

Three inference traps in exploratory space–time analysis: evidence from population ageing in Türkiye

Kerem Yavuz ARSLANLI

arсланli@itu.edu.tr

Istanbul Technical University

Research Article

Keywords: Local indicators of spatial association, Conditional permutation inference, Emerging hot spot analysis, Spatial Markov chains, Population ageing, Türkiye

Posted Date: October 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11120504/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Exploratory local statistics are routinely combined in applied regional analysis, yet three common practices can produce misleading conclusions. Using an annual register-based panel of Türkiye's districts for 2007–2025, with the diffusion of population ageing as a test case, this note documents three inference traps and a diagnostic for each. First, differences between Getis–Ord G_i^* and Local Moran's I maps arise from mixing inference regimes and test sidedness: under conditional permutation both statistics are monotone in the same permuted neighbourhood sum and flag identical locations (all 969 pseudo-p-values coincide), differing only in how spatial outliers are labelled. Second, Emerging Hot Spot Analysis applied to a trending variable records the national trend rather than spatial clustering: the share of significant hot-spot bins rises from 0% to 45% with the raw ageing index but stays near 19% once the index is standardised within year, and agreement with the static map rises from $\kappa = 0.25$ to $\kappa = 0.54$. Third, a spatial Markov chain rejects spatial independence ($LR = 269.6$, $df = 45$, $p < 0.001$), but the apparent neighbour effect reflects units' position within discrete classes: once own ageing is controlled, older neighbourhoods are associated with slower, not faster, upward transitions. These results hold across spatial weights, indicators, sub-periods, population thresholds and the exclusion of earthquake-affected provinces. Substantively, district ageing in Türkiye is strongly clustered and intensifying, but the evidence is consistent with shared regional forcing rather than contagion between neighbouring districts.

JEL Classification C21 · C23 · J11 · R12

1 Introduction

Local indicators of spatial association, space–time hot-spot methods and spatial Markov chains are standard tools of exploratory regional analysis. Local Moran's I (LISA) classifies significant locations into clusters and spatial outliers (Anselin 1995); the Getis–Ord G_i^* statistic identifies hot and cold spots (Getis and Ord 1992; Ord and Getis 1995); Emerging Hot Spot Analysis (EHSA) extends G_i^* to a space–time cube and classifies each location's trajectory (ESRI n.d.); and the spatial Markov chain conditions transitions between classes on the state of a unit's neighbours (Rey 2001). The methods are often combined, and their outputs are increasingly read inferentially: as evidence that two statistics disagree, that clusters are emerging, or that a process diffuses from place to place.

This note shows that three common practices can make such readings misleading. The first is comparing G_i^* and LISA maps evaluated under different inference regimes. The unreliability of analytical approximations for local statistics and the sensitivity of local inference to implementation choices are well documented (Getis and Ord 1996; Sokal et al. 1998; Bivand and Wong 2018; Bivand 2022; Sauer et al. 2021), and software documentation shows the two statistics selecting the same locations under permutation (Anselin 2024). What has not been made explicit is why this must happen, and that test sidedness is a second, separate source of apparent disagreement. The second practice is applying EHSA to a variable with a strong common trend, which is typical of demographic and economic

indicators and of published applications (Purwanto et al. 2021; Morckel and Durst 2023). The third is interpreting a significant spatial Markov test as evidence of diffusion, as is common in convergence studies (Rey 2001; Rey et al. 2011), without checking whether the neighbour effect survives control for a unit's position within its discrete class.

Each trap is demonstrated on the same substantive case: population ageing across Türkiye's districts between 2007 and 2025. Ageing is increasingly studied as a spatial process (Shiode et al. 2014; Wu et al. 2021), and Türkiye combines a rapid national transition with steep regional contrasts between an ageing interior and Black Sea belt and a youthful south-east (Selim and Bilgin 2021; Yakar and Özgür 2022). Existing district-level work maps these patterns descriptively (Özgür and Yakar 2021; Yakar and Özgür 2022, 2026); applications of LISA at this scale are common in regional analysis (Majumder and Cajee 2026). The note adds an inferential layer and shows how conclusions about clustering, emergence and diffusion change once each trap is addressed.

2 Trap 1: G^* and Local Moran's I under conditional permutation

Let x_i be the value at location i among n locations, $\bar{x}$ the global mean, $z_i = x_i - \bar{x}$, $m_2 = \sum z_i^2/n$, and $W = [w_{ij}]$ a spatial weights matrix with $w_{ii} = 0$. Write $S_i = \sum w_{ij} x_j / \sum w_{ij}$ for the (row-standardised) neighbourhood mean. Local Moran's I and the raw G_i^* statistic can be written as

$$I_i = (z_i / m_2)(S_i - \bar{x}), G_i^* = (w_{ii}^* x_i + \lambda_i S_i) / \sum x_i,$$

where $w_{ii}^* > 0$ is the self-weight of the star statistic and $\lambda_i = \sum_{j \neq i} w_{ij}^* > 0$ is the total weight of the neighbours in the G_i^* weights.

