Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Forest cooling is not natural-vegetation cooling: pair-specific and environmentally conditioned land-cover thermal contrasts across India

Jyoti Sharma
Indian Institute of Science Education and Research, Bhopal
Pankaj Kumar
kumarp@iiserb.ac.in

Indian Institute of Science Education and Research, Bhopal

Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11714418/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Forest cooling is not natural-vegetation cooling: pair-specific and environmentally conditioned land-cover thermal contrasts across India

Jyoti Sharma[1] and Pankaj Kumar[1*]

[1]Indian Institute of Science Education and Research, Bhopal, Bhopal, India

*Corresponding Author Email ID: kumarp@iiserb.ac.in

Abstract

Land cover is often treated as a fixed warming or cooling agent, with natural vegetation assumed to be cooler than human-modified surfaces. Here, we challenge this assumption by quantifying land-surface temperature contrasts (ΔLST) among seven major land-cover classes across India using two decades (2003-2022) of satellite observations and a spatially matched moving-window framework complemented by environmental-gradient, statistical and machine-learning analyses. Of 21 possible pairwise contrasts, 17 were sufficiently co-located for analysis. During daytime, forests were the coolest surface, cooler than cropland in 99.6% of matched pairs (mean ΔLST = -2.32 °C) and urban land in 97.6% (-2.57 °C). However, this cooling was forest-specific: savanna, shrubland and grassland were collectively cooler than cropland in only 33.0% of comparisons and averaged 0.73 °C warmer. At night, 11 of 17 contrasts reversed sign, forests became warmer than cropland (+0.40 °C), and urban land was warmer than all vegetated surfaces. Thermal contrasts varied across biogeographic zones and responded nonlinearly to environmental gradients, with breakpoints at background temperatures of 29-44 °C and evapotranspiration-to-potential-evapotranspiration ratios of 0.09-0.40. Forest cooling relative to cropland peaked at intermediate evaporative conditions, whereas water-stressed non-forest vegetation was substantially warmer than neighbouring cropland or urban land. Biophysical attribution identified albedo as the dominant daytime predictor for 12 of 16 contrasts, whereas vegetation structure, productivity and latent heat exchange governed others. These findings establish that land-cover thermal effects are pair-specific, temporally variable and environmentally conditioned, challenging generalized assumptions about natural-vegetation cooling and informing land-surface modelling and nature-based cooling strategies.

Introduction

Natural vegetation is widely associated with land-surface cooling, yet this generalization overlooks substantial thermal differences among vegetation types and environmental settings. Land-cover temperature responses emerge from competing radiative and non-radiative processes, including surface albedo, evapotranspiration, canopy structure and turbulent heat exchange [1–5]. Vegetation can reduce land surface temperature (LST) through enhanced latent heat dissipation while simultaneously altering shortwave radiation absorption through changes in albedo [6,7]. Consequently, vegetation-induced cooling is neither universal nor temporally constant. Global greening has partially offset land warming through increased evapotranspiration [8], while its cooling influence in the Northern Hemisphere is considerably stronger in summer [9]. These findings raise a fundamental question: is land-cover thermal behaviour an intrinsic property of individual vegetation classes, or does it depend on the surfaces being compared and their environmental context?

Evidence increasingly supports the latter interpretation. Tropical forests generally exhibit strong evaporative cooling, whereas forest thermal effects at higher latitudes depend on seasonal changes in evapotranspiration, snow cover and albedo [6,10,11]. Forest loss and degradation can increase surface temperatures, although the magnitude varies with hydroclimatic conditions [12,13]. Cropland thermal effects are similarly context-dependent: approximately 60% of global croplands are warmer than neighbouring natural ecosystems, whereas 40% are cooler [14]. In tropical Africa, croplands are cooler than grasslands at night, but daytime contrasts depend on hydroclimate and differences in leaf area and turbulent heat exchange [15]. Atmospheric dryness and rising CO_2 further modify vegetation-related cooling through changes in water regulation and land-atmosphere interactions [16,17]. Together, these findings suggest that land-cover thermal contrasts depend not only on vegetation identity but also on the comparison surface, environmental conditions and time of observation.

Despite these advances, research has largely examined individual land-cover transitions or selected contrasts, particularly deforestation, agricultural expansion and urbanization [18–22]. A systematic understanding of thermal relationships across multiple coexisting land covers remains limited. Three uncertainties are particularly important. First, whether cooling associated with forests extends to other natural vegetation types remains unresolved. Second, whether changing thermal and evaporative conditions merely alter contrast magnitudes or fundamentally reverse relative warming and

cooling is poorly understood. Third, similar surface-temperature contrasts may emerge through different combinations of albedo, vegetation structure, evapotranspiration and productivity, complicating their physical interpretation [23,24]. Resolving these uncertainties requires a pairwise framework that explicitly links thermal contrasts with temporal variability, environmental gradients and associated biophysical processes.

India provides an opportunity to investigate these relationships because seven major land covers (forest, savanna, shrubland, grassland, cropland, urban and barren land) coexist across pronounced climatic and biogeographic gradients, ranging from humid tropical regions to arid deserts and high-altitude environments [25,26]. Here, we quantify land-surface temperature contrasts among these classes using two decades (2003-2022) of Moderate Resolution Imaging Spectroradiometer (MODIS) observations and a spatially matched moving-window framework. We restrict comparisons to temporally stable, homogeneous land-cover cells with elevation differences ≤100 m, thereby reducing background climatic and topographic differences while avoiding interpretations of spatial contrasts as direct effects of land-cover transitions. We examine the relative thermal hierarchy of land covers, its reorganization across day-night and seasonal cycles, its dependence on biogeographic and environmental gradients, and the dominant biophysical correlates of individual contrasts. Combining environmental-gradient analysis with complementary statistical and machine-learning approaches, we test whether land-cover thermal behaviour is better understood as a fixed class-specific attribute or as a pair-specific, environmentally conditioned property of the land surface.

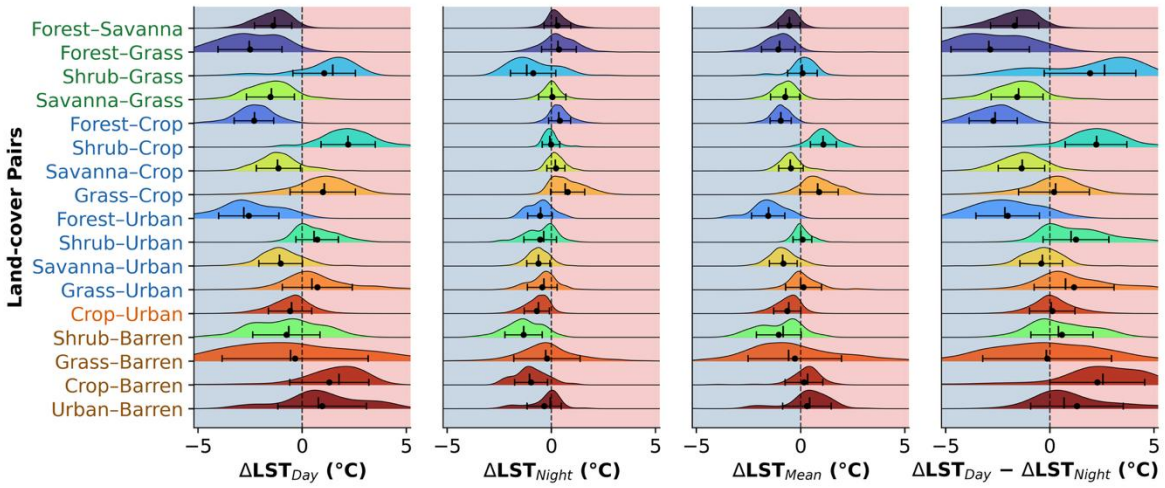
Results

A forest-anchored daytime thermal hierarchy

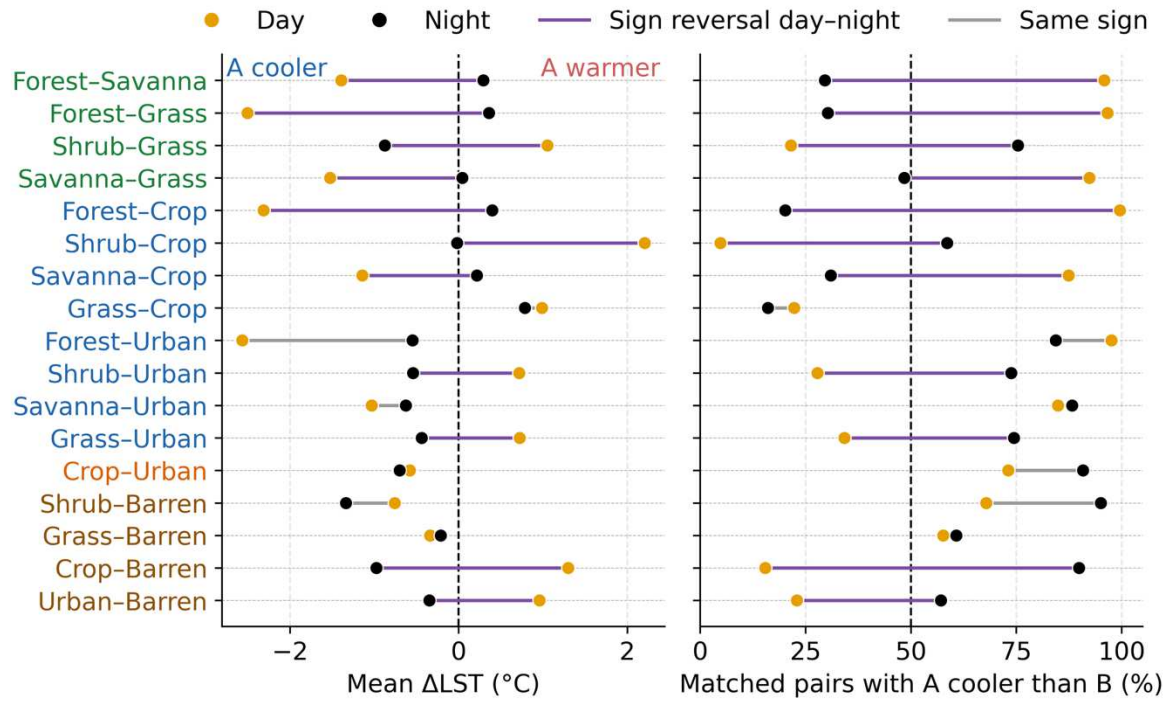
Pairwise comparisons revealed a coherent daytime thermal hierarchy among non-barren land covers: forest < savanna < cropland < urban < grassland < shrubland, where "<" denotes a cooler surface (Fig. 1a-c). Forests anchored the cool end of this hierarchy, with mean daytime land surface temperature (ΔLST) ranging from -1.39 °C relative to savanna to -2.57 °C relative to urban land. This cooling was not only large but remarkably consistent across matched observations: forests were cooler than cropland in 99.6% of pairs (mean ΔLST = -2.32 °C), followed by urban land in 97.6% (-2.57 °C), grassland in 96.6% (-2.50 °C) and savanna in 95.9% (-1.39 °C). Savanna

was also generally cooler than cropland (87.4% of pairs) and urban land (84.9%), whereas shrubland and grassland were mostly warmer than cropland and urban land. Barren land had no fixed position in the hierarchy; the sign of barren-related contrasts depended on the land cover with which it was compared.

a



b-c



d

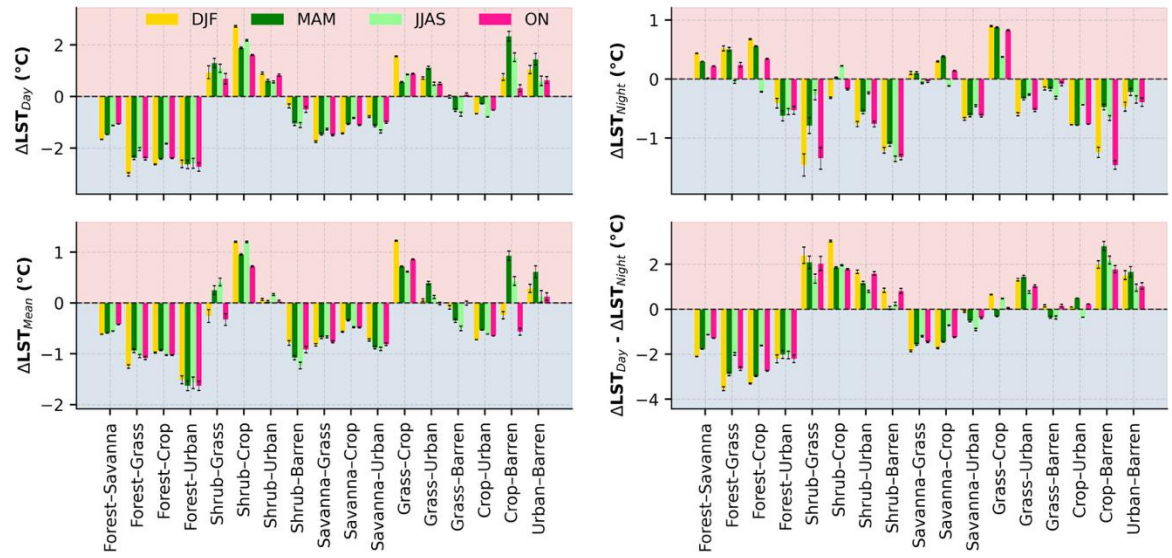


Figure 1. Day-night and seasonal reorganization of pairwise land-cover thermal contrasts across India. (a) Probability-density distributions of ΔLST (= LST of the first class - LST of the second class) for daytime, nighttime and the day-night mean, and of the diurnal difference ΔLST_{day} - ΔLST_{night}; black points and whiskers indicate the mean ±1 standard deviation (SD) while black vertical lines mark the median. **(b)** Mean daytime (orange) and nighttime (black) ΔLST for each pair; purple connectors mark pairs whose sign reverses between day and night. **(c)** Percentage of matched pairs in which the first class is cooler than the second during day and night; the dashed line marks 50%. **(d)** Seasonal mean ΔLST for daytime, nighttime, the day-night mean and the diurnal difference in winter (DJF; December-February), pre-monsoon (MAM; March-May), monsoon (JJAS; June-September) and post-monsoon (ON; October-November); black vertical error bars indicate ±1 standard error (SE). Blue shading and negative values indicate that the first class is cooler; red shading and positive values indicate that it is warmer. Pairs are grouped as natural-natural, natural-anthropogenic, anthropogenic-anthropogenic and barren-related.

