Online Resource 1

Borrowing strength across space in distributed lag non-linear models: a geographically weighted approach

Supplementary Text S1. Construction of the simulated regions

Grid and shapes. Cells have side 1/62, and the region is tiled at random into four-cell pieces (tetrominoes). Each region shape is built from four blocks of 31 × 31 cells in a tetromino arrangement: square (62 × 62 cells), elongated (31 × 124), L-shaped (93 × 62) and T-shaped (62 × 93). Every shape has 3,844 cells and 961 tetrominoes.

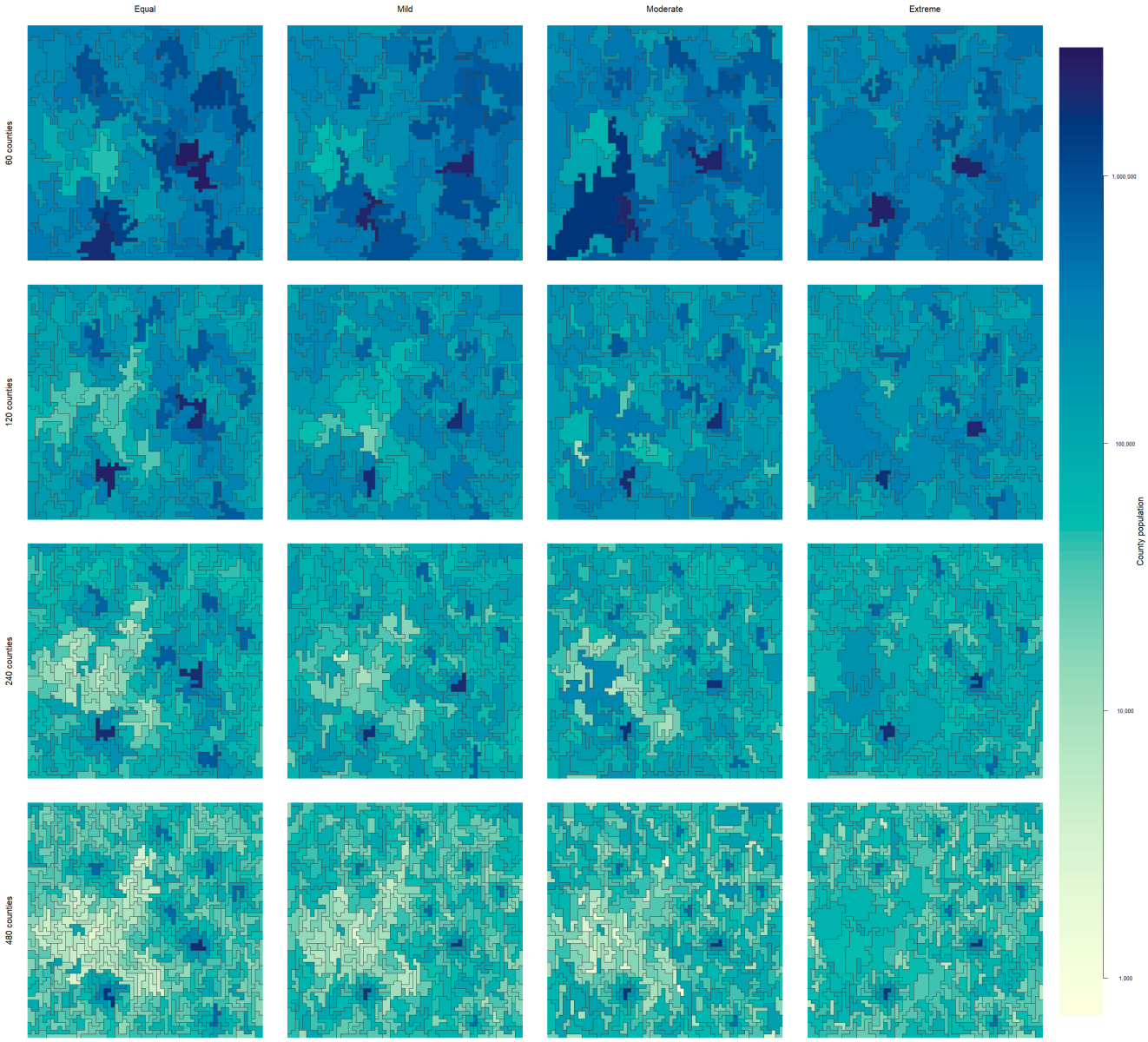
Settlement. Each region has two large cities of 1.5 million and six major cities of 500,000, each with surrounding suburbs, 16 mid-sized cities of 80,000 to 120,000, and rural population elsewhere (lognormal, median 15,000 per tetromino). A sparse inland area covering 18% of the tetrominoes has one fifth of the rural density. Rural populations were rescaled so that all four region shapes have the same total population.

Temperature layouts. The north-south layout is a gradient with a warm south. The cool-ridge layout rotates this gradient by 15° and adds a band 4 °C cooler at its crest, running west to east a third of the way below the northern edge. In the river-valley layout, warm valleys form a Y-shaped network, and temperature declines with distance from the valleys. In the edge layout, temperature is highest along the north and east edges. Each layout spans 10 °C across the region. The urban heat island adds up to 2 °C and fades over the suburbs.

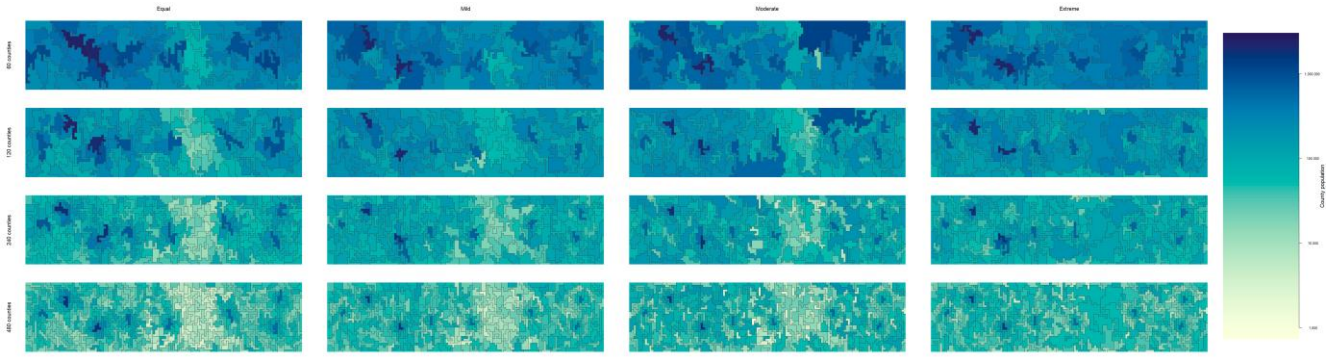
High heterogeneity multiplier. County values of the spatial surface are standardized within the replicate, truncated at ±1.5 and standardized again. They are then multiplied by twice the realized standard deviation of the log multiplier in the matching moderate heterogeneity replicate, exponentiated, and centered so that the multipliers average one over counties.

Supplementary Figures

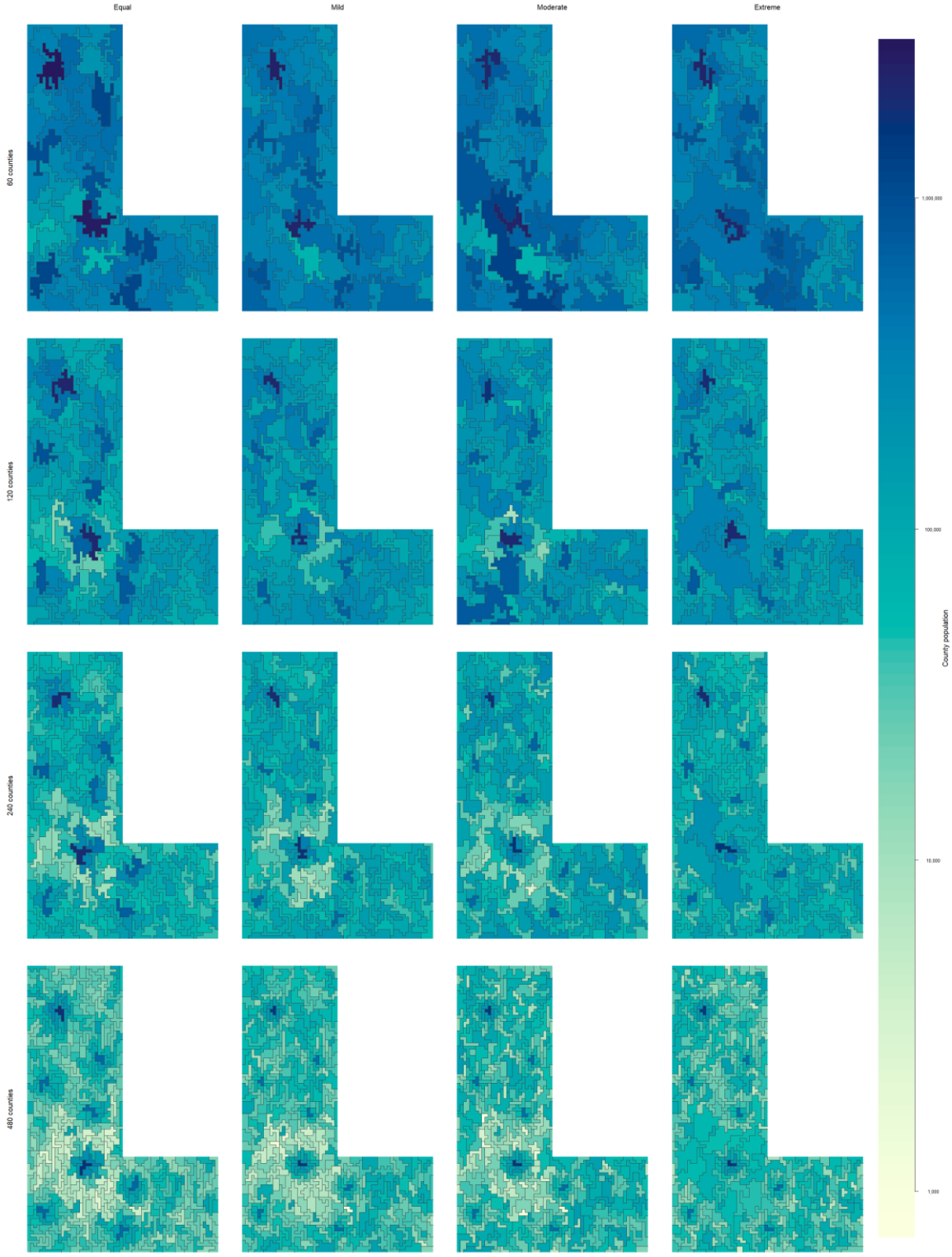
(a)