Proposition

Suppose that (a) both statistics use the same neighbour set for location i and neighbour weights that are proportional to each other within row i ; (b) inference is by conditional permutation, holding x_i fixed and reassigning the other $n - 1$ values; (c) the pseudo-p-value is folded, i.e. it counts permutations at least as extreme as the observed value in the direction of the observed deviation; and (d) both statistics are evaluated on the same permutation draws. Then G_i^* and I_i have identical pseudo-p-values at every location, apart from ties.

Proof sketch. Under conditional permutation $\bar{x}$, m_2 , z_i , x_i , w_{ii}^* and $\sum x_i$ are invariant, and only S_i varies. G_i^* is therefore increasing and affine in S_i , and I_i is affine in S_i with a slope whose sign is that of z_i . Both statistics thus rank the permuted neighbourhood means identically, up to a reversal when $z_i < 0$ that the folding absorbs. ■

Three consequences follow. First, differences between the significant sets of the two statistics can arise only from violations of (a)–(d): different neighbour definitions (for example a distance band for G_i^* and contiguity for LISA, a common software default), independent draws (Monte Carlo error), or analytical

rather than permutation inference. The result holds equally for G_i , but not for local statistics that are not monotone in S_i , such as local Geary. Second, among significant locations the statistics differ only in labelling: G_i^* labels by the level of the neighbourhood including i , LISA by the quadrant of the Moran scatterplot, and the two conflict only for spatial outliers. LISA of course also carries information that G_i^* lacks—its values decompose global Moran's I —but as a device for flagging and classifying locations its only distinct output is the outlier category. Third, test sidedness matters separately: the folded pseudo- p -value is a one-sided test in the observed direction, whereas analytical G_i^* maps are commonly read two-sided ($|z| > 1.96$). Comparing the two at a nominal 5% level therefore compares tests of different size.

3 Data and methods

3.1 Data and harmonised geography

District population counts by sex and five-year age band are taken from the Address-Based Population Registration System of the Turkish Statistical Institute for 2007–2025 (TÜİK 2025). Over the period 73 district codes enter (43 in 2008, 27 in 2013, three later) and 23 exit, so parent districts lose or gain territory even when their codes persist. To obtain a temporally consistent geography, every district involved in a reorganisation is identified by population accounting: in each reorganisation year and province, districts that enter or exit and districts whose population changes by more than 0.05 log points relative to the provincial median are merged into a single unit. This yields 848 units with complete coverage, of which 819 are unchanged districts and 29 merge 176 district codes. Unit boundaries are obtained by dissolving a 969-polygon TÜİK district layer, in which three miscoded polygons were reassigned by location, and all geometries are projected to an azimuthal equidistant projection centred on Türkiye (39°N, 35°E). The static analysis of Trap 1 uses the 969 districts of 2025 directly. Residual year-to-year population jumps are concentrated in small rural districts around local election years, a known feature of address-based registration, and in 2023 in earthquake-affected provinces; both are addressed in the robustness analysis.

The primary indicator is the ageing index, $AI = 100 \times P^{65+} / P^{0-14}$, which combines ageing at the top of the age structure with the thinning of its base. The old-age dependency ratio and the share aged 65 and over are used as alternatives. For the Markov analyses AI is discretised into quintiles of its pooled 2007–2025 distribution; because these classes are invariant to monotone transformations, the Markov results are unchanged by, for example, taking logarithms.

3.2 Local statistics and inference regimes

The baseline weights are first-order queen contiguity (units without contiguous neighbours are linked to their nearest neighbour), with six-nearest-neighbour and binary distance-band weights (the smallest band giving every unit a neighbour) as alternatives. For 2025, G_i^* and LISA are compared under three regimes: (i) the conventional mixed regime—analytical G_i^* with binary weights including i (Ord and Getis 1995), read two-sided at 5%, against LISA with 999 conditional permutations; (ii) the same comparison

with the analytical G_i^* read one-sided, matching the folded pseudo-p-value; and (iii) a common regime in which both statistics use 99,999 conditional permutations with the same random seed. Significance is assessed without correction, after Benjamini–Hochberg false-discovery-rate correction (Benjamini and Hochberg 1995; Caldas de Castro and Singer 2006), and under a Bonferroni bound. Agreement on the scheme {ageing cluster, youthful cluster, spatial outlier, not significant} is measured by observed agreement, Cohen’s κ (Cohen 1960), weighted κ (Cohen 1968) and the Adjusted Rand Index (Hubert and Arabie 1985).

3.3 Space–time hot spots

EHSA computes a space–time G_i^* for each unit–year bin, whose neighbourhood is the unit and its contiguous neighbours in the same and the preceding year, evaluates each bin by its analytical z-score (5%, two-sided), tests each unit’s z-score series with the Mann–Kendall test (Mann 1945; Kendall 1975) and classifies units into the standard categories (ESRI n.d.), without the bin-level false-discovery-rate adjustment applied by default in ArcGIS. Because the space–time G_i^* standardises every bin against the mean of the whole cube, a variable with a strong common trend will appear increasingly ‘hot’ everywhere. EHSA is therefore run twice: on raw AI and on AI standardised within each year (subtracting the annual mean and dividing by the annual standard deviation), which removes the common trend while preserving each year’s spatial configuration. Both versions are compared with the static 2025 G_i^* map.