Strikingly, the daytime cooling of forests did not extend to natural vegetation as a whole. Relative to cropland, non-forest natural vegetation (savanna, shrubland and grassland) was cooler in only 33.0% of matched pairs and was, on average, 0.73 °C warmer; shrubland was warmer than cropland in 95.2% of pairs (+2.21 °C). Pooling forests with the other natural covers therefore almost cancelled the natural-cropland

contrast: natural vegetation as a whole was cooler than cropland in only 45.6% of comparisons, with a mean ΔLST of +0.15 °C (Supplementary Tables S1-S2; Supplementary Fig. S1). The apparent daytime cooling of “natural vegetation” is thus predominantly a forest signal, and broad natural-versus-anthropogenic categories conceal large thermal differences among vegetation types.

Day-night and seasonal reorganization

The daytime hierarchy was substantially reorganized after sunset (Fig. 1a-c). Eleven of the 17 contrasts changed sign between day and night (Fig. 1b), although two of these (savanna-grassland and shrubland-cropland) became near-neutral at night (|ΔLST| ≤ 0.04 °C). The strongest daytime forest contrasts weakened and reversed: forest-cropland shifted from -2.32 to +0.40 °C, forest-grassland from -2.50 to +0.36 °C and forest-savanna from -1.39 to +0.29 °C, and the proportion of pairs in which forests were cooler than cropland fell from 99.6% during the day to 20.2% at night. Reversals in the opposite direction occurred for shrubland-grassland (+1.05 to -0.87 °C), cropland-barren (+1.30 to -0.98 °C), shrubland-urban (+0.72 to -0.54 °C) and grassland-urban (+0.72 to -0.44 °C). At night, every vegetated class was cooler than urban land, in 74-91% of matched pairs (ΔLST -0.44 to -0.70 °C), consistent with a nocturnal urban surface heat island. Forest-urban was the only contrast in which forests remained cooler during both day (-2.57 °C) and night (-0.55 °C). Because forests were cooler by day but warmer by night than most other covers, they also had a markedly narrower diurnal LST range than every land cover with which they could be compared: the diurnal difference ΔLST_{day} - ΔLST_{night} was -1.7 °C relative to savanna, -2.0 °C relative to urban land and -2.7 to -2.9 °C relative to cropland and grassland (Fig. 1a). These sign changes compressed the thermal hierarchy in the day-night mean and show that relative thermal behaviour during daylight does not persist through the full diurnal cycle. Complete sample sizes, cooling frequencies and mean daytime, nighttime and day-night contrasts are provided in Supplementary Tables S1-S2 and Fig. S1.

Seasonality (winter (DJF; December-February), pre-monsoon (MAM; March-May), monsoon (JJAS; June-September) and post-monsoon (ON; October-November)) further modified these relationships, but in a strongly pair-dependent manner (Fig. 1d). Forest-related contrasts were generally strongest in winter (DJF) and weakest during the monsoon: forest-grassland weakened from -3.02 °C in DJF to -2.04 °C in JJAS,

forest-savanna from -1.67 to -1.12 °C and forest-cropland from -2.63 to -1.84 °C. By contrast, forest-urban remained stable across seasons (-2.59 to -2.73 °C), and savanna-urban strengthened towards the monsoon (-0.78 to -1.36 °C). Seasonal reorganization was also evident at night, when forest-cropland changed sign from +0.68 °C in DJF to -0.22 °C in JJAS, with comparable seasonal sign reversals for forest-grassland and savanna-cropland. Thus, the magnitude, and for several pairs the sign, of land-cover thermal contrasts depends on both time of day and season.

Robustness to spatial dependence

The moving-window observations exhibited significant positive spatial autocorrelation, with India-wide Moran's I reaching 0.83 for some well-represented contrasts ($p = 0.001$; Supplementary Table S3). Nevertheless, the thermal contrasts were robust to spatial thinning. For daytime contrasts with >30 observations, reducing the sample to approximately 11.1% of the original pairs via nine non-overlapping spatial phases changed mean ΔLST by only -0.145 to +0.030 °C, corresponding to relative changes of 0.01-16.28% (Supplementary Table S4a; Supplementary Fig. S3). The largest relative changes occurred for sparsely represented contrasts, whereas several well-sampled contrasts changed by <1%; for example, forest-cropland changed from -2.315 to -2.318 °C (0.14%) and grassland-cropland from +0.987 to +0.988 °C (0.07%). The sign of the major contrasts was preserved across all nine phases, although sparsely represented contrasts showed greater phase-to-phase variability (Supplementary Tables S4a-S4c; Supplementary Fig. S3). Equivalence testing confirmed robustness for well-represented contrasts; for forest-cropland, the mean thinning-induced change was -0.003 °C (90% confidence interval (CI): -0.032 to +0.025 °C), within the predefined ±0.05 °C equivalence margin ($P_{TOST} = 0.008$). Despite substantial spatial autocorrelation, therefore, the magnitude and sign of the principal India-wide contrasts are insensitive to overlap in the sampling design.

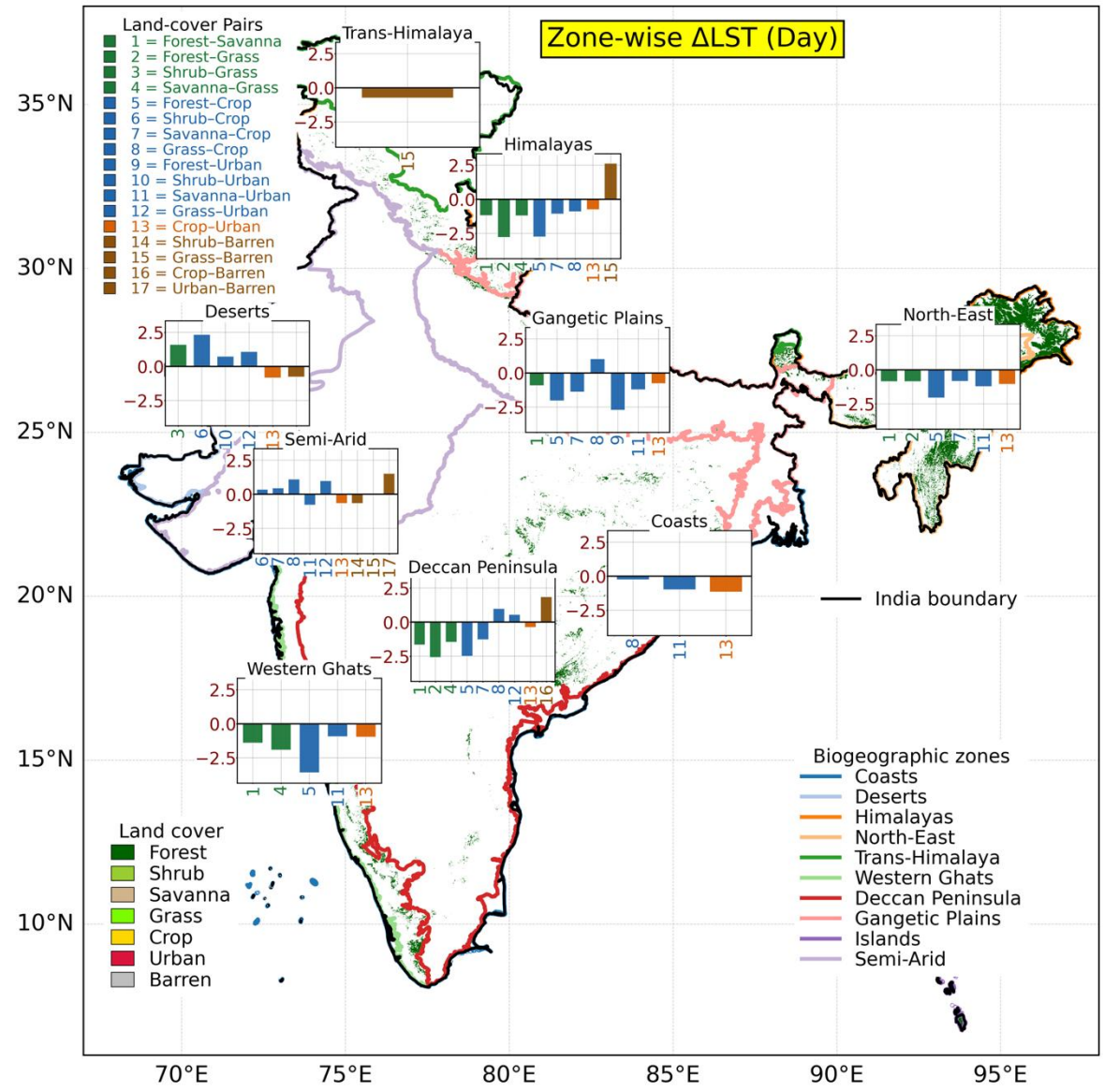
Biogeographic and environmental modulation

Biogeographic setting strongly modified the magnitude and, for several pairs, the sign of thermal contrasts (Fig. 2a,b). Forests were generally cooler than other natural vegetation by day, but the strength of the contrast varied geographically: forest-savanna ΔLST ranged from -0.86 °C in the North-East to -1.66 °C in the Deccan Peninsula, and forest-grassland reached -2.79 °C in the Himalayas. Forests were

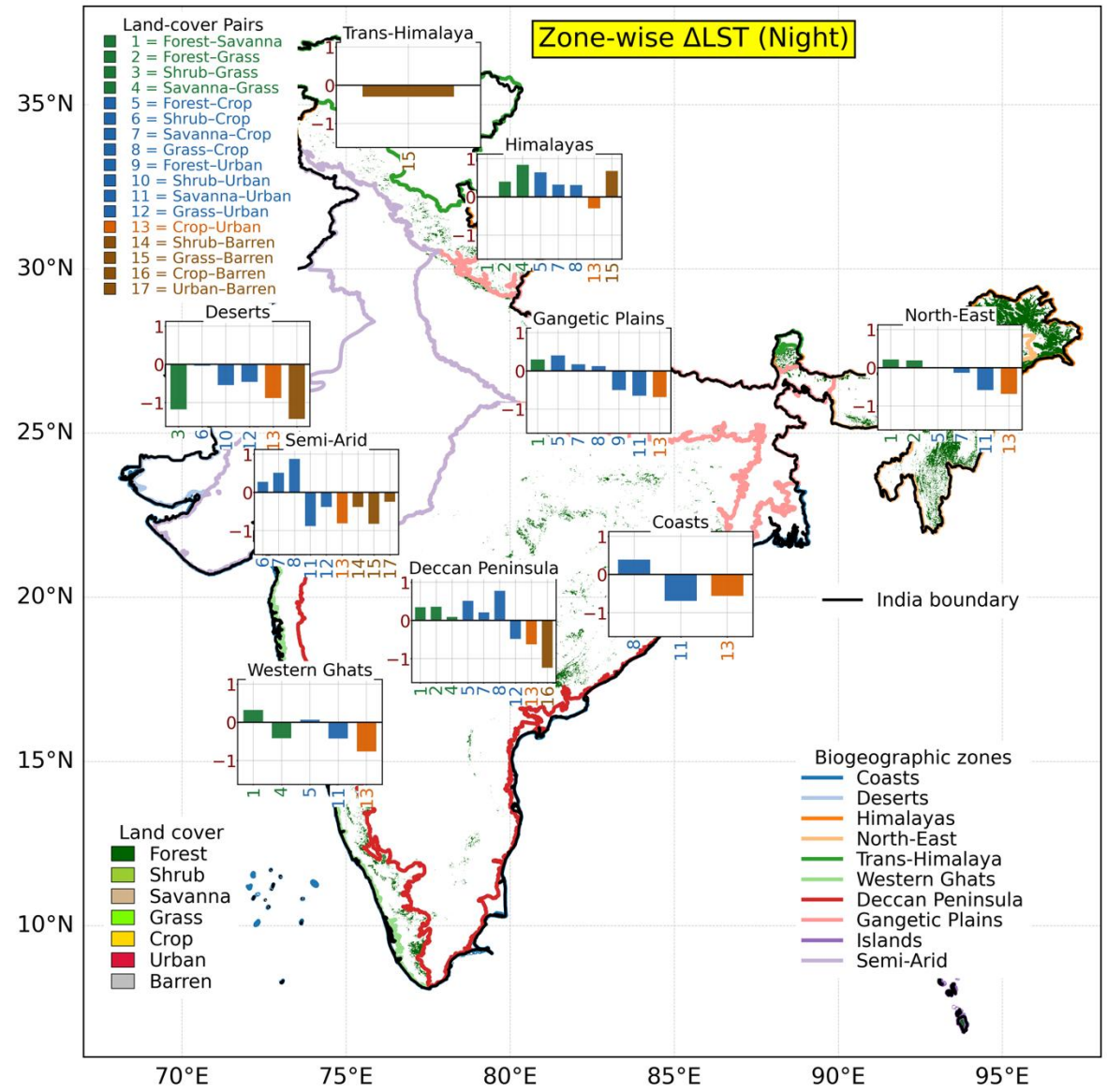
cooler than cropland in every represented zone, with daytime ΔLST from -2.03 °C in the Gangetic Plains to -3.59 °C in the Western Ghats, whereas the sign of grassland-cropland varied from -0.90 °C in the Himalayas to +1.11 °C in the semi-arid zone. Several regional relationships reversed at night; forest-cropland, for example, changed from -2.75 to +0.65 °C in the Himalayas and from -2.50 to +0.52 °C in the Deccan Peninsula.

