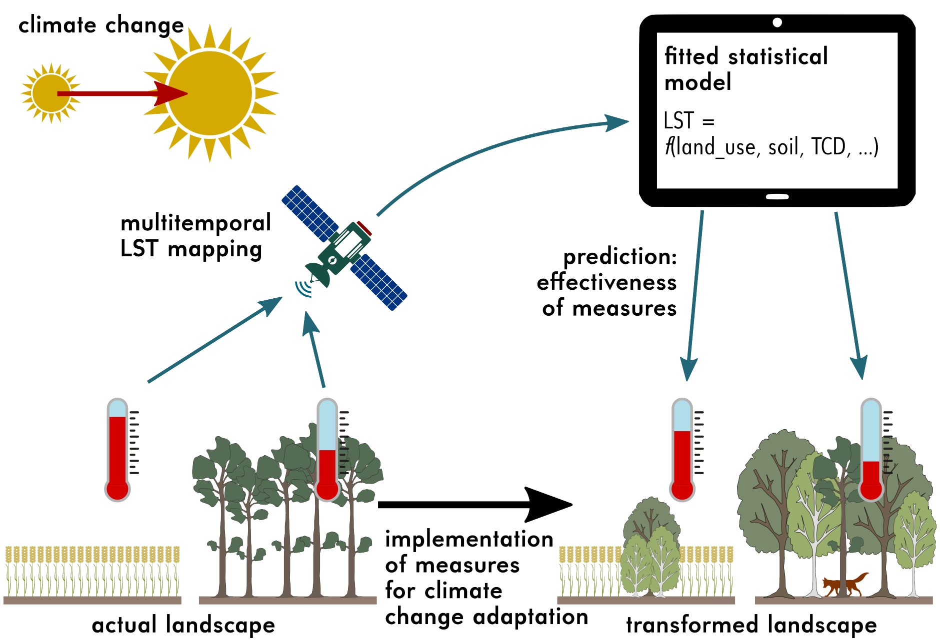
Graphical Abstract

Assessing the cooling potential of climate change adaptation measures in rural areas

Beate Zimmermann, Sarah Kruber, Claas Nendel, Henry Munack, Christian Hildmann



Highlights

Assessing the cooling potential of climate change adaptation measures in rural areas

Beate Zimmermann, Sarah Kruber, Claas Nendel, Henry Munack, Christian Hildmann

- A proxy for the landscape mesoclimate is the land surface temperature
- Land surface temperature can be predicted to assess climate change adaptation
- Measures to increase evapotranspiration and water retention maximize cooling effects
- The determined performance of the measures varies in space and time
- The established ranking of measure performance provides decision support

Assessing the cooling potential of climate change adaptation measures in rural areas

Beate Zimmermann[a,*], Sarah Kruber[a], Claas Nendel[b,c], Henry Munack[d], Christian Hildmann[a]

[a]*Research Institute for Post-Mining Landscapes, Brauhausweg 2, Finsterwalde, 03238, Brandenburg, Germany*
[b]*Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Straße 84, Müncheberg, 15374, Brandenburg, Germany*
[c]*Institute of Biochemistry and Biology, Am Mühlenberg 3, Potsdam, 14476, Brandenburg, Germany*
[d]*Faculty of Science, Medicine and Health, School of Earth, Atmospheric and Life Sciences, and ARC Centre of Excellence for Australian Biodiversity and Heritage, Northfields Ave Wollongong, Wollongong, 2522, NSW, Australia*

Abstract

Evapotranspiration provides local cooling of land surfaces during the vegetation period. Increased water retention in the landscape through nature-based solutions can promote plant transpiration during dry periods and thus contribute to climate change adaptation by stabilising the regional climate. The land surface temperature (LST) can be considered a proxy for landscape mesoclimate. LST data are available through satellite imagery at high spatial and temporal resolution and can be linked to environmental factors.

We established a statistical relationship between LST and environmental predictors in a drought-prone rural study area in Germany. We also mapped

*Corresponding author
Email addresses: `b.zimmermann@fib-ev.de` (Beate Zimmermann), `s.kruber@fib-ev.de` (Sarah Kruber), `claas.nendel@zalf.de` (Claas Nendel), `hmunack@uow.edu.au` (Henry Munack), `c.hildmann@fib-ev.de` (Christian Hildmann)

various climate change adaptation measures to potential areas where they could be implemented. Using the fitted model and the expected expression of predictors before and after a hypothetical measure implementation on the potential areas, we evaluated the measures' probable cooling effect.

The model captured a large part of the LST variation. Measures that promote evapotranspiration and water retention were found to be the most successful in terms of their potential cooling effect. In particular, the afforestation of arable land and urban brownfields and the rewetting of former wet meadows had cooling capacities up to 3.5 K. The effect of the measures varied spatially and also showed temporal dynamics during the year. These aspects can be taken into account with our approach when using the results for decision support.

Keywords: climate change adaptation, land surface temperature, water retention, evapotranspiration, cooling effect, measure evaluation

---

1. Introduction

In 2022, after a continuous 2014–2020 series of warmest years on record (Copernicus, 2022), Europe experienced its hottest summer ever recorded, with temperatures at 2.3 K above the pre-industrial benchmark level (World Meteorological Organization (WMO), 2023). The line-up of adverse impacts of rising temperatures, involving heat and drought extremes, can hardly be overestimated. For Europe, the IPCC (Bednar-Friedl et al., 2023), in light of the global warming peril, has identified four key risks (kr) – heat-related health issues (kr1), heat and drought stress on crops (kr2), water scarcity (kr3), and hydrological extremes (kr4) – which are consistently related to heat and water.

The term “heat” in this context commonly refers to sensible heat of the surrounding air that exceeds a comfort temperature threshold, and that may even become critical for physiological processes of organisms, including humans. Air temperature is a result of (i) the turnover of incoming radiation at the surface into sensible heat of the sunlit material, while a remaining fraction is reflected into the atmosphere and (ii) atmospheric transport processes that exchange air volumes of different heat content dynamically across a large range of temporal and spatial scales. In soils, a considerable fraction of the substrate is liquid water, which, when heated, may turn into vapour and evaporate from the soils’ surface. The change of the state of aggregation requires energy, and this consumption results in a turnover of sensible into latent heat and an immediate temperature decrease of the surrounding media, a local cooling effect. The same principle applies when liquid water turns into vapour in the stomata of plants and then transpires from the plants surface. The cooling effect of transpiration protects plants against extreme heat, as critically high temperatures may cause damage to the plants tissue. This is especially relevant when vulnerable parts are exposed, such as pollen during flowering.

A consequence of the complex interaction of many factors, transpiration is highly dynamic in both space and time. The vegetation’s influence on the surrounding microclimate increases with both the number of plants and their transpiring surface. This causes any vegetation, yet to different degree, to develop its own specific microclimate. Depending on the extent of the vegetation, the combination of evaporation from the soil and transpiration from the vegetation, commonly referred to as evapotranspiration, may even influence meteorological variables at the landscape scale, referred to as the mesoclimate. Landscape cooling by evaporation affects moist atmospheric dynamics and dampens land-ocean temperature contrasts, maintaining the regional hydrologic cycle (Makarieva et al., 2022). Transpired water can be condensed again with a temporal delay and spatial shift (Makarieva et al., 2023). Because long-term temperature contrasts during the warm season in the Northern Hemisphere are close to the threshold where condensation-induced moisture transport ceases, a few degrees of additional land warming can disrupt condensation-induced moisture transport. Especially during heat waves, surface cooling provided by transpiring vegetation can be critical to avoid a tipping point where the climate turns into arid conditions (Makarieva et al., 2022).

While microclimatic effects can only be measured with proximal sensing equipment, mesoclimatic effects can potentially be investigated using remote sensing methods. Three landscape properties are relevant for this: the land surface albedo (i) describes the degree to which incoming radiation is reflected into the atmosphere. It is a proxy for the surface's potential to turn radiation into sensible heat. Surface roughness (ii) is defined as the deviations in the direction of the normal vector of a real surface from its ideal form, usually an ideally flat surface. Land surface roughness usually considers deviations at the scale of decimetres to decametres (Burakowski et al., 2018). Land surface temperature (iii), or LST, represents the transferred energy at the phase interface of the boundary layer and, thus, provides insights into landscape energy turnover. It integrates both albedo and land surface roughness plus the meteorological variables, and is therefore considered a proxy for the landscape's mesoclimate and, within limits, also to the microclimate within vegetation stands (Hesslerová et al., 2018; Ghafarian et al., 2024). It has hence been identified a measurable indicator of ecosystem and landscape functioning (Hesslerová et al., 2013). Albedo and land surface roughness are variable, but follow fairly strict patterns in space and time, while LST is often highly dynamic, responding to many influencing factors. In this study, we will focus solely on LST when investigating how the cooling service of vegetation in rural landscapes can be employed to mitigate potentially harmful effects of landscape overheating.

Using satellite thermal imagery, the LST can be quantified in a high spatial and temporal resolution (e.g. Procházka et al., 2019; Hesslerová et al., 2018). Since the LST represents transferred energy at the boundary layer, it is not directly comparable with the air temperatures measured at weather stations. To date, remotely sensed LST data has been used to map urban heat islands (e.g. Abir et al., 2021; Sultana and Satyanarayana, 2020), but less so with regard to rural regions (Ghafarian et al., 2024). In this context, research as so far focused on disentangling the relationship between LST and environmental factors (Alavipanah et al., 2015; Bertoldi et al., 2010; Das et al., 2020; Seeberg et al., 2022; Wickham et al., 2012), for instance to locate warming and cooling spots in cities (Seeberg et al., 2022). In a next step, climate adaptation strategies can be assessed e.g. based on LST differences between built-up and green spaces (Selim et al., 2023).

Climate adaptation measures resemble each other around the globe. However, due to regionally varying global warming impacts, adaptation measures must be bespoke regional responses to a global problem. Yet many regions have in common that, prior to the advent of industrial land-use practices, water was a ubiquitous climate buffer agent. In many eastern German watersheds, land use became intensive and radical with increasing industrialization (Baessler and Klotz, 2006; Schils et al., 2022). Among other interventions, lowland drainage by means of subsurface and ditch drainage was strongly advanced from the 1960s onward, as a response to a longer period of high rainfall during spring time (Ionita et al., 2020). Landscape drainage, however, leads to accelerated loss of rainwater, rendering the landscape generally drier. Lacking water anyway, and further pressured by advancing climate change, the vegetation in these areas has become increasingly drought-stressed, which has been witnessed especially in the years 2018 and following. Any attempt to re-increase vegetation growth and transpiration-mediated cooling at landscape scale (Sušnik et al., 2022; Monteiro et al., 2016; Hildmann et al., 2022) must therefore address the way pluvial water is handled. Using the landscape itself as an interim water storage is the new goal, reversing the large-scale drainage of the past decades.

A range of measures addressing different spatial scales can be taken to achieve this (Hildmann et al., 2022), starting with the establishment of green roofs in cities, and ending at large-scale measures in rural areas, such as soil organic matter enrichment to increase the water infiltration into soils, or the re-wetting of degraded wetlands. However, a quantification of the effect of these measures for mitigating climate change through evaporative cooling is still lacking, and with this a scientific basis for deriving recommendations to decision makers. To help to fill this knowledge gap, we aimed at quantifying the influence of various environmental variables on LST. Using the estimated relationships, we investigated the cooling effect of climate adaptation measures in a drought-prone rural study area.

2. Materials and Methods

2.1. Study area

Our study area covers the Elbe-Elster-County in the south of the federal state of Brandenburg in Eastern Germany. The county inhabits 100,902 people (Amt für Statistik Berlin-Brandenburg, 2023) on an area of 1,899 km^2. 51% and 36% of the county are agricultural land and forests, respectively (Amt für Statistik Berlin-Brandenburg, 2022). Settlements and traffic areas account for 10% of the total area. Cereal farming dominates the agricultural areas, while about one fourth of the agricultural land is grassland. The forested areas are dominated by pine plantations. The low population density by German standards, the high proportion of agricultural land and forest, and the absence of major cities account for the rural character of the area (Wolff et al., 2021).

Climatically, the Elbe-Elster-County is located in the temperate climate zone with a mean annual air temperature of 9.1 °C and an annual precipitation of 556.9 mm (period 1971–2000, Pfeifer et al., 2021). A comparison of the 30-year periods 1951-1980 versus 1986-2015 revealed in an increase for the annual mean temperature by an average of about 0.9 °C (Pfeifer et al., 2021). The average annual precipitation shows a slight, statistically non-significant increase (Pfeifer et al., 2021). The distribution of precipitation in the course of the year has shifted between the climatic normal period 1961 to 1990 and the current period 1991 to 2020 in the direction of lower spring and higher winter precipitation (Zimmermann and Hildmann, 2021). Already today, the aridity index (AI) pushes the Elbe-Elster-County into the dry sub-humid range (county-wide mean of the annual AI average: 0.64, derived from the Global ET_0 and Aridity Index Database v3 in Zomer et al., 2022) although regional differences exist. The western part in the direction of the river Elbe is drier than the southeastern and eastern areas, which is partly due to the topography.

Particularly in recent years, the region has experienced some severe soil moisture droughts during the growing season, resulting in crop yield losses, forest damage, and wetland degradation. This is consistent with Samaniego et al. (2018)'s projection that the continental climate zone – where the county is located in – will experience negative changes in available soil water content under ongoing climate change in all seasons except winter; an effect of the drastically increasing water vapour deficit of the warming atmosphere and the resulting increase in ET_0.

The soils in the region are predominantly sandy except for some alluvial areas near the rivers Elbe and Schwarze Elster with loamy soils. Their low available water holding capacity, in many areas below 140 mm in the first metre of the soil profile, exacerbate the plant water deficiency during drought periods. Peat and degraded peat soils occur in fens that are often drained for agricultural use. An analysis of historical and current regional maps revealed that the very low natural drainage density is more than doubled by artificial ditches.

2.2. Measures for climate change adaptation

Nature-based solutions for climate adaptation are particularly appropriate in our rural study area. We focused on measures that improve the microclimate and prevent overheating of landscapes via plant evapotranspiration. Since our region often suffers from water deficiency during the growing season, the desired effect of evaporative cooling requires adequate water retention in the landscape. This is particularly important due to the shift of precipitation to the winter months (Chpt. 2.1). In this article, we consider 16 climate adaptation measures (Tab. 1). Their detailed description including some information on costs, implementation hurdles, synergy effects, etc. can be found in Hildmann et al. (2022). Very small-scale measures such as roof and facade greening cannot be considered because their scale does not correspond to the resolution of the satellite data, which we used to assess the measures' effect (Chpt. 2.3.1).

In order to assign the measures to possible areas where they could be implemented, we have developed specific GIS algorithms (App. Appendix C). A small-scale land-use map of our study area with a total of almost 60,000 plots served as the map basis (App. Appendix B).

2.3. Evaluation of the measures

The potential benefits of the 16 selected measures were evaluated by quantifying their influence on land surface temperature (LST). A comparative study, e.g. with a BACI (Before-after-control-impact) design or space-for-time substitution, was not possible due to too few implemented measures in our study area. We therefore derived environmental variables from available geospatial data that are known or suspected to influence LST. By fitting a

Table 1: Considered climate change adaptation measures

Category	Measure	Description
Agriculture	Agroforestry	Combined cultivation of arable crops and perennial woody plants
Agriculture	Deep-rooting crops	Cultivation of deep-rooted arable crops
Agriculture	Landscape structure elements	Planting of woody plants on agricultural land along ditches and trails
Agriculture	Organic fertilisation	Continuous application of organic matter to arable land
Agriculture	Permanent grassland	Conversion of arable land into permanent grassland
Agriculture	Afforestation	Afforestation of marginal arable land with broad-leaved trees
Forestry	Reforestation	Reforestation of degraded forest patches
Forestry	Forest conversion	Conversion of pure conifer stands into mixed or deciduous stands
Nature conservation	Wet meadows	Rewetting of former wet meadows
Settlements	Afforestation	Afforestation of (un)sealed urban brownfields, if necessary after unsealing
Settlements	Partial unsealing	Unsealing and laying of grass pavers
Settlements	Tree groups	Planting or compaction of groups of trees
Settlements	Tree rows	Planting rows of trees
Settlements	Unsealing	Unsealing and subsequent greening of sealed urban brownfields
Water management	Ditch water management	Adapted control of farm dam heights in melioration ditches
Water management	Supporting sills	New construction of supporting sills in melioration ditches

statistical relationship between these predictors and LST, we were able to estimate measure effects via the expected expression of the predictors before and after hypothetical measure implementation. In the following we describe the data and the model used for this purpose. Fig. 1 gives an overview about our methods.