(b)


Equal
Mild
Moderate
Extreme
60 counties
120 counties
240 counties
480 counties
County population

(c)


Equal
Mild
Moderate
Extreme
60 counties
120 counties
240 counties
480 counties
1,000,000
100,000
10,000
1,000
County population

(d)

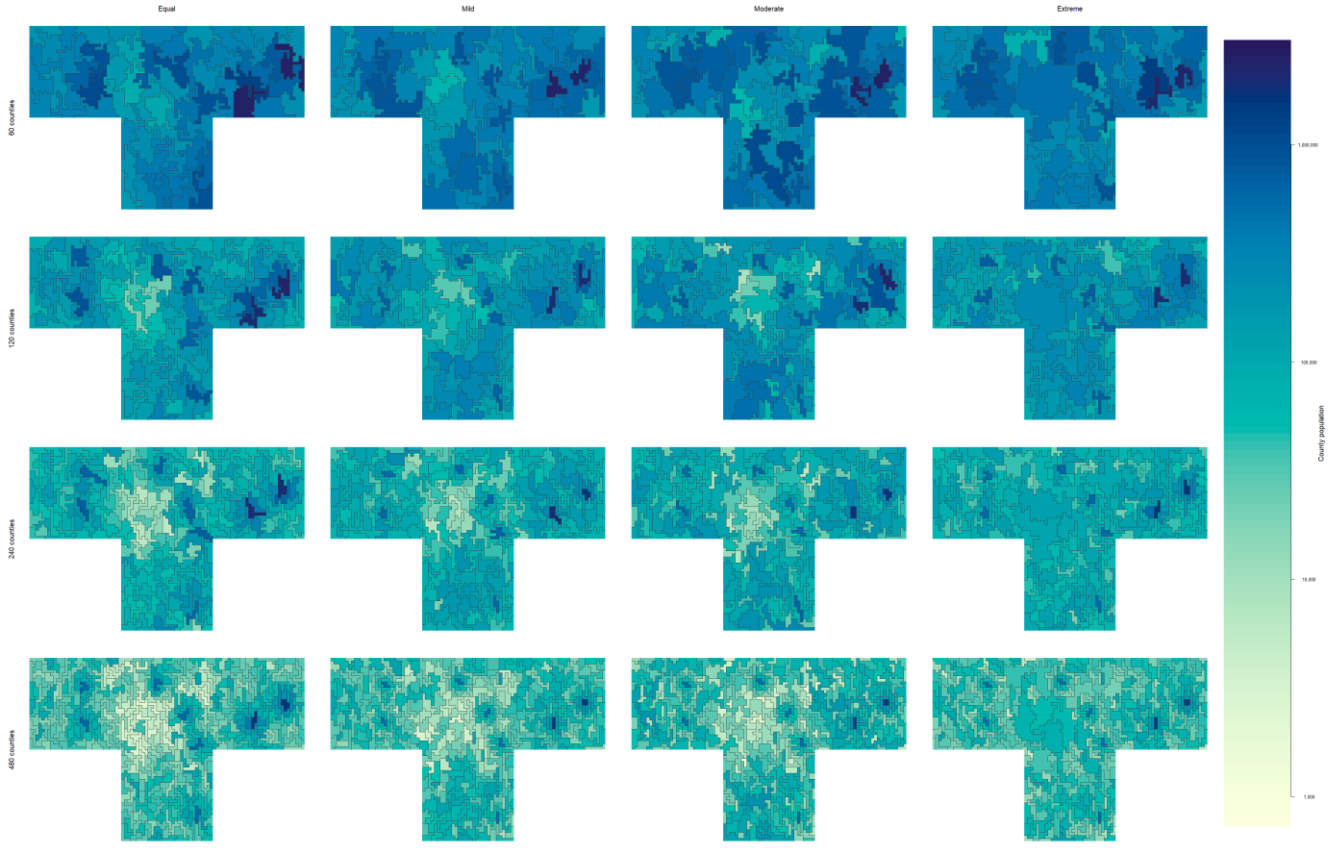
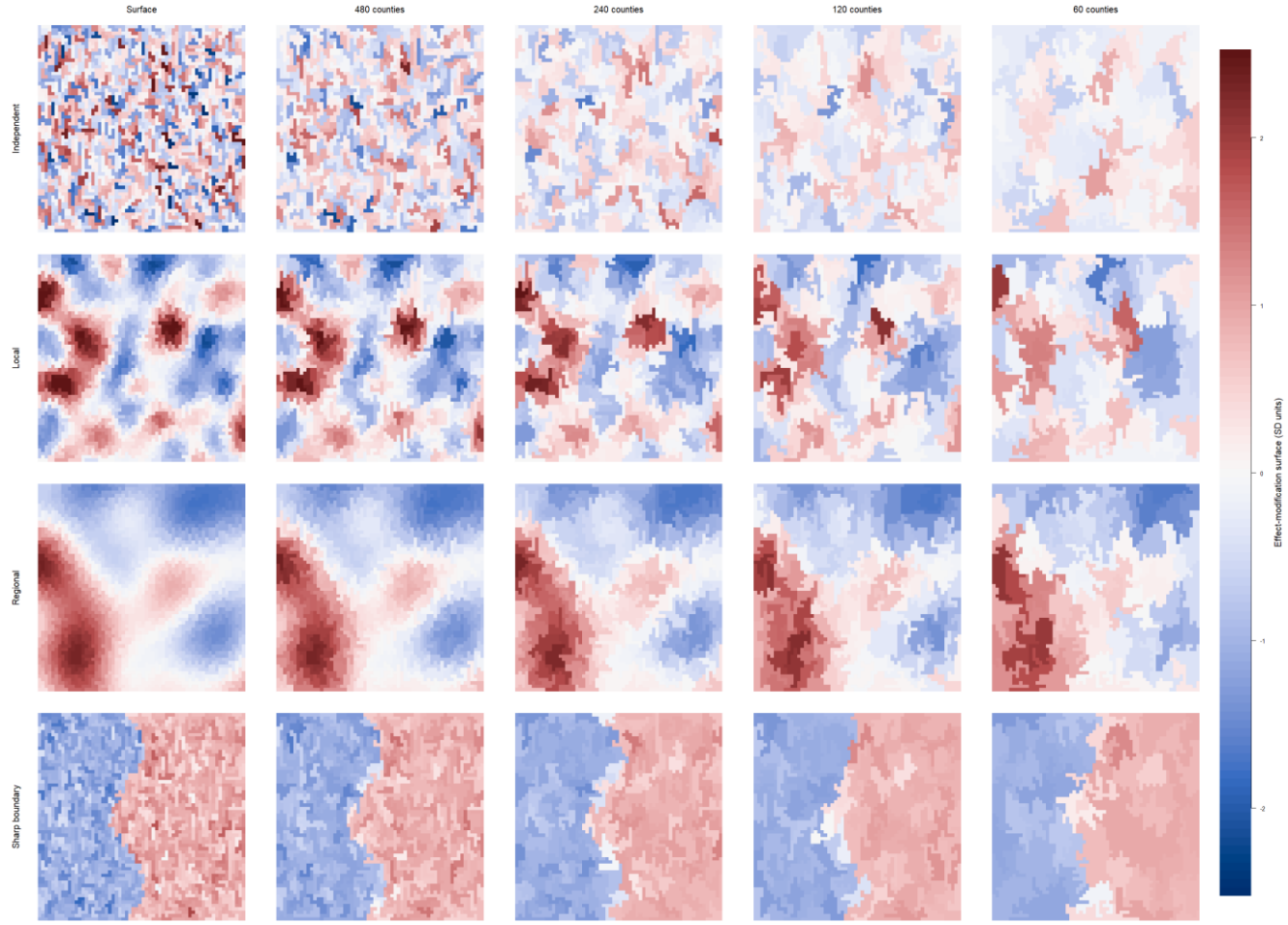
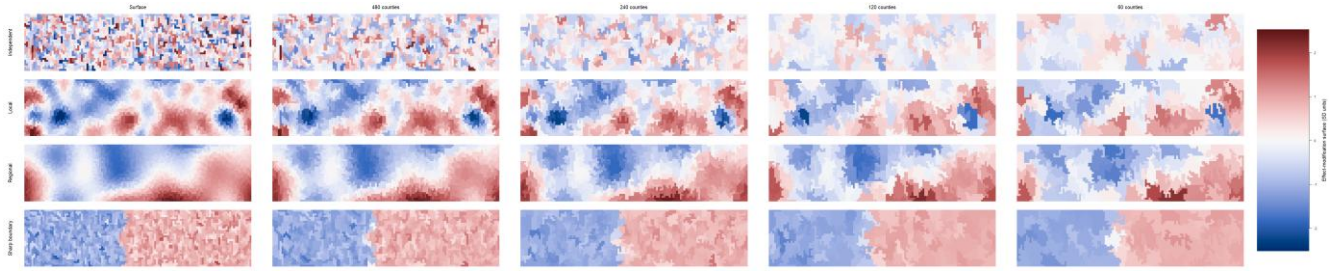


