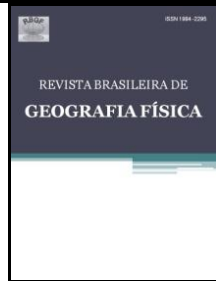




Revista Brasileira de Geografia Física

Homepage: <https://periodicos.ufpe.br/revistas/rbgfe>



Influence of Canopy Cover on Surface Temperature

Vitor Vieira Vasconcelos¹, Helenice Maria Sacht²

¹Assistant Professor, Federal University of ABC (UFABC), Universidade Avenue, Anchieta Road, São Bernardo do Campo, São Paulo, Brazil. E-mail address: vitor.v.v@gmail.com, Phone: +55(31)99331-1593. Orcid: <https://orcid.org/0000-0002-3063-2776>

² Assistant Professor, Federal University of Latin-American Integration (UNILA), 6731 Tancredo Neves Avenue, Building 06, Group B, Foz do Iguaçu, Paraná, Brazil. Orcid: <https://orcid.org/0000-0002-4674-9629>

Artigo recebido em 12/06/2020 e aceito em 03/09/2020

ABSTRACT

Trees affect the microclimate, which influences thermal comfort and ecosystem processes. This study investigated the influence of the canopy cover on daily maximum and minimum temperatures. The data are from a collaborative database, and each measurement consists of the minimum and maximum temperatures under the canopy and in an open adjacent area over a 24-hour period. Paired sample *t*-tests indicated that the canopy decreased the maximum and minimum daily temperatures and narrowed the daily temperature range. Multiple regression showed that the canopy cover percentage decreased the maximum daily temperatures, and this effect was greater in rural areas than in urbanized areas. Another multiple regression indicated that the canopy cover percentage and the distance to the edge of the canopy decreased the daily temperature range. An independent sample *t*-test also indicated that the effect of the canopy on the daily temperature range was higher in rural areas when analysed by parametric and non-parametric tests but not when measured by a robust test. Other independent sample *t*-tests indicated that the distance from a light source also decreased the canopy effect on the minimum daily temperature and the daily temperature range. The main plausible underlying processes include the canopy shade and wind insulation, litter insulation of the ground surface, heat pumps through evapotranspiration and lateral heat fluxes from light bulbs and other anthropogenic sources, especially in urbanized areas. These results provide a greater understanding of the effects of arborization in rural and urban ecosystems, as well as their respective benefits to human communities.

Keywords: Microclimatology; arborisation; temperature; urbanization; canopy; citizen science.

Introduction

Vegetation creates microclimates that are distinct from open environments. The vegetation acts as a windbreak, affecting the heat transfer and transpiration of living beings (Smith and Jarvis, 1998). Plant transpiration also increases humidity and transforms some of the solar radiation as latent heat, which contributes to tissue cooling (Gowing et al., 2008). Canopy and canopy litter have different albedos than that of bare soil, and they prevent direct solar radiation from reaching the ground surface. All these changes can affect the local niches of ecosystems and, furthermore, could be important when planning for more effective thermal comfort in environments inhabited by humans.

Maximum and minimum temperatures are important factors used by plants and other living

beings to control their seasonal cycles (i.e., phenology), such as blossoming and leaf senescence (Ricklefs and Relyea, 2019). Many enzymes and biological processes within living beings also have tolerance ranges and optimal points regarding environment temperature (Brown et al., 2004). Therefore, the microclimates in open areas and in areas under canopy cover may offer distinct conditions in terms of ecosystem development. Along the process of ecological succession, the development of larger trees with a thicker canopy would then change the environmental conditions for different living beings in the ecosystem.

These effects of tree cover on temperature are also important when planning better thermal comfort zones in green areas, such as parks, squares and gardens. Tree shade can be a good

refuge for people and other animals during periods of hot temperatures, such as heat waves. Planning that includes access and equipment (i.e., seats and tables) under trees may add even more value to the ecosystem services of these green areas (Tyrväinen et al., 2005). Additionally, the current trend of global warming may increase the importance of the ecological services provided by trees in human-inhabited areas (Salmond et al., 2016).

The objective of this study was to investigate the relationships between the canopy cover and variations in the surface temperature. These relationships are investigated through the analysis of mean difference tests and multiple regression. The main underlying process under investigation is the reduction in direct solar radiation caused by canopy cover, and the primary hypothesis assumes that the canopy decreases the maximum and minimum daily ground temperatures, as well as the daily temperature range, at the microclimate scale.