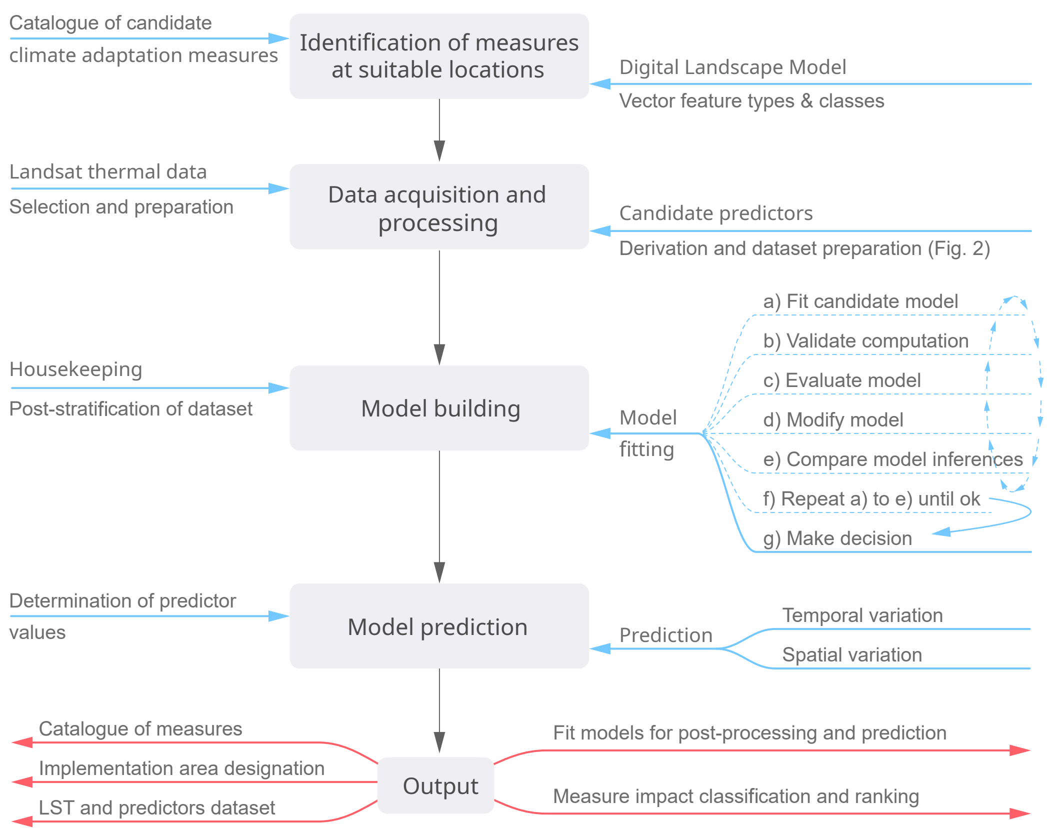


Figure 1: Workflow

2.3.1. Data on land surface temperature

The target variable *LST* was derived from 39 atmospherically corrected Landsat 8 (Collection 2 Level 2) satellite scenes. All 39 scenes came from the main vegetation periods of the years 2013 to 2020, which encompass the months of May to September in our study area. Pre-processing of the scenes included, among other things, removal of cloud shadows (App. Appendix A). As the absolute values of LST fluctuate due to weather conditions, they were converted into relative values. The *min-max feature scaling* was used as the scaling method.

2.3.2. Predictor selection

Drawing from knowledge of land surface temperature and our prediction goals, we chose various environmental variables, encompassing land cover, land use, landscape structure, soil, topography, and climate (Tab. 2).

As the basis for land cover classification, we used the biotope and land-use mapping of the federal state of Brandenburg (Landesamt für Umwelt Brandenburg, 2009). Because the reference year for the map was 2009, we made several updates, e.g. to include newly built solar parks (Chpt. Appendix B). Areas with LULC change between 2013 and 2020 were excluded. The updated map served as the basis for both the localisation of the measures (Chpt. Appendix C) and the derivation of predictors related to LULC and landscape structure. Based on predefined rules, we converted the large number of the original biotope classes into five LULC classes (Tab. 3). We only used classes that are relevant for predicting the efficacy of the climate adaptation measures. As some measures on arable land are associated with specific crop rotations on agricultural land, we have further subdivided arable lands according to crop type (Tab. 3). Forests were divided into coniferous, broad-leaved and mixed forests.

Moreover, the thermal signature of grasslands most likely also depends on

Table 2: Predictors for modelling

Domain	Predictor	Abbreviation	Type[1]	Time reference (year)	Range (unit)/No. of levels
Climate	Potential evapotranspiration	ET	Num.	Scene[2]	1 – 7.9 mm/d
Climate	Climatic water balance	CWB	Num.	Scene[3]	-140 – 51 mm/m
Land cover/use	Land-use/land-cover class	LULC	Cat.	Static (2020)/annual[4]	13
Land cover/use	Tree cover density	TCD	Num.	Static (2018)	0 – 1
Land cover/use	Imperviousness density	IMD	Num.	Static (2018)	0 – 1
Landscape structure	Core area index	CAI	Num.	Static (2020)	0.21 – 0.98
Landscape structure	Related circumscribing circle	CIRCLE	Num.	Static (2020)	0.27 – 0.99
Soil	Available water-holding capacity	AWC	Ord.	Static (2021)	9
Topography	Terrain elevation	ELEV	Num.	Static (2004)	76 – 201 m a.s.l.

[1] Num. = numerical, cat. = categorical, ord. = ordinal

[2] Daily ET on day of satellite overpass

[3] Monthly total before the day of satellite overpass

[4] Crop type from annual agricultural census data

Table 3: LULC classes and corresponding CORINE (CLC) codes (©European Union, Copernicus Land Monitoring Service, European Environment Agency)

LULC class	Subclass	Coding	CLC code[1]
Settlements		1	1.1 to 1.4
Arable land			2.1[2]
	Alfalfa	2	
	Legumes[3]	3	
	Maize[4]	4	
	Sunflower	5	
	Winter cereals[5]	6	
	Winter oil-seed rape	7	
Grassland			2.3
	Fresh grassland[6]	8	
	Dry grassland[7]	9	
Forest	Broad-leaved forest	10	3.1.1
	Coniferous forest	11	3.1.2
	Mixed forest	12	3.1.3
Fens		13	4.1

1 The classes 2.2, 3.2, 3.3, 4.2 and 5 are not used. Class 2.4 is either embedded in class arable land or does not exist in our study area.

2 except 2.1.3

3 peas and lupines

4 silage maize and grain maize

5 spelt, winter barley, winter rye, winter triticale, winter wheat

6 based on NDMI (Chapter 2.3.2), with assumed groundwater influence

7 based on NDMI (Chapter 2.3.2), without assumed groundwater influence

groundwater influence: during drought periods, grassland without groundwater influence tends to dry out more quickly and accordingly heats up more strongly. The scale of the information that the state of Brandenburg provides on groundwater is not suitable for depicting such small-scale patterns. Therefore, we have chosen a remotely sensed moisture index (Normalised Difference Moisture Index, NDMI) that suggests the presence or absence of a groundwater influence on plant water supply (App. Appendix A). With a histogram analysis of the NDMI composite (App. Appendix A), we could distinguish between "fresh" grassland areas influenced by groundwater and those that are remote from groundwater ("dry" grassland, Tab. 3). This enabled us to evaluate the measures "ditch water management" and "supporting sills" (Tab. 1).

Prior to modelling, the LULC classes arable land, forest and grassland were replaced with their subclasses, resulting in the predictor $LULC$ having a total of 13 levels. In total, we included nine predictors in the analysis, which have different time references as well as different data formats (Tab. 2). The variables $ELEV$, ET and CWB were centred and scaled to one unit variance before entering them into the model.

Landscape structure metrics (LSM) were included because it is known that landscape structure influences LST (e.g. Song et al., 2014; Das et al., 2020; Kim et al., 2016). Effects such as urban heat islands (e.g. Kim et al., 2016) or the forest interior climate can thus be mapped. From the large number of available LSMs, we only considered scale-independent metrics at the patch level. We limited the remaining selection to the core area index (CAI) and the related circumscribing circle ($CIRCLE$) because of their assumed importance for LST: CAI is used to describe effects such as heat islands, while $CIRCLE$ can be used to distinguish between linear (e.g. hedges) and clumped landscape elements. Both CAI and $CIRCLE$ are configuration metrics, which have been found to exert a stronger influence on LST than composition metrics (Du et al., 2016).

A summary of the workflow to derive the predictors from the geospatial data is provided in Fig. 2. The common geometry is the raster data format with the LST map resolution of 30 m. Pixels whose land-use class is not homogeneous within a 90 m grid cell (approx. the original resolution of Landsat-derived LST) were classified as mixed pixels and excluded from the subsequent statistical analysis. In addition to the geospatial predictors, we included the pixel id (ID) and the scene number ($SCENE$) as group-level predictors in the modelling (Chpt. 2.3.4)

2.3.3. Generation of subsets of data

The original data set contains nearly 30 million observations. It is neither possible for computational reasons nor is it necessary to transfer the entire data set into the model. In addition, the LST values are auto-correlated: a variogram analysis of the scaled LST averaged over the 39 scenes revealed a very strong and long-range spatial correlation (nugget-to-sill ratio of 0, effective range about 25 km). In order to avoid a correlation of the estimation errors and to reduce the amount of data, we applied a post-stratification routine (App. Appendix D).

The final subsets contained about 50,000 observations. We repeated the sampling procedure three times to get three final calibration subsets. This was done to test the influence of the subset selection on model results. Lastly, we created a validation data subset in the same way, which does not include the three calibration data subsets.

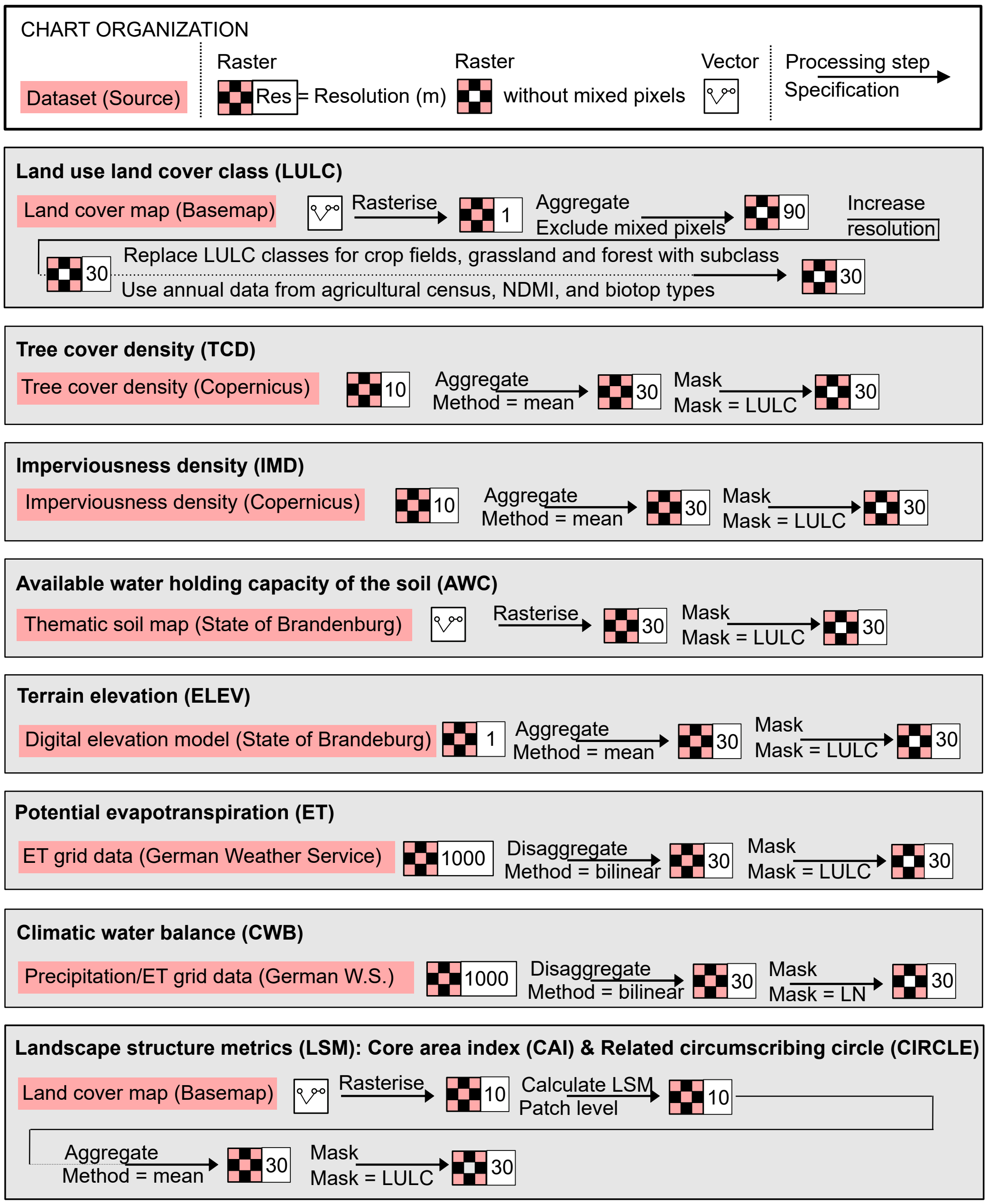


Figure 2: Processing of the predictor variables

2.3.4. Statistical modelling

Each of the 39 Landsat LST scenes was considered individually in modelling. This approach has, besides that it avoids prior data aggregation and the associated loss of information, a number of advantages. Thus, the annually varying crop type on arable land (Chpt. 2.3.2) can be included in the modelling as a time-varying predictor. This avoids the creation of arbitrary predictor variables such as the proportion of deep-rooted plants in crop rotation. The two climatic variables can also be included in the model in this way.

Given this data structure, we have a repeated measures design with censoring, since not all pixels in each scene have an LST value due to clouds. In general, the data set is highly unbalanced because the ranges or levels of the predictors are not evenly distributed. For example there are 400 times as many coniferous forest pixels as fen pixels.

To deal with these aspects, a mixed model was set up, as this type of models is suitable to analyse unbalanced and incomplete data (e.g. Hesselmann, 2018). We chose the Bayesian approach to mixed modelling due to its flexibility and intuitive interpretation of results. As a result of the model-building process (App. Appendix F), we have adapted the following model:

$$LST_{ijk} \sim T\left(\mu_{ijk}, \sigma, \nu\right) \tag{1}$$

where,

$$
\begin{aligned}
\mu_{ijk} &= \beta_{0,LULC[i]} + \alpha_{ID_j,LULC[i]} + \alpha_{SCENE_k,LULC[i]} + \beta_1 TCD_i + \beta_2 IMD_i \\
&+ \beta_3 ELEV_i + \beta_4 ETo_i + \beta_5 CWB_i + \beta_6 CAI_i + \beta_7 CIRCLE_i \\
&+ \beta_8 (CAI_i \times LULC_i) + \beta_9 (CWB_i \times LULC_i) + b \sum_{i=1}^{AWC_n} \zeta_i, \\
&\text{for } ID \text{ j} = 1, \ldots, \text{J and for } SCENE \text{ k} = 1, \ldots, \text{K},
\end{aligned} \tag{2}
$$

with,

$$
\zeta_i \in [0, 1] \tag{3}
$$

Eq. 1 states that the i^{th} value of LST follows a Student distribution with mean μ, scale σ and shape ν. The subscripts j and k refer to the two grouping variables in the model (see below). We chose a Student distribution because of outliers in the residual distribution when using a Gaussian distribution.

The model has 13 population-level intercepts β_0, one for each level of $LULC$ (Eq. 2). We chose this index-variable approach (McElreath, 2020) because it avoids dummy coding of the categorical variable $LULC$. Instead, each of the 13 levels of this categorical variable is given its own intercept.

The population-level intercepts are allowed to vary both by pixel id (ID) and by scene ($SCENE$). Hence, ID and $SCENE$ are the grouping ("random") variables, accounted for in the model by the terms $\alpha_{ID_k,LULC[i]}$ and $\alpha_{SCENE_j,LULC[i]}$. They take into account the repeated measures design and the fact that each scene is subject to certain boundary conditions (e.g.

weather, season) that are not fully reflected by our population-level ("fixed") predictors. The grouping variables were considered as crossed as each pixel occurs in every scene (except missing values due to clouds).