The degree of geographical dependence itself differed among contrasts. Cropland was cooler than urban land by day in every zone where the pair occurred (-0.38 to -1.15 °C), and this direction persisted at night (-0.30 to -0.88 °C), making cropland-urban the most geographically and diurnally consistent contrast. Barren-related contrasts, in contrast, were highly variable: grassland was warmer than barren land in the Himalayas (+2.63 °C) but cooler in the Trans-Himalaya (-0.74 °C), and cropland-barren in the Deccan Peninsula reversed from +1.85 °C by day to -1.24 °C at night. Biogeographic setting therefore modified not only the magnitude but also the sign and diurnal expression of several contrasts, whereas others retained a consistent direction across regions.

a


Zone-wise ΔLST (Day)
Land-cover Pairs
1 = Forest–Savanna
2 = Forest–Grass
3 = Shrub–Grass
4 = Savanna–Grass
5 = Forest–Crop
6 = Shrub–Crop
7 = Savanna–Crop
8 = Grass–Crop
9 = Forest–Urban
10 = Shrub–Urban
11 = Savanna–Urban
12 = Grass–Urban
13 = Crop–Urban
14 = Shrub–Barren
15 = Grass–Barren
16 = Crop–Barren
17 = Urban–Barren
Trans-Himalaya
Himalayas
Deserts
Gangetic Plains
North-East
Semi-Arid
Coasts
Deccan Peninsula
Western Ghats
India boundary
Biogeographic zones
Coasts
Deserts
Himalayas
North-East
Trans-Himalaya
Western Ghats
Deccan Peninsula
Gangetic Plains
Islands
Semi-Arid
Land cover
Forest
Shrub
Savanna
Grass
Crop
Urban
Barren
35°N
30°N
25°N
20°N
15°N
10°N
70°E
75°E
80°E
85°E
90°E
95°E

b


Zone-wise ΔLST (Night)
Land-cover Pairs
1 = Forest–Savanna
2 = Forest–Grass
3 = Shrub–Grass
4 = Savanna–Grass
5 = Forest–Crop
6 = Shrub–Crop
7 = Savanna–Crop
8 = Grass–Crop
9 = Forest–Urban
10 = Shrub–Urban
11 = Savanna–Urban
12 = Grass–Urban
13 = Crop–Urban
14 = Shrub–Barren
15 = Grass–Barren
16 = Crop–Barren
17 = Urban–Barren
Trans-Himalaya
Himalayas
Deserts
Gangetic Plains
North-East
Semi-Arid
Coasts
Deccan Peninsula
Western Ghats
India boundary
Biogeographic zones
Coasts
Deserts
Himalayas
North-East
Trans-Himalaya
Western Ghats
Deccan Peninsula
Gangetic Plains
Islands
Semi-Arid
Land cover
Forest
Shrub
Savanna
Grass
Crop
Urban
Barren
35°N
30°N
25°N
20°N
15°N
10°N
70°E
75°E
80°E
85°E
90°E
95°E

c-d

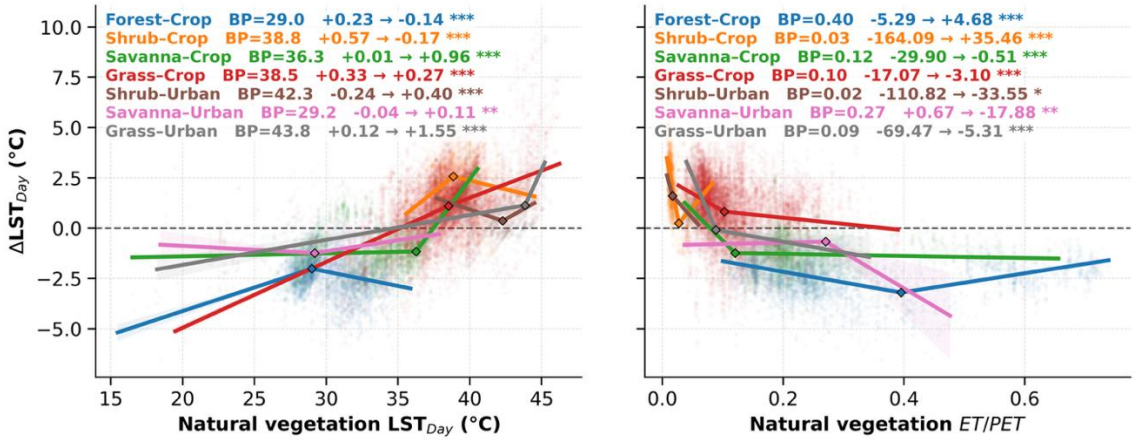

Forest-Crop BP=29.0 +0.23 → -0.14 ***
Shrub-Crop BP=38.8 +0.57 → -0.17 ***
Savanna-Crop BP=36.3 +0.01 → +0.96 ***
Grass-Crop BP=38.5 +0.33 → +0.27 ***
Shrub-Urban BP=42.3 -0.24 → +0.40 ***
Savanna-Urban BP=29.2 -0.04 → +0.11 **
Grass-Urban BP=43.8 +0.12 → +1.55 ***
ΔLST$_{Day}$ (°C)
Natural vegetation LST$_{Day}$ (°C)
Forest-Crop BP=0.40 -5.29 → +4.68 ***
Shrub-Crop BP=0.03 -164.09 → +35.46 ***
Savanna-Crop BP=0.12 -29.90 → -0.51 ***
Grass-Crop BP=0.10 -17.07 → -3.10 ***
Shrub-Urban BP=0.02 -110.82 → -33.55 *
Savanna-Urban BP=0.27 +0.67 → -17.88 **
Grass-Urban BP=0.09 -69.47 → -5.31 ***
Natural vegetation ET/PET

Figure 2. Biogeographic and environmental modulation of land-cover thermal contrasts. (a) Mean daytime ΔLST for each resolved land-cover pair within each biogeographic zone (inset bar charts; pair numbers as in the legend). **(b)** As in a, for nighttime ΔLST. **(c)** Daytime ΔLST of natural vegetation-cropland and natural vegetation-urban pairs as a function of the daytime LST of the natural-vegetation member. **(d)** As in c, as a function of the evapotranspiration-to-potential-evapotranspiration ratio (ET/PET) of the natural-vegetation member. In c,d, lines show fitted continuous one-breakpoint hinge regressions (Eq. 1), diamonds mark the estimated breakpoints (BP), and annotations give the slopes below → above the breakpoint (°C per °C in c; °C per unit ET/PET in d); asterisks denote the significance of the change in slope (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). Shading indicates bootstrap 95% confidence bands. Negative ΔLST indicates that the natural-vegetation member is cooler.

Continuous environmental gradients revealed that daytime natural-anthropogenic contrasts were frequently nonlinear, but that the form and strength of nonlinearity differed among pairs (Fig. 2c). Along the background-temperature gradient, strong hinge responses were supported for forest-cropland, shrubland-cropland, savanna-cropland, shrubland-urban and grassland-urban, with breakpoints spanning 29.0-43.8 °C. Below 29.0 °C, the forest-cropland contrast weakened as background temperature increased (slope +0.23 °C $°C^{-1}$), whereas above this breakpoint it strengthened again (-0.14 °C $°C^{-1}$), indicating that the forest cooling advantage widens with further warming. Shrubland-cropland likewise switched from a positive to a negative slope around 38.8 °C. In contrast, savanna-cropland increased sharply above 36.3 °C (+0.96 °C $°C^{-1}$), so that in the hottest settings savanna lost its cooling advantage over cropland, and shrubland-urban switched from a negative to a positive slope around 42.3 °C. Grassland-urban also showed strong support for a hinge near 43.8 °C, although its breakpoint and post-break slope were comparatively uncertain. Savanna-urban showed only partial support, and grassland-cropland was better described by a quadratic response despite a significant slope change. Rising background temperature therefore neither uniformly strengthened nor uniformly weakened land-cover contrasts; the direction of thermal sensitivity was pair-specific.

Evaporative state, represented by natural-vegetation evapotranspiration-to-potential-evapotranspiration ratio (ET/PET), showed similarly strong but pair-specific

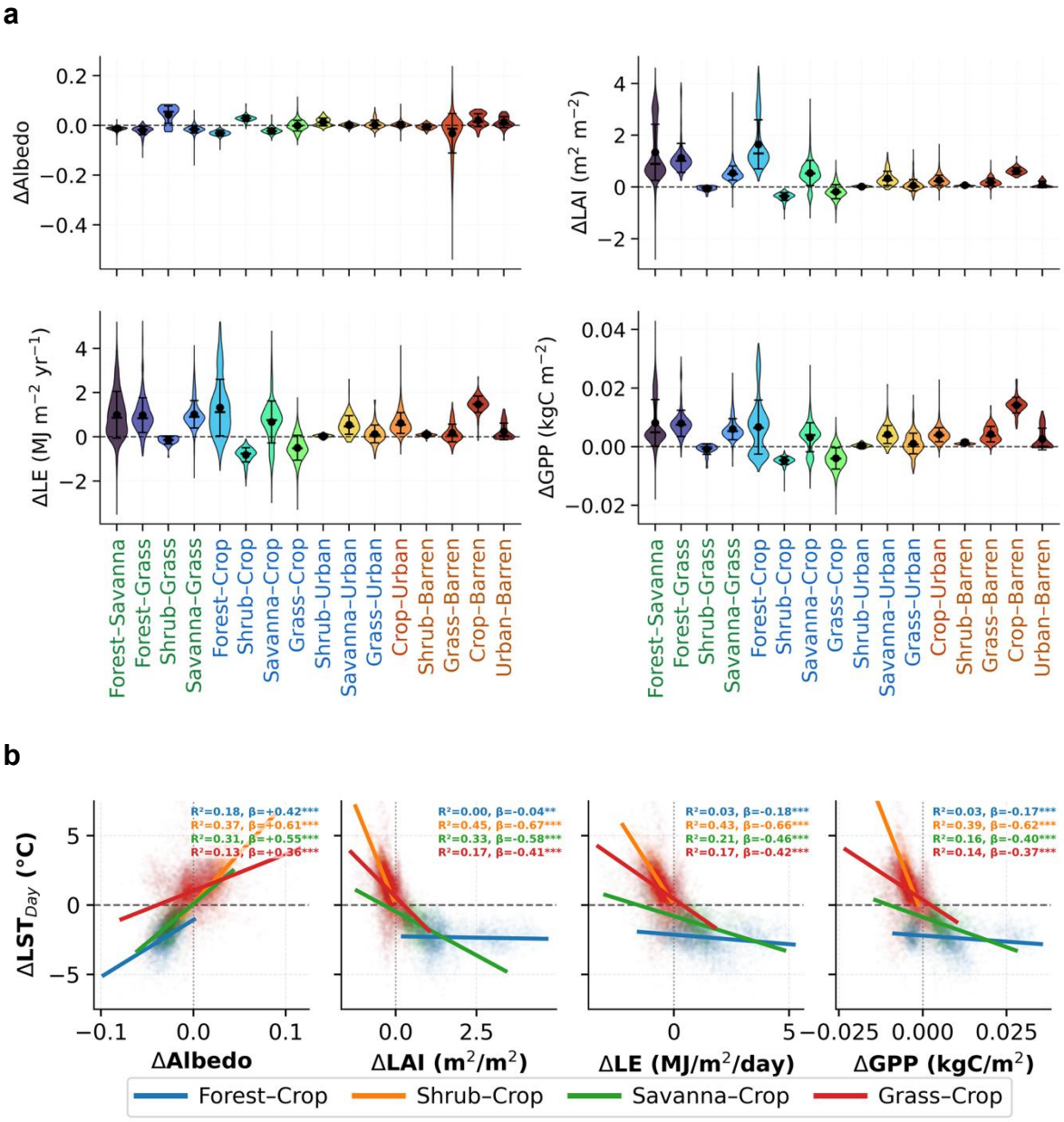
nonlinearity (Fig. 2d). Strong hinge support was found for forest-cropland, savanna-cropland, grassland-cropland and grassland-urban, with breakpoints at ET/PET = 0.395, 0.121, 0.103 and 0.089, respectively. Forest-cropland showed the clearest reversal: the contrast deepened as ET/PET increased up to ≈0.40 and weakened beyond it, so that the forest cooling advantage over cropland peaked at intermediate evaporative states. For savanna-, grassland- and shrubland-cropland and for grassland-urban, steeply negative slopes at the driest conditions (ET/PET $\lesssim$ 0.1) indicate that water-stressed non-forest vegetation was substantially warmer than adjacent cropland or urban land, a penalty that diminished rapidly once ET/PET exceeded the breakpoint. Shrubland-cropland also showed pronounced nonlinearity, but a quadratic model was preferred to the hinge, whereas shrubland-urban and savanna-urban showed only partial hinge support and larger breakpoint uncertainty. Forest-urban was excluded because only three matched ET/PET observations were available. Collectively, these results show that thermal contrasts vary nonlinearly with both background temperature and evaporative state, but with no single breakpoint shared across land-cover pairs.

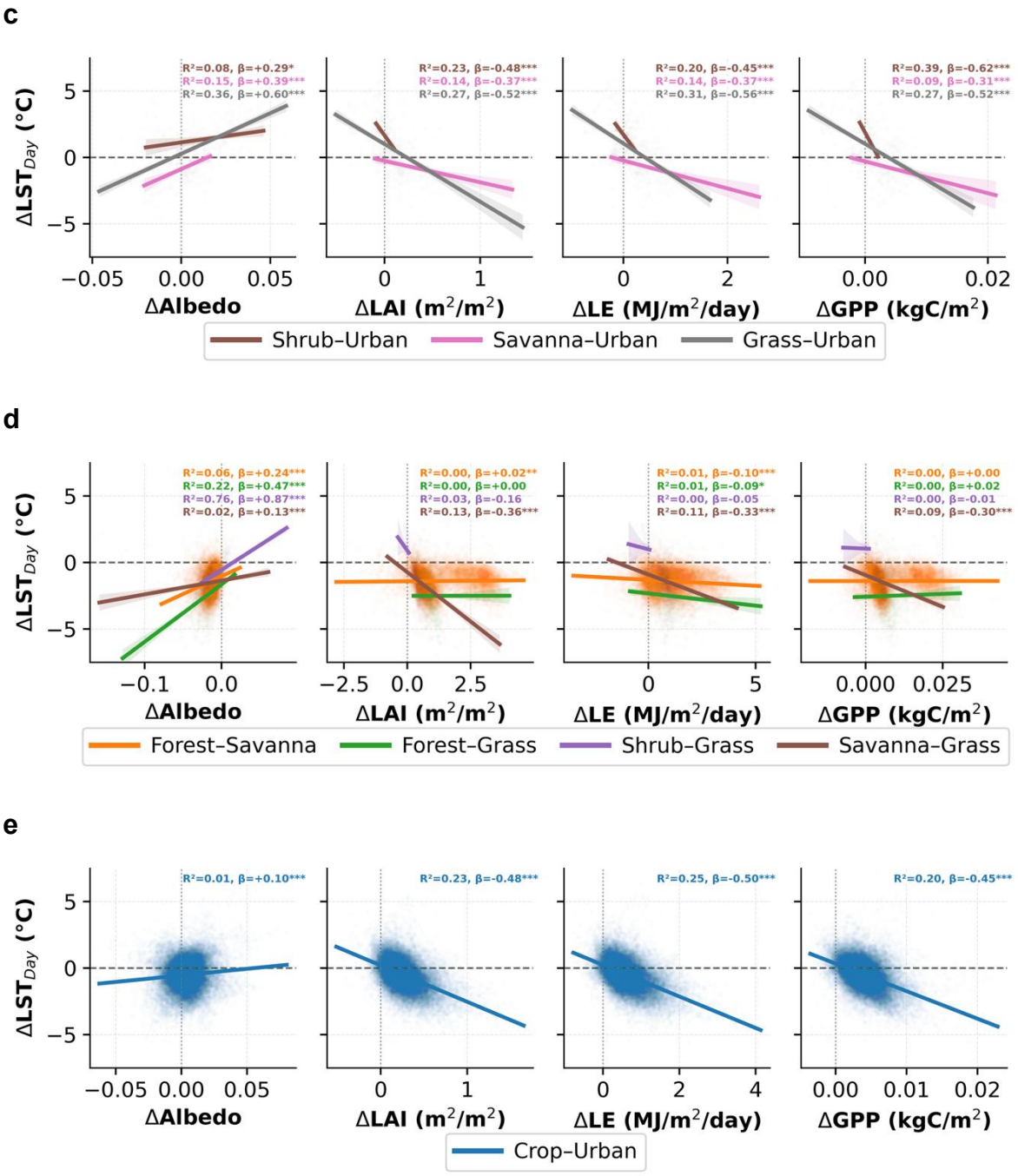
Support for nonlinear environmental dependence was robust for most, though not all, relationships. Of the 14 temperature and ET/PET relationships, nine showed strong support for a hinge response, AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) agreed on the preferred functional form in 11, and bootstrap intervals supported a non-zero slope change in 10 (Supplementary Tables S5-S10; Supplementary Fig. S4).