Materials and Methods

Data Collection

The Open University of the United Kingdom organized a collaborative data

measurement project that collected data from different places around the world between 29 October and 14 December 2016. In total, the collaborators obtained 363 measurements, each consisting of the maximum and minimum surface temperatures in an open area and an area under the canopy (i.e. 1 metre from the trunk) during a continuous 24-hour period. Each collaborator used two Tmin/Tmax thermometers to record the maximum and minimum air temperatures at the surface level, unshielded and un aspirated (i.e. measurement under natural ventilation).

Most measurements (91.5%) were collected in the United Kingdom; however, 6.3% of measurements were collected in other European countries, and 2.2% were collected elsewhere in the world (including four sites in the Southern Hemisphere) (Figure 1). Therefore, while the deciduous and semi-deciduous trees in the Northern Hemisphere were in the process of losing their leaves during autumn, the deciduous and semi-deciduous trees in the Southern Hemisphere were sprouting their new leaves during spring.

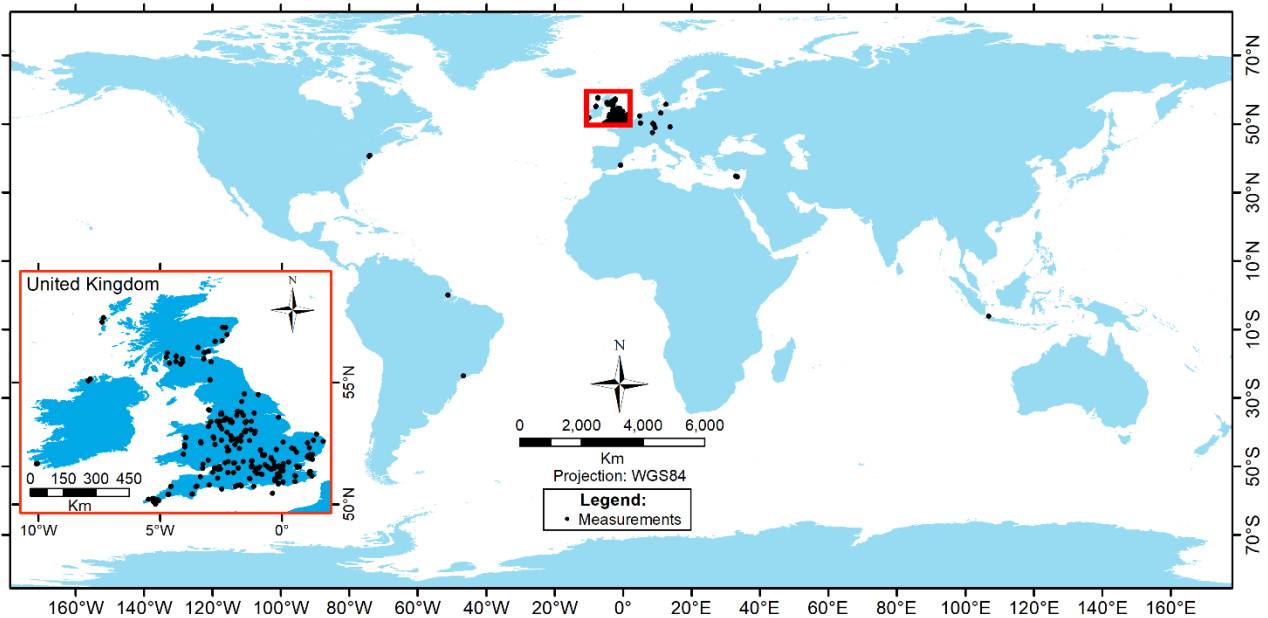


Figure 1. Location of the measurements

The possible influencing environmental factors were noted on a standardized data sheet, which is available as a supplement to this article. Canopy cover was estimated using a reference sheet (Vasconcelos (2021a)). Canopy cover was estimated in intervals of 10%, and the values were transformed into a canopy index that represented the mid-values of each interval. Each collaborator

was instructed to select trees that were located as far as possible from buildings and light sources (preferably 20 metres or more). Additionally, collaborators were instructed to select trees that were mature and had well-formed deciduous canopies. Davies and Gowing (2016) described the prescribed methods for collection and analysis in more detail. After collection, the data were

synthesized by The Open University team and then returned to the collaborators, who could then perform their own statistical analysis.