Figure S1. All 64 simulation maps, one panel set per region shape: (a) square, (b) elongated, (c) L-shaped, (d) T-shaped. Rows show 60, 120, 240 and 480 counties, and columns show county size inequality from equal to extreme. Color gives county population (log scale, common across all maps). Lines are county boundaries.

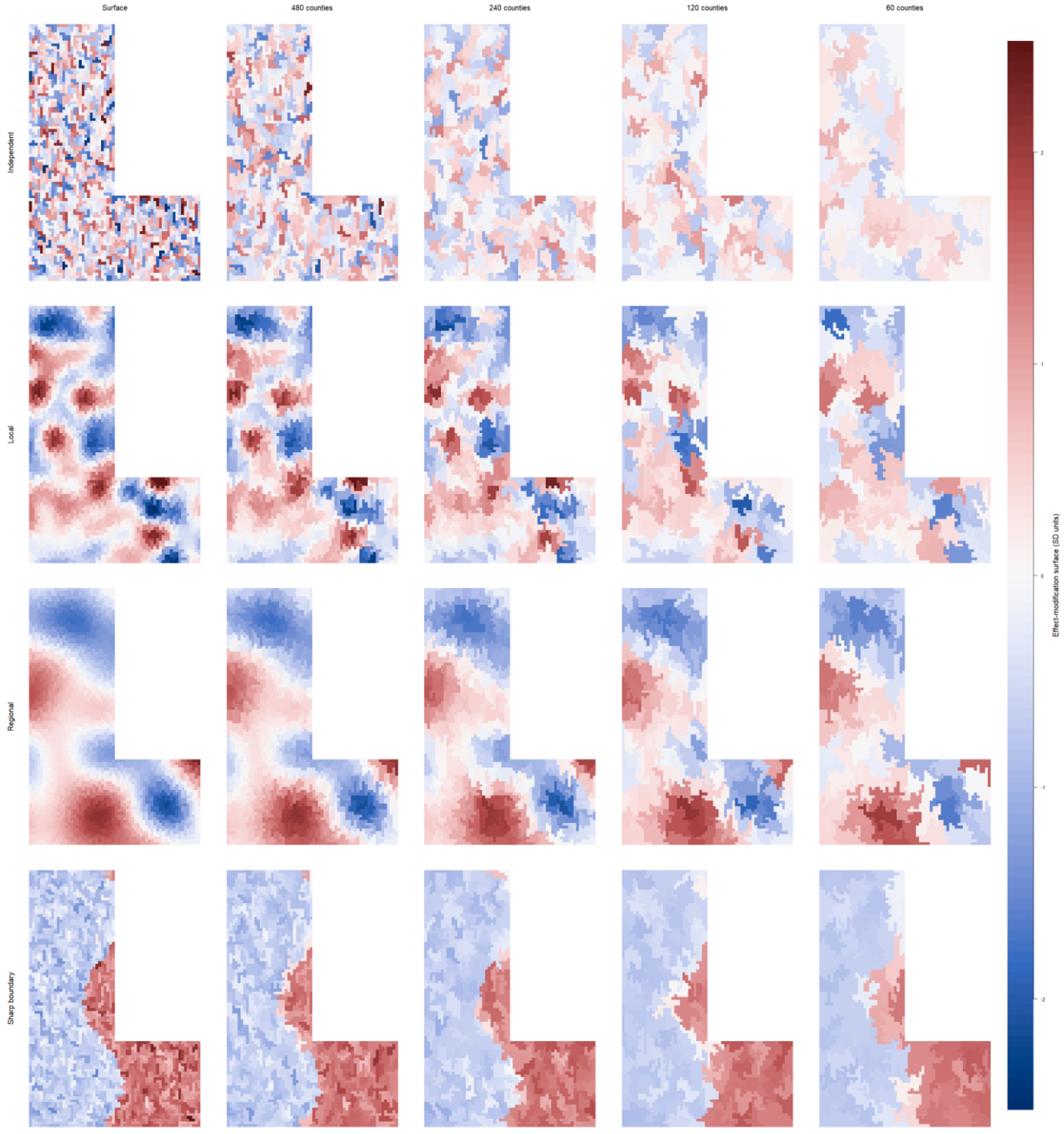
(a)


Surface
480 counties
240 counties
120 counties
60 counties
Independent
Local
Regional
Sharp boundary
Effect-modification surface (SD units)

(b)



(c)


Surface
480 counties
240 counties
120 counties
60 counties
Independent
Local
Regional
Sharp boundary
Effect-modification surface (SD units)
2
1
0
-1
-2

(d)

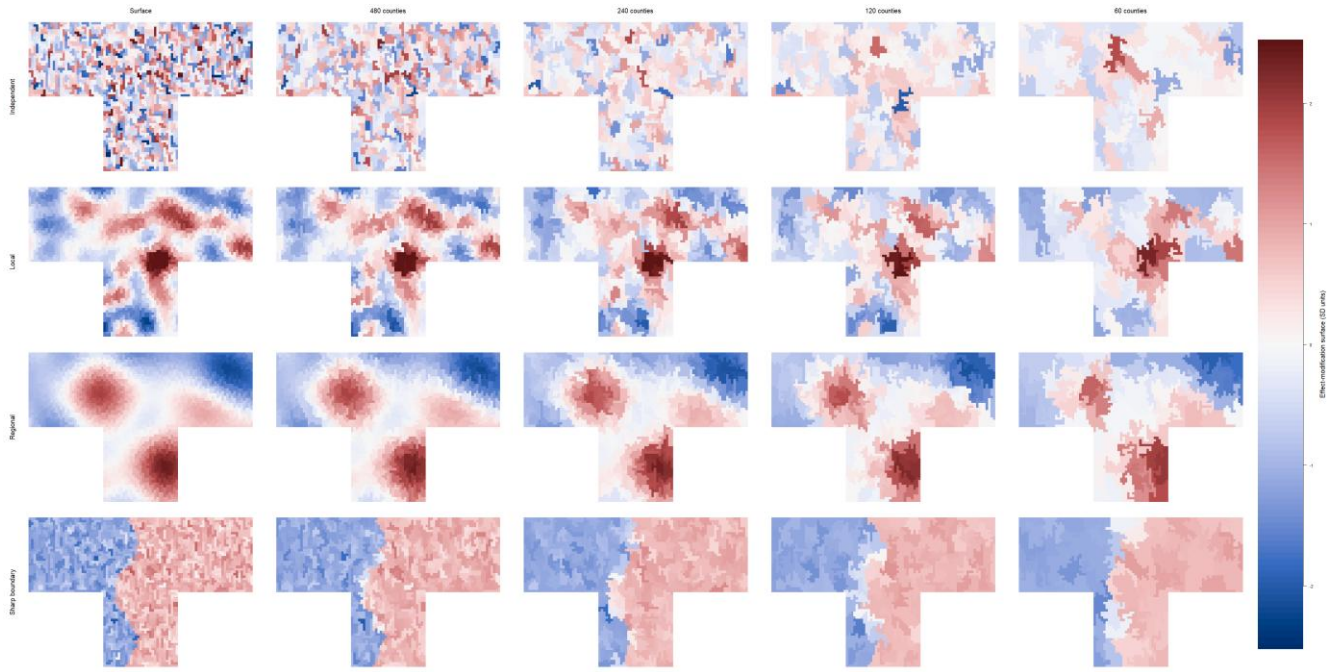


Figure S2. Spatial patterns of between-county variation in the temperature effect, one panel set per region shape: (a) square, (b) elongated, (c) L-shaped, (d) T-shaped. Rows show the four patterns (spatially independent, local, regional, and a sharp boundary). The first column shows the underlying effect-modification surface, and the remaining columns show its county averages for 480, 240, 120 and 60 counties of equal size. Color gives the surface in standard-deviation units, on a common scale.

