (9576 - 28728\sqrt{3})LN(3) + 86048 + 4704\sqrt{3}\} / 129600 \quad (2).$$

The relative efficiency (i.e., an unbiased parameter estimate's comparative degree of concentration reflecting its precision as sample size increases) is $\sigma^2_{SQ} / \sigma^2_{HEX} \approx (0.38811/0.39963)^2 \approx 0.94$, consigning a square-for-hexagon substitution to the

same exchange category as classical nonparametric, vis-à-vis parametric, statistics. Simplicity together with this somewhat modest quintessential theoretical statistical property difference is a primary reason Griffith and Plant (2022) chose the square, rather than the hexagon, tessellation for their simulation experiments.

A relative efficiency exploration of a Polish administrative polygon surface partitioning

Bäsel (2021) presents the necessary statistical moments to posit formulae like (1) and (2) for a sequence of regular polygons, beginning with a triangle, then going to a square, a pentagon, a hexagon, and so forth through a dodecagon, finally ending with the circle as a convergent limit to the number of regular polygon sides going to infinity. Now the remaining unanswered question asks whether or not relative efficiency insights are possible about TSTRS built upon an even more convenient, but almost always involving an irregular polygon shapes tiling, administrative districts surface partitioning.

Table 3 tabulates selected illustrative summary statistics to help illuminate this situation. Bhardwaj and Kumar (2019) discuss a simple shape index – whose formula is $\frac{\text{perimeter}}{\sqrt{4\pi(\text{aea})}}$, which equals 1 for a circle, $\frac{2}{\sqrt{\pi}} \approx 1.13$ for a square, and approximately 1.36 and 2.10, respectively, for a rectangular shaped region displaying the equivalent silhouette of 3.5 juxtaposed squares and a trapezoidal shaped region exhibiting the equivalent silhouette of 2.5 juxtaposed squares sandwiched between two half-square right triangles having the same orientation, both housing the same aforementioned area. Figure 2a divulges that the Polish voivodeships do not approximate either circles or squares, a contention corroborated by their respective shape indices appearing in Table 3. In general, they most closely resemble distorted trapezoids, which, once more, their respective shape indices signal, after appropriately adjusting their empirical boundary length

Table 3. A preliminary strata assessment of a coarse partitioning of Poland into 16 provincial level administrative units

Voivodeship	Area	Shape index	Relative efficiency†	Voivodeship	Area	Shape index	Relative efficiency†
Dolnośląskie	19356	1.40	1.31	Podlaskie	20506	1.10	1.27
Kujawsko-pomorskie	18246	1.25	1.09	Pomorskie	18893	1.34	1.35
Lubelskie	24410	1.24	1.19	Warmińsko-Mazurskie	24797	1.07	1.23
Lubuskie	13965	1.29	1.34	Wielkopolskie	29692	1.44	1.37
Mazowieckie	35514	1.49	1.36	Zachodniopomorskie	22993	1.45	1.54
Małopolskie	14355	1.28	1.28	Śląskie	11788	1.49	1.40
Opolskie	9030	1.31	1.17	Świętokrzyskie	11220	1.18	1.16
Podkarpackie	16988	1.22	1.33	Łódzkie	17878	1.24	1.16

†this comparison is between the voivodeship and trapezoid tessellation sampling variances

measurements for excess sinuosity (e.g., see Chen 2020, Banaszak et al. 2023). Accordingly, their within location average random distances increase by a factor of nearly 2.2, perceptibly diminishing SA impacts upon attribute sampling distributions vis-à-vis regular square and hexagon strata, whereas their variance increases by slightly more than a factor of 6. Table 3 reports relative efficiency calculations for these sixteen coarse geographic resolution areal units when compared with their trapezoidal-shaped irregular polygon counterparts. Its arithmetic average is about 0.90, again, comparable to the preceding square-to-hexagon strata comparison. Of course, a trapezoid-square or hexagon comparison would expose much lower relative efficiency scores.