The model has 45 population-level effects, i.e. all terms in Eq. 2 except the α terms. An interaction term with $LULC$ is included for both CAI and CWB. We hypothesised that the influence of CAI on LST would vary by LULC class; for example, large settlements might be expected to be hotter than small settlements, or large forests cooler than small forest patches. Regarding prewetness, we assumed that the effect of CWB on LST may also depend on land cover or land use, e.g., due to the different rooting depths of arable crops and forest trees.

The predictor AWC is ordinal with nine levels (Tab. 2). To parameterise its regression coefficient, we used a technique called monotonic effects (Bürkner and Charpentier, 2020). The parameterisation is done in terms of a scale parameter b which represents the direction and size of the effect of AWC on LST like a "normal" regression coefficient. The second parameter, ζ, is a simplex (Eq. 3). It describes the expected difference between the levels i and $i-1$ in the form of a proportion of the overall difference b (Bürkner and Charpentier, 2020). The advantage of this approach over coding AWC as a nominal predictor is that the unknown differences between the nine levels of AWC are estimated by the model.

The priors for the group-level effects, the family-specific parameters σ and ν and for the simplex were taken from the default in the library brms. That is a student_t prior (3, 0, 2.5) for the standard deviation of the group-level effects, a gamma(2, 0.1) prior for ν and a dirichlet(1) prior for the simplex.

For the population-level effects, we used weakly informative normal (0,1) priors because it is very unlikely that these effects lie outside the 0 ± 2 range.

2.3.5. Model prediction for evaluation of measures

We used the fitted model to predict the cooling effect of the measures during the vegetation period. For this purpose, the predictors were assigned values that characterise the state of an area before and after (hypothetical) implementation of the measures. We based this on reasonable assumptions or observed conditions from our study area.

To parameterise the measures in the agriculture category, we defined five-year crop rotations that maximise profits in our study area. For fields with average and above average quality (German soil quality index $\geq$ 25), the crop rotation 1 (CR1) consists of winter cereal – winter oil-seed rape – winter cereal – maize – legume. For fields with low soil quality (German soil quality index $<$ 25), the crop rotation 2 (CR2) was defined as winter cereal – maize – winter cereal – maize – legume. The measure “deep-rooting crops” received specific crop rotations (SCR). SCR1 consists of winter oil-seed rape – winter cereal – sunflower – alfalfa – alfalfa (German soil quality index $\geq$ 25) and SCR2 was composed of winter oil-seed rape – winter cereal – maize – alfalfa – alfalfa (German soil quality index $<$ 25).

For measures „afforestation of marginal arable land” and „afforestation of urban brownfields”, the proposed composition of the new forest corresponds to a balanced mixture of sessile oak, red beech and hornbeam.

Both forest measures replace coniferous forest with mixed or broad-leaved forest. For the parameterisation, we followed the common forestry practice in our study area, which makes silvicultural decisions based on nutrient and moisture supply. Depending on the moisture and nutrient conditions, several types of deciduous or mixed forests can often be considered for a site. Furthermore, different mixing ratios are possible in mixed stands. Based on the Stand Target Type Ordinance (Luthardt, 2006), a type and a mixing ratio is defined for each site (App. Appendix E).

Measure "organic fertilisation" was parameterised by increasing the soil water holding capacity. If we assume a 3% mass increase of soil organic carbon over 30 years of organic fertilisation, which is already at the upper end of the achievable effects in Germany, the resulting increase in AWC is limited to about 3% volumetric water content (Minasny and McBratney, 2017). Translated to the AWC classes of the soil map used (Tab. 2), this means an increase by a maximum of two classes, for example from class low/very low to class low/medium.

The effectiveness of the measures is subject to both temporal and spatial variation. The first is due to differences in meteorological conditions and seasonal effects. Spatial variation arises partly because predictor values vary spatially among the nearly 60,000 plots (Chpt. 2.2). Second, not all measures affect the entire area of a plot (Fig. 3).

To deal with temporal variation, we can predict LST for each of the 39 scenes and observe the spread of the predictions. Thus, because $SCENE$ was a predictor in our model, scene-specific influences of environmental conditions not captured in our population-level predictors can also be mapped to LST. These include, for example, seasonal differences in the degree of cover of arable land or in the foliage of deciduous trees. The predictors ET and CWB were set to scene-specific values. Their spatial differences were not

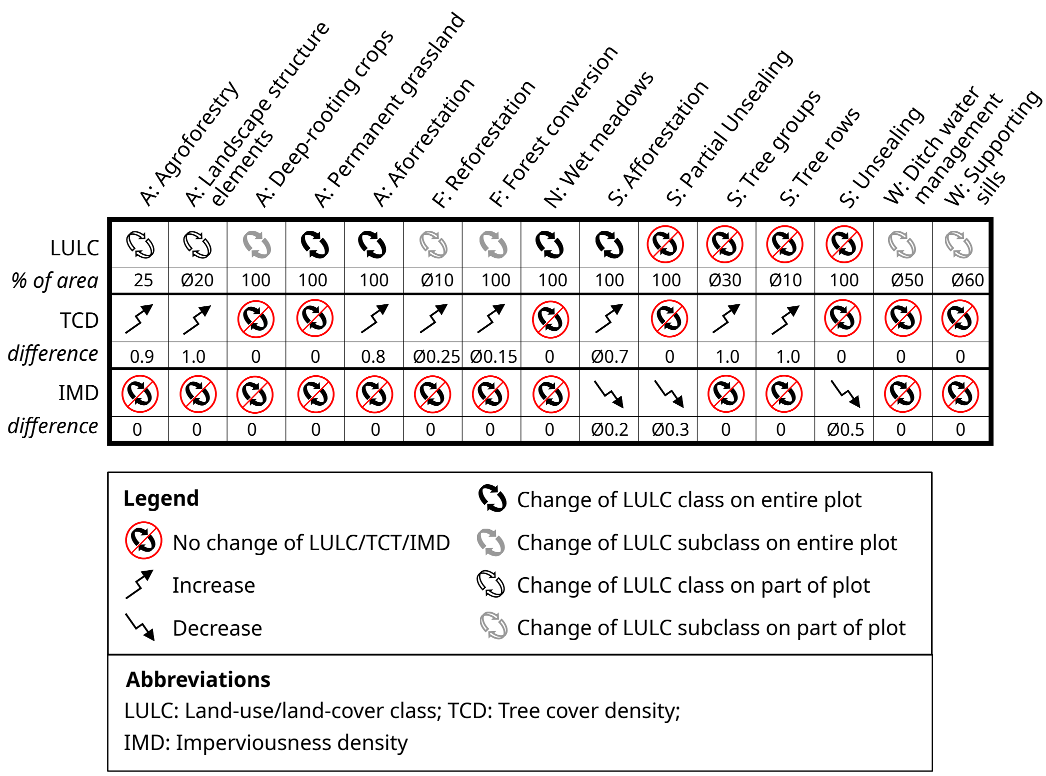


Figure 3: Change in predictor values before and after hypothetical measure implementation. The landscape structure metrics CAI and $CIRCLE$ were also specified on the basis of their current values and the values that can be expected after implementation. Measure “organic fertilization” does not fit into the scheme; its parameterisation is described in chapter 2.3.5.

considered, i.e., we passed spatial averages into the predictions. All other predictors were assigned to fixed (static or mean) values (Fig. 3, second and third row). Since we predict each scene in the temporal approach, backscaling to actual LST values is possible. Accordingly, the differences in LST before and after hypothetical measure implementation can be expressed in Kelvin.

We also assessed differences in measure effects between individual plots, i.e., spatial variation. We therefore determined the values of the spatial predictors for each plot, where the respective measure could be implemented (App. Appendix C), both for the status quo and the expected state after implementation of the measures. These values were then transferred to the model prediction. This time, the values of the two climatic variables were fixed at their temporal averages for each grid cell. Both group level variables were not included.

In the final step, the predicted grid cell values were transferred to the plots (spatial vectors) by calculating (area-weighted) averages. For measures not covering the entire plot, weights were set according to the conditions of the specific plot (App. Appendix C, spatial variation approach) or to averages for our study area (Fig. 3, temporal variation approach). For measure "agroforestry", we assumed that 25% of the cropland is planted with woody vegetation (Fig. 3).

2.3.6. Software

All calculations were performed with R (R Core Team, 2021). The GIS algorithms required to localise the measures were programmed in GRASS GIS (GRASS Development Team, 2022) with an R interface enabled by the R library rgrass7 (Bivand et al., 2018). The LST data were processed also using

GRASS GIS with R. For geospatial data processing, we used the libraries rgdal (Bivand et al., 2023) and terra (Hijmans, 2023). Landscape structure metrics were calculated with the library landscapemetrics (Hesselbarth et al., 2019). For model building and prediction, we used the library brms (Bürkner, 2021). Plotting was done with the library ggplot2 (Wickham, 2016).

3. Results

3.1. Model performance

The model was run using eight chains, each with 2000 iterations, of which 1000 were used for warm-up. We fitted nearly 100 models using iterative model building (Fig. 1, Gelman et al., 2020). Most earlier candidate models had convergence problems ($\hat{R} > 2$, divergent transitions). Model modification included the switch from a Gaussian to a Student response distribution (with identity link), the inclusion of several interaction terms, the switch from correlated to uncorrelated group-level effects, the choice between the index-variable approach and dummy coding of $LULC$ and the use of several sub-data sets from earlier post-stratification routines. The model used for post-processing and prediction was chosen by posterior predictive checking and model comparison among valid candidate models.

Convergence diagnostics for the chosen model indicate convergence with $\hat{R}$ values smaller than 1.01 and effective sample sizes (much) greater than 400 (Tab. F.6, F.7, F.8). Only for two variables, sd(LULC class 10) (F.6) and ET (F.7), the $\hat{R}$ was 1.01. Since this deviation from the target $\hat{R}$ value is very small, the effective sample size is large enough and the trace plots do not show any peculiarities, we do not consider this to be overly problematic.

Posterior predictive checks using the validation data also reveal good model fit: the model generated data resemble the observed validation data (Fig. F.7). However, due to the predictive uncertainty, the scatter in the model simulations is considerable (the y_{rep} in Fig. F.7). Therefore, we performed the predictions with at least 100 draws to reduce uncertainty in the estimated means.

The goodness of model fit can further be determined by the Bayesian version of the R^2, which is defined as the variance of the predicted values divided by the variance of predicted values plus the expected variance of the errors (Gelman et al., 2019). The Bayes R^2 value of 0.74 for the selected model confirms a good fit, but a quarter of the variance remains unexplained.

Lastly, the use of the two alternative calibration subsets (Chpt. 2.3.4) became noticeable only in the second decimal place of the estimated coefficients. Hence, the sampling approach we chose to reduce the overall data set has resulted in consistent results.

3.2. LST as a function of the environmental conditions

The desired evaluation of adaptation measure performance is related to the influence of each predictor on the target variable LST. To estimate variable importance, we primarily use conditional effect plots, which visualise the change in the target variable for given changes in predictor values. In interpreting the results, we consider not only the effect size but also the estimation uncertainty, which is described by the confidence intervals for the regression coefficients and the width of the posterior predictive distribution.

All population-level predictors affect LST, as indicated by confidence intervals of the estimated coefficients that do not include 0 (Tab. F.7). An important variable is the potential evapotranspiration, which is positively correlated with LST (Fig. 4). The tree cover density has a similarly large effect on LST, but the correlation is negative: with larger tree cover, LST decreases (Fig. 4). The other population-level predictors affect LST much less (Fig. 4). The direction of their effects is negative, i.e. an increase in the predictor value leads to a decrease in LST except for the imperviousness density, which increases LST as expected.

Regarding the effect of the land-use/land-cover class on LST, the ranking is as follows: Settlements and arable land appear warmer than forests and fens, while grassland areas are in between. Dry grassland was predicted to have a higher LST than fresh grassland and broad-leaved forest were the coolest class among forest types. For arable crops, winter crops as well as alfalfa were simulated slightly cooler than the summer crops. At the same time, the simulations come with wide confidence intervals. For example, the upper part of the posterior distribution of forests overlaps with the lower part of the posterior distributions of the warmer classes.

The relationship between landscape structure and LST was considered including the core area index (CAI) and the related circumscribing circle ($CIRCLE$) into the model. CAI describes area and shape of a landscape patch simultaneously: large and compact patches have larger core area. The parameter $CIRCLE$ characterises the compactness of a patch. The LST response to CAI differs among the land-use classes (Tab. F.8, Fig. 5). While in settlements and agricultural areas a larger core area coincides with an increase of the predicted LST, the opposite is true for forests and fens (Fig. 5). Particularly pronounced influences of CAI on LST are predicted for

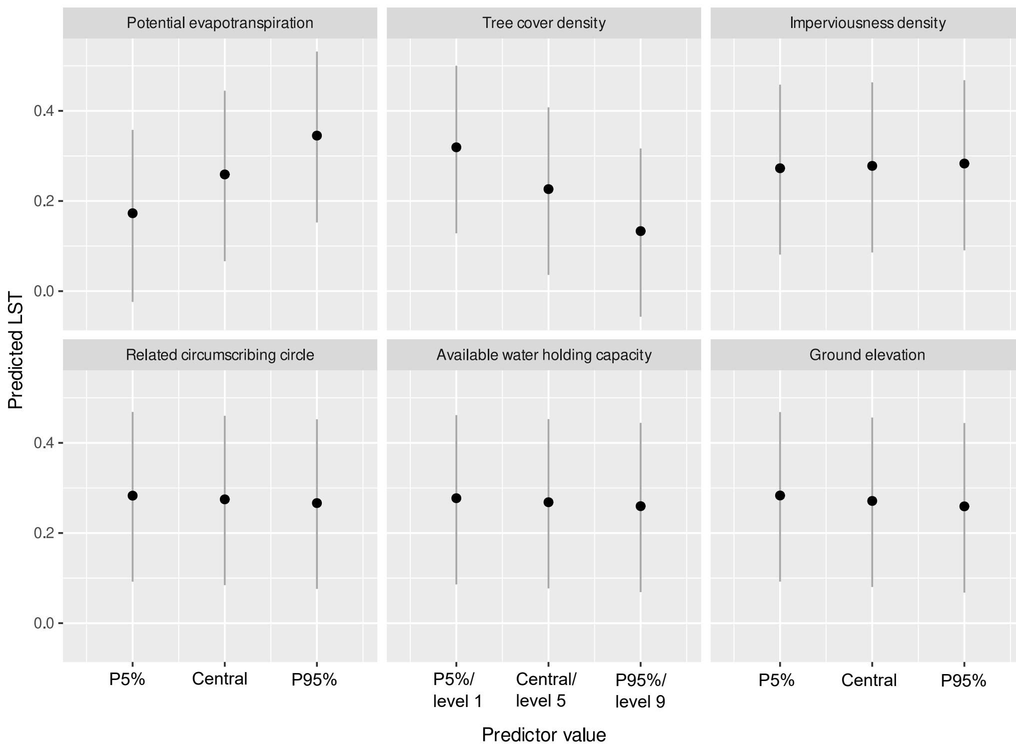


Figure 4: Change of the predicted scaled LST where the predictors shown are varied between their 5th and 95th percentile. Central is mid-range, which we used as a measure of central tendency. All predictors except the one shown were fixed at their mean. The predicted values are drawn from the posterior distribution, taking into account scene- and ID-specific deviations in the intercepts. The bars indicate the confidence intervals of the predicted LST.

the land-use classes fens and settlements. For legumes, winter grain and dry grassland, the uncertainty of the estimation does not allow a clear assessment (Tab. F.8). The influence of $CIRCLE$ on the LST is much smaller than that of CAI (Fig. 4).

We used the climatic water balance of the month before a satellite overflight (CWB), as a simple drought indicator. We expected that the LULC class modulates the LST response to this variable (Chpt. 2.3.2). In fact, LST in settlements and fresh grassland does not seem to respond to drought situations, while for all other LULC classes, LST increases with lower water availability, Fig. 5). It is noteworthy that especially the arable crops react very differently to this factor, with legumes and sunflowers being particularly sensitive (Fig. 5).

Considering all population-level predictors together and given the data and model at hand, a ranking of predictor importance can be established (Tab. 4). This ranking is confirmed by valid earlier model candidates. For example, ET and TCD were important in all of them.