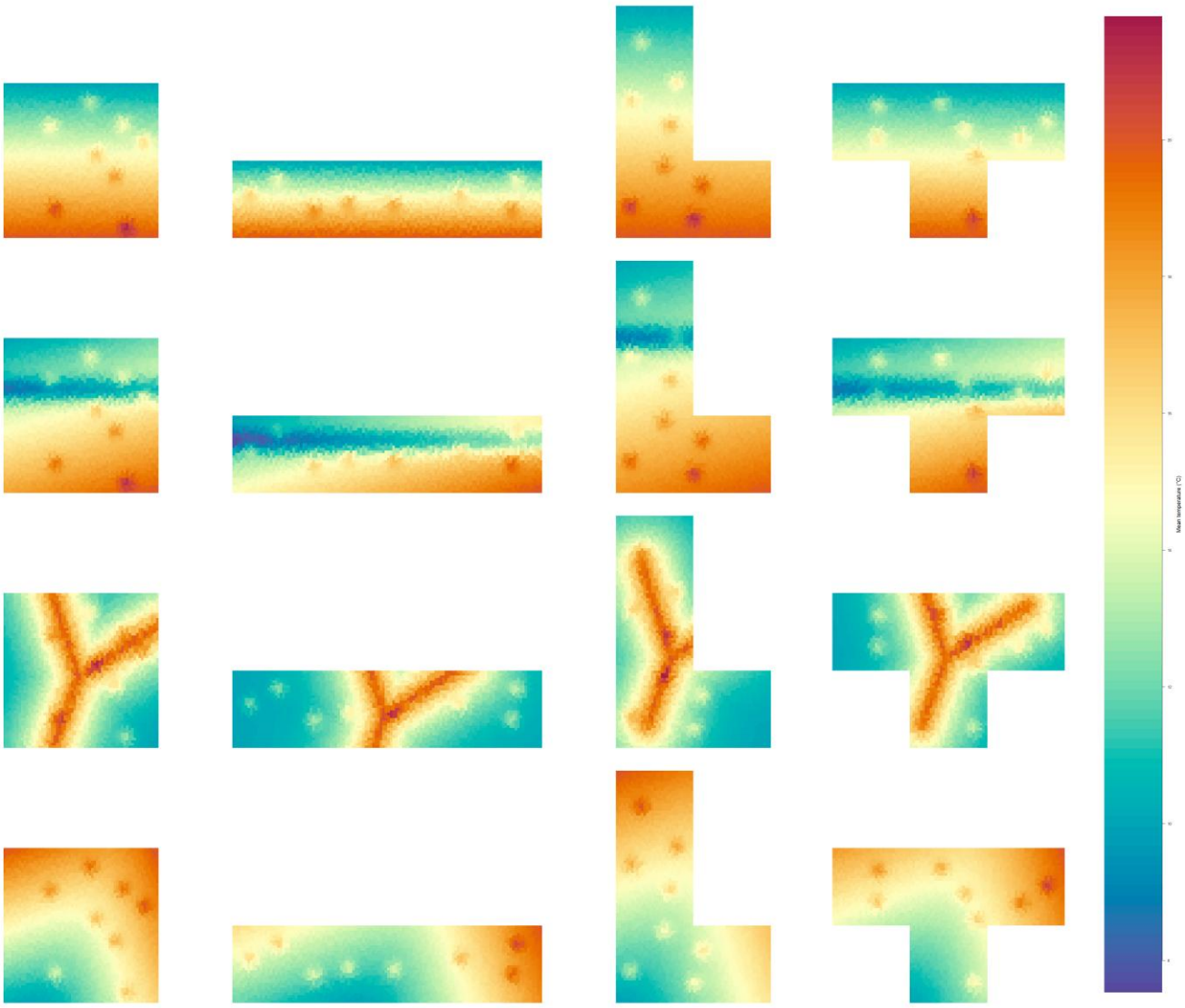


Figure S3. Mean-temperature layouts used in the simulation. Rows show the four layouts: a north-south gradient, the same gradient with a cool ridge, warm river valleys, and warm north and east edges. Columns show the four region shapes. Each layout spans 10 °C across the region, with urban heat islands added. Color gives mean temperature (°C).

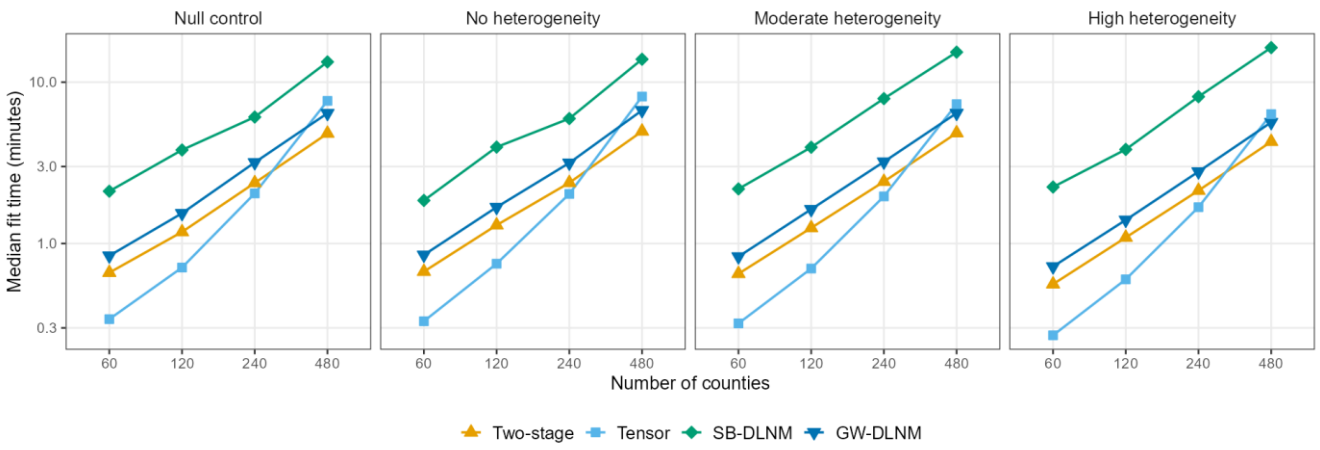
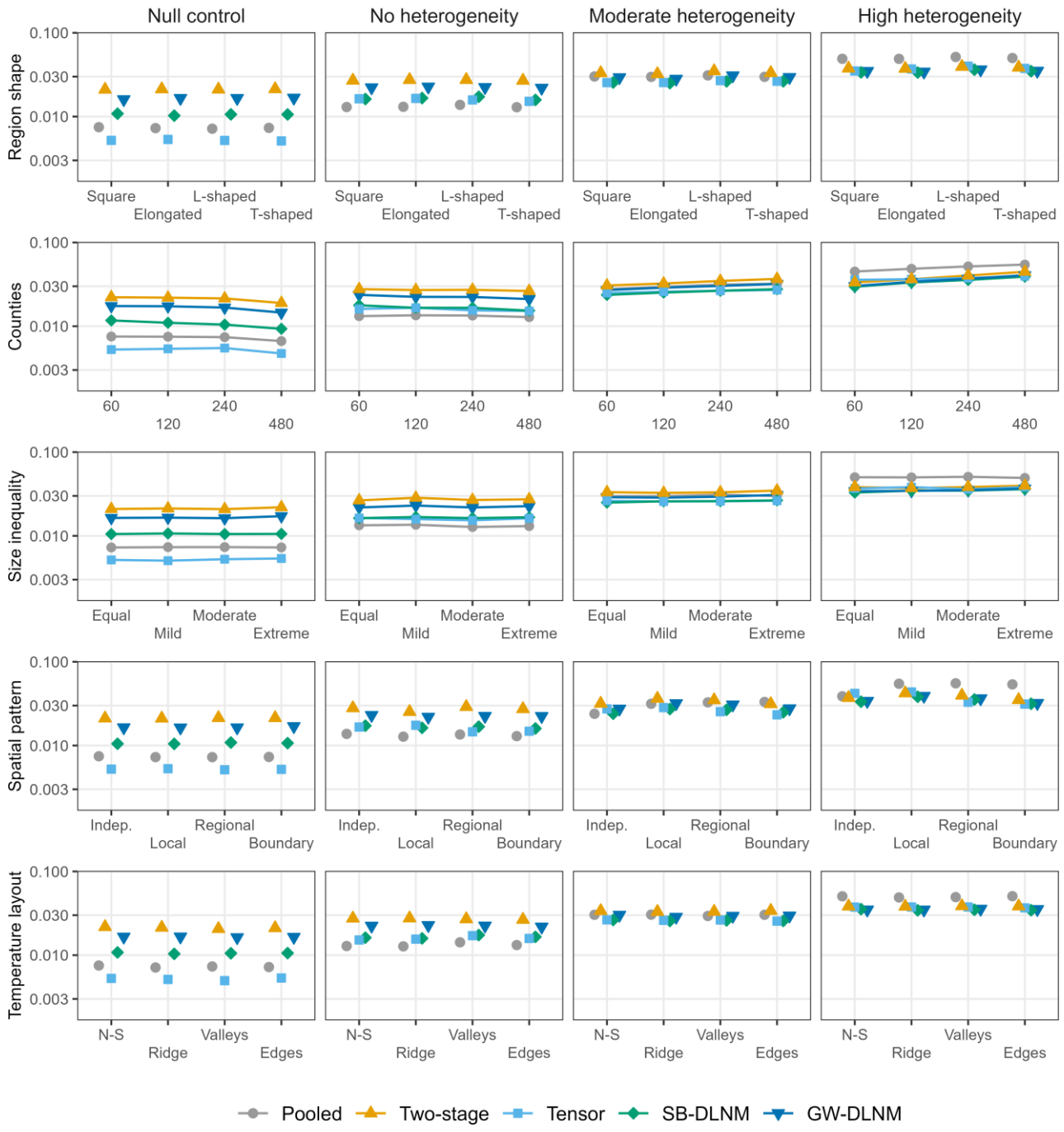