Design-based regional arithmetic average sampling variance estimation: pseudo-equiprobability sampling

A critical automatic benefit of a regular polygon TSTRS plan is equiprobability. This is why the preceding squares, hexagons, rectangles, and trapezoids all need to have the same area, as well as why some polygons in and of themselves are incapable of exhaustively partitioning a surface. Thus, any shape alterations (e.g., replacing squares with trapezoids) impact only geographic coverage, and hence SA effects, essentially maintaining the goal of more precise information about, and hence more accurate estimates for, a parameter of interest (e.g., a geographic landscape-wide global arithmetic mean). If this uniform distribution of sampling probabilities is unattainable, then legitimate statistical computations can incorporate sampling weights (the inverse selection probabilities for each observation) to convert varying choice probabilities to a single constant. In other words, an application of sampling weights enables a reconfiguration of a sample so that it behaves as if it were a simple random draw from its parent population. After all, the principal objective of sampling is to calculate a descriptive statistic that accurately measures its true parameter value in its parent population. In other words, a paramount purpose of adopting sampling weights is to overturn any sampling distribution distortions imposed by using differential sampling probabilities.

If polygons have varying areas, A_i (i = 1, 2, …, n), then the probability of sample inclusion when drawing one observation from each polygon is the known fixed measure $p_i = \frac{A_i}{\sum_{i=1}^{i=n} A_i} > 0$ for areal unit i. This fraction reduces to a single value for regular polygons, satisfying the aforementioned equally likely principle. In contrast, if irregular polygon strata vary in size, the ensuant nonconstant probabilities require usage of a weighting scheme in which the weights are inversely

proportional to these sampling probabilities, namely, $\left(\frac{A_i}{\sum_{i=1}^{i=n} A_i}\right)^{-1}$. Accordingly, the weighted mean for some random variable Y is given by

$$\hat{\mu}_w = \sum_{i=1}^{n} \left[\frac{y_i}{\left(\frac{A_i}{\sum_{i=1}^{i=n} A_i}\right)^{-1}} \right] = \sum_{i=1}^{n} \left[\frac{A_i y_i}{\sum_{i=1}^{i=n} A_i} \right], \quad (3)$$

which has an expected value of μ because each of the n A_i weights is a fixed amount (i.e., they do not change between repeated sampling exercises with a given geographic landscape). This statistic's affiliated pertinent quantity that is in keeping with the goals of this paper is its sampling variance, $\sigma^2_{\hat{\mu}_w}$, an expression that differs from the weighted variance companion of estimator (1) often given by

$$\hat{\sigma}_w = \sum_{i=1}^{n} \left[\frac{A_i (y_i - \hat{\mu}_w)^2}{\sum_{i=1}^{i=n} A_i} \right] \frac{n}{n-1}, \quad (4)$$

which has an expected value of σ^2.

Meanwhile, the sampling variance of the estimated weighted mean [i.e., equation (3)] is given by

$$\sigma^2_{\hat{\mu}_w} = \frac{\sigma^2 \sum_{i=1}^{i=n} A_i^2}{\left(\sum_{i=1}^{i=n} A_i\right)^2} = \frac{\sigma^2}{\left[\frac{\left(\sum_{i=1}^{i=n} A_i\right)^2}{\sum_{i=1}^{i=n} A_i^2} \right]}, \quad (5)$$

where $\frac{\left(\sum_{i=1}^{i=n} A_i\right)^2}{\sum_{i=1}^{i=n} A_i^2}$ denotes a more traditional effective sample size, n_{eff}. This factor certifies that an irregular polygon TSTRS plan in the presence of SA incurs a modification to Figure 1b (hence, Fig. 1c).

Polish 2023 PM10 resampling experiments: implications for effective geographic sample size n*

Table 4 tabularizes the most germane simulation experiment output. First, as already commented with regard to formula (3), the weighted arithmetic average equals its unweighted mean population parameter (except for rounding error). Second, formula (5) implies $n_{eff} \leq n$ (the inequality part being attributable to weighting to compensate for nonconstant selection probabilities), which materializes here by a sizeable amount (i.e., 70% to 89%). The estimated formula (4) result, after being adjusted for unequal probabilities sampling, still reflects an overwhelming SA impact (see Table 3). Sampling only one observation from each tessellation strata helps mediate this complication within a sample, but not between samples, in practice.