3.4 Spatial Markov chain and a within-class diagnostic

The spatial Markov chain (Rey 2001) conditions annual transitions between the five AI classes on the class of each unit’s spatial lag; spatial independence is tested with the likelihood-ratio test of conditional homogeneity (Bickenbach and Bode 2003; Rey et al. 2016). A significant test is commonly read as evidence of diffusion. However, when a variable is spatially autocorrelated, units with older neighbours also tend to lie higher within their own class and so need a smaller change to move up. As a diagnostic, the probability of an upward transition is modelled by a logit on the log of the unit’s own AI and the log of its spatial lag, with year fixed effects (and class fixed effects when all classes are pooled) and standard errors clustered by unit; a growth regression of the annual change in log AI on the same variables provides a continuous counterpart. A diffusion interpretation requires a positive lag coefficient conditional on own level. Temporal homogeneity of the transition matrix is tested between 2007–2016 and 2016–2025 (Bickenbach and Bode 2003).

3.5 Robustness and implementation

All results are re-estimated under the alternative weights and indicators, for units with at least 10,000 inhabitants, for the two sub-periods, excluding the eleven provinces affected by the February 2023 earthquakes, and under year-specific and fixed-threshold (AI 40, 70, 110, 180) discretisations. All analyses are implemented in Python with libpysal, esda and giddy (Rey and Anselin 2007) and statsmodels; the space–time G_i^* and EHSA classification are coded directly from the definitions above. Data, spatial weights and code are available at https://osf.io/n4ufb/overview?view_only=f347bde9de8a4f41abafa670bff4651f. A generative AI tool (Claude, Anthropic) assisted with

writing the analysis and plotting code and with editing the text; all code, results and references were verified by the author.

4 Results

4.1 Clustering and intensification

The mean ageing index across the 848 units rose from 47.5 in 2007 to 136.1 in 2025 (2025 median 104.1). The distribution is strongly right-skewed, ranging in 2025 from 890 in Rize-Hemşin to below 7 in Şanlıurfa-Harran and Akçakale. Ageing is clearly clustered: the 2025 global Moran's I is 0.362 (pseudo- p = 0.001) under queen contiguity, 0.362–0.392 across weights, and higher for the old-age dependency ratio (0.469) and the 65 + share (0.518).

4.2 Trap 1: inference regime and sidedness, not the statistic

Under the conventional mixed regime the two statistics appear to disagree (Table 1; Fig. 1): analytical G_i^* flags 133 hot and 135 cold spots, LISA 79 high–high and 211 low–low clusters and 37 outliers (κ = 0.69). Most of the 77 LISA clusters not flagged by G_i^* are low–low districts at the edge of the youthful south-east. Reading the analytical G_i^* one-sided, to match the folded pseudo- p -value, reduces these to 19 and raises agreement to κ = 0.77; the remaining disagreement reflects the normal approximation, which is poorly suited to a variable with a skewness of 1.8. Under common permutation all disagreement in significance disappears: all 969 pseudo- p -values coincide, and 348 districts are significant without correction, 204 after false-discovery-rate correction and 60 under the Bonferroni bound for both statistics. The residual κ of 0.92 reflects labelling alone: of LISA's 40 spatial outliers, the 36 low–high outliers—relatively young districts inside ageing surroundings—are G_i^* hot spots and the four high–low outliers G_i^* cold spots, exactly as the Proposition implies. The equivalence holds under six-nearest-neighbour (398/253/69 significant; 44 outliers; κ = 0.92) and distance-band weights (448/339/89; 60 outliers; κ = 0.90).

Table 1
Agreement between G_i^* and LISA classifications of the ageing index, 2025, under three inference regimes

Regime	G_i^* hot / cold	Observed agreement	Cohen's κ	Weighted κ	ARI	Disagreement in significance
Analytical G_i^* two-sided vs LISA (999 permutations)	133 / 135	0.85	0.69	0.72	0.63	28 G_i^* only; 77 LISA only
Analytical G_i^* one-sided vs LISA (999 permutations)	169 / 196	0.88	0.77	0.79	0.73	61 G_i^* only; 19 LISA only
Both 99,999 conditional permutations, same draws	122 / 226	0.96	0.92	—	—	None (labels differ for 40 outliers)

969 districts; queen contiguity; 5% level without multiple-comparison correction. LISA (999 permutations): 79 high–high, 211 low–low, 37 outliers.

4.3 Trap 2: EHSA on a trending variable

With the raw ageing index, the share of units that are significant space–time hot spots rises almost linearly from 0% in 2007 to 45% in 2025 (Fig. 2a). The resulting classification is dominated by 303 ‘consecutive’ and 48 ‘new’ hot spots (Table 2; Fig. 2b) units that were not hot early in the period and became so as they crossed the mean of the whole cube. These categories record the national ageing trend rather than the emergence of spatial clusters, and their agreement with the static 2025 map is low ($\kappa = 0.25$).