Figure S4. Median fitting time per replicate by method, number of counties and scenario block, for successful fits (log scale). GW-DLNM times include the Two-stage meta-regression that GW-DLNM uses as its starting value. Pooled, whose fits took under a second, is not shown.

(a)


Null control
No heterogeneity
Moderate heterogeneity
High heterogeneity
Region shape
Counties
Size inequality
Spatial pattern
Temperature layout
0.100
0.030
0.010
0.003
Square
L-shaped
Elongated
T-shaped
60
120
240
480
Equal
Moderate
Mild
Extreme
Indep.
Regional
Local
Boundary
N-S
Valleys
Ridge
Edges
Pooled
Two-stage
Tensor
SB-DLNM
GW-DLNM

(b)

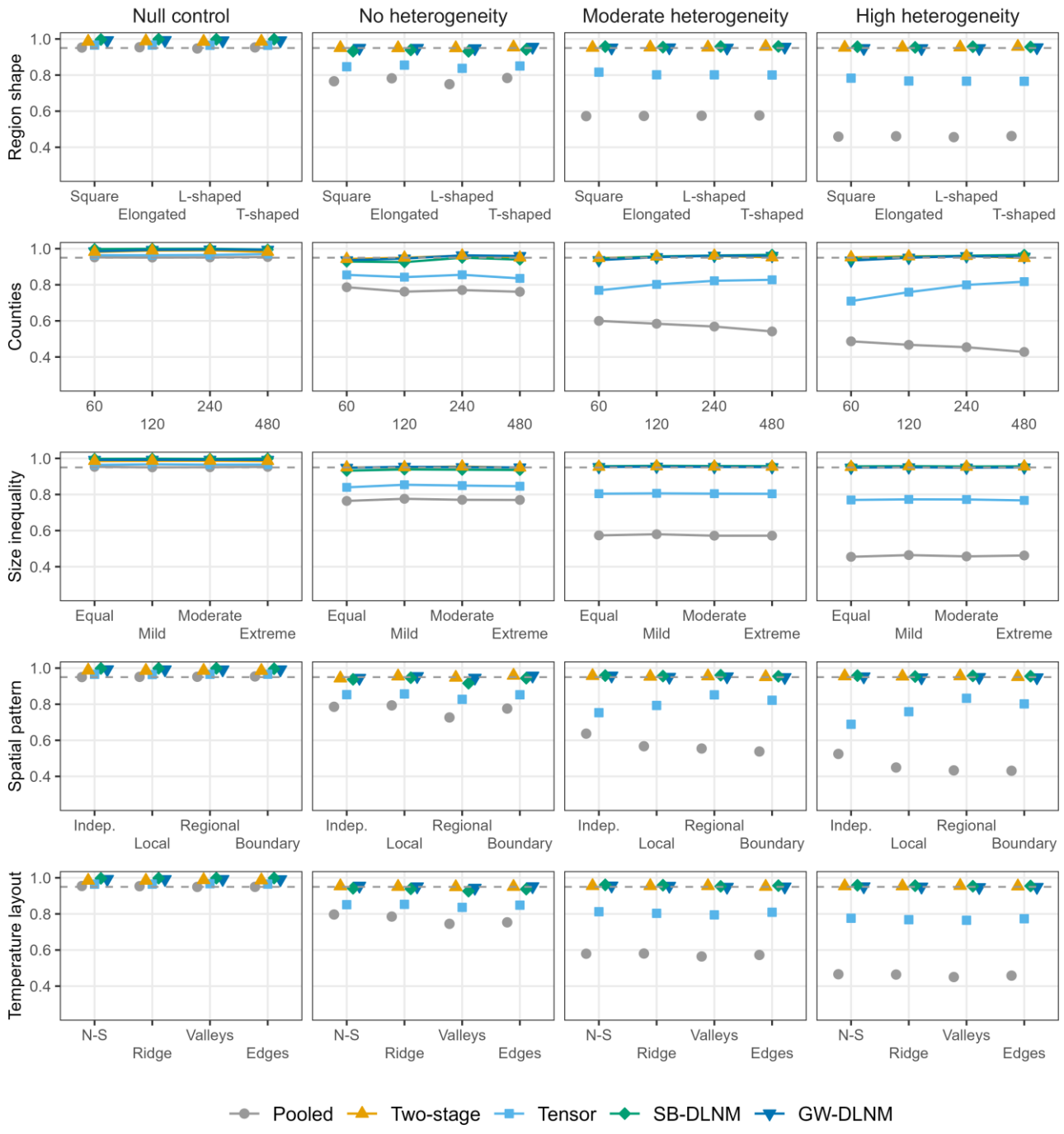
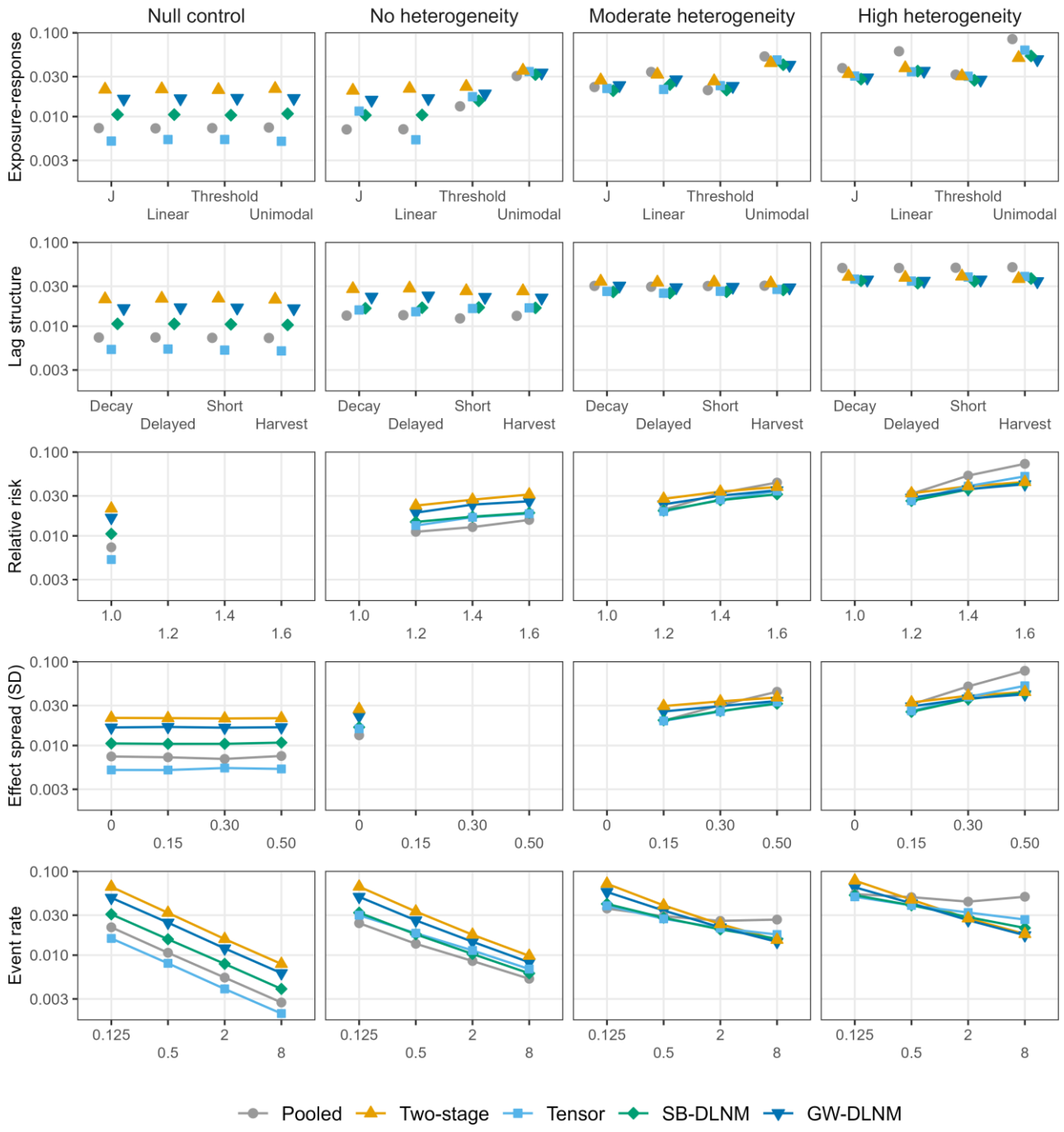