Table 4. Simulation output for an irregular polygon mesh TSTRS plan; 10,000 replications

Geographic resolution	n	μ & $\tilde{\mu}_w$	Central limit theorem: $\sigma / \sqrt{n}$	Median # km² points	n_{eff}	% of n	$\hat{\sigma}_{\tilde{\mu}_w}$
Voivodeships	16		2.244/4 ≈ 0.5610	18,570	14.3	89.1	0.2433
Poviats	369	24.7	2.244/19.2 ≈ 0.1168	768	258.5	70.0	0.0241
Gminas	2.489		2.243/49.9 ≈ 0.0450	110	1761.0	70.8	0.0050

In conclusion, Figure 1a implies a conjecture stating that weighting to rectify design-based unequal sample selection probabilities leaves the relationship between SA and effective geographic sample size, n*, fundamentally untouched, although the resulting trendline curvature may be steeper. Clearly, considerably more future research needs to address this topic. Nevertheless, the cardinal finding is that concocting idealized tessellations to execute a proper spatial sampling design is needless and preventable, offering considerable civic budgetary savings when performing sustainability monitoring. Existing administrative units, accompanied by their legislative authority, are easily adaptable for this purpose. The air pollution case furnishes a powerful exemplification of this contention.

Conclusions and Implications

Fruitful sustainable regional development compels continuous monitoring and surveillance of the human condition, natural resources (especially energy) compiling/consumption, agronomy, and environmental quality, all endeavors amenable to spatial sampling because it is bursting with SA. Accordingly, successful sustainability strategies warrant taking advantage of this latent SA, particularly to minimize the number and complexity – and thus operating costs – of geographic sampling networks. This paper encapsulates such an approach for pollution monitoring, a most relevant sustainability illustration, expounding upon ways that SA can stimulate efficiencies and furnish cost savings when spatially sampling

a geographic landscape with the use of TSTRS techniques. One recorded finding is that administrative unit point sampling behavior may resemble that of trapezoidal tessellations. Verification of this claim could inspire numerous theoretical studies cloning those already existing for regular hexagon and square tessellations. Another is that SA impacts, which instigate very similar numerical results when repeatedly sampling attributes in a compact local area, are about twice as pronounced as those produced by resorting to classical weighted statistics to counterbalance a practicality-obliged implementation of varying sample selection probabilities. New evidence reviewed in this paper fortifies that already provided by Griffith and Plant (2022), solidifying a general conclusion about the importance of SA in both design- and model-based inference. This composition also is another testimony to the authenticity of a near-universal inclination by georeferenced attributes to possess positive-negative SA mixtures. Because its use of interpolated/imputed (i.e., krigged) data is somewhat obfuscating in this investigation (e.g., suppression of natural variation), far more environmental datasets other than air pollution merit such scrutiny before appending their kind to the already uncovered and steadily growing cluster of mixture variables. On the one hand, air pollution is a natural phenomenon for having unmistakable positive SA because of its physical fluidity. On the other hand, wind-forced circulation, remediating interventions, and sundry anthropomorphic activities can cause it to compete for territory within itself as well as with nearby lesser-polluted regions, the hallmark of negative SA. All in all, then, the perceptive conclusion reached through this air pollution narrative argues that SA does matter for the broader context of sustainable regional development planning and ongoing evaluation.

References

Banaszak M., Górnisiewicz K., Nijkamp P., Ratajczak W. 2023. Fractal dimension complexity of gravitation fractals in central place theory. Scientific Reports, 13(1): 2343 (12 pp.).

Bäsel, U. 2021. The moments of the distance between two random points in a regular polygon (https://arxiv.org/abs/2101.03815).