Standardising the index within each year removes the trend while preserving each year’s spatial configuration: the hot-spot share is then stable at about 19% (Fig. 2a), agreement with the static map rises to $\kappa = 0.54$ (0.68 for the final-year bins), and the classification changes character (Fig. 2c). The ageing core of the western Black Sea and north-central Anatolia now appears as 38 intensifying and 21 persistent hot spots, while 49 diminishing hot spots mark districts whose relative ageing is easing. On the youthful side, 148 diminishing cold spots in the east and south-east survive standardisation: the relative youthfulness of this region is genuinely weakening, not merely rising with the national tide.

Table 2

Emerging Hot Spot Analysis of the ageing index, 2007–2025, raw and standardised within year

Share of hot-spot bins, 2007 / 2025	Raw ageing index	Standardised within year
	0.00 / 0.45	0.14 / 0.19
Hot-spot units (intensifying, persistent, consecutive, new, other)	378 (1, 0, 303, 48, 26)	221 (38, 21, 33, 2, 127)
Cold-spot units (of which diminishing)	330 (111)	224 (148)
No pattern	140	403
κ with static 2025 G_i^* (classification / final-year bins)	0.25 / 0.44	0.54 / 0.68

848 harmonised units; queen contiguity; space–time neighbourhood of the current and preceding year; 5% level. ‘Other’ hot spots are diminishing, oscillating, sporadic and historical.

4.4 Trap 3: neighbour dependence or position within class?

The spatial Markov chain is highly persistent and almost exclusively upward (annual upward probability 0.089, downward 0.004; the oldest class is effectively absorbing at 0.991), and it rejects spatial independence decisively (LR = 269.6, df = 45, $p < 0.001$). The probability that a C_2 unit moves to C_3 rises from 0.044 when its neighbours are in the youngest lag quintile to 0.194 when they are in the oldest (Fig. 3a)—the pattern usually read as diffusion. But C_2 units with older neighbours also lie higher within C_2 ; their median own AI rises from 34.6 to 44.5 across the lag quintiles, closer to the C_2/C_3 threshold of 50.5 (Fig. 3b).

Once own level is controlled the neighbour effect reverses (Table 3). In a logit of the $C_2 \rightarrow C_3$ transition, the lag coefficient is + 0.94 on its own but – 1.09 (SE 0.25) conditional on own AI; pooling all classes with class fixed effects gives – 0.53 (SE 0.14), and the growth regression gives a negative lag coefficient as well (– 0.006, SE 0.001). Holding own AI fixed at 48, the predicted probability of moving up falls from 0.74 with the youngest neighbourhoods to 0.42 with the oldest (Fig. 3c). The negative conditional effect holds under both alternative weights, in both sub-periods and for units of at least 10,000 inhabitants. The transition matrix itself is temporally homogeneous (LR = 13.9, df = 12, $p = 0.31$). The significant spatial Markov test is therefore real, but it reflects the spatial sorting of units within discrete classes, not neighbours accelerating each other’s ageing; conditional on its own level, a unit surrounded by older districts ages somewhat more slowly, which is consistent with regional convergence rather than contagion.

Table 3
Spatial Markov test and within-class diagnostic across specifications

Specification	Global Moran's I (2025)	Spatial Markov LR (df)	Lag coefficient, logit with own AI (SE)
Baseline (queen, AI, pooled quintiles)	0.362	269.6 (45)	-0.53 (0.14)
Six nearest neighbours	0.392	278.1 (48)	-0.41 (0.14)
Distance band	0.373	278.4 (48)	-0.34 (0.14)
2007–2016 / 2016–2025	—	123.0 (33) / 156.8 (43)	-0.70 (0.21) / -0.32 (0.16)
Units $\geq 10,000$ inhabitants (657)	0.400	171.9 (29)	-0.62 (0.16)
Excluding earthquake provinces (102 units)	0.334	237.9 (40)	—
Old-age dependency ratio / share 65+	0.469 / 0.518	233.5 (48) / 235.5 (44)	—
Year-specific / fixed-threshold classes	—	102.9 (40) / 334.3	—

848 harmonised units unless stated. All likelihood-ratio statistics significant at $p < 0.001$. Logit: upward transition on log own AI and log spatial lag with class and year fixed effects, standard errors clustered by unit. C_2 -only logit: lag + 0.94 (0.10) without and - 1.09 (0.25) with own AI. Growth regression of $\Delta \log \text{AI}$: lag - 0.006 (0.001).

5 Discussion and conclusion

Each trap has a simple diagnostic. For comparisons of local statistics, fix the inference regime and the sidedness of the test before attributing differences to the statistic; under conditional permutation, G_i^* and LISA flag the same locations by construction, and LISA's distinct contribution to mapping is its outlier category. For EHSA, check whether the share of significant bins drifts over time and, if the variable trends, standardise it within period before building the cube; otherwise 'consecutive' and 'new' hot spots largely record the trend. For spatial Markov chains, add a unit's own level to a simple transition or growth regression before interpreting neighbour dependence as diffusion; with coarse classes and spatially autocorrelated data, a significant conditional-homogeneity test is expected even without any neighbour influence.