Biophysical correlates of daytime thermal contrasts

Biophysical differences between paired land covers were strongly but highly pair-specifically associated with daytime ΔLST (Fig. 3). In all natural vegetation-cropland pairs, higher albedo differences (ΔAlbedo) were associated with higher ΔLST, whereas larger differences in leaf area index (ΔLAI), latent heat flux (ΔLE) and gross primary productivity (ΔGPP) were predominantly associated with lower ΔLST. The relative strength of these associations, however, differed markedly among vegetation types. Forest-cropland was associated mainly with albedo ($r = +0.42$, $R^2 = 0.18$), whereas shrubland-cropland was much more strongly associated with LAI ($r = -0.67$, $R^2 = 0.45$), LE ($r = -0.66$, $R^2 = 0.43$) and GPP ($r = -0.62$, $R^2 = 0.39$), alongside an opposing positive association with albedo ($r = +0.61$). Savanna-cropland was

associated with both canopy structure and albedo, and grassland-cropland showed comparable associations with LE and LAI. Even among vegetation-cropland contrasts, therefore, similar thermal differences arose from different combinations of radiative and vegetation-mediated properties.




c
R²=0.08, β=+0.29*
R²=0.15, β=+0.39***
R²=0.36, β=+0.60***
R²=0.23, β=-0.48***
R²=0.14, β=-0.37***
R²=0.27, β=-0.52***
R²=0.20, β=-0.45***
R²=0.14, β=-0.37***
R²=0.31, β=-0.56***
R²=0.39, β=-0.62***
R²=0.09, β=-0.31***
R²=0.27, β=-0.52***
ΔLST_Day (°C)
ΔAlbedo
ΔLAI (m²/m²)
ΔLE (MJ/m²/day)
ΔGPP (kgC/m²)
Shrub–Urban
Savanna–Urban
Grass–Urban
d
R²=0.06, β=+0.24***
R²=0.22, β=+0.47***
R²=0.76, β=+0.87***
R²=0.02, β=+0.13***
R²=0.00, β=+0.02**
R²=0.00, β=+0.00
R²=0.03, β=-0.16
R²=0.13, β=-0.36***
R²=0.01, β=-0.10***
R²=0.01, β=-0.09*
R²=0.00, β=-0.05
R²=0.11, β=-0.33***
R²=0.00, β=+0.00
R²=0.00, β=+0.02
R²=0.00, β=-0.01
R²=0.09, β=-0.30***
ΔLST_Day (°C)
ΔAlbedo
ΔLAI (m²/m²)
ΔLE (MJ/m²/day)
ΔGPP (kgC/m²)
Forest–Savanna
Forest–Grass
Shrub–Grass
Savanna–Grass
e
R²=0.01, β=+0.10***
R²=0.23, β=-0.48***
R²=0.25, β=-0.50***
R²=0.20, β=-0.45***
ΔLST_Day (°C)
ΔAlbedo
ΔLAI (m²/m²)
ΔLE (MJ/m²/day)
ΔGPP (kgC/m²)
Crop–Urban

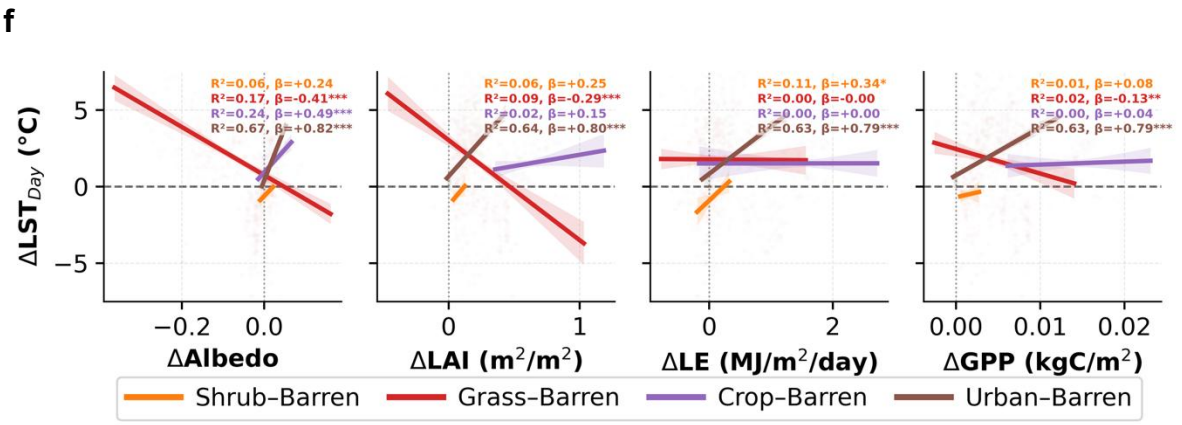


Figure 3. Pair-specific biophysical correlates of daytime land-cover thermal contrasts. (a) Distributions of pairwise differences in shortwave black-sky albedo (ΔAlbedo), leaf area index (ΔLAI), latent heat flux (ΔLE) and gross primary productivity (ΔGPP) for each land-cover pair. **(b-f)** Bivariate relationships between daytime ΔLST and each biophysical difference for **(b)** natural vegetation-cropland, **(c)** natural vegetation-urban, **(d)** natural-natural, **(e)** cropland-urban and **(f)** barren-related pairs. Lines show ordinary least-squares regression fits, and annotations report standardized regression coefficients (equivalent to Pearson's r for simple linear regression), R^2 and statistical significance (***, $p < 0.001$). Different pairs reach similar ΔLST through different combinations of radiative, structural, turbulent-energy and productivity differences.

Natural-urban contrasts were similarly heterogeneous. Shrubland-urban was most strongly associated with GPP, followed by LAI and LE, whereas savanna-urban showed more balanced associations with albedo, LAI and LE. Grassland-urban exhibited strong associations with all four predictors: positive with albedo and negative with LE, LAI and GPP. Forest-urban was excluded from the complete-case attribution because only three matched observations had all four predictors. No single biophysical variable thus explains natural-anthropogenic contrasts: radiative differences were more prominent for some pairs, and canopy structure, turbulent energy exchange and productivity for others.

The same pair specificity extended to the remaining contrasts (Fig. 3d-f). Among natural-natural pairs, albedo showed the strongest bivariate association for forest-savanna, forest-grassland and especially shrubland-grassland ($r = +0.87$, $R^2 = 0.76$), whereas savanna-grassland was more strongly associated with LAI. Cropland-urban

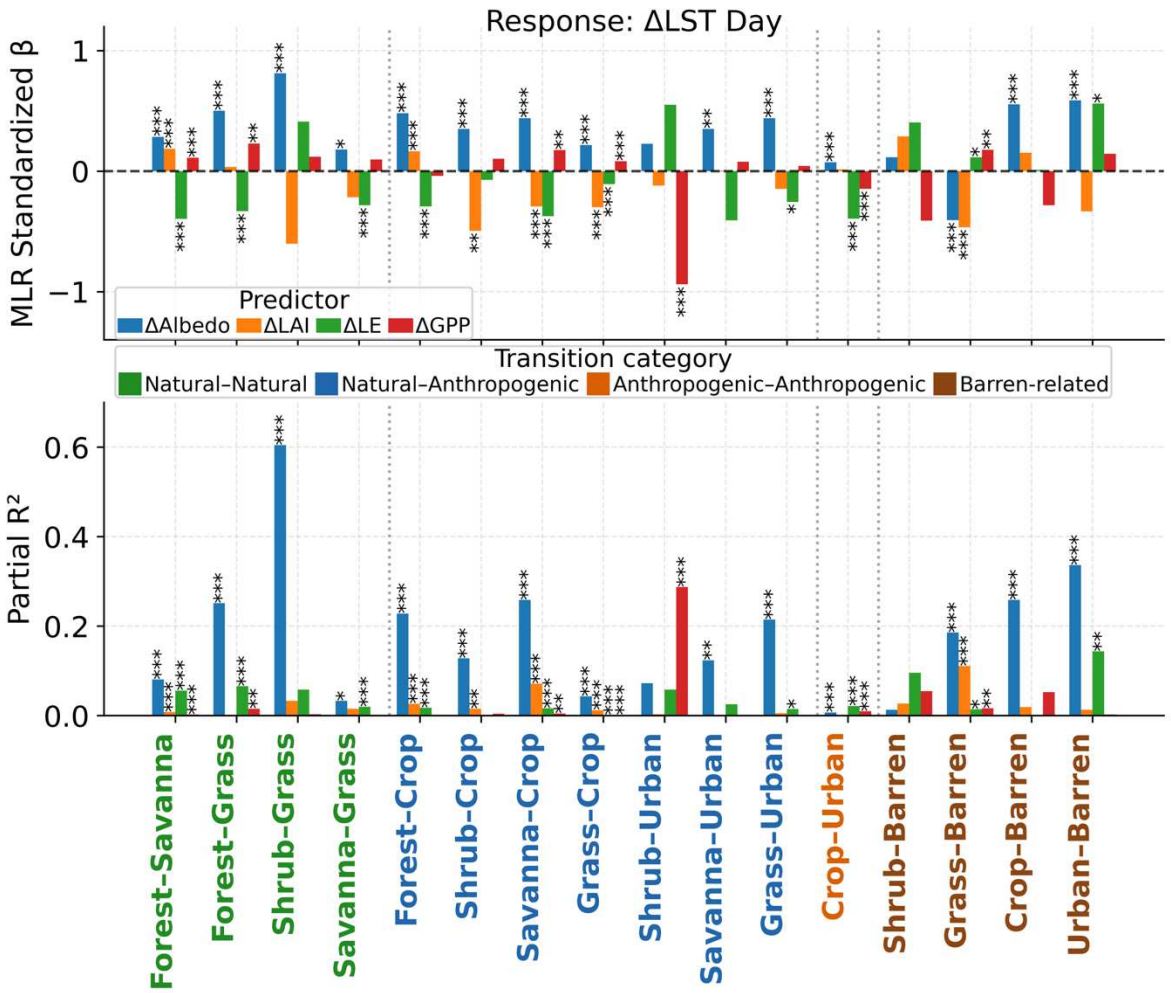
was most strongly associated with LE (r = -0.50, R^2 = 0.25), with similar negative associations for LAI and GPP. Barren-related contrasts were heterogeneous, although albedo was strongly associated with grassland-barren, cropland-barren and urban-barren. The distributions of biophysical differences, together with their annual and seasonal structure, are shown in Supplementary Figs. S5-S6.

Multivariate and machine-learning attribution

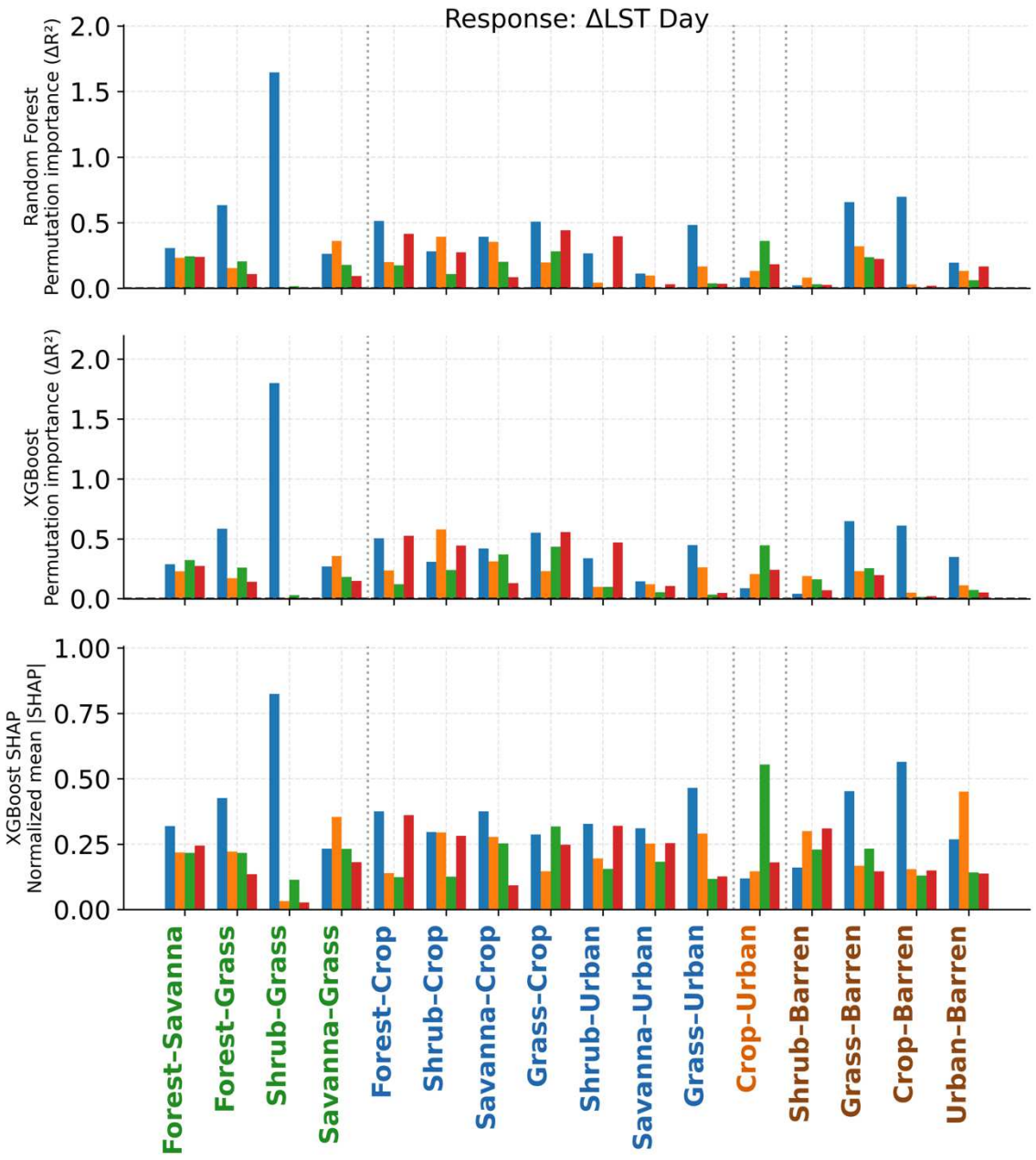
Multivariate and machine-learning analyses confirmed that the biophysical variables associated with daytime ΔLST differed substantially among pairs (Fig. 4). Standardized multiple linear regression (MLR) coefficients and partial R^2 revealed strong pair-specific differences in the conditional associations of albedo, LAI, LE and GPP with ΔLST (Fig. 4a). This heterogeneity extended to predictive skill: Random forest (RF) regression and extreme gradient boosting (XGBoost) achieved high out-of-sample skill for shrubland-grassland (R^2_{RF} = 0.83; R^2_{XGB} = 0.81), shrubland-cropland (0.62; 0.61), savanna-cropland (0.62; 0.62) and grassland-urban (0.58; 0.56), but much lower skill for other pairs, notably savanna-urban (0.02; 0.05; Supplementary Tables S11-S12, Fig. S8). The four biophysical predictors thus captured markedly different fractions of ΔLST variability among contrasts (Supplementary Tables S13-S16).