Figure S5. Whole-curve performance across the geographic design settings in the four scenario blocks (columns): (a) median RMSE of the cumulative log RR over each county's own 5th to 95th percentile temperature range (log scale) and (b) mean pointwise coverage of 95% confidence intervals over the same range (line at 0.95). Summaries are over the replicates in which all five methods produced estimates: null control (7,508 replicates), no heterogeneity (5,647), moderate heterogeneity (16,896) and high heterogeneity (16,754).

(a)


Null control
No heterogeneity
Moderate heterogeneity
High heterogeneity
Exposure-response
Lag structure
Relative risk
Effect spread (SD)
Event rate
0.100
0.030
0.010
0.003
J
Linear
Threshold
Unimodal
Decay
Delayed
Short
Harvest
1.0
1.2
1.4
1.6
0
0.15
0.30
0.50
0.125
0.5
2
8
Pooled
Two-stage
Tensor
SB-DLNM
GW-DLNM

(b)

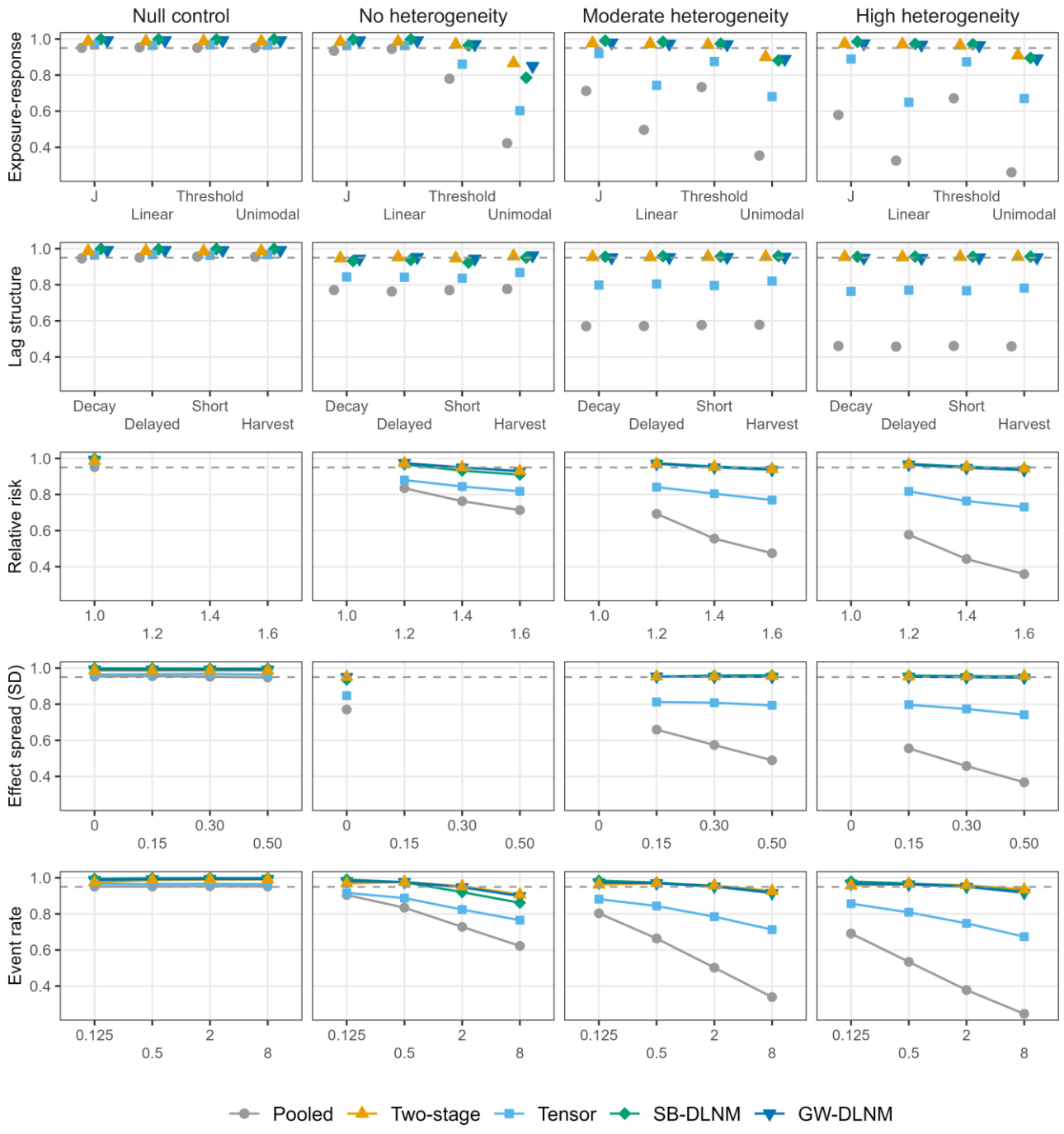
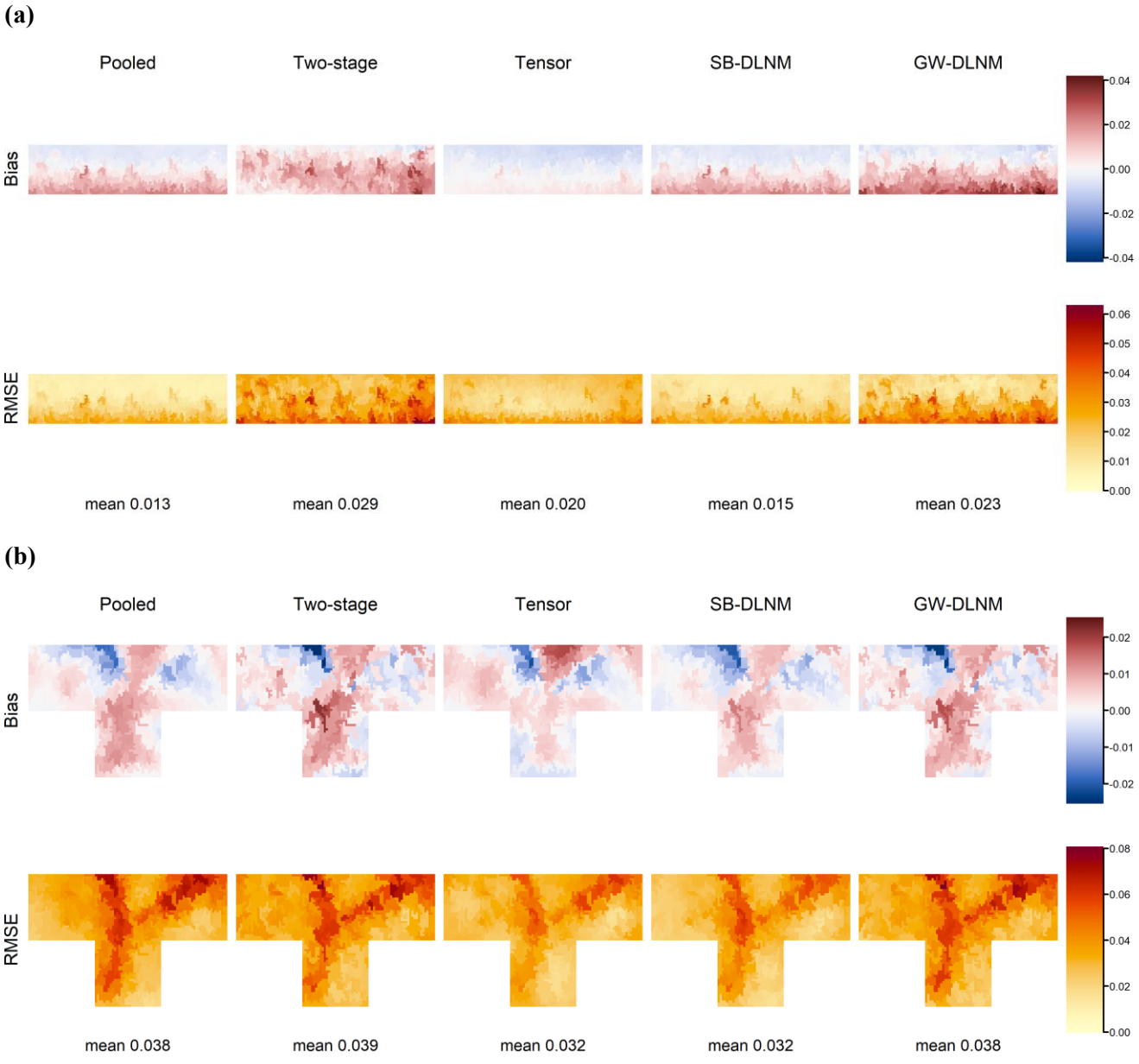


Figure S6. Whole-curve performance across the epidemiological design settings, laid out as in Figure S5: (a) median RMSE and (b) mean pointwise coverage.