For the substantive case, the corrected picture is clear. District ageing in Türkiye is strongly clustered and intensifying, with a relatively ageing core in the western Black Sea and north-central Anatolia, and the relative youthfulness of the east and south-east is weakening even after the national trend is removed. The evidence does not, however, support contagious diffusion between neighbouring districts:

conditional on their own level, units surrounded by older districts age more slowly. This is consistent with shared regional drivers fertility regimes, migration fields and labour markets that span district boundaries and with convergence within regions. The analysis complements descriptive mapping of Turkish ageing (Yakar and Özgür 2022, 2026) with these inferential checks.

Several limitations apply. The harmonisation relies on population accounting with a fixed threshold and may merge some units unnecessarily; results are conditional on the district scale (Openshaw 1984); address-based registration contains election-related fluctuations in small rural districts; and the diagnostics in Sect. 3.4 are descriptive rather than causal. Spatial panel models that include fertility, migration and economic covariates are the natural next step to separate shared regional forcing from genuine neighbour effects.

Declarations

Author Contribution

K.Y.A.: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing – original draft; Writing – review & editing.

References

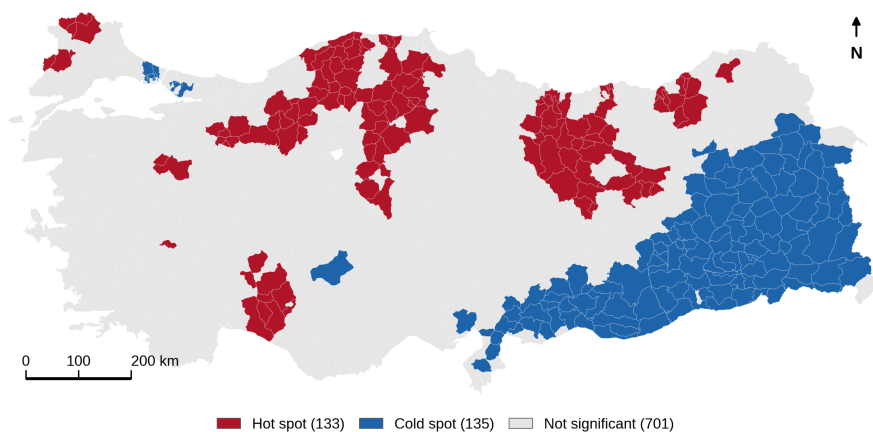
1. Anselin, L.: Local indicators of spatial association—LISA. *Geogr. Anal.* **27**(2), 93–115 (1995). <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
2. Anselin, L.: An Introduction to Spatial Data Science with GeoDa, Volume 1: Exploring Spatial Data. CRC, Boca Raton (2024). <https://doi.org/10.1201/9781003274919>
3. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R Stat. Soc. Ser. B.* **57**(1), 289–300 (1995). <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
4. Bickenbach, F., Bode, E.: Evaluating the Markov property in studies of economic convergence. *Int. Reg. Sci. Rev.* **26**(3), 363–392 (2003). <https://doi.org/10.1177/0160017603253789>
5. Bivand, R.: R packages for analyzing spatial data: a comparative case study with areal data. *Geogr. Anal.* **54**(3), 488–518 (2022). <https://doi.org/10.1111/gean.12319>
6. Bivand, R.S., Wong, D.W.S.: Comparing implementations of global and local indicators of spatial association. *TEST.* **27**(3), 716–748 (2018). <https://doi.org/10.1007/s11749-018-0599-x>
7. Caldas de Castro, M., Singer, B.H.: Controlling the false discovery rate: a new application to account for multiple and dependent tests in local statistics of spatial association. *Geogr. Anal.* **38**(2), 180–208 (2006). <https://doi.org/10.1111/j.0016-7363.2006.00682.x>
8. Cohen, J.: A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **20**(1), 37–46 (1960). <https://doi.org/10.1177/001316446002000104>