Despite this heterogeneity, albedo emerged as the most pervasive daytime predictor across all approaches. Held-out permutation importance from both RF and XGBoost most frequently ranked albedo first (Supplementary Tables S17-S19), and SHapley Additive exPlanations (SHAP) analysis showed that albedo strongly influenced fitted predictions for natural-natural, natural-anthropogenic and barren-related contrasts alike (Fig. 4b,c; Supplementary Tables S20-S22). For most vegetated pairs, high ΔAlbedo values were associated with positive SHAP contributions, i.e. the brighter member of a pair tended to be the warmer one (Fig. 4c). The dominance of albedo was nonetheless not universal: canopy structure, productivity and latent heat exchange became more important for particular pairs. Sensitivity to alternative predictor formulations and to multicollinearity is documented in Supplementary Tables S23-S26.

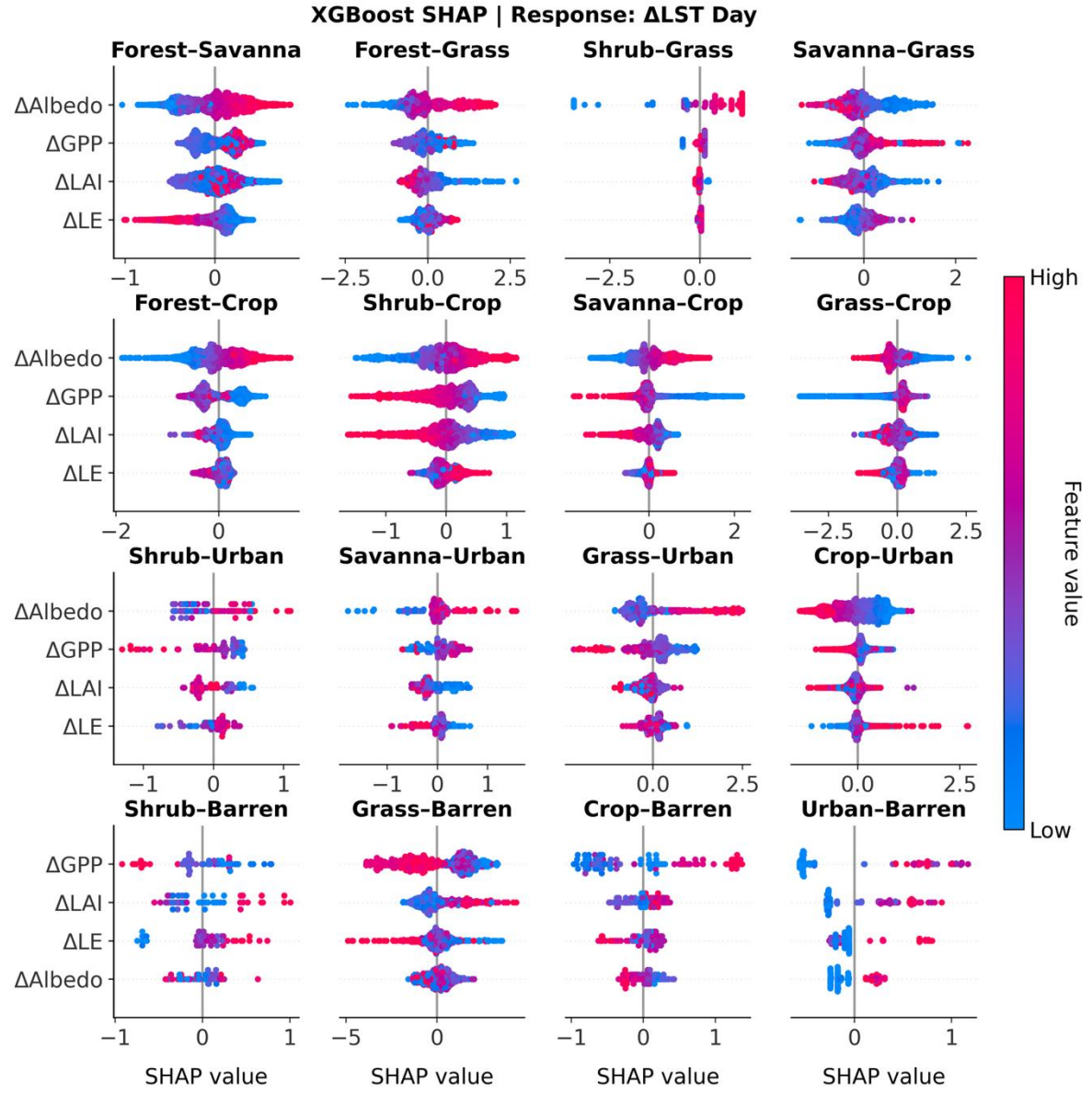
a

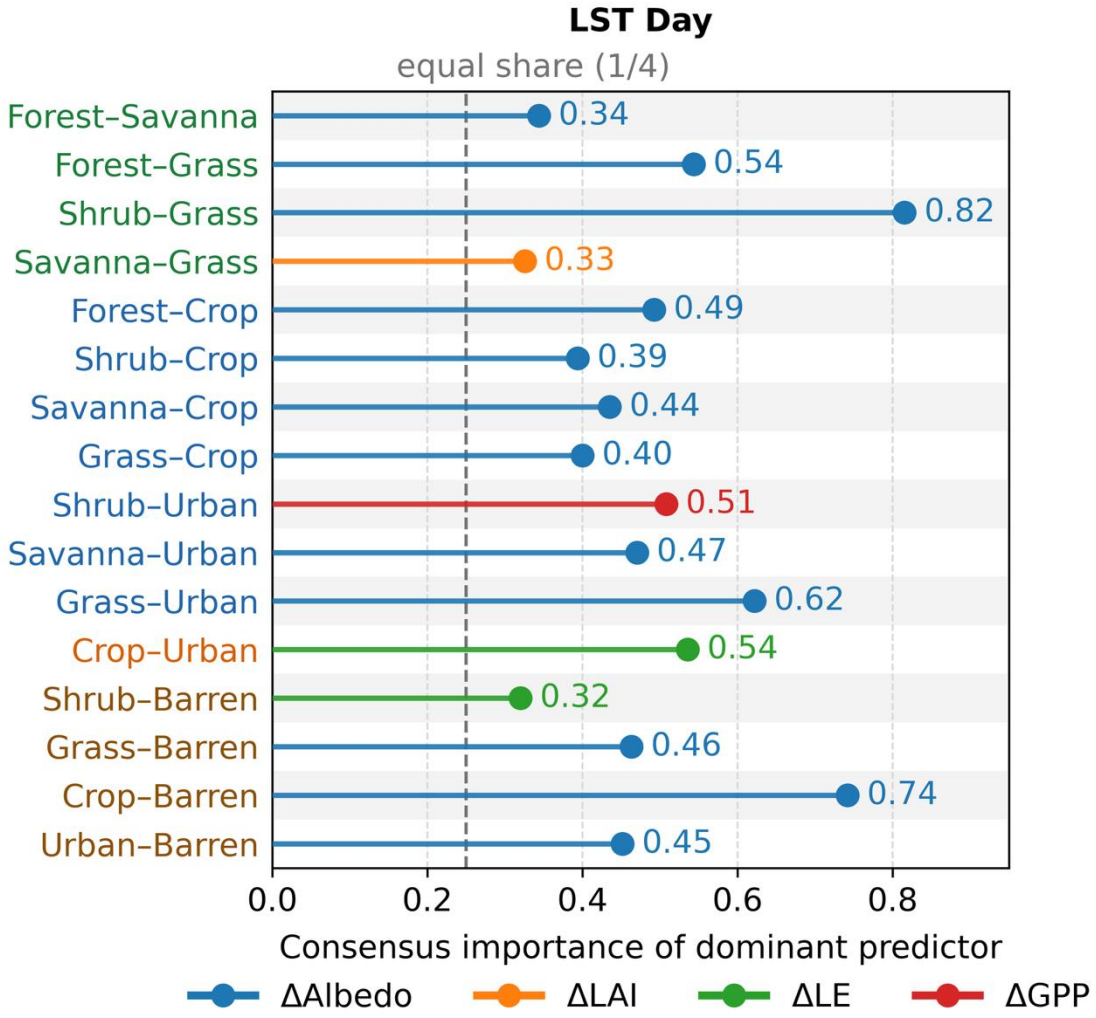

Response: ΔLST Day
MLR Standardized β
Predictor
ΔAlbedo
ΔLAI
ΔLE
ΔGPP
Transition category
Natural–Natural
Natural–Anthropogenic
Anthropogenic–Anthropogenic
Barren-related
Partial R²
Forest-Savanna
Forest-Grass
Shrub-Grass
Savanna-Grass
Forest-Crop
Shrub-Crop
Savanna-Crop
Grass-Crop
Shrub-Urban
Savanna-Urban
Grass-Urban
Crop-Urban
Shrub-Barren
Grass-Barren
Crop-Barren
Urban-Barren

b


Response: ΔLST Day
Random Forest
Permutation importance (ΔR²)
XGBoost
Permutation importance (ΔR²)
XGBoost SHAP
Normalized mean |SHAP|
Forest–Savanna
Forest–Grass
Shrub–Grass
Savanna–Grass
Forest–Crop
Shrub–Crop
Savanna–Crop
Grass–Crop
Shrub–Urban
Savanna–Urban
Grass–Urban
Crop–Urban
Shrub–Barren
Grass–Barren
Crop–Barren
Urban–Barren

c


XGBoost SHAP | Response: ΔLST Day
Forest–Savanna
Forest–Grass
Shrub–Grass
Savanna–Grass
Forest–Crop
Shrub–Crop
Savanna–Crop
Grass–Crop
Shrub–Urban
Savanna–Urban
Grass–Urban
Crop–Urban
Shrub–Barren
Grass–Barren
Crop–Barren
Urban–Barren
ΔAlbedo
ΔGPP
ΔLAI
ΔLE
SHAP value
High
Low
Feature value



d

Figure 4. Convergent multivariate and machine-learning attribution of daytime land-cover thermal contrasts. (a) Standardized multiple linear regression (MLR) coefficients (top) and partial R^2 (bottom) for ΔAlbedo, ΔLAI, ΔLE and ΔGPP; asterisks denote significance based on HC3 standard errors (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). **(b)** Held-out permutation importance from random forest (RF) regression (top) and extreme gradient boosting (XGBoost) (middle) and normalized mean absolute SHapley Additive exPlanations (SHAP) values from XGBoost (bottom). **(c)** XGBoost SHAP distributions for each pair; each point is one observation, its horizontal position the predictor's contribution to modelled ΔLST and its colour the predictor value (blue, low; red, high). **(d)** Dominant consensus predictor and its consensus importance (Eq. 5) for each pair; the dashed line marks an equal share among the four predictors (0.25). Consensus importance combines standardized MLR coefficients, partial R^2, RF and XGBoost permutation importance

and mean absolute SHAP. Albedo is the dominant predictor for 12 of 16 daytime contrasts, whereas LAI, GPP or LE dominate the remaining four. Forest-urban is omitted because of insufficient complete-case data.

Cross-method synthesis made this pair specificity explicit (Fig. 4d; Supplementary Table S27). Albedo was the dominant consensus predictor in 12 of the 16 evaluated daytime contrasts, with particularly strong dominance for shrubland-grassland (0.82), cropland-barren (0.74), grassland-urban (0.62) and forest-grassland (0.54), i.e. two to three times the equal-share expectation of 0.25. The four exceptions identified distinct non-radiative pathways: LAI dominated savanna-grassland (0.33), GPP dominated shrubland-urban (0.51) and LE dominated cropland-urban (0.54) and shrubland-barren (0.32). For savanna-grassland and shrubland-barren, consensus importance was only modestly above the equal share, indicating co-dominance among several predictors rather than control by a single variable. Nighttime attribution was similarly pair-specific, with the relative roles of albedo, LAI, LE and GPP varying among contrasts (Supplementary Fig. S7; Supplementary Tables S11-S12, S14, S16, S17-S22 and S27).