Bhardwaj G., Kumar A. 2019. The comparison of shape indices and perimeter interface of selected protected areas especially with reference to Sariska Tiger Reserve, India. Global Ecology and Conservation, 17: e00504 (10 pp.). https://doi.org/10.1016/j.gecco.2018.e00504

Birch C., Oom S., Beecham J. 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecological Modelling, 206(3–4): 347–359.

Bradley J., Zaucha J. (eds.) 2017. Territorial Cohesion: A Missing Link Between Economic Growth and Welfare – Lessons from the Baltic Tiger. Uniwersytet Gdański, Gdańsk.

Brown T., Wood J., Griffith D. 2017. Using spatial autocorrelation analysis to guide mixed methods survey sample design decisions. Journal of Mixed Methods Research, 11(3): 394–414.

Brundtland G. 1987. Our Common Future: Report of the World Commission on Environment and Development.: United Nations UN – Dokument A/42/427, Geneva.

Carr D., Olsen A., White D. 1992. Hexagon mosaic maps for display of univariate and bivariate geographical data. Cartography and Geographic Information Systems, 19(4): 228–236.

Chen Y. 2020. Two sets of simple formulae to estimating fractal dimension of irregular boundaries. Mathematical Problems in Engineering: Article ID 7528703 (15 pp.). https://doi.org/10.1155/2020/7528703

Devuyst D. 2000. Linking impact assessment and sustainable development at the local level: the introduction of sustainability assessment systems. Sustainable Development, 8(2): 67–78.
Dočekalová M., Kocmanová A., Koleňák J. 2015. Determination of economic indicators in the context of corporate sustainability performance. Business: Theory and Practice/Verslas: Teorija ir Praktika, 16(1): 15–24.
Fages D., Cerda M. 2022. Migration and social preferences. Economics Letters, 218: 110773 (6 pp.).
Frank B., Monleon V. 2021. Comparison of variance estimators for systematic environmental sample surveys: considerations for post-stratified estimation. Forests, 12(6): 772 (20 pp.). https://www.mdpi.com/1999-4907/12/6/772
Graymore M., Sipe N., Rickson R. 2008. Regional sustainability: How useful are current tools of sustainability assessment at the regional scale? Ecological Economics, 67(3): 362–372.
Griffith D. 2005. Effective geographic sample size in the presence of spatial autocorrelation. Annals of the Association of American Geographers, 95(4): 740–760.
Griffith D. 2013. Establishing qualitative geographic sample size in the presence of spatial autocorrelation. Annals of the Association of American Geographers, 103(5): 1107–1122.
Griffith D. 2015. Approximation of Gaussian spatial autoregressive models for massive regular square tessellation data. International Journal of Geographical Information Science, 29(12): 2143–2173.
Griffith D. 2020. A family of correlated observations: From independent to strongly interrelated ones. Stats, 3(3): 166–184.
Griffith D. 2023. Spatial autocorrelation in geospatial disease data: An important global epidemiologic/public health assessment ingredient. Transactions in GIS, 27(3): 730–751.
Griffith D., Agarwal K., Chen M., Lee C., Panetti E., Rhyu K., Venigalla L., Yu X. 2022a, b. Geospatial socio-economic/demographic data: the masking of negative by, and existence of, mixtures of spatial autocorrelation in georeferenced data: Part I and Part II. Transactions in GIS, 26(1): 72–87 & 26(1): 88–99.
Griffith D., Morris E., Thakar V. 2016. Spatial autocorrelation and qualitative sampling: The case of snowball type sampling designs. Annals of the American Association of Geographers, 106(4): 773–787.
Griffith D., Liau Y. 2021. Imputed spatial data: cautions arising from response and covariate imputation measurement error. Spatial Statistics, 42: 100419 (12 pp.).
Griffith D., Plant R. 2022. Statistical analysis in the presence of spatial autocorrelation: selected sampling strategy effects. Stats, 5(4): 1334–1353.
Hart J.E., Källberg H., Laden F., Costenbader K.H., Yanosky J.D., Klareskog L., Alfredsson L., Karlson E.W. 2013. Ambient air pollution exposures and risk of rheumatoid arthritis. Arthritis Care & Research, 65(7): 1190–1196.
Martin G., Webster S. 2020. Does residential sorting explain geographic polarization? Political Science Research and Methods, 8(2): 215–231.
Mocănașu D. 2020. Determining the sample size in qualitative research. Proceedings of the 4th International Multidisciplinary Scientific Conference on the Dialogue Between Sciences & Arts, Religion & Education, October 26–27. Târgoviște, Romania: Ideas Forum International Academic and Scientific Association, p. 181–187.
Mthuli S., Ruffin F., Singh N. 2022. 'Define, Explain, Justify, Apply' (DEJA): An analytic tool for guiding qualitative research sample size. International Journal of Social Research Methodology, 25(6): 809–821.
Nasheeda A. 2022. Sampling, sample size, and data saturation in qualitative research. VC Research Digest, 10 (April): 7–9.
Olea R. 1984. Sampling design optimization for spatial functions. Mathematical Geology, 16(4): 369–392.
Overton W., Stehman S. 1993. Properties of designs for sampling continuous spatial resources from a triangular grid. Communications in Statistics: Theory and Methods, 22: 2641–2660.
Rutkauskas A., Raudeliuniene J., Racinskaja I. 2014. Integral knowledge, innovation and technology cluster formation nurturing the universal development sustainability in the context of globalization. Economics & Sociology, 7(4): 41.
Sebele-Mpofu F. 2021. The sampling conundrum in qualitative research: Can saturation help alleviate the controversy and alleged subjectivity in sampling? International J. of Social Science Studies, 9(5): 11–25.