9. Cohen, J.: Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol. Bull.* **70**(4), 213–220 (1968). <https://doi.org/10.1037/h0026256>
10. ESRI: How Emerging Hot Spot Analysis works. ArcGIS documentation, Space Time Pattern Mining toolbox. <https://doc.arcgis.com/en/allsource/1.5/analysis/geoprocessing-tools/space-time-pattern-mining/learnmoreemerging.htm> (n.d.). Accessed 22 September 2026
11. Getis, A., Ord, J.K.: The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **24**(3), 189–206 (1992). <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>
12. Getis, A., Ord, J.K.: Local spatial statistics: an overview. In: Longley, P., Batty, M. (eds.) *Spatial Analysis: Modelling in a GIS Environment*, pp. 261–277. GeoInformation International, Cambridge (1996)
13. Hubert, L., Arabie, P.: Comparing partitions. *J. Classif.* **2**(1), 193–218 (1985). <https://doi.org/10.1007/BF01908075>
14. Kendall, M.G.: *Rank Correlation Methods*, 4th edn. Griffin, London (1975)
15. Majumder, S., Cajee, L.: Unveiling socio-economic inequalities and spatial disparities across districts in Meghalaya: insights from PCA, Moran's I and LISA. *Lett. Spat. Resour. Sci.* **19**, 3 (2026). <https://doi.org/10.1007/s12076-025-00424-z>
16. Mann, H.B.: Nonparametric tests against trend. *Econometrica.* **13**(3), 245–259 (1945). <https://doi.org/10.2307/1907187>
17. Morckel, V., Durst, N.: Using Emerging Hot Spot Analysis to explore spatiotemporal patterns of housing vacancy in Ohio metropolitan statistical areas. *Urban Aff Rev.* **59**(1), 309–328 (2023). <https://doi.org/10.1177/10780874211065014>
18. Openshaw, S.: *The Modifiable Areal Unit Problem*. CATMOG 38. Geo Books, Norwich (1984)
19. Ord, J.K., Getis, A.: Local spatial autocorrelation statistics: distributional issues and an application. *Geogr. Anal.* **27**(4), 286–306 (1995). <https://doi.org/10.1111/j.1538-4632.1995.tb00912.x>
20. Özgür, E.M., Yakar, M.: Coğrafya önemlidir! Nüfus yaşlanmasına mekânsal bir perspektiften bakmak. *Senex: Yaşlılık Çalışmaları Dergisi.* **5**(2), 19–45 (2021)
21. Purwanto, P., Utaya, S., Handoyo, B., Bachri, S., Astuti, I.S., Utomo, K.S.B., Aldianto, Y.E.: Spatiotemporal analysis of COVID-19 spread with Emerging Hotspot Analysis and space–time cube models in East Java, Indonesia. *ISPRS Int. J. Geo-Inf.* **10**(3), 133 (2021). <https://doi.org/10.3390/ijgi10030133>
22. Rey, S.J.: Spatial empirics for economic growth and convergence. *Geogr. Anal.* **33**(3), 195–214 (2001). <https://doi.org/10.1111/j.1538-4632.2001.tb00444.x>
23. Rey, S.J., Anselin, L.: PySAL: a Python library of spatial analytical methods. *Rev. Reg. Stud.* **37**(1), 5–27 (2007). <https://doi.org/10.52324/001c.8285>
24. Rey, S.J., Kang, W., Wolf, L.: The properties of tests for spatial effects in discrete Markov chain models of regional income distribution dynamics. *J. Geogr. Syst.* **18**(4), 377–398 (2016). <https://doi.org/10.1007/s10109-016-0234-x>

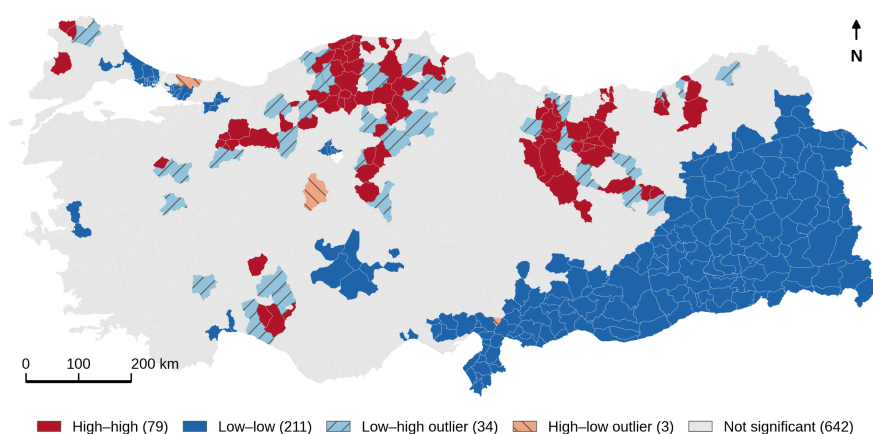
25. Rey, S.J., Murray, A.T., Anselin, L.: Visualizing regional income distribution dynamics. *Lett. Spat. Resour. Sci.* **4**(1), 81–90 (2011). <https://doi.org/10.1007/s12076-010-0048-2>
26. Sauer, J., Oshan, T.M., Rey, S., Wolf, L.J.: The importance of null hypotheses: understanding differences in local Moran's I under heteroskedasticity. *Geogr. Anal.* (2021). <https://doi.org/10.1111/gean.12304>
27. Selim, S., Bilgin, D.: Türkiye'de doğurganlık, göç ve mekânsal etkileşim: seçicilik hipotezi kapsamında bir analiz. *Coğrafya Dergisi.* **43**, 111–125 (2021). <https://doi.org/10.26650/JGEOG2021-893270>
28. Shiode, N., Morita, M., Shiode, S., Okunuki, K.: Urban and rural geographies of aging: a local spatial correlation analysis of aging population measures. *Urban Geogr.* **35**(4), 608–628 (2014). <https://doi.org/10.1080/02723638.2014.905256>
29. Sokal, R.R., Oden, N.L., Thomson, B.A.: Local spatial autocorrelation in a biological model. *Geogr. Anal.* **30**(4), 331–354 (1998). <https://doi.org/10.1111/j.1538-4632.1998.tb00406.x>
30. TÜİK (Turkish Statistical Institute): Address-Based Population Registration System (ADNKS) results, 2007–2025. TÜİK, Ankara: (2025). <https://www.tuik.gov.tr> Accessed 22 September 2026
31. Wu, L., Huang, Z., Pan, Z.: The spatiality and driving forces of population ageing in China. *PLoS ONE.* **16**(1), e0243559 (2021). <https://doi.org/10.1371/journal.pone.0243559>
32. Yakar, M., Özgür, E.M.: Türkiye'de nüfus yaşlanması, yerel düzeyde tehlike çanları çalıyor! *Coğrafya Dergisi.* **44**, 231–250 (2022). <https://doi.org/10.26650/JGEOG2022-1059215>
33. Yakar, M., Özgür, E.M.: An analysis of the spatial pattern of demographic ageing in Türkiye using bivariate mapping. *Coğrafya Dergisi.* **52**, 1–8 (2026). <https://doi.org/10.26650/jgeog2026-1659196>