The model also includes the group-level predictors $SCENE$ and ID. They account for the fact that some of the variance can be attributed to the temporal and spatial groups in our data. The importance of these predictors is given by the intra-class correlation (ICC), which can be interpreted as the proportion of explained variance that is due to the group-level effects. Hence, if the ICC is close to 0, the grouping contains no information. If, on the other hand, it tends towards 1, all observations in a group are identical (Gelman and Hill, 2006). In our model, the ICC is 0.22 with a 95% confidence interval of [0.1 0.31] – the grouping definitely contains information. When we

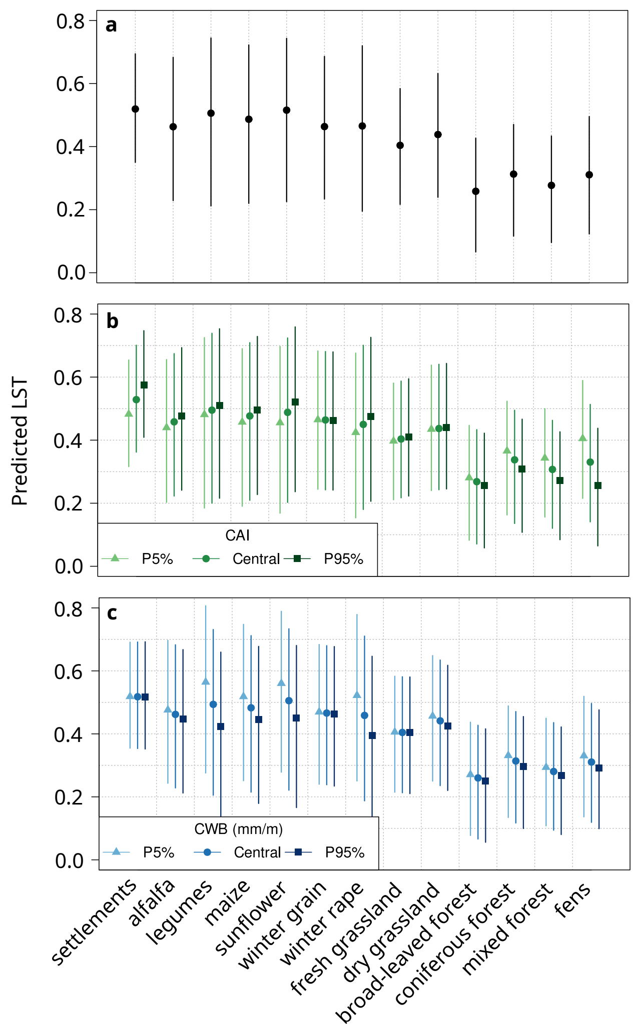


Figure 5: Conditional effects plots for $LULC$ (a), the interaction between LULC and CAI (b), and the interaction between LULC and CWB (c). Plot (a) shows the predicted scaled LST for each LULC class, while plots (b) and (c) illustrate the change of the predicted scaled LST where CAI and CWB, respectively, are varied between their 5th percentile, central, and 95th percentile. All other predictors were fixed to LULC-specific means. The predicted values are drawn from the posterior distribution, taking into account scene- and ID-specific deviations in the intercepts. The bars indicate the confidence intervals of the predicted LST.

Table 4: Classification of predictor importance based on the magnitude of the LST change (Figs. 5, 4) and on the credible intervals of the parameter estimates (App. Appendix F)

Predictor	Land-use/land-cover class	Importance		
		H[1]	M[2]	L[3]
Potential evapotranspiration		×		
Climatic water balance	Settlements			×
	Arable land: alfalfa			×
	Arable land: legumes	×		
	Arable land: maize		×	
	Arable land: sunflower	×		
	Arable land: winter cereals			×
	Arable land: winter rape	×		
	Grassland: fresh			×
	Grassland: dry			×
	Forest: broad-leaved			×
	Forest: coniferous		×	
	Forest: mixed			×
	Fens			×
Land-use/land-cover class		×		
Tree cover density		×		
Imperviousness density			×	
Core area index	Settlements	×		
	Arable land: alfalfa		×	
	Arable land: legumes			×
	Arable land: maize		×	
	Arable land: sunflower		×	
	Arable land: winter cereals			×
	Arable land: winter rape		×	
	Grassland: fresh		×	
	Grassland: dry			×
	Forest: broad-leaved		×	
	Forest: coniferous		×	
	Forest: mixed		×	
	Fens	×		
Related circumscribing circle			×	
Aval. water holding capacity			×	
Terrain elevation			×	

[1] H: High - scaled LST difference ≥ 0.1 and confidence intervals do not include 0

[2] M: Medium - scaled LST difference $[>0, <0.1]$ and confidence intervals do not include 0

[3] L: Low - scaled LST difference $\rightarrow 0$ and confidence intervals include 0

calculate the ICC separately for each group-level predictor, their contribution is similar: an ICC of 0.13 (CI 95% [0.03 0.21]) for ID and 0.11 (CI 95% [-0.02 0.22]) for $SCENE$. Note that the ICC is more uncertain for $SCENE$ than for ID as indicated by the confidence interval for the ICC containing 0. The importance of the two predictors is further supported by the confidence intervals for their standard deviation estimates (Tab. F.6).

3.3. Assessment of adaptation measures' performance

3.3.1. Temporal variation of cooling effects

We first consider the measure effect, taking into account its temporal variation. We therefore predicted the status quo and the new state after hypothetical measure implementation for each scene with its specific climatic conditions, but with spatial averages for all other predictors. The resulting ranking of measures is based on their average effect in time, ignoring spatial variation among plots (Fig. 6).

The three most effective measures, with average cooling capacities of 3 to 3.5 K, require a change in land use - from cropland or urban wasteland to forest or from grassland to fens (Fig. 3; note that the measure "wet meadows" is parameterised as fens). The remaining measures achieve an average cooling effect of up to 2 K (Fig. 6). They are associated with a reduction in surface sealing, an increase in soil water holding capacity, the installation of tree rows or groups and a change from dry to fresh grassland (Fig. 6).

Measure "reforestation of degraded forest patches" ranks in the lower end, which is most likely due to the low average portion of degraded forest patches within a forest stand (Fig. 3). The effect of organic fertilisation also appears

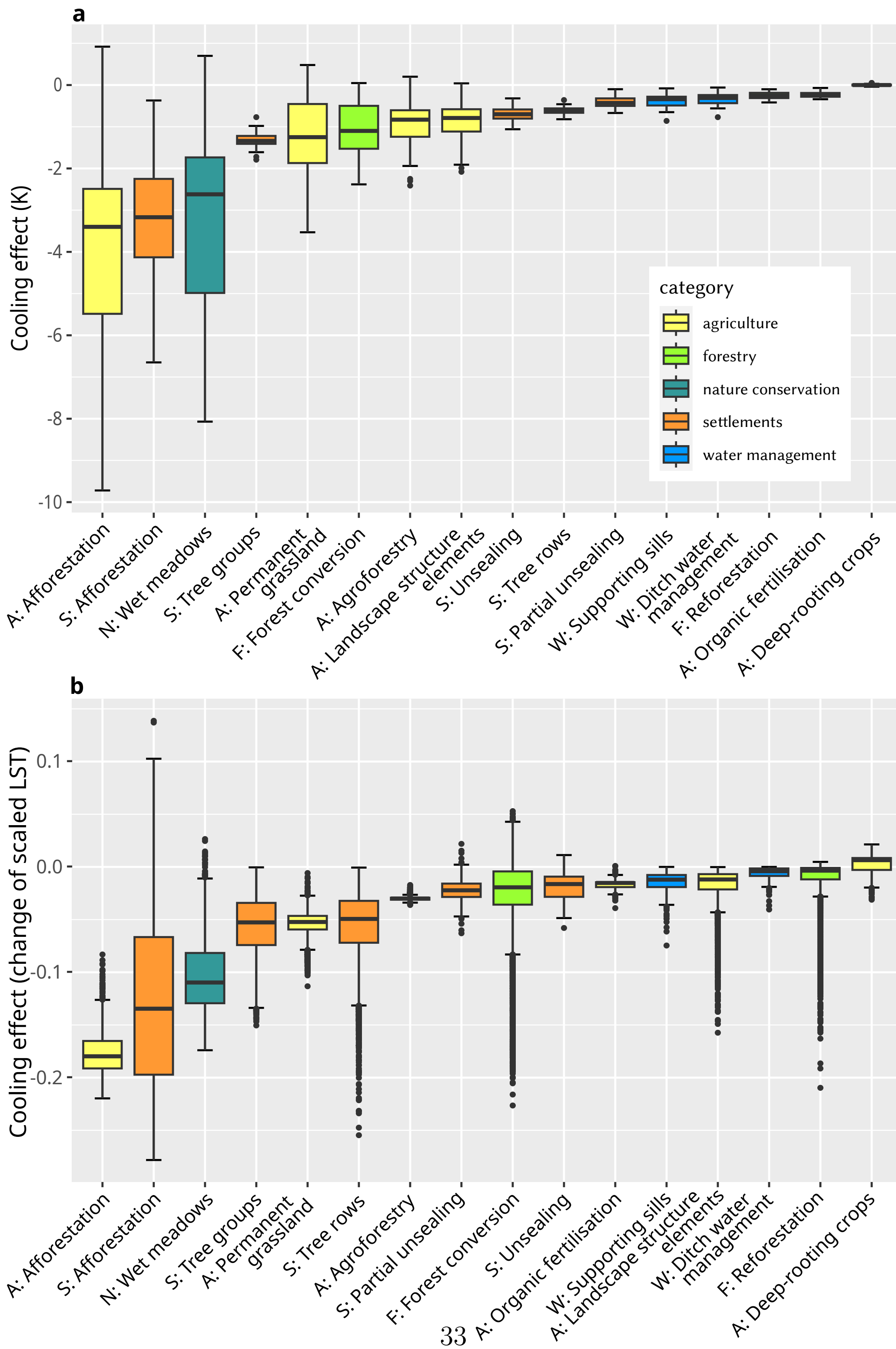


Figure 6: Predicted cooling effects of the adaptation measures, temporal variation (a) and spatial variation (b).

to be quite limited, since the associated increase in water holding capacity is limited (Chpt. 2.3.5) and, in addition, the influence of the predictor AWC on the LST is also rather moderate (Fig. 4). The measure "cultivation of deep-rooting arable crops" shows no effect at all, which can be explained by our choice of crops in the simulated crop rotations: both the conventional and the measure-specific crop rotations contain crops with higher LST (legumes in the conventional and sunflowers in the specific rotation, see Fig. 5).

3.3.2. Comparison with spatial variation of cooling effects

We also assessed differences in measure effects between individual plots, i.e., spatial variation. This time, we set the climatic parameters to their spatially varying temporal averages and implemented the other predictors into the predictions according to their plot-specific expressions before and after implementation of the measures (Chpt. 2.3.5).

The performance ranking for the three most efficient measures remains unchanged for the spatial approach compared to the temporal approach (Fig. 6). Some slight differences in ranking characterise the remaining measures. Particularly noteworthy are the numerous outliers in the direction of both higher and lower effectiveness, especially for the measures "afforestation of urban brownfields", "tree rows in settlements", "forest conversion", "landscape structure elements" and "reforestation of degraded forest patches". This wide variation in the impact of these measures is the result of the different sources of spatial variation (Tab. G.9). For instance, the range of area share of the measure landscape structure elements is extremely wide, between slightly above 0 and nearly 100% of a plot that could accommodate this measure.

The spatial approach to measure evaluation accounts for plot-specific differences in predictor values for mean climatic conditions. It is therefore most suitable for deciding where which measure should be implemented most appropriately. The time-varying prediction can then be used to incorporate possible seasonal and climatic effects into the evaluation.

When comparing the contributions of both sources of variation, the following picture emerges: for most measures, the contribution of spatial variation is larger than that of temporal variation (Tab. G.9). However, in the case of measures related to agriculture, the opposite is true, which is already indicated by the large importance of the group-level effect $SCENE$ (App. F.6) on the LST footprint of many crops. Seasonal differences in crop growth and cover could explain this observation.

4. Discussion

4.1. Drivers of land surface temperature

Our analysis revealed settlements as the warmest LULC class, warmer than forests or cropland. This is a widely supported finding, with corresponding examples from all around the world (e.g. Alavipanah et al., 2015; Das et al., 2020; Seeberg et al., 2022; Wickham et al., 2012). Here, the domination of surfaces from which only ponding water can evaporate leads to a significant turnover of radiation into sensible heat, with increasing temperature as a consequence. Using light colours on building walls and roofs is an ancient technique to increase the albedo and thus reduce the turn-over towards higher reflectance in regions that suffer from high temperatures during summer.

In vegetated areas, evapotranspiration governs the energy conversion at the land surface (Sheil, 2018). Also in our study, ET was identified as the dominating factor influencing LST. This relation finds a counterpart in the close relationship between solar radiation and LST found by Bertoldi et al. (2010), as solar radiation is one of the key factors affecting potential evapotranspiration (Priestley and Taylor, 1972). The ability of a land surface to evaporate or transpire water is a primary characteristic that governs LST. Consequently, sealing or unsealing of land has a significant influence on LST (e.g. Song et al., 2014; Seeberg et al., 2022). In our study, this is evident in the positive correlation between imperviousness density and LST. Similarly, research has shown that densely clustered vegetation patterns are more effective at reducing LST than scattered ones (Fan et al., 2015; Kim et al., 2016). We found that for most land-use classes, the core area index, which describes both patch size and shape, exerts a strong effect on LST. The additional inclusion of $CIRCLE$, which characterises patch compactness, in the modelling appears less important.

Using the group-level variables $SCENE$ and ID, we included additional error terms in the model. They account for the hierarchical nature of the data: each pixel belongs to a particular unit in space (e.g., all pixels in a field or forest stand) and to one of the 39 satellite scenes. Their proven significance suggests that the sometimes substantial differences in LST between LULC classes do not occur everywhere (predictor ID) or at all times (predictor $SCENE$). Using them in the model also helps to compensate for the fact that not all factors affecting LST could be included in the model. For example, the predictor "tree cover density" is actually variable over time, but was implemented as a static variable (Tab. 2). Since $SCENE$ was used to estimate a scene-specific intercept for each LULC class, this temporal variance can thus be accounted for to some extent. In general, the inclusion of the group-level predictors resulted in a more realistic representation of uncertainty in the model estimates of population-level effects. For instance, if we had calculated the confidence intervals in the conditional effects plots without taking group-level effects into account, the overlap among the LULC classes (Fig. 5) would not have emerged.

4.2. Adaptation measure efficiency

In many studies devoted to understanding LST, the focus is on mitigation of urban heat islands (e.g. Alavipanah et al., 2015; Seeberg et al., 2022; Selim et al., 2023). In our study, we transferred these approaches to rural areas to predict adaptation measure effectiveness. Seeberg et al. (2022)'s approach to evaluating the success of heat stress interventions and Selim et al. (2023)'s work on the cooling potential of green spaces took a similar approach in the urban context, albeit retro- rather than prospective.

The temperature regime of forests is known to be cooler during summer as compared to the open agricultural landscape (Müttrich, 1897; Ghafarian et al., 2022). Due to the large influence of tree cover density and LULC class on LST, those measures that significantly increase tree cover or result in a change in LULC class are ranked highest. Consequently, the effectiveness of measures "agroforestry" and "landscape structure elements" increases with a greater density of these elements. For example, if the agroforestry strips were placed every 24 m, this would result in an increase of the cooling effect to 1.1 K on average. Ghafarian et al. (2024) analysed small woody features in an agricultural landscape close to our study site and concluded that a significant

LST decrease was identified when trees or woody vegetation patches had a distance smaller than 75 m. Conversely, reducing tree cover densities – in our parameterisation, tree cover was set to the maximum possible values (Fig. 3) – would significantly reduce the cooling effect. This suggests that a closed canopy and a high density of the woody vegetation elements are particularly important for these measures to be effective.

In contrast, measures have a lower effectiveness if the importance of the predictors was only moderate, such as IMD and AWC (Tab. 4). Among those are measures "organic fertilisation" and "unsealing" with subsequent establishment of a grass cover (Fig. 6). Moreover, if the only change in predictor values is among subclasses of $LULC$ on parts of the plots (measures "ditch water management" and "supporting sills", Fig. 3), the cooling effect appears limited (Fig. 6). A somewhat different picture could emerge if these water management measures were located on arable land and could contribute to better water supply for crops, e.g. by controlled drainage systems that allow water retention when dry periods are expected, and drainage when high soil moisture levels hamper field traffic or crop growth. However, this assessment was not possible in this study, since the NDMI, which we used as proxy for groundwater influence, depends not only on the water supply, but also on the crop type. Our database was too limited to further elaborate these aspects.