There was also a prescribed list of tree species, which included the following, in order of preference: *Quercus robur* (pedunculated oak, comprising 56.7% of the trees), *Quercus petraea* (sessile or durmast oak, comprising 19.3%), other oaks (8%), other deciduous trees (13.2%), other trees (1.7%) and other shrubs (1.1%, including woody plants with no clear trunk and foliage within 0.5 m of the ground). The preference was given to oak species because they retain their leaves for a longer period of time in autumn, providing more uniformity within the measurement period. The ground coverage also followed a prescribed preference: grass (53.1%), litter (20.8%), undergrowth (17.5%), bare soil (7.1%) and artificial surface (1.5%). Among other environmental variables, the collaborators noted whether the environment was rural (44.6%) or urban (55.4%, which included villages and suburban areas), the geographical coordinates, the distance from the thermometer under the canopy to the edge of the canopy, and the distance to the nearest light source.

Modelling Framework

This article follows a theoretical framework that is based on King and Roberts (2014); specifically, rather than selecting what would be the best model among conventional parametric, robust or non-parametric alternatives,

these modelling approaches could be interpreted as complementary perspectives of the same phenomena under analysis. Therefore, if similar patterns emerged from these distinct modelling approaches that used the same database, there was stronger evidence that these patterns were valid. On the other hand, if there was significant disagreement among these distinct modelling approaches, it was usually an indicator that there was a need for further analysis and modelling improvement. The significance values (p-levels) were interpreted in conjunction with the effect sizes, as proposed by Henson and Smith (2000). The power was evaluated using the recommendations of Hair et al. (2018) when measuring the risk of type II errors.

The individual daily maximum and minimum temperature measurements would undoubtedly diverge because the collaborators collected data from various locations under distinct weather and other environmental conditions. To address this variation, a paired analysis was used to compute temperature differences between each pair of open area and canopy cover measurements. The difference between the maximum and minimum temperatures in the open area and the area under the canopy was calculated using Equations 1 and 2, respectively. The difference between the temperature range in the open area and the area under the canopy, termed D, was calculated using Equation 3. Table 1 presents these three equations.

Table 1 - Equations used to calculate differences in maximum temperature, minimum temperature and temperature range in the open area and the area under the canopy

Equation number	Variable (result)	Equation (calculus)
1	Difference between maximum temperatures in the open area and the area under the canopy	$T_{max_{open}} - T_{max_{canopy}}$
2	Difference between minimum temperatures in the open area and the area under the canopy	$T_{min_{open}} - T_{min_{canopy}}$
3	Difference between temperature range in the open area and the area under the canopy (D)	$(T_{max_{open}} - T_{min_{open}}) - (T_{max_{canopy}} - T_{min_{canopy}})$

where $T_{max_{open}}$ is the maximum temperature in the open area; $T_{min_{open}}$ is the minimum temperature in the open area; $T_{max_{canopy}}$ is the maximum temperature under the canopy; and $T_{min_{canopy}}$ is the minimum temperature under the canopy.

A paired t-test was conducted to evaluate the overall effect of the tree canopy on the three variables listed in Table 3. The parametric assumption of normality was verified using the Shapiro and Wilk (1965) W test. The assumption of homogeneity of variance was checked using Levene's (1960) test. The significance of the paired t-test was re-checked using the Yuen (1974) robust paired t-test on 20% trimmed means and the

Wilcoxon (1945) non-parametric signed rank test with continuity correction.

In addition to significant differences obtained in the t-tests, two multiple linear regression analyses were performed to understand the effects of possible causes on these processes. One multiple regression evaluated the effect of the percentage of canopy cover and urbanization (i.e., the independent variables) on the difference in the

maximum temperature values (i.e., the dependent variable). Urbanization was coded as a dummy variable (rural = 1, urban = 0).