Stehman S., Overton W. 1996. Spatial sampling. [In:] S. Arlinghaus, D. Griffith (eds.), Practical Handbook of Spatial Statistics. CRC Press Boca, Raton, FL, p. 31–63.
Stough T., Cressie N., Kang E., Michalak A., Sahr K. 2020. Spatial analysis and visualization of global data on multi-resolution hexagonal grids. Japanese J. of Statistics and Data Science, 3: 107–128. https://doi.org/10.1007/s42081-020-00077-w
Straková J. 2015. Sustainable value added as we do not know it. Business: Theory and Practice Verslas: Teorija ir Praktika, 16(2): 168–173.
Subedi K. 2021. Determining the sample in qualitative research. Scholars' Journal, 4: 1–13.
Webster R., Oliver M. 2007. Geostatistics for Environmental Scientists. 2nd ed. Wiley, Chichester, UK.
White D., U.S. Environmental Protection Agency 1992. EPA 40 km Hexagons for Conterminous United States. U.S. Geological Survey, Washington, DC. https://doi.org/10.5066/P9C56AY1

Czy autokorelacja przestrzenna ma znaczenie w kontekście planowania i oceny zrównoważonego rozwoju regionalnego?

Zarys treści: Dążenie do osiągnięcia różnych wymiarów zrównoważonego rozwoju zobowiązuje władze społeczne do zaangażowania się w bardziej gruntowne monitorowanie zbiorowej podaży i popytu, m.in. w sferze ekonomicznej, szczególnie w kontekście geograficznym. W rezultacie, nakłady i wydajność na które ma to wpływ, jak również zasoby/towary/usługi do wykorzystania oraz generowane odpady, które występują i są oznaczone pośrednio lub bezpośrednio w przestrzeni geograficznej, są wyraźnymi nośnikami autokorelacji przestrzennej. Wykorzystanie tej prawie wszechobecnej właściwości danych georeferencyjnych pociąga za sobą możliwość wspierania wydajnych i skutecznych przedsięwzięć w zakresie zrównoważonego rozwoju. Losowy dobór próby metodą tesalacji warstwowej w celu monitorowania zanieczyszczenia środowiska nawiązuje do jednego z przykładów tego twierdzenia. Artykuł ilustruje ten przykład poprzez analizę jakości powietrza w Polsce w 2023 roku. W ten sposób struktura oparta na wyidealizowanych tesalacjach zostaje przełożona na strukturę polskich okręgów administracyjnych; to przekształcenie metodologiczne umożliwia organizacjom rządowym uczestniczenie w każdym planowanym monitorowaniu oraz jego nadzorowaniu bez dodatkowych komplikacji prawnych. Przypadkowe odkrycia naukowe obejmują wstępne rozszerzenie zbioru standardowych kształtów wielokątów (np. kwadratów i sześciokątów) o trapezy w celu pobrania próbek przestrzennych oraz ewentualność, że wpływ autokorelacji przestrzennej na statystyki oparte na projektach może mieć znaczną przewagę nad naruszeniem konwencjonalnego przykazania zrównoważonego losowego pobierania próbek. Wniosek jaki się nasuwa w trakcie analiz streszczonych w niniejszej publikacji dowodzi, że autokorelacja przestrzenna ma znaczenie w planowaniu i ocenie zrównoważonego rozwoju regionalnego.