The predicted cooling effects per measure (Fig. 6) cover a wide range, which furthermore depends on the season. For example, the cooling effect of measures that increase the proportion of trees is lower in May and/or in September than in the summer months (Fig. H.8). Reasons for this observation are the seasonality of the drivers for evapotranspiration and of the presence of leaves – the primary transpiring organs – on the deciduous trees and shrubs. Similar to the day/night conversion of the cooling effect (Peng et al., 2014), dense vegetation turns into a warmer environment in winter, due to heat storage in the tree mass, limited long-wave radiation, and ceased transpiration (Zellweger et al., 2019). The effect of improved water retention with the measures "ditch water management" and "supporting sills" seems to generally decrease towards September, which is again related to the decrease in evapotranspiration, but most likely also to the decrease in ditch- and groundwater levels as the summer progresses. Due to the limited number of available thermal scenes, these seasonal results should be interpreted with caution. However, they support that the cooling effect of a measure varies over time and that this time component varies for individual measures.

The plot-specific assessment can indicate where which measures should be prioritized in terms of benefits. If the aim is to achieve an even distribution of cooling effects over the course of the year, or if a cooling effect is particularly relevant in the summer months, for example, the seasonal differences in the measure performance would also be of interest for practical implementation. We leave it to further research to explore this approach in more detail.

Of course, other criteria need to be taken into account in addition to effectiveness, such as costs, synergies with nature conservation, or local constraints. In a cost-benefit analysis of water retention measures, which also covered eastern Germany, it was determined that measures in croplands pay off the most (Sušnik et al., 2022). The results we have presented focus on the benefit side, while the inclusion of costs represents the next step.

4.3. Reliability of statistical LST-based measure evaluation

The use of satellite data of LST has limitations. The Landsat 8 Collection 2 Level 2 thermal data are transferred into LST using emissivity data by the ASTER Global Emissivity data set (ASTER GED). The ASTER GED, in turn, is a product that gives a long-term emissivity per pixel over the period 2000 to 2008, which was derived from vegetation cover, among other things. If there was a land cover change between this period and 2013, when the first LST data sets became available, then incorrect emissivities were used. This may then cause erroneous LST data. Masking this data beforehand, e.g. by identifying changes in the NDVI, could remove this inaccuracy but was beyond the scope of our analysis. Another limitation of the satellite data is their fixed return interval, which is at 10 AM in our study area. This overflight time does not correspond to the time of day when evaporation normally reaches its maximum. Therefore, the reported LST differences in response to the measures are strictly valid only for this time window, while they will be lower or higher at other times.

The success and reliability of the statistical modelling heavily depend on the decisions made throughout the process from data acquisition, model development, to model prediction. Among those are the choice of the population- and group-level predictors, interaction terms, response families, priors etc. It is particularly important to emphasise that a whole range of options is already available when selecting the model tool. For example, a machine learning (ML) approach could have been chosen as an alternative way for predicting adaptation measure performance. We decided against ML because we put emphasis in model interpretability and uncertainty understanding (Gross et al., 2020). However, it is quite possible that the predictive performance can be improved by choosing ML, because e.g. it omits the specification of interaction terms. It would certainly be interesting to compare alternative approaches in a separate study.

Another point regarding the reliability of the results addresses the parameterisation of model predictions. Due to the inevitable simplification of reality, parameterisation comes with numerous assumptions and limitations. For instance, existing hedgerows and trees on arable land are not included in our land cover database (App. Appendix B). Consequently, the LST for the status quo of arable land might be overestimated in the presence of linearly shaped woody vegetation. Furthermore, we assigned a core area index of 0 for measures “agroforestry” and “landscape structure elements”, which represents an extrapolation since the model data set does not contain such low CAI values (Tab. 2). The resulting potential error is 0.4 K at maximum, as indicated by predictions for these measures using the lowest CAI in the model data set. Given these and other uncertainties, we have avoided making sharp distinctions in interpreting the measures’ ranking.

To conclude, the presented model is not a final one and can and should be updated when more information becomes available. This relates both to the satellite scenes as well as to the geospatial data that serve for predictor derivation. The limited importance of the water-holding capacity of the soil on LST might also be related to the coarse map resolution and classification of originally numerical values in the thematic soil map, which was used to derive AWC (Fig. 2). Still missing information such as seasonality of tree cover density and more finely resolved data on soil water retention could be incorporated in future model candidates.

5. Conclusion

Our analyses confirmed that the process of evapotranspiration by vegetation and enhanced water retention are major contributors to local land surface cooling. The quality of the model output was shown to be suitable for the desired ex-ante assessment of the impact of climate adaptation measures. With the included predictors and the fitted model, a large part of the variance in land surface temperature could be explained. The most important predictors were the potential evapotranspiration, the tree cover density, and the land-use/land-cover (LULC) class. Both landscape structure and drought severity (predictor CWB) influenced LST but only for some LULC classes. The other predictors are of moderate to minor importance in our rural study region.

By implementing nature-based climate adaptation measures, evapotranspiration and water retention can be influenced if the measures are designed in such a way that they strengthen these processes. This includes, firstly, the introduction of plants into land use systems that have a high evapotranspiration capacity. These include woody plants in particular. Since evapotranspiration is an area-based process, the amount of area that can be provided with woody plants is crucial for their cooling effect. Adequate water supply is also critical to evapotranspiration. If water retention in the landscape can be improved, for example by wetland restoration, largely unrestricted evaporation is ensured even in summer, which also results in high cooling performance. Some of the measures we have considered show only a minor effect, which is mainly due to only a gradual change of a plot's condition (e.g. change of a LULC subclass compared to a LULC class), a low area coverage of the measure, or the subordinate importance of the predictors that can be influenced by the measure.

With our approach, climate change adaptation measures can be evaluated before their implementation and in a spatially differentiated manner. By including multiple satellite scenes in the analysis, it is also possible to take into account the temporal dynamics of the cooling effects within the course of the year. The parallel evaluation of different measures is a valuable tool for decision support on climate adaptation. The results can be applied to the different land use sectors to select the most promising measures for a given area. In addition, they can be used in spatial planning and in the management of subsidy allocations.

CRediT authorship contribution statement

Beate Zimmermann: Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization, Funding acquisition. **Sarah Kruber**: Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing. **Claas Nendel**: Writing - Review & Editing. **Henry Munack**: Conceptualization, Methodology, Writing - Review & Editing, Visualization, Funding acquisition. **Christian Hildmann**: Conceptualization, Writing - Review & Editing, Supervision, Visualization, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Satellite images of Landsat 8 the Copernicus data are available for free use (https://earthexplorer.usgs.gov/, https://land.copernicus.eu/en). Weather data from the German Weather Service are available via https://www.dwd.de/EN/ and also without use restriction. Data for the local allocation of measures is made freely available by the Federal State of Brandenburg. In addition, freely available data of the tracks from Deutsche Bahn and roads and paths from OpenStreetMap have been used. Data with use restriction are the land cover model of Germany, locations of ditches and farm dams, the mean groundwater level map, building outlines and forest site characteristics map of the federal state of Brandenburg.

Acknowledgments

We would like to thank our project partners from the Elbe-Elster county for their cooperation and the provision of data through their GIS system. We also thank Dr. Dirk Knoche, Research Institute for Post-Mining Landscapes, for his support in parameterising the forestry measures.

Funding

This work was supported by the German Federal Ministry of Education and Research (BMBF, Grant number 01LR2004B).

References

Abir, F.A., Ahmmed, S., Sarker, S.H., Fahim, A.U., 2021. Thermal and ecological assessment based on land surface temperature and quantifying multivariate controlling factors in bogura, bangladesh. Heliyon 7, e08012. doi:https://doi.org/10.1016/j.heliyon.2021.e08012.

Alavipanah, S., Wegmann, M., Qureshi, S., Weng, Q., Koellner, T., 2015. The role of vegetation in mitigating urban land surface temperatures: A case study of Munich, Germany during the warm season. Sustainability 7, 4689–4706. doi:https://doi.org/10.3390/su7044689.

Amt für Statistik Berlin-Brandenburg, 2022. Statistischer Bericht, Flächenerhebung nach Art der tatsächlichen Nutzung im Land Brandenburg 2021. URL: https://download.statistik-berlin-brandenburg.de/4bb421915ed644bc/1b39ca17cfc6/SB_A05-03-00_2021j01_BB.pdf. assessed 24 July 2023.

Amt für Statistik Berlin-Brandenburg, 2023. Statistischer Bericht, Bevölkerungsentwicklung und Bevölkerungsstand im Land Brandenburg Dezember 2022. URL: https://download.statistik-berlin-brandenburg.de/8ee0bad9b1168256/a3df42d855eb/SB_A01-07-00_2022m12_BB.pdf. assessed 24 July 2023.

Baessler, C., Klotz, S., 2006. Effects of changes in agricultural land-use on landscape structure and arable weed vegetation over the last 50 years. Agric. Ecosyst. Environ. 115, 43–50. doi:https://doi.org/10.1016/j.agee.2005.12.007.

Bednar-Friedl, B., Biesbroek, R., Schmidt, D.N., Alexander, P., Børsheim, K.Y., Carnicer, J., Georgopoulou, E., Haasnoot, M., Le Cozannet, G., Lionello, P., Lipka, O., Möllmann, C., Muccione, V., Mustonen, T., Piepenburg, D., Whitmarsh, L., 2023. Europe, in: Climate Change 2022 – Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, pp. 1817–1928. doi:https://doi.org/10.1017/9781009325844.015.

Bertoldi, G., Notarnicola, C., Leitinger, G., Endrizzi, S., Zebisch, M., Chiesa, S.D., Tappeiner, U., 2010. Topographical and ecohydrological controls on land surface temperature in an alpine catchment. Ecohydrology 3, 189–204. doi:https://doi.org/10.1002/eco.129.

Bivand, R., Keitt, T., Rowlingson, B., 2023. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. Http://rgdal.r-forge.r-project.org, https://gdal.org, https://proj.org, https://r-forge.r-project.org/projects/rgdal/.

Bivand, R., Krug, R., Neteler, M., Jeworutzki, S., 2018. rgrass7: Interface Between GRASS 7 Geographical Information System and R. URL: https://CRAN.R-project.org/package=rgrass7. r package version 0.1-12.

Burakowski, E., Tawfik, A., Ouimette, A., Lepine, L., Novick, K., Ollinger, S., Zarzycki, C., Bonan, G., 2018. The role of surface roughness, albedo, and bowen ratio on ecosystem energy balance in the eastern united states. Agric. For. Meteorol. 249, 367–376. doi:https://doi.org/10.1016/j.agrformet.2017.11.030.

Bürkner, P.C., 2021. Bayesian item response modeling in R with brms and Stan. J. Stat. Softw. 100, 1–54. doi:https://doi.org/10.18637/jss.v100.i05.

Bürkner, P.C., Charpentier, E., 2020. Modelling monotonic effects of ordinal predictors in Bayesian regression models. Br. J. Math. Stat. Psychol. 73, 420–451. doi:https://doi.org/10.1111/bmsp.12195.

Copernicus, 2022. Copernicus: Globally, the seven hottest years on record were the last seven; carbon dioxide and methane concentrations continue to rise. URL: https://climate.copernicus.eu/copernicus-globally-seven-hottest-years-record-were-last-seven. assessed 18 September 2023.

Das, D.N., Chakraborti, S., Saha, G., Banerjee, A., Singh, D., 2020. Analysing the dynamic relationship of land surface temperature and landuse pattern: A city level analysis of two climatic regions in India. City Environ. Interact. 8, 100046. doi:https://doi.org/10.1016/j.cacint.2020.100046.

Du, S., Xiong, Z., Wang, Y.C., Guo, L., 2016. Quantifying the multilevel effects of landscape composition and configuration on land surface temperature. Remote Sensing of Environment 178, 84–92. doi:https://doi.org/10.1016/j.rse.2016.02.063.

Fan, C., Myint, S., Zheng, B., 2015. Measuring the spatial arrangement of urban vegetation and its impacts on seasonal surface temperature. Prog. Phys. Geogr.: Earth Environ. 39, 199–219. doi:https://doi.org/10.1177/0309133314567583.

Gelman, A., Goodrich, B., Gabry, J., Vehtari, A., 2019. R-squared for bayesian regression models. Am. Stat. 73, 307–309. doi:https://doi.org/10.1080/00031305.2018.1549100.

Gelman, A., Hill, J., 2006. Data Analysis Using Regression and Multilevel/Hierarchical Models. Analytical Methods for Social Research, Cambridge University Press. doi:https://doi.org/10.1017/CBO9780511790942.

Gelman, A., Vehtari, A., Simpson, D., Margossian, C.C., Carpenter, B., Yao, Y., Kennedy, L., Gabry, J., Bürkner, P.C., Modrák, M., 2020. Bayesian workflow. doi:https://doi.org/10.48550/arXiv.2011.01808. preprint at 10.48550/arXiv.2011.01808.

Ghafarian, F., Ghazaryan, G., Wieland, R., Nendel, C., 2024. The impact of small woody features on the mesoclimate in an agricultural landscape. Under review.

Ghafarian, F., Wieland, R., Lüttschwager, D., Nendel, C., 2022. Application of extreme gradient boosting and Shapley Additive explanations to predict temperature regimes inside forests from standard open-field meteorological data. Environ. Model. Softw. 156, 105466. doi:https://doi.org/10.1016/j.envsoft.2022.105466.

GRASS Development Team, 2022. Geographic Resources Analysis Support System (GRASS GIS) Software, Version 8.2. Open Source Geospatial Foundation. URL: https://grass.osgeo.org.

Gross, D.P., Steenstra, I.A., Harrell, F.E., Bellinger, C., Zaïane, O., 2020. Machine learning for work disability prevention: Introduction to the special series. J. Occup. Rehabil. 30, 303–307. doi:https://doi.org/10.1007/s10926-020-09910-1.

Hesselbarth, M.H.K., Sciaini, M., With, K.A., Wiegand, K., Nowosad, J., 2019. landscapemetrics: an open-source R tool to calculate landscape metrics. Ecography 42, 1648–1657. doi:https://doi.org/10.1111/ecog.04617.

Hesselmann, G., 2018. Applying linear mixed effects models (LMMs) in within-participant designs with subjective trial-based assessments of awareness—a caveat. Front. Psychol. 9. doi:https://doi.org/10.3389/fpsyg.2018.00788.

Hesslerová, P., Huryna, H., Pokorný, J., Procházka, J., 2018. The effect of forest disturbance on landscape temperature. Ecol. Eng. 120, 345–354. doi:https://doi.org/10.1016/j.ecoleng.2018.06.011.

Hesslerová, P., Pokorný, J., Brom, J., 2013. Daily dynamics of radiation surface temperature of different land cover types in a temperate cultural landscape: Consequences for the local climate. Ecol. Eng. 54, 145–154. doi:https://doi.org/10.1016/j.ecoleng.2013.01.036.

Hijmans, R.J., 2023. terra: Spatial Data Analysis. URL: https://CRAN.R-project.org/package=terra. r package version 1.7-39.

Hildmann, C., Zimmermann, B., Schlepphorst, R., Rösel, L., Kleinschmidt, F., Kruber, S., Lukas, S., Joshi, D.C., Hecker, L.P., Witt, J.C., Sturm, A., Wätzold, F., 2022. Measures for adaptation to climate change through water retention and cooling by transpiration: A catalogue of measures for a drought-prone area in eastern Germany. doi:https://doi.org/10.5281/zenodo.6811079. preprint at 10.5281/zenodo.6811079.

Ionita, M., Nagavciuc, V., Kumar, R., Rakovec, O., 2020. On the curious case of the recent decade, mid-spring precipitation deficit in central Europe. NPJ Clim. Atmos. Sci. 3, 49. doi:https://doi.org/10.1038/s41612-020-00153-8.

Kim, J.H., Gu, D., Sohn, W., Kil, S.H., Kim, H., Lee, D.K., 2016. Neighborhood landscape spatial patterns and land surface temperature: An empirical study on single-family residential areas in Austin, Texas. Int. J. Environ. Res. Public Health 13, 880. doi:https://doi.org/10.3390/ijerph13090880.

Landesamt für Umwelt Brandenburg, 2009. Flächendeckende Biotop- und Landnutzungskartierung (btln) im Land Brandenburg - CIR-Biotoptypen 2009. URL: https://lfu.brandenburg.

de/lfu/de/aufgaben/natur/biotopschutz/biotopkartierung/biotop-und-landnutzungskartierung/. assessed 21 September 2023.

Luthardt, M., 2006. Bestandeszieltypen für die Wälder des Landes Brandenburg. Ministerium für Ländliche Entwicklung, Umwelt und Verbraucherschutz des Landes Brandenburg, Abteilung Forst und Naturschutz. Potsdam. URL: https://mlul.brandenburg.de/cms/media.php/lbm1.a.4595.de/bzt_brdb.pdf.