The principal attribution patterns were robust to spatial-block cross-validation. Although predictive skill generally declined as the spatial separation between training and test data increased, dominant-predictor rankings remained stable for the major, well-represented contrasts and across alternative block-grid origins (Supplementary Tables S28-S32), indicating that the attribution is not merely a product of short-range spatial dependence.

Discussion

Our nationwide analysis reveals that land-cover thermal behaviour cannot be reduced to a simple natural-versus-anthropogenic distinction. Forests consistently occupied the coolest end of the daytime thermal spectrum, whereas shrublands and grasslands were frequently warmer than cropland, and barren land showed no fixed thermal position. Forest cooling is consistent with enhanced transpiration and turbulent heat exchange associated with deep-rooted, aerodynamically rough canopies [6,12,13]. Crucially, however, this cooling did not extend to natural vegetation generally. Previous studies have reported contrasting cropland thermal responses depending on the surrounding vegetation and hydroclimatic setting [14,27]. Our results help reconcile

these differences: aggregating forests with savanna, shrubland and grassland nearly eliminated the natural-cropland contrast (+0.15 °C; cooler in 45.6% of pairs), whereas separating forests from other natural vegetation revealed opposing signals of -2.32 °C and +0.73 °C, respectively. These differences likely reflect contrasting canopy structures, surface reflectance, rooting characteristics and water-use strategies. Irrigation may further enhance cropland cooling relative to non-forest vegetation in India [28], although this mechanism requires explicit evaluation. Thus, aggregating ecologically distinct vegetation types can obscure thermal contrasts of opposite sign.

The daytime hierarchy was substantially reorganized at night and varied across seasons. Eleven of 17 contrasts reversed sign, with forests becoming warmer than cropland, grassland and savanna after sunset. This reversal is consistent with a shift from daytime radiative and evaporative controls towards nocturnal heat storage, longwave exchange and canopy-atmosphere coupling. Forest canopies may reduce radiative heat loss and promote turbulent mixing, whereas open vegetation cools more rapidly, contributing to a 1.7-2.9 °C reduction in the diurnal LST range of forests relative to co-occurring land covers consistent with previously reported forest thermal buffering [25,26]. Urban surfaces, conversely, remained warmer than all vegetated classes at night, consistent with heat release from built materials, while forests remained cooler than urban land throughout the observed day-night cycle. Forest cooling also weakened towards the monsoon, when increased moisture availability may reduce differences in evaporative cooling between forests and open surfaces. These patterns align with evidence that radiative and non-radiative contributions vary seasonally [6,9] and that cropland-grassland contrasts depend on diurnal and hydroclimatic conditions [15]. Together, they demonstrate that temporal variability is itself pair-specific rather than a uniform response across land covers.

Environmental context further reshaped these relationships. Thermal contrasts varied markedly among India's humid, arid, monsoonal and mountainous regions, with several vegetation- and barren-related pairs reversing sign across biogeographic zones, whereas cropland-urban contrasts retained a consistent direction. Such spatial heterogeneity accords with established influences of precipitation, water availability, snow cover and atmospheric demand on vegetation-temperature relationships [6,12,14,15,27,29]. Continuous-gradient analyses revealed nonlinear, pair-specific responses, with strongly supported breakpoints at approximately 29-44 °C and

ET/PET ratios of 0.09-0.40. Forest cooling relative to cropland was strongest near intermediate evaporative conditions (ET/PET ≈ 0.4) and increased with temperature above approximately 29 °C. Conversely, savanna, grassland and shrubland were substantially warmer than neighbouring cropland and urban surfaces under very dry conditions (ET/PET $\lesssim$ 0.1). These associations are consistent with differences in vegetation water regulation and evaporative limitation, although the identified breakpoints represent statistical changes in response rather than universal ecological thresholds [2,6]. Emerging evidence that atmospheric dryness and rising CO_2 can modify vegetation cooling further suggests that present-day contrasts may shift under future climate conditions [16,17].

Biophysical attribution revealed that similar thermal contrasts can be associated with distinct surface properties. Albedo emerged as the dominant daytime predictor in 12 of 16 contrasts across statistical and machine-learning approaches, generally showing a positive association with ΔLST. Although higher albedo directly reduces absorbed shortwave radiation, this positive relationship suggests that surface brightness may also indicate sparse vegetation, exposed soil or moisture limitation, where reduced evaporative cooling can outweigh the radiative advantage. This interpretation is consistent with the competing radiative and non-radiative controls on land-surface temperature, particularly in warm regions [2,6,7]. However, predictor importance varied among pairs: LAI dominated savanna-grassland, GPP dominated shrubland-urban, and latent heat exchange dominated cropland-urban and shrubland-barren contrasts. These differences agree with evidence linking vegetation cooling to evapotranspiration, canopy development and turbulent energy exchange [8,15]. Although statistical associations cannot isolate causal contributions, the convergence across analytical approaches indicates that land-cover thermal contrasts reflect different combinations of radiative, structural and evaporative properties. Land-cover thermal behaviour is therefore best understood as a pair-specific, temporally dynamic and environmentally conditioned relationship rather than a fixed warming or cooling attribute of individual land-cover classes.

Implications, limitations and future directions

Our findings establish that land-cover thermal behaviour is not an intrinsic property of individual classes but depends on the surfaces compared, their environmental setting, temporal context and underlying biophysical processes. This has important

implications for land-surface modelling and climate adaptation. Models relying on broad vegetation classes or fixed thermal responses may reproduce observed temperature contrasts without adequately representing the mechanisms responsible, potentially limiting their reliability under changing climatic conditions. Our pair-specific, seasonally and environmentally resolved contrasts provide observational benchmarks for evaluating these representations. For nature-based adaptation, the findings challenge the assumption that all natural vegetation provides comparable cooling. Forests exhibited the strongest daytime cooling relative to urban land (-2.57 °C) and remained cooler at night, whereas grasslands and shrublands were warmer than urban surfaces during daytime. These patterns highlight the potential importance of tree-based interventions while emphasizing that cooling benefits depend on vegetation type and local hydroclimatic conditions.

Several limitations warrant consideration. Land-cover classification uncertainty and mixed-pixel effects were reduced through temporal-stability and spatial-homogeneity criteria but cannot be eliminated. Moreover, the matched moving-window approach constrains local environmental differences without fully accounting for variations in soil moisture, topography, irrigation, phenology and land management. Spatial thinning and spatial-block cross-validation support the robustness of the principal patterns, although spatial dependence reduces the effective number of independent observations. Importantly, the estimated ΔLST values represent observational contrasts between persistent land covers rather than causal responses to land-cover conversion. Satellite-derived LST also differs from near-surface air temperature [30], and Aqua's daytime and nighttime observations do not resolve the complete diurnal cycle. Furthermore, statistical and machine-learning attribution identifies associations rather than causal mechanisms, particularly because radiative, structural and evaporative processes interact with atmospheric and subsurface feedbacks not explicitly resolved here [31,32].

Future research should combine land-cover transition observations, finer representations of vegetation and land management, and controlled land-surface or regional climate-model experiments to establish causal responses and atmospheric feedbacks. Explicit consideration of soil moisture, irrigation and vegetation phenology, together with interannual climate variability, would help determine whether the observed relationships persist during extreme conditions. Testing their stability under

increasing atmospheric dryness and CO_2 would further clarify how changing radiative and evaporative processes reshape land-cover thermal contrasts. Together, these advances could support more physically realistic land-surface modelling and context-specific climate adaptation strategies.

Conclusion

Land-cover thermal behaviour across India cannot be reduced to a simple distinction between natural and human-modified surfaces. By resolving all locally co-occurring pairs among seven major land covers, we show that thermal contrasts are pair-specific and environmentally conditioned, varying with the land covers compared, the time of day, the season, the biogeographic setting and the background thermal and evaporative state. The pronounced daytime cooling of forests does not extend to other natural vegetation, which is on average warmer than cropland; most contrasts reverse between day and night; and contrasts respond nonlinearly, with pair-specific breakpoints, to temperature and evaporative state. The biophysical basis of these contrasts is equally heterogeneous: albedo is the most pervasive daytime correlate, acting largely as a marker of sparse, dry surfaces, but canopy structure, productivity and latent heat exchange dominate specific pairs, so similar temperature contrasts arise through different pathways. Land-cover thermal behaviour is therefore better represented as a relational property of the land surface than as a fixed warming or cooling attribute of individual classes. Accounting for the specific land-cover pair, environmental state, temporal context and biophysical pathway provides a more physically grounded basis for representing land-cover heterogeneity in land-surface models and for designing land-based climate-adaptation strategies across environmentally diverse regions.

Methods

Datasets

We used MODIS Collection 6.1 products to characterize land-cover thermal contrasts and associated biophysical properties across India. Annual land cover was obtained from MCD12Q1 at 500-m resolution for 2001-2022. Daytime and nighttime LST was derived from the Aqua MODIS MYD11A2 8-day product at 1-km resolution for 2003-2022, corresponding to nominal overpass times of approximately 13:30 and 01:30 local solar time. Shortwave black-sky albedo was obtained from MCD43A3; LAI and

fraction of absorbed photosynthetically active radiation (FPAR) from MYD15A2H; and LE, potential latent heat flux (PLE), Evapotranspiration (ET) and Potential Evapotranspiration (PET) from MYD16A2GF. Vegetation productivity was characterized using GPP, net primary productivity (NPP) and net photosynthesis (PsnNet) from the MYD17 product suite. Biophysical variables covered 2003-2022, except GPP, which extended through 2021. Elevation was obtained from the 90-m Shuttle Radar Topography Mission (SRTM) digital elevation model to constrain topographic differences between paired land-cover cells. Product specifications, spatial and temporal resolutions, and data availability are summarized in Supplementary Table S33.

Data preprocessing and land-cover harmonization

MODIS variables were extracted and quality-screened separately within bounding boxes enclosing each biogeographic zone of India (geographical extents in Supplementary Table S34). Product-specific quality-assurance (QA) flags were applied following the corresponding MODIS user guides. For ET, PET, PLE, LE, FPAR and LAI, we retained retrievals with MODLAND_QC = 0 (main algorithm, good quality) or 1 (backup algorithm or other quality) together with cloud-state flags of 0 (clear) or 3 (not defined, assumed clear). GPP, PsnNet and NPP were screened using the corresponding product QA, and values outside the valid data range, including land-cover fill values, were excluded. For daytime and nighttime LST, we retained retrievals with mandatory QA flags of 0 (LST produced, good quality) or 1 (LST produced, other quality) and data-quality flags of 0 (good) or 1 (other quality). For MCD43A3 shortwave black-sky albedo, mandatory QA classes 0, 2, 4 and 6, representing good-quality full bidirectional reflectance distribution function (BRDF) inversions, were retained, whereas magnitude-inversion retrievals and missing values were excluded. Values rejected by quality screening were not reconstructed. Pixel-retention fractions for each variable and biogeographic zone are reported in Supplementary Table S35.

Quality-screened observations were aggregated separately within each biogeographic zone into seasonal climatological means for winter (DJF; December-February), pre-monsoon (MAM; March-May), monsoon (JJAS; June-September) and post-monsoon (ON; October-November), together with an overall climatological mean. The analysis combined LST climatologies for 2003-2022 with GPP climatologies for 2003-2021, reflecting differences in product availability. Aggregation used only valid observations,

so that missing or quality-rejected retrievals did not contribute to the climatologies. All variables other than LST were then conservatively remapped to the 1-km LST grid, providing a common spatial framework for pairwise comparison. After remapping, all fields were clipped to the boundaries of their respective biogeographic zones.

Annual land-cover maps were derived from MCD12Q1 using the International Geosphere-Biosphere Programme (IGBP) classification (class definitions in the MCD12Q1 C6.1 User Guide; https://lpdaac.usgs.gov/documents/1409/MCD12_User_Guide_V61.pdf). The IGBP classes were aggregated into seven major categories: forest (IGBP classes 1-5), shrubland (6-7), savanna (8-9), grassland (10), cropland (12), urban land (13) and barren land (16). To minimize contamination from land-cover change, a cell was classified as temporally stable only if it was assigned to the same aggregated category in every year of the study period. The resulting single stable land-cover map was used for all subsequent comparisons; the analysis therefore quantifies spatial contrasts among persistent land covers and excludes observed land-cover transitions by design.

For harmonization with the 1-km LST grid, each stable category was converted to a binary 500-m mask (1 = target category; 0 = any other category) and conservatively remapped to the LST grid. Only 1-km cells with a remapped value of exactly 1 were retained. This criterion ensured that every LST cell entering the pairwise analysis was fully covered by a single stable land-cover category, thereby minimizing mixed-pixel contamination.

Spatially matched moving-window framework

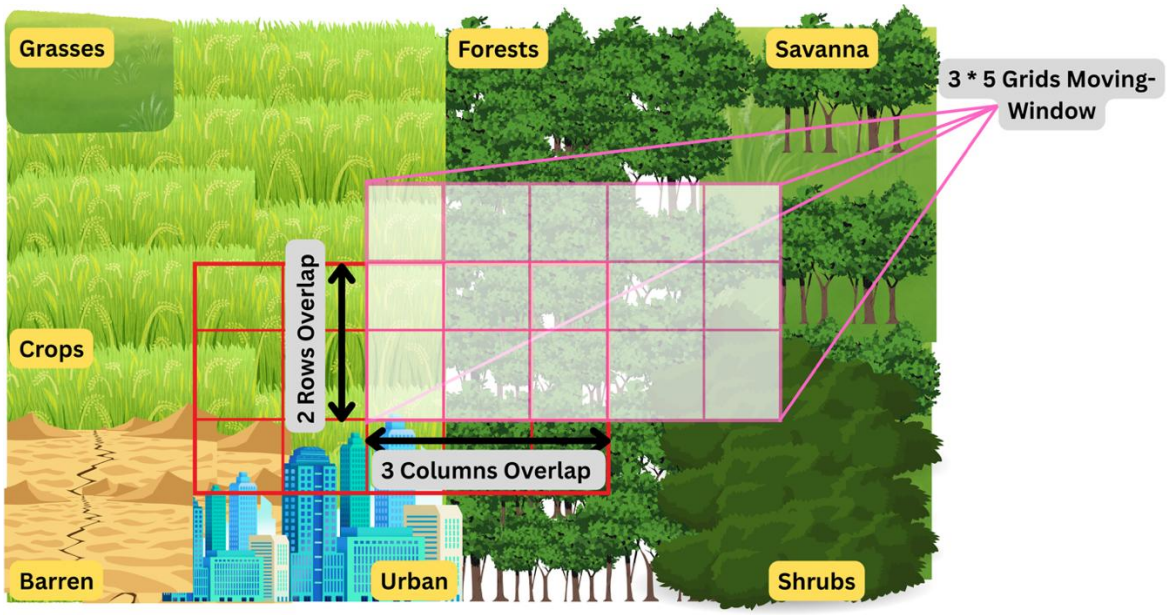
We applied a spatially matched moving-window approach adapted from [6] to quantify land-cover thermal contrasts while minimizing background environmental differences. A 3 × 5-cell window (approximately 3 × 5 km) was selected based on sensitivity tests of alternative window sizes (2 × 4, 3 × 5, 3 × 7 and 4 × 6 km; [26]. Windows were scanned across each biogeographic zone, overlapping by three columns horizontally and two rows vertically (Fig. 5a). The calculation of the number of windows is given in Supplementary Text S1.