The other multiple regression analysis evaluated the effects of the canopy index and the distance to the edge of the canopy (i.e. independent variables) on D (i.e. a dependent variable). Non-linear relationships were verified between the independent and dependent variables. The residual autocorrelation was verified using the Durbin and Watson (1971) test. The homoscedasticity was evaluated using the score test for non-constant error variance of Breusch and Pagan (1979). Multicollinearity was evaluated by the variance inflation factor (VIF), according to the guidelines of Hair et al. (2018) The outliers were evaluated by counting the cases in which Cook’s distance was higher than 1, as advised by Cook and Weisberg (1982), and the number of hat (i.e., leverage) values that were higher than $3(k+1)/n$, where k is the number of parameters and n is the number of cases, as advised by Pituch and Stevens (2015). In addition to the conventional ordinary least squares (OLS) method, the multiple linear regressions were evaluated using the robust standard deviations (White,1980), the robust iterated reweighted least squares (IRWLS) with a method-of-moments (MM) estimator (Yohai and Zamar, 1988; Koller and Stalhel, 2011), and the non-parametric quantile regression (Koenker and Bassett, 1978) with pseudo-R² (Cragg and Uhler, 1970; Nagelkerke,

1991). The MM-estimator for IRWLS was preferred over the M-estimator (Holland and Welsch, 1977) because it can generate robust R² and significance values (Renaud and Victoria-Feser, 2010), which makes it easier to compare the results with those of the other modelling approaches.

Three independent sample *t*-tests, with the Welch (1938) adjustment for heterogeneity of variance, were conducted to evaluate the effects of urbanization and the distance from light sources on the independent variables. The analogous diagnostics as well as the robust (with bootstrapping) and non-parametric tests used in the paired samples were also used in the independent sample *t*-tests. The goal of performing the independent *t*-tests was to separately evaluate the effects of potential explanatory variables whose coefficients were not significant when incorporated into the multiple regression analyses together with other variables.

The null and alternate hypotheses for each model are described in Table 2. The parametric assumptions of each model were tested, and the statistical significance of the results was double-checked using robust and non-parametric alternatives. The code in R used for the statistical analysis and the respective datasets are provided respectively in Vasconcelos (2018) and Vasconcelos (2021b).

Table 2 - Null and alternative hypotheses analysed in this article

Models	Variable	H ₀ (null hypothesis)	H ₁ (alternate hypothesis)
Paired sample <i>t</i> -tests	T _{max_{open}} – T _{max_{canopy}}	Tree canopy has no effect on daily maximum temperature	Tree canopy changes daily maximum temperature
	T _{min_{open}} – T _{min_{canopy}}	Tree canopy has no effect on daily minimum temperature	Tree canopy changes daily minimum temperature
	D	Tree canopy has no effect on the daily temperature range (D)	Tree canopy changes daily temperature range (D)
Multiple regressions	T _{max_{open}} – T _{max_{canopy}}	The canopy cover and urbanization have no effect on the difference between the daily maximum temperatures in the open area and the area under the canopy	The canopy cover and urbanization affect the difference between the daily maximum temperatures in the open area and the area under the canopy
	D	The canopy cover and distance from the edge of the canopy have no effect on the difference between the daily temperature ranges (D) in the open area and the area under the canopy	The canopy cover and distance from the edge of canopy affect the difference between the daily temperature ranges (D) in the open area and the area under the canopy
Independent sample <i>t</i> -tests	T _{min_{open}} – T _{min_{canopy}}	The distance from the light source has no effect on the difference between the daily minimum	The distance from the light source has an effect on the difference between the daily minimum temperatures in

		temperatures in the open area and the area under the canopy	the open area and the area under the canopy
	D	The distance from the light source has no effect on the difference between the daily temperature ranges (D) in the open area and the area under the canopy	The distance from the light source has an effect on the difference between the daily temperature ranges (D) in the open area and the area under the canopy
		The urbanization has no effect on the difference between the daily temperature ranges (D) in the open area and the area under the canopy	The urbanization has an effect on the difference between the daily temperature ranges (D) in the open area and the area under the canopy

Results

Table 3 shows the results of the paired t-tests. The Shapiro-Wilk normality test on the three variables shown in Table 1 were significant with 99.9% confidence, indicating that the data were significantly different from that of a normal distribution. However, even if the sample size was large enough to generate a normal distribution based on the central limit theorem, Table 3 also presents the p-values for the robust and non-parametric alternatives to the paired sample tests. Levene’s test also showed that, at least for D, the variance in the open area and the area under the canopy was significantly distinct, violating another parametric assumption. Nevertheless, the p-values of the parametric, non-parametric and robust tests for the three variables were significant at the 99% confidence level. The areas under the canopy exhibited lower maximum and minimum daily temperatures and presented narrower daily temperature ranges. Nevertheless, the average difference and effect size (i.e., the Cohen “d”) of the daily maximum temperatures were lower than those of the other two variables.