Figures

(a) Getis-Ord G_i^* (analytical inference)



(b) Local Moran's I (999 conditional permutations)



(c) Agreement between (a) and (b)

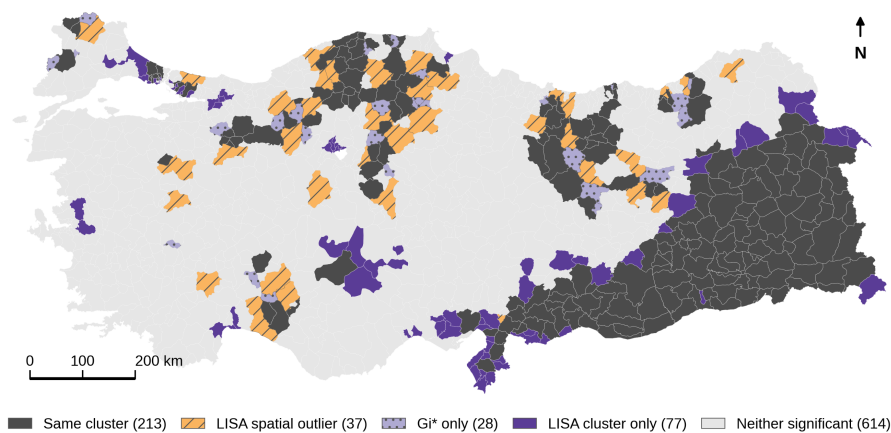


Figure 1

Static classifications of the district ageing index, 2025, under queen contiguity (969 districts; 5% significance): (a) Getis-Ord G_i^* hot and cold spots, analytical inference read two-sided; (b) Local Moran's I clusters and spatial outliers, 999 conditional permutations; (c) agreement between (a) and (b). Under common permutation inference the two statistics flag identical districts (Table 1)