Słowa kluczowe: PM10, Polska, autokorelacja przestrzenna, zrównoważony rozwój regionalny, losowy dobór próby metodą tesalacji warstwowej

Appendix A

Peripheral locations of excluded gminas resolution krigged points

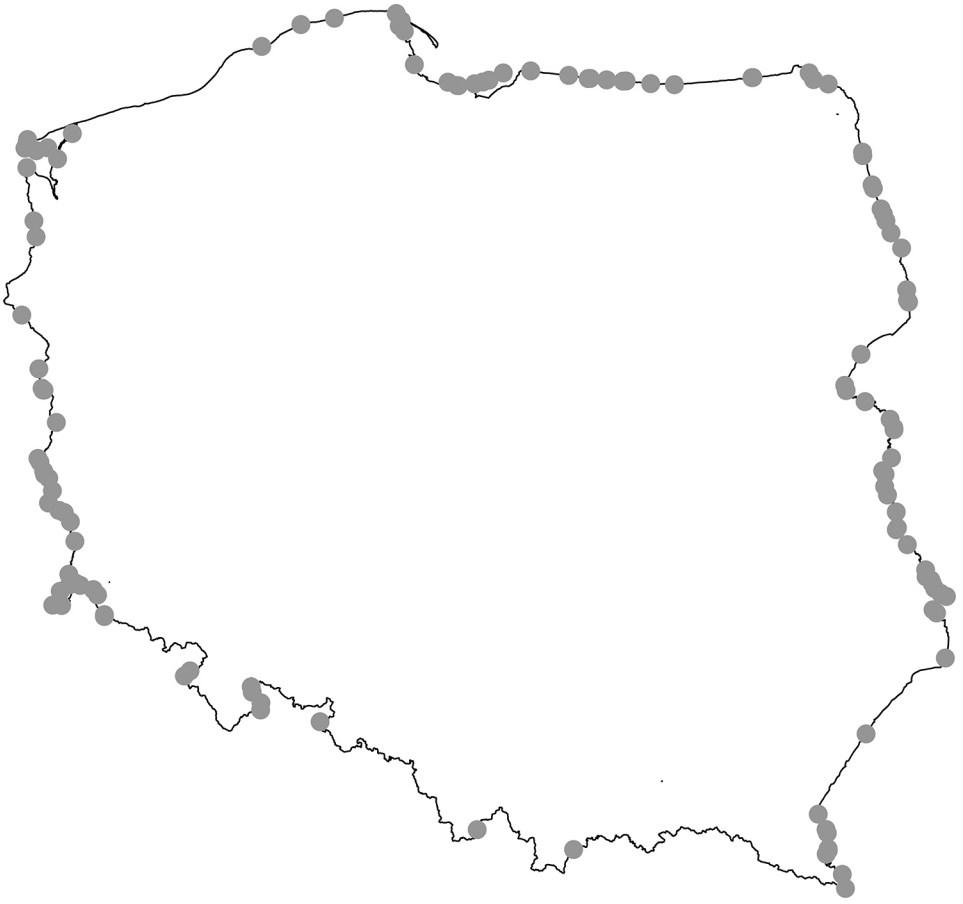

Fig. A1. The 125 points located outside of gminas, but not voivodeships or poviats, boundaries