Within each window, all co-occurring land-cover pairs were identified, yielding up to 21 unique combinations among seven classes (Fig. 5b). Pairs were retained only when both cells had valid land-cover, LST and elevation data and differed in elevation by

≤100 m. Duplicate cell pairs arising from overlapping windows were removed, while pairs sharing only one cell were retained as distinct observations. The spatial distribution of the resulting cells is shown in Fig. 5c. Thermal contrasts were calculated as $\Delta LST = LST_A - LST_B$, with positive and negative values indicating that class A was warmer or cooler than class B, respectively. Only combinations with more than 30 valid pairs were analysed at the zonal or national scale. Seventeen combinations met this criterion; forest-shrubland, savanna-shrubland, forest-barren and savanna-barren were excluded owing to insufficient local co-occurrence.

For India-wide estimates, all valid matched pairs were pooled without equal weighting across biogeographic zones, reflecting their heterogeneous land-cover distributions (Supplementary Fig. S2; Supplementary Table S36). Thus, national estimates represent the available matched-pair distribution rather than equally weighted zonal averages. The same matching procedure was applied to derive pairwise differences in albedo, LAI, FPAR, LE, PLE, ET, PET, GPP, PsnNet and NPP.

a



b

21 Possible Land-Cover Pairs

Land Cover	Forest (1)	Savanna (2)	Shrubland (3)	Grassland (4)	Cropland (5)	Urban (6)	Barren (7)
Forest (1)	×	×	×	×	×	×	×
Savanna (2)	1-2	×	×	×	×	×	×
Shrubland (3)	1-3	2-3	×	×	×	×	×
Grassland (4)	1-4	2-4	3-4	×	×	×	×
Cropland (5)	1-5	2-5	3-5	4-5	×	×	×
Urban (6)	1-6	2-6	3-6	4-6	5-6	×	×
Barren (7)	1-7	2-7	3-7	4-7	5-7	6-7	×

c

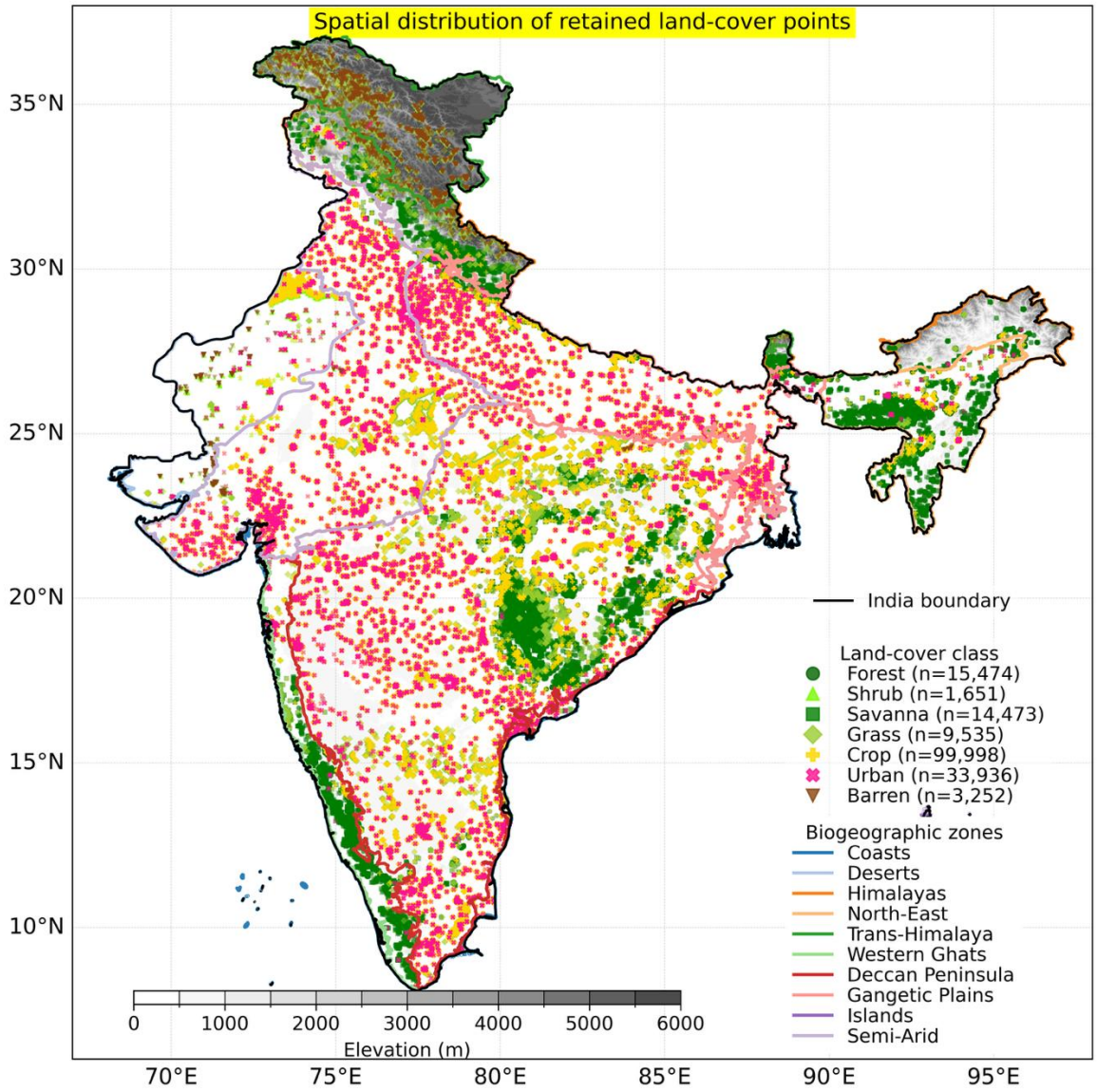


Figure 5. Spatially matched framework for resolving pairwise land-cover thermal contrasts across India. **(a)** Schematic of the moving-window design. A 3 × 5 grid-cell (≈3 km × 5 km) window is moved across each biogeographic zone with an overlap of three columns and two rows; within each window, every pair of cells belonging to different stable land-cover classes is extracted and retained if both cells have valid data and their elevation difference is ≤100 m. **(b)** The 21 possible pairwise combinations among the seven land-cover classes; ΔLST is always computed as the first (column) class minus the second (row) class. **(c)** Locations of the retained stable, homogeneous land-cover cells (n denotes the number of 1-km cells per class), overlaid on elevation and the ten biogeographic zones of India.

Because overlapping windows can introduce spatial dependence among neighbouring observations, we explicitly evaluated the robustness of the thermal contrasts to spatial redundancy. Spatial autocorrelation in ΔLST was quantified using global Moran's I [33] with an eight-nearest-neighbour spatial-weights matrix, and significance was assessed with 999 permutations. Sensitivity to overlapping sampling was further evaluated by partitioning the observations into nine complete, non-overlapping spatial phases, each retaining approximately 11.1% of the original pairs, and recalculating India-wide ΔLST for each phase. Agreement between the original and thinned estimates was assessed with two one-sided tests (TOST; $\alpha = 0.05$; [34]) using a predefined equivalence margin of ±0.05 °C. Spatial-autocorrelation results are reported in Supplementary Table S3, and spatial-thinning sensitivity for daytime, nighttime and day-night mean ΔLST in Supplementary Tables S4a-S4c and Supplementary Fig. S3.

Environmental modulation of thermal contrasts

Environmental dependence was assessed through biogeographic stratification and continuous environmental-gradient analyses. Daytime and nighttime ΔLST were examined across biogeographic zones for natural-natural, natural-anthropogenic, anthropogenic-anthropogenic and barren-related pairs. Daytime natural vegetation-cropland and natural vegetation-urban contrasts were further analysed against the natural-vegetation cell's LST and ET/PET, representing local thermal and evaporative conditions, respectively. ET and PET were matched to LST observations by land-cover class and grid-cell coordinates, and analyses required more than 30 valid pairs.

Nonlinear responses were evaluated using continuous one-breakpoint segmented regression [35]:

$$\Delta\mathrm{LST} = \beta_0 + \beta_1(\mathrm{x} - \mathrm{BP}) + \beta_2 \max(0, \mathrm{x} - \mathrm{BP}) \qquad (1)$$

where (x) is the natural-vegetation cell's daytime LST or ET/PET, BP is the breakpoint, and β_1 and $\beta_1 + \beta_2$ represent slopes below and above BP, respectively. Breakpoint searches were restricted to the predictor's 10th-90th percentiles, with at least 15 observations and 10% of the sample required on either side. We estimated 95% confidence intervals for breakpoints and slopes using 1,000 non-parametric pair-bootstrap replicates.

Segmented models were compared with linear and quadratic alternatives using the AIC and BIC. Evidence for nonlinearity was evaluated jointly from model selection and bootstrap uncertainty in the slope change (β_2). Breakpoint estimates and uncertainty are reported in Supplementary Tables S5-S6, model comparisons in Supplementary Tables S7-S8 and Supplementary Fig. S4, evidence classifications in Supplementary Table S9, and robustness assessments in Supplementary Table S10.

Biophysical attribution of thermal contrasts

To identify the biophysical variables associated with thermal contrasts, daytime and nighttime ΔLST were analysed separately for each land-cover pair using ΔAlbedo, ΔLAI, ΔLE and ΔGPP as predictors. These four variables represent complementary radiative, structural, turbulent-energy and productivity dimensions of the surface while limiting redundancy among correlated MODIS variables. Only observations with complete ΔLST and predictor data were used, and pair-specific models were fitted. Because no single method provides a complete measure of biophysical importance, we combined linear inference, partial association and nonlinear predictive modelling, and synthesized them into a cross-method consensus.

MLR [36,37] was used to quantify the conditional linear association of each predictor with ΔLST while accounting for the others:

$$\Delta\mathrm{LST} = \beta_0 + \beta_1\Delta\mathrm{Albedo} + \beta_2\Delta\mathrm{LAI} + \beta_3\Delta\mathrm{LE} + \beta_4\Delta\mathrm{GPP} + \varepsilon \quad (2)$$

ΔLST and all predictors were standardized (z-scores) within each land-cover pair before fitting, so that standardized coefficients are comparable among predictors. Inference was based on heteroscedasticity-consistent (HC3) standard errors and 95% confidence intervals. Model fit was evaluated with R^2 and adjusted R^2, and multicollinearity was assessed with variance inflation factors (VIFs). Pair-specific coefficients, HC3 confidence intervals and fit statistics are provided for daytime and nighttime ΔLST in Supplementary Tables S13 and S14, respectively.

Partial regression [36] was used to quantify the unique linear association between each predictor and ΔLST after removing covariance with the other three predictors. For each focal predictor, ΔLST and the focal predictor were each regressed on the remaining three predictors, and the two sets of residuals were related by linear regression. The partial correlation $r_{partial}$ and the partial coefficient of determination $R^2_{partial} = r_{partial}^2$ quantified the strength of each unique association, with HC3 standard errors and 95%

589 confidence intervals for inference. Pair-specific partial slopes, partial correlations and
590 partial R^2 are provided for daytime and nighttime ΔLST in Supplementary Tables S15
591 and S16. Robustness to alternative predictor formulations was evaluated by replacing
592 LE with ET, replacing GPP with PsnNet or NPP, and adding the elevation difference
593 (ΔDEM); the tested specifications, multicollinearity diagnostics and resulting sensitivity
594 are documented in Supplementary Tables S23-S26.

595 RF [38] and XGBoost [39] were used to capture nonlinear relationships and interactions
596 not represented by the linear models. Separate models were fitted for each land-cover
597 pair and for daytime and nighttime ΔLST using the same four predictors. Predictive
598 skill was evaluated with repeated nested cross-validation: an outer loop of five folds
599 repeated twice was used exclusively for evaluation, while five-fold inner cross-
600 validation within each outer training set was used for hyperparameter optimization by
601 randomized search, minimizing root-mean-square error (RMSE). This design ensured
602 that observations used to evaluate skill were never used for tuning. Skill on held-out
603 outer folds was quantified with R^2, RMSE and mean absolute error (MAE).
604 Permutation importance was computed exclusively on held-out outer-fold
605 observations: each predictor was permuted ten times and the resulting deterioration
606 in predictive skill quantified its out-of-sample importance. Predictor rankings were
607 recorded for every outer fold to evaluate the stability of importance across partitions.
608 The hyperparameter search space and the pair-specific optimized hyperparameters
609 are given in Supplementary Tables S37-S39, held-out skill in Supplementary Tables
610 S11-S12, permutation importance in Supplementary Tables S17-S18 and cross-
611 algorithm agreement in dominant-predictor rankings in Supplementary Table S19;
612 observed versus out-of-fold predictions are shown in Supplementary Fig. S8.

613 SHAP [40,41] analysis was used to interpret the XGBoost models and characterize the
614 contribution of each predictor to modelled ΔLST. After cross-validated evaluation, a
615 final XGBoost model was tuned and fitted to all available observations for each eligible
616 pair and response, and SHAP values were computed with the TreeSHAP algorithm
617 for up to 5,000 observations per model. The global importance of predictor j was
618 quantified as its mean absolute SHAP value,

619 $$I_j^{SHAP} = (1/\mathrm{n})\ \Sigma_i|\varphi_{ij}| \qquad (3)$$

where φ_{ij} is the contribution of predictor j to the prediction for observation i and n is the number of observations evaluated. Mean absolute SHAP values were normalized across the four predictors to allow comparison within each pair, and SHAP-value distributions were examined to characterize how the magnitude and direction of each predictor were associated with modelled ΔLST. Permutation importance and SHAP were treated as complementary: the former measures out-of-sample predictive dependence, the latter the contribution of each predictor to fitted predictions. Neither was interpreted as a causal contribution of a biophysical process to ΔLST. Dominant SHAP attribution and complete SHAP importance are provided in Supplementary Tables S20-S22, and the corresponding nighttime MLR, partial-R^2, permutation-importance and SHAP patterns are summarized in Supplementary Fig. S7.