Makarieva, A.M., Nefiodov, A.V., Nobre, A.D., Baudena, M., Bardi, U., Sheil, D., Saleska, S.R., Molina, R.D., Rammig, A., 2023. The role of ecosystem transpiration in creating alternate moisture regimes by influencing atmospheric moisture convergence. Glob. Chang. Biol. 29, 2536–2556. doi:https://doi.org/10.1111/gcb.16644.

Makarieva, A.M., Nefiodov, A.V., Nobre, A.D., Sheil, D., Nobre, P., Pokorný, J., Hesslerová, P., Li, B.L., 2022. Vegetation impact on atmospheric moisture transport under increasing land-ocean temperature contrasts. Heliyon 8, e11173. doi:https://doi.org/10.1016/j.heliyon.2022.e11173.

McElreath, R., 2020. Statistical Rethinking. Chapman and Hall/CRC. doi:https://doi.org/10.1201/9780429029608.

Minasny, B., McBratney, A.B., 2017. Limited effect of organic matter on soil available water capacity. Eur. J. Soil Sci. 69, 39–47. doi:https://doi.org/10.1111/ejss.12475.

Monteiro, M.V., Doick, K.J., Handley, P., Peace, A., 2016. The impact of greenspace size on the extent of local nocturnal air temperature cooling in

London. Urban For. Urban Green. 16, 160–169. doi:https://doi.org/10.1016/j.ufug.2016.02.008.

Müttrich, A., 1897. Beobachtungen. Springer, Berlin.

Peng, S.S., Piao, S., Zeng, Z., Ciais, P., Zhou, L., Li, L.Z., Myneni, R.B., Yin, Y., Zeng, H., 2014. Afforestation in China cools local land surface temperature. Proc. Natl. Acad. Sci. U.S.A. 111, 2915–2919. doi:https://doi.org/10.1073/pnas.1315126111.

Pfeifer, S., Bathiany, S., Rechid, D., 2021. Klimaausblick Elbe-Elster. Technical Report. Climate Service Center Germany (GERICS). URL: https://www.gerics.de/klimaausblick-landkreise.

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100, 81–92.

Procházka, J., Pokorný, J., Vácha, A., Novotná, K., Kobesová, M., 2019. Land cover effect on water discharge, matter losses and surface temperature: Results of 20 years monitoring in the Šumava mts. Ecol. Eng. 127, 220–234. doi:https://doi.org/10.1016/j.ecoleng.2018.11.030.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E.F., Marx, A., 2018. Anthropogenic warming exacerbates European soil moisture droughts. Nat. Clim. Chang. 8, 421–426. doi:https://doi.org/10.1038/s41558-018-0138-5.

Schils, R.L., Bufe, C., Rhymer, C.M., Francksen, R.M., Klaus, V.H., Abdalla, M., Milazzo, F., Lellei-Kovács, E., ten Berge, H., Bertora, C., Chodkiewicz, A., Dămătîrcă, C., Feigenwinter, I., Fernández-Rebollo, P., Ghiasi, S., Hejduk, S., Hiron, M., Janicka, M., Pellaton, R., Smith, K.E., Thorman, R., Vanwalleghem, T., Williams, J., Zavattaro, L., Kampen, J., Derkx, R., Smith, P., Whittingham, M.J., Buchmann, N., Price, J.P.N., 2022. Permanent grasslands in Europe: Land use change and intensification decrease their multifunctionality. Agric. Ecosyst. Environ. 330, 107891. doi:https://doi.org/10.1016/j.agee.2022.107891.

Seeberg, G., Hostlowsky, A., Huber, J., Kamm, J., Lincke, L., Schwingshackl, C., 2022. Evaluating the potential of landsat satellite data to monitor the effectiveness of measures to mitigate urban heat islands: A case study for Stuttgart (Germany). Urban Sci. 6, 82. doi:https://doi.org/10.3390/urbansci6040082.

Selim, S., Eyileten, B., Karakuş, N., 2023. Investigation of green space cooling potential on land surface temperature in Antalya city of Turkey. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. XLVIII-M-1-2023, 107–114. doi:https://doi.org/10.5194/isprs-archives-XLVIII-M-1-2023-107-2023.

Sheil, D., 2018. Forests, atmospheric water and an uncertain future: the new biology of the global water cycle. For. Ecosyst. 5. doi:https://doi.org/10.1186/s40663-018-0138-y.

Song, J., Du, S., Feng, X., Guo, L., 2014. The relationships between landscape compositions and land surface temperature: Quantifying their resolution sensitivity with spatial regression models. Landsc. Urban Plan. 123, 145–157. doi:https://doi.org/10.1016/j.landurbplan.2013.11.014.

Sultana, S., Satyanarayana, A.N.V., 2020. Assessment of urbanisation and urban heat island intensities using landsat imageries during 2000 – 2018 over a sub-tropical Indian city. Sustain. Cities Soc. 52, 101846. doi:https://doi.org/10.1016/j.scs.2019.101846.

Sušnik, J., Masia, S., Kravčík, M., Pokorný, J., Hesslerová, P., 2022. Costs and benefits of landscape-based water retention measures as nature-based solutions to mitigating climate impacts in eastern Germany, Czech Republic, and Slovakia. Land Degrad. Dev. 33, 3074–3087. doi:https://doi.org/10.1002/ldr.4373.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., Bürkner, P.C., 2021. Rank-normalization, folding, and localization: An improved $\widehat{R}$ for assessing convergence of MCMC (with discussion). Bayesian Anal. 16, 667–718. doi:https://doi.org/10.1214/20-BA1221.

Webster, R., Lark, M., 2012. Field Sampling for Environmental Science and Management. Routledge. doi:https://doi.org/10.4324/9780203128640.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. URL: https://ggplot2.tidyverse.org.

Wickham, J.D., Wade, T.G., Riitters, K.H., 2012. Comparison of cropland and forest surface temperatures across the conterminous united states. Agric. For. Meteorol. 166-167, 137–143. doi:https://doi.org/10.1016/j.agrformet.2012.07.002.

Wolff, S., Hüttel, S., Nendel, C., Lakes, T., 2021. Agricultural landscapes in Brandenburg, Germany: an analysis of characteristics and spatial patterns. Int. J. Environ. Health Res. 15, 487–507. doi:https://doi.org/10.1007/s41742-021-00328-y.

World Meteorological Organization (WMO), 2023. State of the climate in Europe 2022. URL: https://library.wmo.int/idurl/4/66206. accessed 18 September 2023.

Zellweger, F., Coomes, D., Lenoir, J., Depauw, L., Maes, S.L., Wulf, M., Kirby, K.J., Brunet, J., Kopecký, M., Máliš, F., Schmidt, W., Heinrichs, S., den Ouden, J., Jaroszewicz, B., Buyse, G., Spicher, F., Verheyen, K., De Frenne, P., 2019. Seasonal drivers of understorey temperature buffering in temperate deciduous forests across europe. Glob. Ecol. Biogeogr. 28, 1774–1786. doi:https://doi.org/10.1111/geb.12991.

Zimmermann, B., Hildmann, C., 2021. Klimawandel im Landkreis Elbe-Elster ganz konkret: Eine Analyse von Daten aus der Wetterstation Doberlug-Kirchhain. URL: https://www.lkee.de/media/custom/2112_8805_1.PDF?1621336033. accessed 25 July 2023.

Zomer, R.J., Xu, J., Trabucco, A., 2022. Version 3 of the global aridity index and potential evapotranspiration database. Sci. Data 9. doi:https://doi.org/10.1038/s41597-022-01493-1.

Appendix A. Processing of satellite data

This part of the appendix presents the general processing steps of Landsat 8 satellite scenes to land surface temperature (LST) and normalised difference moisture index (NDMI) scenes. They serve as a basis for determining the cooling effect of individual measures in the form of a statistical model.

At the beginning, the objective of the processing is defined and the different input products are named. The processing steps are numbered. If more profound steps are necessary, they are underlined with further key points. The final step is the description of the data output. All data was processed with GRASS GIS 8.0.2 (incl. addons).

Appendix A.1. Thermal images

- **Target:** Thermal images with scaled temperature as input for statistical modelling
- **Data input:**
  - Vector: Study area
  - Raster: Landsat 8 Collection 2 Level 2 band 10 (30x30 m)
  - Raster: Landsat 8 Collection 2 Level 2 Pixel Quality Assessment Band (QA band) (30x30 m)
- **Processing steps**:
  1. Removing cloud shadows:

    - Classify the QA band according to corresponding pixel quality: cloud_confidence: Medium, High; cloud_shadow_confidence: Medium, High; cirrus_confidence: Medium, High (GRASS GIS Addon: i.landsat.qa)
    - Reclassification of the Quality Assessment Band with the created category values of the previous step
    - Mask satellite scene
    - Median (3x3) and average (27x27) filters to remove missing classification (cloud shadows etc.) and to account for the original 100 x 100 m resolution of the thermal scene

2. Converting the cell values to temperature (°C) using the included meta file
3. Clip of the satellite scene by study area
4. Rescaling of the temperature with min-max feature scaling $LST_{scale} = \frac{LST - min(LST)}{max(LST) - min(LST)}$

- **Data Output:**
    - Raster: 39 satellite scenes with LSTscale values (30x30 m)

Appendix A.2. Normalised Difference Moisture Index (NDMI)

- **Target:** Processing of spectral bands in NDMI scenes for categorisation of grassland.
- **Data input:**
    - Vector: Study area

- Raster: Landsat 8 Collection 1 Level 2 band 5 (30x30 m)
- Raster: Landsat 8 Collection 1 Level 2 band 6 (30x30 m)
- Raster: Landsat 8 Collection 1 Level 1 Pixel Quality Assessment Band (QA band) (30x30 m)

- **Processing steps**:
  1. Removing cloud shadows:
     - Creating a cloud mask from QA band using the USGS unpack_collection_qa software (–cloud_shadow=med –cirrus=med –cloud_confidence=med)
     - Mask satellite scene
     - Median (3x3) and average (5x5) filters to remove missing classification (cloud shadows etc.) and aiming to fill small voids and add 2-cell margin to to-be-masked areas
  2. Calculation of the NDMI:
     $NDMI = \frac{NIR-SWIR}{NIR+SWIR} = \frac{Band5-Band6}{Band5+Band6}$
  3. Clip of the satellite scene by study area

- **Data Output:**
  - Raster: 39 satellite scenes with NDMI values (30x30 m)

Appendix B. Updated land cover map

A uniform area basis is required to determine potential implementation areas for individual measures. The subdivision into plots is based on uniform land cover or land use (e.g. arable land). The biotope mapping of Brandenburg from 2009 was available as a data basis. Due to its age, this map was updated with further (mostly spatially coarser or spatially incomplete) data. Linear features such as hedgerows are not included in the updated land cover map.

As in chapter Appendix A, this part is divided into target, data input, processing steps and created output. The data was processed with GRASS GIS 8.0.2 (incl. addons)

- **Target:** Area basis for subsequent localization and measure evaluation. Partial areas are derived from uniform land cover.

- **Data input:**

  - Vector: Study area
  - Vector: Biotop and land-use mapping of the federal state of Brandenburg (Germany)
  - Vector: Land cover model of Germany 2018
  - Vector: Agricultural land-use data of the federal state of Brandenburg (Germany) 2020
  - Vector: Soil score for agricultural land of the federal state of Brandenburg (Germany)
  - Raster: Copernicus high resolution layer imperviousness density (IMD) 2018 (10x10 m)
  - Raster: Copernicus high resolution layer tree cover density (TCD) 2015 (20x20 m)

  - Raster: Copernicus high resolution layer tree cover density (TCD) 2018 (10x10 m)
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)
  - Raster: Mean groundwater level map of the federal state of Brandenburg (Germany) (250x250 m)

- **Processing steps**:
  1. Processing biotope and land-use mapping
     - Tailoring biotope mapping to study area
     - Translate biotopes to *BIOTOP_CODE*
  2. Processing land cover model of Germany 2018
     - Tailoring land cover model to study area
     - Extract solar parks: Select energy production plants and extract solar parks by visual matching and manual correction
     - Update of the attribute column *BIOTOP_CODE* for solar plants
     - Inserting the solar plants in basemap
     - Transfer resulting splinter polygons into larger neighbour polygons (Threshold area < 1000 m^2)
     - Extract water surfaces
     - Update of the attribute column *BIOTOP_CODE* for water surfaces
     - Inserting the water surfaces in basemap

    - Transfer resulting splinter polygons into larger neighbour polygons (Threshold area < 1000 m^2)

3. Processing agricultural land-use data 2020
    - Tailoring agricultural land-use data to study area
    - Standardise the area information on arable land (value=1) and permanent grassland (2)
    - Translation of $BIOTOP\_CODE$ into arable land (1) and permanent grassland (2)
    - Rasterise of the basemap und agricultural land-use data according to arable land and permanent grassland
    - Calculate of the difference of agricultural use of both rasters: $Agrar_{diff} = 10^{Basemap} - 10^{Agrar}$
    - assign the 50th percentile of the $Agrar_{diff}$ value to polygons of the basemap
    - Translation of $Agrar_{diff}$ to final $BIOTOP\_CODE$ und land-use change
4. Deriving the land use/land cover classes 1 - 22 from $BIOTOP\_CODE$ (LUC)
5. Determine the average degree of sealing from IMD 2018 for each subarea of the basemap and convert to polygons ($SEALING$)
6. Transfer average German soil quality index into basemap ($AGRAR\_INDEX$)
    - Rasterise soil score for agricultural land by soil quality index
    - Determine the mean soil quality index per polygon for the arable and grassland areas of the basemap

    - If there is no soil quality index, determine and use the modal value of arable and grassland areas of the district separately

7. Transfer mean depth of groundwater level to basemap (GW_DEPTH)
    - Resample DEM to 250x250 m with the method average (resolution of the groundwater map)
    - Calculate depth of groundwater table: $GW_{depth} = DEM - GW_{Trend}$
    - Transfer median of groundwater depth into basemap for all polygons

- **Data Output:**
    - Vector: Basemap with attribute columns: $AREA$, $BIOTOP\_CODE$, LUC, $SEALING$, $AGRAR\_INDEX$, GW_DEPTH

Appendix C. Localisation of measures

To determine the cooling effects of individual measures, potential areas must be detected. Measure-specific processing steps and threshold values were defined on the resulting partial areas of chapter Appendix B. Localisation is based on technical and not on economic considerations.

As in chapter Appendix A and Appendix B, the sections are divided into target, data input, processing steps and created output. The data was processed in GRASS GIS 8.0.2 (incl. addons) and R 4.2.0 (R-library: rgrass7).

Appendix C.1. Localisation: arable land

Agroforestry systems (alley cropping)

- **Data input:**

    - Vector: Land cover map (Appendix Appendix B Updated land cover map)

- **Processing steps**:
    - Extract all arable land from basemap with *BIOTOP_CODE*
- **Data Output:**
    - Vector: Potential areas for agroforestry systems

Cultivating deep-rooted crops

- **Data input:**
    - Vector: Land cover map (Appendix Appendix B Updated land cover map)
- **Processing steps**:
    - Extract all arable land from basemap with *BIOTOP_CODE*
- **Data Output:**
    - Vector: Potential areas for deep-rooted crops

Landscape structure element

Since the resolution of the Corine Tree Cover density grids are too coarse, this does not allow conclusions to be drawn about existing hedge structures. With the help of a digital surface model (bdom) this circumstance could be counteracted.