(a)
Pooled
Two-stage
Tensor
SB-DLNM
GW-DLNM
Bias
RMSE
0.04
0.02
0.00
-0.02
-0.04
0.06
0.05
0.04
0.03
0.02
0.01
0.00
mean 0.013
mean 0.029
mean 0.020
mean 0.015
mean 0.023
(b)
Pooled
Two-stage
Tensor
SB-DLNM
GW-DLNM
Bias
RMSE
0.02
0.01
0.00
-0.01
-0.02
0.08
0.06
0.04
0.02
0.00
mean 0.038
mean 0.039
mean 0.032
mean 0.032
mean 0.038

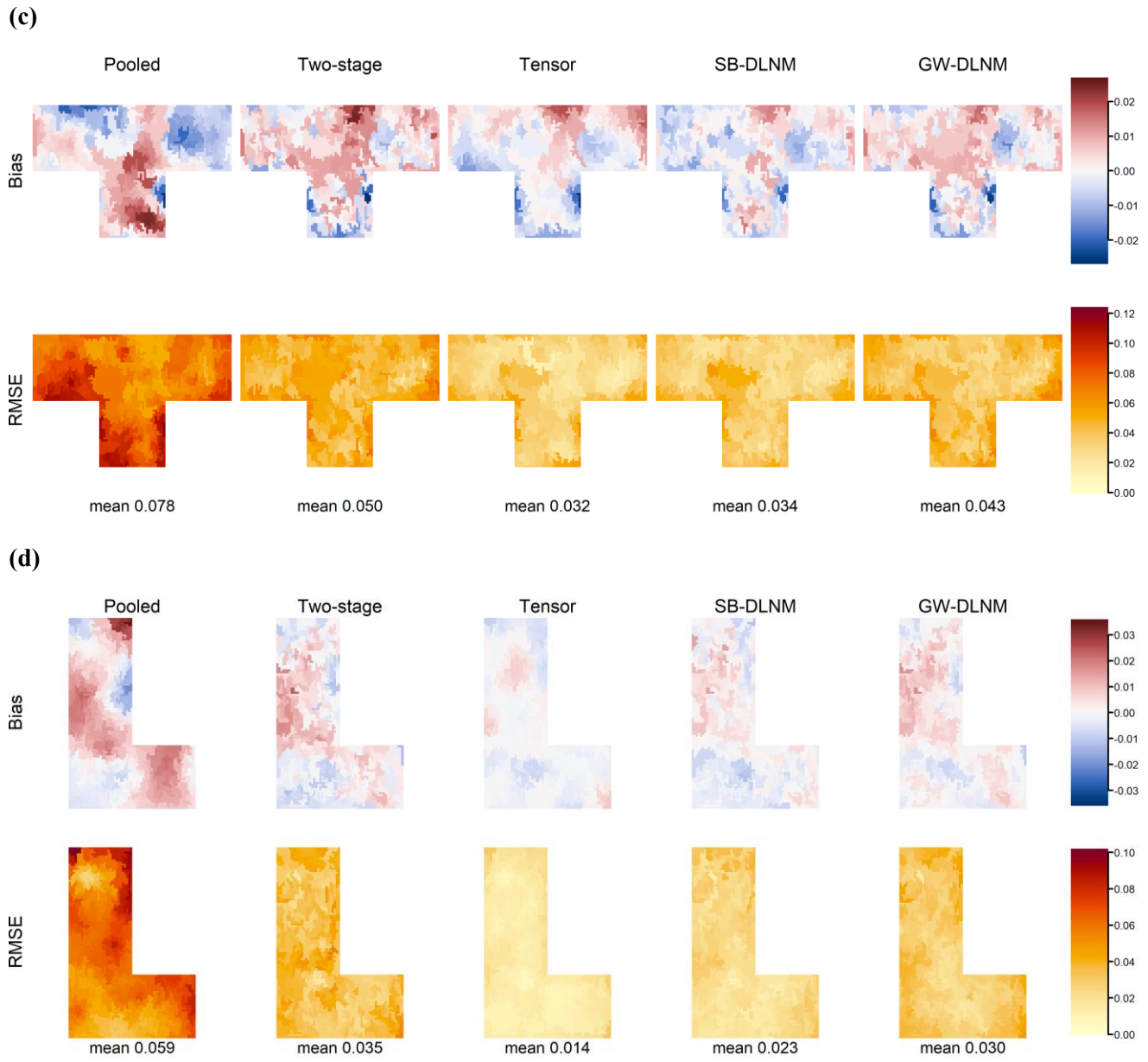


Figure S7. Further examples of county-level errors, laid out as in Figure 6, for scenarios with median I^2 of (a) 0.03, (b) 0.20, (c) 0.40 and (d) 0.61. Examples were chosen by a fixed rule: 240 counties, a median of 300 to 5,000 total events per county, a true RR above one, and the median I^2 closest to each target value. County errors are averaged over each method's successful replicates.

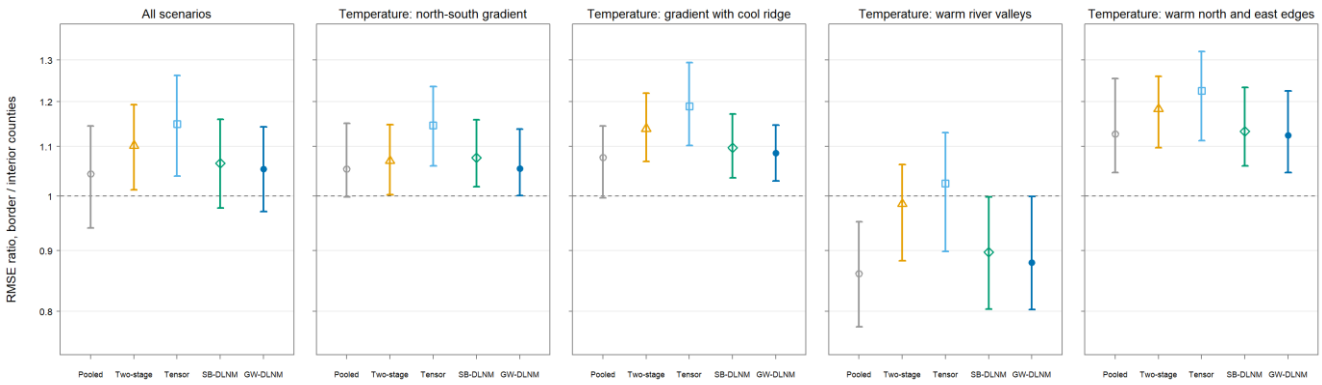


Figure S8. Accuracy in border counties relative to interior counties. A border county has at least one grid cell adjacent to the edge of the region. For each scenario and method, the ratio is the mean RMSE of the cumulative log RR at each county's own 99th-percentile temperature over border counties divided by that over interior counties. Values above one mean larger errors at the border. Points show the median and bars the interquartile range over scenarios with a true effect, for all of them and by mean-temperature layout. County errors are averaged over each method's successful replicates. Log scale.

Supplementary Tables

Table S1. Null control block (RR = 1.0, 256 scenarios, 7,508 replicates in which all five methods produced estimates): share of counties whose 95% confidence interval for the log RR at the county's own 99th-percentile temperature excludes zero, overall and by event rate and number of counties, and the median RMSE of the spurious effects.

Replicates	n	Pooled	Two-stage	Tensor	SB-DLNM	GW-DLNM
All replicates	7,508	0.049	0.035	0.039	0.001	0.014
Event rate 0.125	1,767	0.040	0.084	0.040	0.004	0.031
Event rate 0.5	1,907	0.047	0.031	0.039	0.001	0.011
Event rate 2	1,915	0.052	0.015	0.039	0.001	0.008
Event rate 8	1,919	0.057	0.012	0.039	0.001	0.008
Number of counties 60	1,916	0.046	0.021	0.041	0.001	0.015
Number of counties 120	1,913	0.050	0.018	0.040	0.001	0.008
Number of counties 240	1,907	0.053	0.031	0.039	0.000	0.009
Number of counties 480	1,772	0.047	0.071	0.036	0.004	0.024
Median RMSE of spurious effects	7,508	0.011	0.040	0.011	0.018	0.029

Rows other than the last give the share of counties whose 95% interval excludes zero.