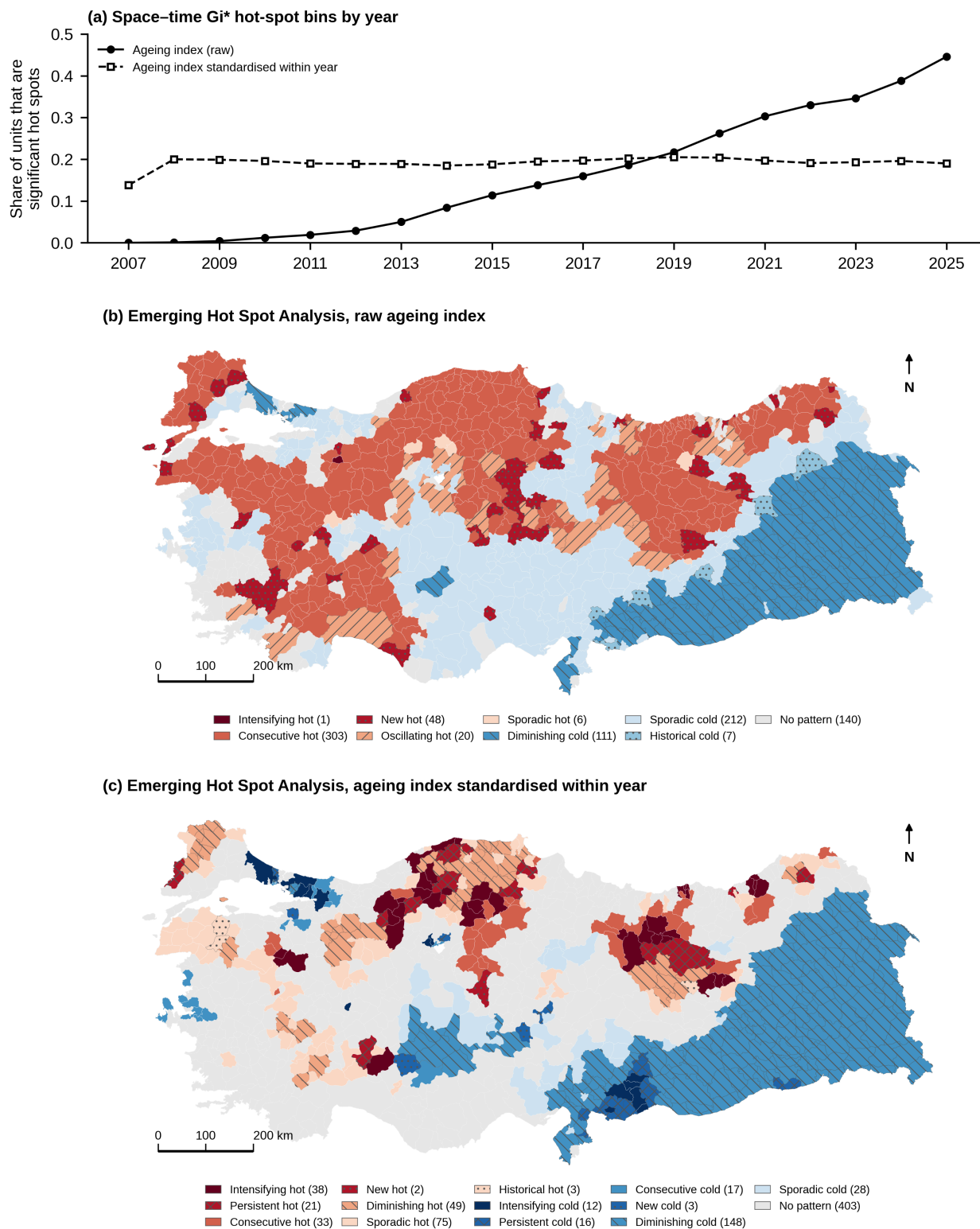


Figure 2

Emerging Hot Spot Analysis of the ageing index, 2007–2025 (848 harmonised units): (a) share of units that are significant space–time hot spots in each year; (b) classification of the raw index; (c) classification of the index standardised within year. Unfilled polygons in (b) and (c) are water bodies or areas without harmonised coverage

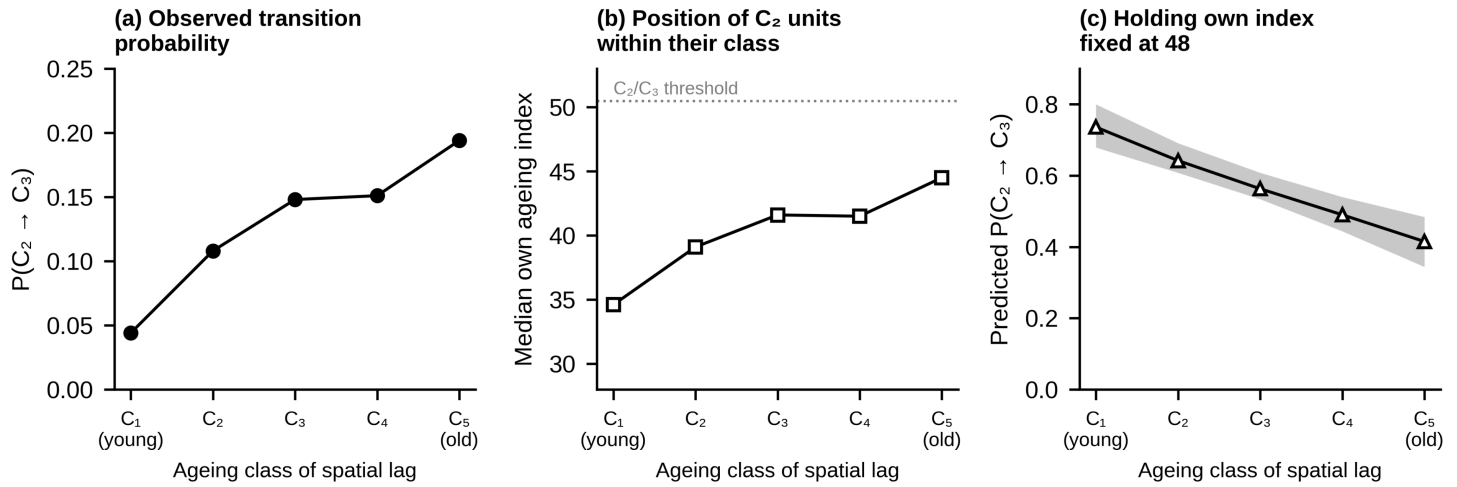


Figure 3

The within-class confound in the spatial Markov chain, $C_2 \rightarrow C_3$ transitions, 848 units, 2007–2025: (a) observed transition probability by quintile of the spatial lag; (b) median own ageing index of C₂ units by lag quintile, with the C₂/C₃ threshold; (c) predicted transition probability from a logit with own and lag ageing index and year fixed effects, holding own index fixed at 48 (90th percentile within C₂); the band is a 95% cluster-bootstrap interval