- **Data input:**

  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Building outlines
  - Vector: Location of wind turbines
  - Vector: Roads and paths
  - Vector: Ditches
  - Vector: Study area
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)
  - Raster: Image based digital surface model (DSM) of the federal state of Brandenburg (Germany) (1x1m)

- **Processing steps**:
  1. Extract all arable and grassland areas from Basemap with $BIOTOP\_CODE$
  2. Identification of possible hedge locations
     - Removing short ditches (length $< 30m$)
     - Buffer roads/paths/ditches to 10 m (equivalent to removing hedges from roads/paths/ditches)
     - Overlap of roads/paths/ditches areas with cropland and grassland areas from 1st step
     - Convert roads/paths/ditches areas into lines
     - Remove lines at borders to forest polygons
     - Remove lines at borders to study area

3. Determination of vegetation heights $> 2m$ in the study area
   - Calculation of the height of objects:
     $High_{Object} = DEM - DSM$
   - Remove houses (with buffer of 2m to counteract inaccuracies) from the $\mathrm{High}_{\mathrm{Object}}$ map
   - Removal of wind turbines (with buffer of rotor radius and 5m for inaccuracies)
   - Extract of all object heights $\geq 2m$
   - Resample (method = sum) and normalise to 30x30 m raster
4. Identifying existing hedges
   - Extract all arable, grassland and woodland from Basemap with $BIOTOP\_CODE$
   - Rasterise of subareas with values 1 (arable & grassland) and 2 (forest areas)
   - Resampling to 30x30 m according to minimum
   - Resampling to 30x30 m according to maximum
   - Create a new raster without arable/grassland to forest border:
     $Min - Max == 0$
   - Extract of cells from arable/grassland without forest edge with vegetation height $\geq 2m$
5. Removal of existing hedges
6. Removing hedge lengths $< 30m$
7. Sum hedge lengths per polygon

8. Determine proportion of hedge length to total area (assumption hedge width is 7.5 m): $HEDGE\_REL = 7.5 * hedgelength/area$

- **Data Output:**
  - Vector: Potential areas for landscape structure element with attribute column: $HEDGE\_REL$

Organic fertilisation

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
- **Processing steps**:
  - Extract all arable land from basemap with $BIOTOP\_CODE$
- **Data Output:**
  - Vector: Potential areas for organic fertilisation

Converting arable land into permanent grassland

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
- **Processing steps**:
  - Extract all arable land from basemap with $BIOTOP\_CODE$

- **Data Output:**
  - Vector: Potential areas for permanent grassland

Afforestation of marginal arable land

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
- **Processing steps**:
  - Extract all arable land with $AGRAR\_INDEX < 23$ from basemap with $BIOTOP\_CODE$
- **Data Output:**
  - Vector: Potential areas for afforestation

Appendix C.2. Localisation: forestry

Ecological forest conversion

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Forest site characteristics map of the federal state of Brandenburg (Germany)
- **Processing steps**:

1. Extract areas of coniferous and mixed forest (excluding natural coniferous and mixed forests) from basemap with $BIOTOP\_CODE$ (deciduous forests have no conversion potential)
2. Rasterise Forest site characteristics map by nutrient level
3. Determine the modal value of the nutrient level ($FOREST\_NUTRIENT$) per subarea
4. If no value for nutrient level can be determined, modal value of the project region is used
5. Comparison of actual state ($BIOTOP\_CODE$) with target type $FOREST\_TARGET$ (derived from nutrient level), extract of areas worse than target type

- **Data Output:**
  - Vector: Potential areas for ecological forest conversion with attribute column: $FOREST\_NUTRIENT$, $FOREST\_TARGET$

Reforestation

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Forest site characteristics map of the federal state of Brandenburg (Germany)
  - Raster: Copernicus high resolution layer tree cover density (TCD) 2015 (20x20 m)

- Raster: Copernicus high resolution layer tree cover density (TCD) 2018 (10x10 m)

- **Processing steps**:

  1. Determine thinned forest area
     - Creating a grid where the difference of TCD is at least 25%.: $25Diff = TCD_{2015} - TCD_{2018} \geq 25$
  2. Extract forest areas from basemap with $BIOTOP\_CODE$
  3. Determination of the thinned area per subarea ($FOREST\_THIN$)
  4. If thinned area > 0 then tree species selection for reforestation as in forest conversion
     - Rasterise Forest site characteristics map by nutrient level
     - Determine the modal value of the nutrient level ($FOREST\_NUTRIENT$) per subarea
     - If no value for nutrient level can be determined, modal value of the project region is used
     - Comparison of actual state ($BIOTOP\_CODE$) with target type $FOREST\_TARGET$ (derived from nutrient level), extract of areas worse than target type
  5. Determine tree cover density in thinned forest area: if $25Diff$ then TCD_{2018}
  6. Transferring the tree cover density into partial areas TCD_THIN
  7. Determine tree cover density in remaining forest stand: if not $25Diff$ then TCD_{2018}
  8. Transferring the tree cover density into partial areas TCD_REMAIN

- **Data Output:**
  - Vector: Potential areas for reforestation with attribute column: $FOREST\_THIN$, $FOREST\_NUTRIENT$, $FOREST\_TARGET$, TCD_THIN, TCD_REMAIN

Appendix C.3. Localisation: nature conservation

Restoration of wet meadows

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Ditches
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)
- **Processing steps**:
  1. Extract grassland from basemap with $BIOTOP\_CODE$
  2. Extract polygons with $GW\_STAND \leq 2m$
  3. Determine areas with ditches
     - Buffering the ditches with 5 m
     - Select partial areas overlaid with buffer of ditches
  4. Determination of areas where at least one seal (closure of the ditch) must be placed
     - Extract points with 10 m spacing from ditches (line)
     - Assign height from DEM for each point

    - Determination of subsets of points overlapping with grassland areas
    - Determine points with minimum and maximum height for each ditch section per grassland area
    - Determine the number of seals per ditch section. 0.2 m is the tolerable fluctuation range of the groundwater level
      $N\_SEAL = \frac{max(height) - min(height)}{0.2}$
    - Sum up the number of seals per subarea
    - Extract of the partial areas with $N\_SEAL > 0$
5. Determination of the area on which the measure has an effect
    - Rasterise the ditches and assign height from the DEM to each cell through which a ditch passes
    - Determine the 90th percentile from the ditch grid for each partial area
    - Create raster of effective area per subarea:
      $DEM - Percentile_{90} < 0.2$
    - Resample the grid to target resolution of 30x30 m with the method average

- **Data Output:**
    - Vector: Potential areas for restoration of wet meadows with attribute column: N_SEAL
    - Raster: Impact area of the measure (30x30 m)

Appendix C.4. Localisation: settlements

Afforestation of urban brownfield sites

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Study area
  - Vector: Land cover model of Germany 2018
  - Vector: Rail network
  - Vector: Roads
  - Vector: Building outlines
  - Vector: Location of wind turbines
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)
  - Raster: Image based digital surface model (DSM) of the federal state of Brandenburg (Germany) (1x1 m)
- **Processing steps**:
  1. Extract brownfields and ruderal areas from basemap with $BIOTOP\_CODE$
  2. Extract of settlement areas from land cover model in study area
  3. Select brownfields and ruderal areas that overlap with settlement areas
  4. Since ruderal areas in settlement areas are often located along roads and railroad tracks, these small areas must be removed

        - Tailoring rails and roads to study area
        - Buffering of these line vectors by 10 m
        - Remove polygons that are in the buffer
    5. Removal of other narrow brownfields and ruderal areas, as they are not suitable for reforestation
        - Buffer of the partial areas to the inside (-10 m)
        - Remove areas that are completely filled by the buffer
    6. Removing areas with already existing trees
        - Calculation of the height of objects: $High_{Object} = DEM - DSM$
        - Remove houses (with buffer of 2m to counteract inaccuracies) from the $\mathrm{High}_{\mathrm{Object}}$ map
        - Removal of wind turbines (with buffer of rotor radius and 5m for inaccuracies)
        - Extract of all object heights $\geq 3m$
        - Resample (method = sum) and normalise to 30x30 m raster
        - Determine the median of the tree cover density (TCD) per polygon
        - Extract of partial areas with $TCD < 0.5$

- **Data Output:**
    - Vector: Potential areas for afforestation of urban brownfield sites

Partial unsealing with grass stones

- **Data input:**

- Vector: Land cover map (Appendix Appendix B Updated land cover map)
- Vector: Building outlines

- **Processing steps**:
  1. Extract sealed parking spaces from basemap with $BIOTOP\_CODE$
  2. Subtract building area from sealed area
     - Determination of all buildings that overlap with polygons
     - Determination of the building area
     - Summing up the building areas per parking polygon
     - Subtraction of building areas from sealed area:
       $Area_{parking} = AREA \cdot SEALING - Areabuilding$
  3. Extract of polygons whose area sealed by parking lots is > 0

- **Data Output:**
  - Potential areas for partial unsealing with grass stones

Tree groups in settlement areas

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Building outlines
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)

    - Raster: Image based digital surface model (DSM) of the federal state of Brandenburg (Germany) (1x1 m)

- **Processing steps**:

    1. Extract urban green areas (e.g. lawns, gardens, campsites, playgrounds), residential areas, industrial and commercial areas, Publicly owned land (e.g. schools, hospitals) from basemap
    2. Determination of existing trees (vegetation heights $> 2m$)
        - Calculation of the height of objects:
          $High_{Object} = DEM - DSM$
        - Remove houses (with buffer of 2m to counteract inaccuracies) from the $\mathrm{High}_{\mathrm{Object}}$ map
        - Extract of all object heights $\geq 2m$
        - Determine the summed tree canopy area per partial area ($TREEAREA\_ACTUAL$)
    3. Allocation of additional tree cover (assumption: 60% of the unsealed area should be covered with trees):
       $TREEAREA\_NEW = (AREA - AREA * SEALING/100) * 0.6 - SUM\_TREEPIXEL$
    4. Calculation of the number of additional trees (assumption: 1 tree $\hat{=}$ 80m^2)
    5. Extract partial areas with at least one tree planting
    6. Determine the current tree cover density (TCD_ACTUAL) and tree cover density after implementation (TCD_AFTER):

$$TCD\_ACTUAL = \frac{(TREEAREA\_ACTUAL)}{AREA}$$

$$TCD\_AFTER = \frac{(TREEAREA\_ACTUAL+TREEAREA\_NEW)}{AREA}$$

- **Data Output:**
  - Vector: Potential areas for tree groups in settlement areas with attribute columns: $TREEAREA\_ACTUAL$, $TREEAREA\_NEW$, TCD_ACTUAL, TCD_AFTER

Tree rows along roads out of town & Tree rows in urban areas

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Roads
  - Vector: Building outlines
  - Vector: Tree avenues and rows
- **Processing steps**:
  1. Identify possible areas for tree rows
     - Buffering roads with 8 m
     - Convert buffer area to line
     - Remove building areas from lines
     - Extract lines longer than 8 m (8 m $\hat{=}$ 1 tree)
     - Summing up the length of possible rows of trees per partial area (POT_LENGTH)

2. Intersect existing tree rows/alley with partial areas and determine summed length per partial area ($ACTUAL\_LENGTH$)
3. Calculation of the number of new trees to be planted: $N\_TREE = \frac{(POT\_LENGTH - ACTUAL\_LENGTH)}{8}$
4. extract polygons with $N\_TREE > 0$
5. Determine the current tree cover density (TCD_ACTUAL) and tree cover density after implementation (TCD_AFTER): $TCD\_ACTUAL = \frac{ACTUAL\_LENGTH \cdot 8 \cdot \pi}{4}$ $TCD\_AFTER = \frac{(ACTUAL\_LENGTH + POT\_LENGTH) \cdot 8 \cdot \pi}{4}$

- **Data Output:**
  - Vector: Potential areas for tree rows with attribute columns: N_TREE, POT_LENGTH, $ACTUAL\_LENGTH$, TCD_ACTUAL, TCD_AFTER

Unsealing of urban brownfields

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Roads
- **Processing steps**:
  1. Extract brownfields from basemap with $BIOTOP\_CODE$
  2. Subtract road areas from sealed areas
     - Buffering the roads by 10 m

- Summarise the area of the roads in the polygons
- Subtraction of road areas from sealed area:
  $Area_{parking} = AREA \cdot SEALING - Arearoads$

3. Extract of polygons whose area sealed by brownfields is > 0

- **Data Output:**
  - Potential areas for unsealing of urban brownfields

Appendix C.5. Localisation: water management

Active ditch water level management

- **Data input:**
  - Vector: Land cover map (Appendix Appendix B Updated land cover map)
  - Vector: Farm dams
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)

- **Processing steps**:
  1. Extract grassland and arable land from basemap with $BIOTOP\_CODE$
  2. Extract polygons with $GW\_STAND \leq 2m$
  3. Extract areas with farm dams
     - Buffering of the farm dams with 10 m
     - Select partial areas overlaid with buffer of farm dams
     - Extract polygons with number of farm dams > 0

4. Determination of the area on which the measure has an effect
    - Buffering of the dams with 2.5 m
    - Read height values from DEM for buffer area
    - Rasterise the farm dams and assign height from the DEM to each cell through which the buffer area is located
    - Rasterise of the 10 m buffer areas of the farm dams from step 3
    - Determine and transfer the 90th percentile to the raster for each farm dam
    - Calculation of maximum height for each arable and grassland polygon from raster
    - Create raster of effective area per subarea: $DEM - Max < 0.2$
    - Resample the grid to target resolution of 30x30 m with the method average

- **Data Output:**
    - Vector: Potential areas for active ditch water level management
    - Raster: Impact area of the measure (30x30 m)

Installing supporting sills in inland ditches

- **Data input:**
    - Vector: Land cover map (Appendix Appendix B Updated land cover map)

  - Vector: Ditches
  - Raster: Digital elevation model (DEM) of the federal state of Brandenburg (Germany) (1x1 m)

- **Processing steps**:

  1. Extract grassland and arable land from basemap with $BIOTOP\_CODE$
  2. Extract polygons with $GW\_STAND \leq 2m$
  3. Determine areas with ditches
     - Buffering the ditches with 5 m
     - Select partial areas overlaid with buffer of ditches
  4. Determination of areas where at least one seal (closure of the ditch) must be placed
     - Extract points with 10 m spacing from ditches (line)
     - Assign height from DEM for each point
     - Determination of subsets of points overlapping with grassland areas
     - Determine points with minimum and maximum height for each ditch section per grassland area
     - Determine the number of sills per ditch section. 0.2 m is the tolerable fluctuation range of the groundwater level $N\_SILL = \frac{max(hight)-min(height)}{0.2}$
     - Sum up the number of sills per subarea
     - Extract of the partial areas with $N\_SILL > 0$
  5. Determination of the area on which the measure has an effect

    - Rasterise the ditches and assign height from the DEM to each cell through which a ditch passes
    - Determine the 90th percentile from the ditch grid for each partial area
    - Create raster of effective area per subarea: $DEM - Percentile_{90} < 0.2$
    - Resample the grid to target resolution of 30x30 m with the method average

- **Data Output:**
  - Vector: Potential areas for restoration of wet meadows with attribute column: N_SILL
  - Raster: Impact area of the measure (30x30 m)

Appendix D. Post-stratification

Post-stratifictaion was performed to reduce data set size and autocorrelation. First, we assigned each pixel id (variable ID in the model, Chpt. 2.3.4) to a landscape group. The groups represent spatial clusters of land cover or land use derived from the smallest administrative units in each case: the forest stand in forest areas (2,126 groups), the field block in agricultural landscapes (7,201 groups), the community in settlement areas (181 groups), and the catchment for fens (397 groups). The groups are considered as strata in a stratified simple random sample. Within each stratum, we next selected two pixel ids. This preliminary subset still contained more than 700,000 observations. To reduce the subset size further, we sampled the selected ids such that each predictor occurs over its entire range in ten equal intervals, with $n = 20$ per interval. This equal spreading of predictor values was shown to be optimal for regression (Webster and Lark, 2012). Lastly, we assured that each pixel id occurred at least five times (i.e. five scenes) in the final data subset to provide a sufficient sample size for estimation of the effect of the variable ID.

Appendix E. Forest types and mixing ratios

The two forest measures were predicted using the forest composition as given in Tab. E.5.

Table E.5: Forest types and mixing ratios for the forest measures

Nutrition level[1]	Meaning	Area share	Current forest composition[2]	New forest composition[3]
Z	Quite poor	0.65	81% Coniferous forest, 19% Mixed forest	Mixed forest (33% Red oak, 33% Birch, 33% Pine)
M	Moderately nutritious	0.25	68% Coniferous forest, 32% Mixed forest	Deciduous forest (66% Pedunculate Oak, 11% Red beech, 11% Winter lime, 11% Hornbeam)

1 Nutrition levels A (poor) and Z+ (between Z and M) are not considered because they together cover only 10% of the area

2 In areas with conversion potential, see App. Appendix C

3 Forest composition of the replacement forest, for details see Hildmann et al. (2022)

Appendix F. Specific model results

Apart from the posterior predictive check plot (Fig. F.7), this appendix contains regression tables with coefficient estimates, the estimation uncertainty and convergence measures.