To synthesize evidence across the attribution approaches, we computed a consensus importance from five measures: the absolute standardized MLR coefficient, partial R^2, RF permutation importance, XGBoost permutation importance and XGBoost mean absolute SHAP. Because these measures differ in scale and meaning, importance was normalized within each method, land-cover pair and day/night response as

$$\mathrm{NI}_{j,m} = \mathrm{I}_{j,m} / \Sigma_j \mathrm{I}_{j,m} \qquad (4)$$

where $\mathrm{I}_{j,m}$ and $\mathrm{NI}_{j,m}$ are the raw and normalized importance of predictor j under method m. Negative permutation importance was set to zero before normalization because it provides no evidence of a positive predictive contribution. Consensus importance was then calculated as the mean normalized importance across methods,

$$\mathrm{CI}_j = (1/\mathrm{M}) \Sigma_m \mathrm{NI}_{j,m} \qquad (5)$$

where M is the number of methods providing a valid estimate. The predictor with the largest CI_j for a given pair and response was designated the dominant consensus predictor; with four predictors, CI_j = 0.25 corresponds to an equal share, providing a natural reference for the strength of dominance. The standard deviation of $\mathrm{NI}_{j,m}$ across methods was retained as a measure of cross-method agreement. Consensus importance is interpreted as a synthesis of relative predictor importance, not as a partitioning of causal contributions to ΔLST. Complete consensus importance and between-method variability are provided in Supplementary Table S27.

Spatial-block cross-validation

A spatial-block analysis evaluated whether RF and XGBoost skill and predictor importance were sensitive to spatial dependence. For each pair, observations were assigned to spatial blocks according to the midpoint latitude and longitude of the paired cells. Five block sizes (0.5°, 1°, 2°, 5° and 10°) progressively increased the separation between training and test data. For each block size, spatially grouped five-fold nested cross-validation kept all observations from a block within the same fold, and hyperparameters were tuned using spatially grouped inner folds, reducing spatial leakage during both tuning and evaluation. To test sensitivity to the arbitrary placement of block boundaries, each block size was evaluated for two grid origins: the original grid and a grid shifted by half a block width in latitude and longitude. Skill on held-out blocks (R^2, RMSE, MAE) was summarized across outer folds and grid origins, and its change with block size was used to assess the geographical transferability of the models. Block configurations and RF/XGBoost sensitivity, including comparisons between random and spatially blocked cross-validation, are documented in Supplementary Tables S28-S32.

Acknowledgments

The authors gratefully acknowledge institutional support from the Indian Institute of Science Education and Research (IISER) Bhopal for providing the computational and research infrastructure essential to this study. Financial support through the Prime Minister's Research Fellowship (PMRF) and the University Grants Commission (UGC) is sincerely acknowledged. We also thank the providers of the MODIS datasets and our colleagues and collaborators for their valuable feedback.

Data Availability Statement

All MODIS datasets used in this study are publicly available through NASA's Land Processes Distributed Active Archive Center (LP DAAC) and were accessed using the Application for Extracting and Exploring Analysis Ready Samples (AppEEARS)[43]. Land-cover data were obtained from the MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid V061 product (MCD12Q1.061; https://doi.org/10.5067/MODIS/MCD12Q1.061)[44]. Land-surface temperature data were obtained from the MODIS/Aqua Land Surface Temperature/Emissivity 8-Day L3 Global 1 km SIN Grid V061 product (MYD11A2.061; https://doi.org/10.5067/MODIS/MYD11A2.061)[45]. Surface albedo was obtained from

683 the MODIS/Terra+Aqua BRDF/Albedo Albedo Daily L3 Global 500 m V061 product
684 (MCD43A3.061; https://doi.org/10.5067/MODIS/MCD43A3.061)[46]. Leaf area index
685 (LAI) and fraction of photosynthetically active radiation (FPAR) were obtained from the
686 MODIS/Aqua Leaf Area Index/FPAR 8-Day L4 Global 500 m SIN Grid V061 product
687 (MYD15A2H.061; https://doi.org/10.5067/MODIS/MYD15A2H.061)[47].
688 Evapotranspiration (ET), potential evapotranspiration (PET), latent heat flux (LE), and
689 potential latent heat flux (PLE) were obtained from the MODIS/Aqua Net
690 Evapotranspiration Gap-Filled 8-Day L4 Global 500 m SIN Grid V061 product
691 (MOD16A2GFMYD16A2GF.061;
692 https://doi.org/10.5067/MODIS/MYD16A2GF.061)[48]. Gross primary productivity
693 (GPP) and net photosynthesis (PsnNet) were obtained from the MODIS/Aqua Gross
694 Primary Productivity Gap-Filled 8-Day L4 Global 500 m SIN Grid V061 product
695 (MYD17A2HGF.061; https://doi.org/10.5067/MODIS/MYD17A2HGF.061)[49], while net
696 primary productivity (NPP) was obtained from the MODIS/Aqua Net Primary
697 Production Gap-Filled Yearly L4 Global 500 m SIN Grid V061 product
698 (MYD17A3HGF.061; https://doi.org/10.5067/MODIS/MYD17A3HGF.061)[50]. Elevation
699 data were obtained from the NASA Shuttle Radar Topography Mission Global 3 Arc-
700 Second V003 dataset (SRTMGL3.003), produced by the NASA Jet Propulsion
701 Laboratory, with an approximate spatial resolution of 90 m
702 (https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL3.003)[51].

703 Data processing and analysis were performed using Climate Data Operators (CDO)
704 version 2.4.4 (https://mpimet.mpg.de/cdo) and Python version 3.13.1 (Python
705 Software Foundation; https://www.python.org/downloads/release/python-3131/). CDO
706 was used for command-line processing and manipulation of gridded climate and
707 remote-sensing datasets, whereas Python was used for statistical analysis, data
708 handling, and visualization. The data generated and analysis scripts used in this study
709 are publicly available through Zenodo at https://doi.org/10.5281/zenodo.23672848[42].

710 **Conflict of Interest Disclosure**

711 The authors declare there are no conflicts of interest for this manuscript.

712 **Reference**

713 1. Devaraju, N., Bala, G. & Nemani, R. Modelling the influence of land-use changes
714 on biophysical and biochemical interactions at regional and global scales. *Plant*
715 *Cell Environ.* **38**, 1931–1946 (2015).

2. Duveiller, G., Hooker, J. & Cescatti, A. The mark of vegetation change on Earth's surface energy balance. *Nat. Commun.* **9**, 679 (2018).
3. Lawrence, D., Coe, M., Walker, W., Verchot, L. & Vandecar, K. The Unseen Effects of Deforestation: Biophysical Effects on Climate. *Front. For. Glob. Change* **5**, (2022).
4. Mahmood, R. *et al.* Land cover changes and their biogeophysical effects on climate. *Int. J. Climatol.* **34**, 929–953 (2014).
5. Perugini, L. *et al.* Biophysical effects on temperature and precipitation due to land cover change. *Environ. Res. Lett.* **12**, (2017).
6. Li, Y. *et al.* Local cooling and warming effects of forests based on satellite observations. *Nat. Commun.* **6**, 6603 (2015).
7. Zhang, Z., Zhang, F., Wang, L., Lin, A. & Zhao, L. Biophysical climate impact of forests with different age classes in mid- and high-latitude North America. *For. Ecol. Manag.* **494**, 119327 (2021).
8. Zeng, Z. *et al.* Climate mitigation from vegetation biophysical feedbacks during the past three decades. *Nat. Clim. Change* **7**, 432–436 (2017).
9. Lian, X. *et al.* Biophysical impacts of northern vegetation changes on seasonal warming patterns. *Nat. Commun.* **13**, 3925 (2022).
10. Li, Y. *et al.* Amplified local cooling effect of forestation in warming Europe. *Nat. Commun.* **16**, 8412 (2025).
11. Massey, R. *et al.* Forest composition change and biophysical climate feedbacks across boreal North America. *Nat. Clim. Change* **13**, 1368–1375 (2023).
12. Alkama, R. & Cescatti, A. Biophysical climate impacts of recent changes in global forest cover. *Science* **351**, 600–604 (2016).
13. Zhu, L. *et al.* Comparable biophysical and biogeochemical feedbacks on warming from tropical moist forest degradation. *Nat. Geosci.* **16**, 244–249 (2023).
14. Chen, C. *et al.* Biophysical effects of croplands on land surface temperature. *Nat. Commun.* **15**, 10901 (2024).
15. Luo, H. & Quaas, J. Cropland biophysical impacts on land surface temperature show diurnal differences across tropical Africa. *Commun. Earth Environ.* **7**, 309 (2026).
16. Kan, F. *et al.* Diminished biophysical cooling benefits of global forestation under rising atmospheric CO2. *Nat. Commun.* **16**, 4410 (2025).
17. Zhang, C. *et al.* Globally constrained forest biophysical cooling benefits under rising atmospheric dryness. *Nat. Clim. Change* **16**, 956–962 (2026).
18. Bren d'Amour, C. *et al.* Future urban land expansion and implications for global croplands. *Proc. Natl. Acad. Sci.* **114**, 8939–8944 (2017).
19. Huang, B. *et al.* Predominant regional biophysical cooling from recent land cover changes in Europe. *Nat. Commun.* **11**, 1066 (2020).
20. Sharma, J. & Kumar, P. *Quantification of Potential Forestation Induced Change in Daytime Land Surface Temperature in India*. (2025). doi:10.5194/egusphere-egu25-5008.
21. Syktus, J. I. & McAlpine, C. A. More than carbon sequestration: Biophysical climate benefits of restored savanna woodlands. *Sci. Rep.* **6**, 29194 (2016).

22. You, N. *et al.* Climate-driven global cropland changes and consequent feedbacks. *Nat. Geosci.* **18**, 639–645 (2025).
23. Bonan, G. B. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science* **320**, 1444–1449 (2008).
24. Foley, J. A. *et al.* Global Consequences of Land Use. *Science* **309**, 570–574 (2005).
25. Sharma, J. Land-Cover Dynamics and Their Influence on Surface Temperature Across India: Evidence From Forest–Open Thermal Contrasts. *J. Geophys. Res. Biogeosciences* **131**, (2026).
26. Sharma, J. & Kumar, P. *Biophysical Controls on Climate Responses to Tree-Cover Change Across Land-Use Types*. (2026). doi:10.21203/rs.3.rs-9409460/v1.
27. Si, M. *et al.* Global non-uniformity in biophysical surface temperature responses to cropland expansion over non-forest vegetation. *Commun. Earth Environ.* **6**, 546 (2025).
28. Kuttippurath, J. & Kashyap, R. Greening of India: Forests or Croplands? *Appl. Geogr.* **161**, 103115 (2023).
29. Pitman, A. J. *et al.* Importance of background climate in determining impact of land-cover change on regional climate. *Nat. Clim. Change* **1**, 472–475 (2011).
30. Li, Y. *et al.* Observed different impacts of potential tree restoration on local surface and air temperature. *Nat. Commun.* **16**, 2335 (2025).
31. Duveiller, G. *et al.* Revealing the widespread potential of forests to increase low level cloud cover. *Nat. Commun.* **12**, 4337 (2021).
32. Luo, H., Quaas, J. & Han, Y. Decreased cloud cover partially offsets the cooling effects of surface albedo change due to deforestation. *Nat. Commun.* **15**, 7345 (2024).
33. MORAN, P. A. P. NOTES ON CONTINUOUS STOCHASTIC PHENOMENA. *Biometrika* **37**, 17–23 (1950).
34. Schuirmann, D. J. A comparison of the Two One-Sided Tests Procedure and the Power Approach for assessing the equivalence of average bioavailability. *J. Pharmacokinet. Biopharm.* **15**, 657–680 (1987).
35. Muggeo, V. M. R. Estimating regression models with unknown break-points. *Stat. Med.* **22**, 3055–3071 (2003).
36. Frisch, R. & Waugh, F. V. Partial Time Regressions as Compared with Individual Trends. *Econometrica* **1**, 387–401 (1933).
37. Pearson, K. VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philos. Trans. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* 253–318 (1896) doi:10.1098/rsta.1896.0007.
38. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
39. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, New York, NY, USA, 2016). doi:10.1145/2939672.2939785.
40. Lundberg, S. M. *et al.* From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2**, 56–67 (2020).

41. Lundberg, S. M. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. in (2017).
42. Sharma, J. & Kumar, P. Data and code for: Forest cooling is not natural-vegetation cooling: pair-specific and environmentally conditioned land-cover thermal contrasts across India. https://doi.org/10.5281/zenodo.23672848 (2026) doi:10.5281/zenodo.23672848.
43. AppEEARS Team. (2026). Application for Extracting and Exploring Analysis Ready Samples (AppEEARS). Ver. 3.125.1. NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, USA. Accessed August 21, 2026. https://appeears.earthdatacloud.nasa.gov
44. Friedl, M., & Sulla-Menashe, D. (2022). MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MCD12Q1.061
45. Wan, Z., Hook, S., & Hulley, G. (2021). MODIS/Aqua Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MYD11A2.061
46. Schaaf, C., & Wang, Z. (2021). MODIS/Terra+Aqua BRDF/Albedo Daily L3 Global - 500m V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MCD43A3.061
47. Myneni, R., Knyazikhin, Y., & Park, T. (2021). MODIS/Aqua Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MYD15A2H.061 Date Accessed: 2026-10-26
48. Running, S., Mu, Q., Zhao, M., & Moreno, A. (2021). MODIS/Aqua Net Evapotranspiration Gap-Filled 8-Day L4 Global 500m SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MYD16A2GF.061
49. Running, S., & Zhao, M. (2021). MODIS/Aqua Gross Primary Productivity Gap-Filled 8-Day L4 Global 500m SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MYD17A2HGF.061
50. Running, S., & Zhao, M. (2021). MODIS/Aqua Net Primary Production Gap-Filled Yearly L4 Global 500m SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MYD17A3HGF.061
51. NASA JPL. (2013). NASA Shuttle Radar Topography Mission Global 3 arc second [Dataset]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MEASURES/SRTM/SRTMGL3.003 Date Accessed: 2026-10-26

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTexts.pdf
- SupplementaryTables.pdf
- SupplementaryFigures.pdf