Table S2. Meta-model of simulation performance: share of the variation in log scenario-mean RMSE, at the own 99th percentile and over the whole curve, explained by each design setting, the heterogeneity block, their interactions with method and all two-way interactions among the settings and the block, over the 1,344 scenarios with a true effect (no-heterogeneity control, moderate and high heterogeneity blocks). Shares are from type II sums of squares. The spread-by-block interaction has 2 degrees of freedom because the high heterogeneity block has no scenarios with an effect spread of 0. The upper part gives the shares summed by group of terms. The lower part lists every term with a share of at least 0.5% for either response. All terms are given in Online Resource 4.

Group of terms	**Own 99th percentile (%)**	**Whole curve (%)**
Epidemiological settings	61.3	69.7
Epidemiological × epidemiological	9.9	4.6
Epidemiological settings × method	8.1	10.1
Residual	6.9	4.6
Heterogeneity block	4.9	3.9
Geographic settings	2.3	1.9
Method	1.9	1.6
Geographic × epidemiological	1.6	1.0
Geographic settings × method	1.1	0.9
Geographic × geographic	0.9	0.5
Heterogeneity block × method	0.6	0.8
Block × epidemiological	0.5	0.3
Block × geographic	0.0	0.0

Term	**df**	**Own 99th percentile (%)**	**Whole curve (%)**
Event rate	3	21.8	32.0
Effect spread	3	19.9	13.0
Effect size (RR)	2	11.2	7.9
Exposure-response shape	3	8.3	16.8
Heterogeneity block	1	4.9	3.9
Effect spread × Exposure-response shape	9	3.9	1.5
Method × Event rate	12	3.6	4.5
Method × Effect spread	12	2.6	2.7
Exposure-response shape × Event rate	9	2.5	1.1
Method	4	1.9	1.6
Effect spread × Event rate	9	1.5	1.0
Number of counties	3	1.5	1.3
Exposure-response shape × Effect size (RR)	6	1.1	0.3
Method × Effect size (RR)	8	1.1	1.1
Method × Spatial pattern	12	0.8	0.8
Method × Exposure-response shape	12	0.7	1.7
Spatial pattern	3	0.7	0.6
Method × Heterogeneity block	4	0.6	0.8

Table S3. Accuracy and coverage by exposure-response shape in the three blocks with an effect. Cells give the median RMSE of the cumulative log RR, with the mean coverage of 95% intervals in parentheses, at each county's own 99th percentile of temperature, at its own 95th percentile, and over the whole curve between its own 5th and 95th percentiles. Summaries are over the replicates in which all five methods produced estimates.

Own 99th percentile

Block	Shape	Pooled	Two-stage	Tensor	SB-DLNM	GW-DLNM
No-heterogeneity control	J-shaped	0.011 (0.94)	0.036 (0.97)	0.023 (0.93)	0.018 (1.00)	0.026 (0.98)
No-heterogeneity control	Linear	0.011 (0.95)	0.036 (0.97)	0.011 (0.96)	0.017 (1.00)	0.025 (0.99)
No-heterogeneity control	Threshold	0.022 (0.75)	0.041 (0.94)	0.028 (0.86)	0.027 (0.97)	0.033 (0.96)
No-heterogeneity control	Unimodal	0.089 (0.44)	0.094 (0.82)	0.077 (0.67)	0.089 (0.79)	0.085 (0.81)
Moderate heterogeneity	J-shaped	0.065 (0.39)	0.066 (0.94)	0.059 (0.71)	0.050 (0.95)	0.057 (0.94)
Moderate heterogeneity	Linear	0.071 (0.36)	0.061 (0.94)	0.046 (0.72)	0.048 (0.98)	0.054 (0.95)
Moderate heterogeneity	Threshold	0.067 (0.40)	0.069 (0.93)	0.060 (0.69)	0.055 (0.93)	0.061 (0.93)
Moderate heterogeneity	Unimodal	0.096 (0.36)	0.099 (0.83)	0.087 (0.70)	0.096 (0.83)	0.091 (0.83)
High heterogeneity	J-shaped	0.124 (0.23)	0.083 (0.94)	0.089 (0.63)	0.072 (0.94)	0.075 (0.93)
High heterogeneity	Linear	0.130 (0.21)	0.076 (0.95)	0.075 (0.62)	0.068 (0.98)	0.071 (0.94)
High heterogeneity	Threshold	0.120 (0.25)	0.091 (0.93)	0.091 (0.62)	0.077 (0.91)	0.080 (0.92)
High heterogeneity	Unimodal	0.112 (0.28)	0.112 (0.84)	0.103 (0.70)	0.109 (0.85)	0.104 (0.83)

Own 95th percentile

Block	Shape	Pooled	Two-stage	Tensor	SB-DLNM	GW-DLNM
No-heterogeneity control	J-shaped	0.009 (0.94)	0.027 (0.99)	0.017 (0.95)	0.015 (1.00)	0.020 (0.99)
No-heterogeneity control	Linear	0.008 (0.95)	0.027 (0.99)	0.010 (0.96)	0.014 (1.00)	0.020 (0.99)
No-	Threshold	0.016	0.028	0.020	0.020	0.023

heterogeneity control		(0.75)	(0.98)	(0.88)	(0.98)	(0.97)
No-heterogeneity control	Unimodal	0.034 (0.53)	0.039 (0.96)	0.030 (0.80)	0.033 (0.93)	0.035 (0.95)
Moderate heterogeneity	J-shaped	0.049 (0.43)	0.048 (0.96)	0.043 (0.76)	0.038 (0.97)	0.043 (0.96)
Moderate heterogeneity	Linear	0.059 (0.36)	0.049 (0.96)	0.039 (0.72)	0.039 (0.98)	0.044 (0.96)
Moderate heterogeneity	Threshold	0.047 (0.46)	0.047 (0.96)	0.042 (0.76)	0.038 (0.96)	0.042 (0.95)
Moderate heterogeneity	Unimodal	0.057 (0.39)	0.049 (0.95)	0.048 (0.80)	0.046 (0.95)	0.045 (0.94)
High heterogeneity	J-shaped	0.091 (0.26)	0.060 (0.96)	0.065 (0.68)	0.053 (0.96)	0.054 (0.95)
High heterogeneity	Linear	0.109 (0.20)	0.062 (0.96)	0.064 (0.61)	0.055 (0.98)	0.057 (0.95)
High heterogeneity	Threshold	0.084 (0.30)	0.061 (0.96)	0.063 (0.69)	0.053 (0.95)	0.055 (0.94)
High heterogeneity	Unimodal	0.093 (0.28)	0.059 (0.95)	0.067 (0.75)	0.059 (0.95)	0.055 (0.93)

Whole curve

Block	**Shape**	**Pooled**	**Two-stage**	**Tensor**	**SB-DLNM**	**GW-DLNM**
No-heterogeneity control	J-shaped	0.007 (0.93)	0.020 (0.98)	0.012 (0.96)	0.010 (1.00)	0.016 (0.99)
No-heterogeneity control	Linear	0.007 (0.95)	0.022 (0.99)	0.005 (0.96)	0.010 (1.00)	0.017 (0.99)
No-heterogeneity control	Threshold	0.013 (0.78)	0.023 (0.97)	0.017 (0.86)	0.015 (0.96)	0.019 (0.97)
No-heterogeneity control	Unimodal	0.031 (0.42)	0.035 (0.87)	0.034 (0.60)	0.032 (0.79)	0.033 (0.85)
Moderate heterogeneity	J-shaped	0.023 (0.71)	0.027 (0.98)	0.022 (0.92)	0.020 (0.99)	0.024 (0.98)
Moderate heterogeneity	Linear	0.034 (0.50)	0.032 (0.97)	0.021 (0.74)	0.024 (0.98)	0.028 (0.97)
Moderate heterogeneity	Threshold	0.021 (0.73)	0.027 (0.97)	0.023 (0.88)	0.021 (0.97)	0.023 (0.97)
Moderate heterogeneity	Unimodal	0.052 (0.35)	0.044 (0.90)	0.047 (0.68)	0.042 (0.88)	0.041 (0.89)
High heterogeneity	J-shaped	0.038 (0.58)	0.032 (0.97)	0.030 (0.89)	0.028 (0.98)	0.029 (0.97)
High	Linear	0.060	0.038	0.034	0.035	0.035

heterogeneity		(0.33)	(0.97)	(0.65)	(0.97)	(0.97)
High heterogeneity	Threshold	0.032 (0.67)	0.030 (0.96)	0.030 (0.87)	0.027 (0.97)	0.027 (0.96)
High heterogeneity	Unimodal	0.084 (0.26)	0.050 (0.91)	0.062 (0.67)	0.053 (0.89)	0.048 (0.89)

Additional Online Resources

Online Resource 2. County-level error maps for every scenario, laid out as in Figure 6, one scenario per page.

Online Resource 3. Index of the scenarios in Online Resource 2: page number, scenario id and the level of every design setting.

Online Resource 4. Full meta-model table (Table S2): degrees of freedom, share of variation, F statistic and p-value for every term, for the own 99th percentile and for the whole curve.