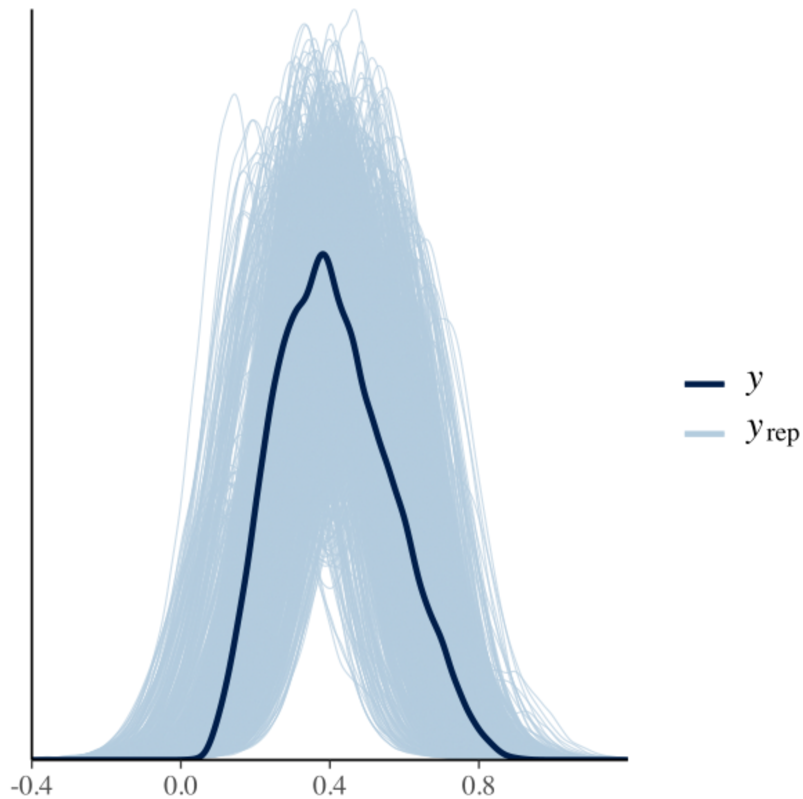


Figure F.7: Kernel density estimate for the observations in the validation data subset (y) and for 1000 draws from the posterior predictive distribution (y_{rep}).

Table F.6: Group-level effects. $\hat{R}$, Bulk-ESS and Tail-ESS are convergence diagnostics for assessing convergence of the Markov chain Monte Carlo (MCMC) algorithm.

Group	Variable[1]	Estimate	Estimation error	l-95% CI[2]	u-95% CI[3]	$\hat{R}$[4]	Bulk-ESS[5]	Tail-ESS[6]
ID	sd(LULC class 1)	0.06	0.00	0.05	0.06	1.001	2811	4527
ID	sd(LULC class 2)	0.08	0.00	0.07	0.08	1.001	1787	4124
ID	sd(LULC class 3)	0.08	0.00	0.07	0.08	1.002	2465	4503
ID	sd(LULC class 4)	0.07	0.00	0.07	0.08	1.001	2318	3996
ID	sd(LULC class 5)	0.07	0.00	0.07	0.08	1.003	2307	4649
ID	sd(LULC class 6)	0.06	0.00	0.06	0.07	1.003	2572	4759
ID	sd(LULC class 7)	0.06	0.00	0.06	0.07	1.002	2593	4749
ID	sd(LULC class 8)	0.05	0.00	0.05	0.05	1.002	2952	5257
ID	sd(LULC class 9)	0.06	0.00	0.06	0.06	1.002	3145	4850
ID	sd(LULC class 10)	0.03	0.00	0.03	0.04	1.003	2797	5122
ID	sd(LULC class 11)	0.03	0.00	0.03	0.03	1.001	3468	5395
ID	sd(LULC class 12)	0.03	0.00	0.03	0.03	1.002	3381	5065
ID	sd(LULC class 13)	0.05	0.00	0.05	0.06	1.002	2194	3833
SCENE	sd(LULC class 1)	0.07	0.01	0.05	0.09	1.004	1494	3437
SCENE	sd(LULC class 2)	0.09	0.01	0.07	0.12	1.005	1801	3442
SCENE	sd(LULC class 3)	0.12	0.02	0.09	0.16	1.006	1297	2562
SCENE	sd(LULC class 4)	0.11	0.02	0.09	0.15	1.005	1517	2898
SCENE	sd(LULC class 5)	0.12	0.02	0.09	0.16	1.002	1701	2458
SCENE	sd(LULC class 6)	0.10	0.01	0.08	0.13	1.004	1413	2912
SCENE	sd(LULC class 7)	0.13	0.02	0.10	0.17	1.003	1423	2286
SCENE	sd(LULC class 8)	0.08	0.01	0.06	0.11	1.005	1608	3258
SCENE	sd(LULC class 9)	0.09	0.01	0.07	0.11	1.004	1600	3383
SCENE	sd(LULC class 10)	0.09	0.01	0.07	0.11	1.010	1114	2957
SCENE	sd(LULC class 11)	0.09	0.01	0.07	0.11	1.004	1317	2799
SCENE	sd(LULC class 12)	0.09	0.01	0.07	0.11	1.009	1245	2762
SCENE	sd(LULC class 13)	0.09	0.01	0.07	0.11	1.002	1237	2308

1 LULC class : land-use/land-cover class as given in Tab. 3; sd: standard deviation of group-level effects

2 lower limit of the 95% confidence interval for the estimate

3 upper limit of the 95% confidence interval for the estimate

4 $\hat{R}$ is the potential scale reduction factor on split chains (at convergence, $\hat{R} = 1$), should be smaller than 1.01 (Vehtari et al., 2021)

5 Bulk effective sample size: sampling efficiency in the bulk of the posterior, should be greater than 400 (Vehtari et al., 2021)

6 Tail effective sample size: minimum of the effective sample sizes of the 5% and 95% quantiles, should be greater than 400 (Vehtari et al., 2021)

Table F.7: Population-level effects for the population-level predictors except LULC and the two interaction terms. $\hat{R}$, Bulk-ESS and Tail-ESS are convergence diagnostics for assessing convergence of the Markov chain Monte Carlo (MCMC) algorithm.

Variable	Estimate	Estimation error	l-95% CI[1]	u-95% CI[2]	$\hat{R}$[3]	Bulk-ESS[4]	Tail-ESS[5]
Tree cover density	-0.22	0.01	-0.23	-0.20	1.001	3941	5570
Imperviousness density	0.05	0.01	0.04	0.07	1.001	2597	4470
Available water holding capacity[6]	-0.00	0.00	-0.00	-0.00	1.001	4043	5925
Elevation	-0.01	0.00	-0.01	-0.01	1.001	4606	5774
Reference crop evapotranspiration	0.05	0.00	0.05	0.06	1.011	788	1804
Related circum-scribing circle	-0.04	0.00	-0.05	-0.03	1.001	3351	5083

1 lower limit of the 95% confidence interval for the estimate

2 upper limit of the 95% confidence interval for the estimate

3 $\hat{R}$ is the potential scale reduction factor on split chains (at convergence, $\hat{R} = 1$), should be smaller than 1.01 (Vehtari et al., 2021)

4 Bulk effective sample size: sampling efficiency in the bulk of the posterior, should be greater than 400 (Vehtari et al., 2021)

5 Tail effective sample size: minimum of the effective sample sizes of the 5% and 95% quantiles, should be greater than 400 (Vehtari et al., 2021)

6 Expected average difference between two adjacent categories of the ordinal predictor AWC (Bürkner and Charpentier, 2020)

Table F.8: Population-level effects for the interaction terms. $\hat{R}$, Bulk-ESS and Tail-ESS are convergence diagnostics for assessing convergence of the Markov chain Monte Carlo (MCMC) algorithm.

Variable[1]	Estimate[2]	Estimation error	l-95% CI[3]	u-95% CI[4]	$\hat{R}$[5]	Bulk-ESS[6]	Tail-ESS[7]
LULC class 1:CAI	0.34	0.05	0.23	0.45	1.001	1566	2456
LULC class 2:CAI	0.14	0.04	0.05	0.23	1.002	1337	2310
LULC class 3:CAI	0.11	0.06	-0.01	0.22	1.000	1746	2718
LULC class 4:CAI	0.14	0.03	0.08	0.21	1.001	1355	2554
LULC class 5:CAI	0.24	0.08	0.08	0.40	1.002	2218	4092
LULC class 6:CAI	-0.01	0.02	-0.06	0.04	1.001	1348	2330
LULC class 7:CAI	0.19	0.06	0.08	0.30	1.002	1633	3208
LULC class 8:CAI	0.05	0.02	0.01	0.08	1.001	1308	2431
LULC class 9:CAI	0.02	0.02	-0.02	0.05	1.001	1233	2465
LULC class 10:CAI	-0.09	0.04	-0.16	-0.01	1.001	1605	3046
LULC class 11:CAI	-0.21	0.03	-0.27	-0.15	1.001	1536	3494
LULC class 12:CAI	-0.27	0.03	-0.32	-0.21	1.001	1504	2793
LULC class 13:CAI	-0.55	0.05	-0.65	-0.46	1.001	1200	2235
LULC class 1:CWB	0.00	0.00	-0.01	0.01	1.004	1453	2864
LULC class 2:CWB	-0.01	0.01	-0.02	0.00	1.003	2159	3755
LULC class 3:CWB	-0.05	0.01	-0.06	-0.03	1.003	2176	3563
LULC class 4:CWB	-0.02	0.00	-0.03	-0.02	1.002	1829	3517
LULC class 5:CWB	-0.04	0.01	-0.05	-0.02	1.005	2083	3623
LULC class 6:CWB	0.00	0.00	-0.01	0.01	1.005	1454	3474
LULC class 7:CWB	-0.04	0.01	-0.06	-0.03	1.006	1801	3179
LULC class 8:CWB	0.00	0.00	-0.01	0.01	1.005	1444	2999
LULC class 9:CWB	-0.01	0.00	-0.02	0.00	1.005	1432	3141
LULC class 10:CWB	-0.01	0.00	-0.01	0.00	1.006	1498	2620
LULC class 11:CWB	-0.01	0.00	-0.02	-0.01	1.005	1315	2710
LULC class 12:CWB	-0.01	0.00	-0.01	0.00	1.005	1378	2716
LULC class 13:CWB	-0.01	0.01	-0.03	0.00	1.005	1203	2277

1 LULC class: land-use/land-cover class as given in Tab. 3; CAI: core area index, CWB: climatic water balance

2 the estimates were obtained by adding draws of the posterior distributions of the baseline category and the remaining categories

3 lower limit of the 95% confidence interval for the estimate, based on the added draws of the estimates

4 upper limit of the 95% confidence interval for the estimate, based on the added draws of the estimates

5 $\hat{R}$ is the potential scale reduction factor on split chains (at convergence, $\hat{R} = 1$), should be smaller than 1.01 (Vehtari et al., 2021)

6 Bulk effective sample size: sampling efficiency in the bulk of the posterior, should be greater than 400 (Vehtari et al., 2021)

7 Tail effective sample size: minimum of the effective sample sizes of the 5% and 95% quantiles, should be greater than 400 (Vehtari et al., 2021)

Appendix G. Variation of measures' performance

The cooling effect varies both in time and space. Variation in time is captured with the predictor $SCENE$ and most likely reflects meteorological conditions and seasonality. Spatial variation occurs both because of spatially varying predictors and differences in the area coverage by a measure.

Table G.9: Contribution of spatial and temporal variation to variation in measure performance, based on the scaled LST

Measure	Interquartile range[1]		Range[1]		Source of spatial variation[2]
	Temporal approach	Spatial approach	Temporal approach	Spatial approach	
A: Agroforestry	0.03	0.00	0.10	0.02	*CAI* (s.q.), *CWB*, *LULC* (s.q.; crop roation 1 or 2)
A: Deep-rooting crops	0.02	0.01	0.10	0.05	*CAI*, *CWB*
A: Landscape structure elements	0.02	0.01	0.07	0.16	*CAI* (s.q.), *CWB*, LULC (s.q.; crop rotation 1 or 2, fresh/dry grassland), area share
A: Permanent grassland	0.05	0.01	0.13	0.11	*CAI*, *CWB*, *LULC* (relative portion of fresh/dry grassland in a plot)
A: Organic fertilisation	0.00	0.00	0.01	0.04	Distribution of AWC classes (raster) in the plots (vector with averaged raster values)
A: Afforestation	0.10	0.03	0.37	0.14	*CAI*, *CWB*
F: Reforestation	0.00	0.01	0.01	0.21	*TCD* (four possible forest types), area share, *CAI*, *CWB*
F: Forest conversion	0.03	0.03	0.10	0.28	*TCD* (four possible forest types), *CWB*
N: Wet meadows	0.12	0.05	0.32	0.20	*CAI*, *CWB*, *LULC* (s.q.; relative portion of fresh/dry grassland in a plot)
S: Afforestation	0.07	0.13	0.26	0.42	*IMD*, *TCD* (s.q.), *CAI*, *CWB*
S: Partial unsealing	0.00	0.01	0.03	0.08	*IMD* (s.q.)
S: Tree groups	0.00	0.04	0.01	0.15	*TCD* (s.q.), area share
S: Tree rows	0.00	0.05	0.00	0.26	*TCD* (s.q.), area share
S: Unsealing	0.01	0.02	0.01	0.07	*IMD*, *TCD* (s.q.)
W: Ditch water management	0.00	0.01	0.02	0.04	Area share
W: Supporting sills	0.00	0.01	0.02	0.07	Area share

1 of the predictions

2 s.q.: status quo before measure implementation; area share: percentage of area on which measure can be implemented on a plot; *LULC*: land-use/land-cover class, *TCD*: tree cover density, *CAI* core area index, *IMD*: imperviousness density, *AWC*: available water-holding capacity

Appendix H. Seasonal effects

The cooling effect seems to be subject to certain patterns depending on the month of satellite overflight.

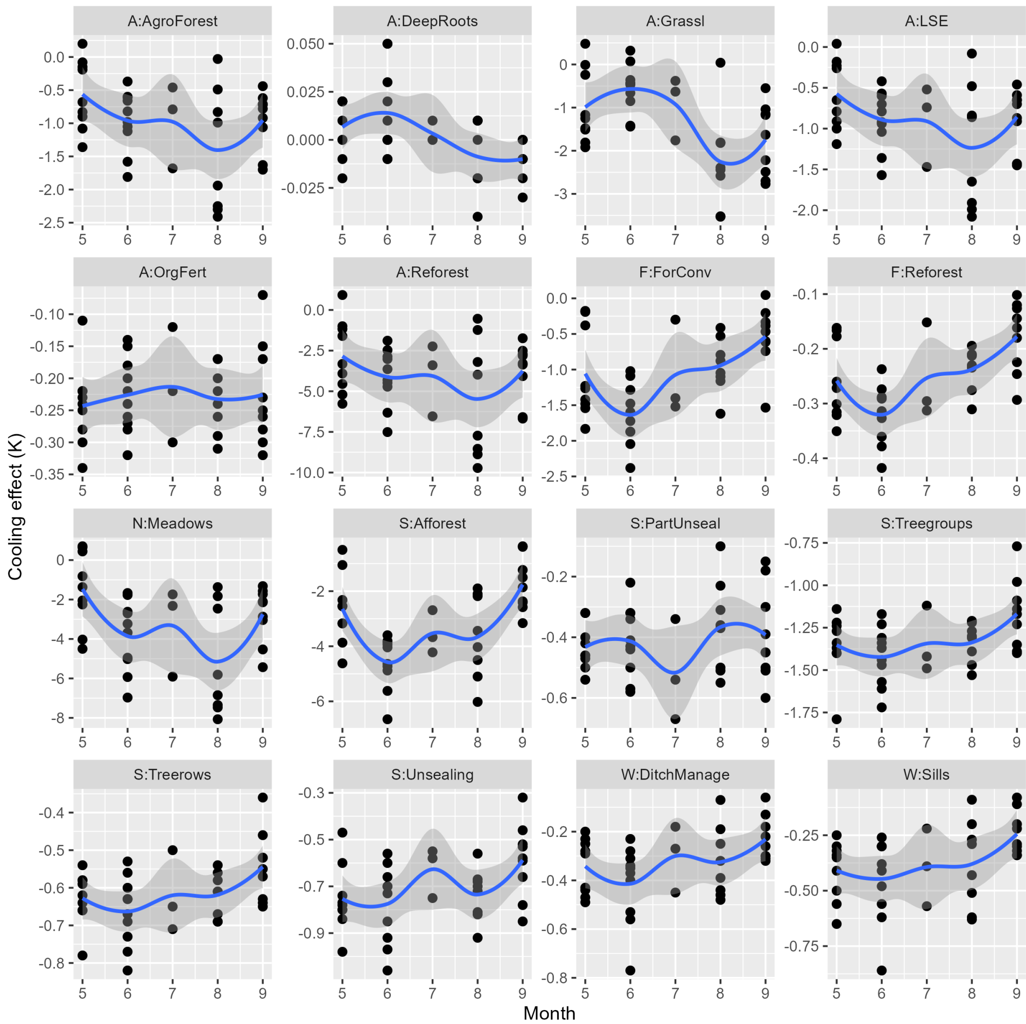


Figure H.8: Plots of the cooling effects against the months in which the thermal scenes were recorded.