Tables 4 and 5 present the data for the multiple regressions of the two formulas ($T_{max_{open}} - T_{max_{canopy}}$ and D as dependent variables). The two formulas were significant at the 99.9% confidence level using OLS (including the robust standard deviation correction) and IRWLS, and the formulas were significant at the 99% confidence level using quantile regression. The coefficients of the dependent variables were also significant at the 95% confidence for OLS (even including the robust standard deviation correction), IRWLS and quantile regression. However, all models had very low R² values, from 3.3% to 5.9%, indicating that there may have been many other environmental factors influencing the behaviour of the temperature both under the canopy and in the open. The Durbin-Watson tests with W values near 2 indicated that autocorrelation between residuals

would not adversely affect the outcome of either model. The VIF also indicated low multicollinearity for both models. However, the Breusch-Pagan test showed that the hypothesis of homoscedasticity of the residuals was not significant for both models; thus, paying special attention to the robust and non-parametric models is recommended.

For each additional 1% of canopy cover, the difference between the maximum daily temperatures in the open area and the area under the canopy increased by 0.026°C (OLS), 0.016°C (IRWLS), or 0.012°C (quantile regression). Therefore, a tree with 95% canopy cover would have a maximum daily temperature that was 2.94°C (OLS), 1.52°C (IRWLS), or 1.14°C (quantile regression) lower than that in the adjacent open areas. In rural areas, the difference between the maximum daily temperatures in the open area and the area under the canopy was 0.782°C (OLS), 0.937°C (IRWLS), or 0.675°C (quantile regression) higher than those in the urban areas. Comparing the standardized coefficients, the canopy cover had more influence than did urbanization on the difference between the maximum daily temperatures in the open area and the area under the canopy using OLS (even including the correction for robust standard deviation); however, the relative influence of urbanization was higher using quantile regression and even higher using IRWLS. In OLS, there were no outliers identified by the Cook’s distance method; furthermore, only one outlier was identified using the leverage values, and it had a small influence (0.8%) relative to all the leverage values of the model.

Table 3 - Paired sample t-tests

Paired sample t-test	Average difference (°C)	Levene test (p)	Shapiro-Wilk normality test (p)	Cohen d effect size	Paired t-test (p)	Power ($\alpha = 0.05$)	Yuen's paired test on trimmed means for dependent samples (p) Tr = 0.2	Wilcoxon signed rank test with continuity correction (p)
T min open - T min canopy	-2.11	0.168	3.88E-09	-0.85	2.2E-16	1.00	0.000	2.20E-16
T max open - T max canopy	0.51	0.321	8.17E-08	0.16	0.003	0.86	0.009	0.0009
D = (T max open - T min open) - (T max canopy - T min canopy)	2.62	0.017	6.92E-07	0.66	2.2E-16	1.00	0.000	2.20E-16

Table 4 - General results of the multiple regressions

Equation	Model	Durbin-Watson test for residual autocorrelation	Score test for non-constant error variance - Breusch-Pagan (p)	Variance inflation factor (multicollinearity)	Number of cases with Cook's distance >1	Hat (leverage) values higher than $3(k+1)/n = 0.0248$	
						Number of cases	% of total leverage
"T max open - T max canopy" = $\beta_1 + (\beta_2 * \text{Canopy_index}) + (\beta_3 * \text{Rural})$	OLS	1.83	0.82	1.002	0	1	0.8
"D (T max open - T min open) - (T max canopy - T min canopy)" = $\beta_1 + (\beta_2 * \text{Canopy_index}^2) + (\beta_3 * \text{distance_to_thermometer_under_edge_of_canopy})$	OLS	1.89	0.74	1.000	0	5	12.5

Table 5 - Coefficient values, significance (p) and standardized coefficients of the multiple regression models

Equation	Model	R ²	Power (α = 0.05)	Coefficients			Standardized coefficients	
				β ₁	β ₂	β ₃	β ₂	β ₃
“T max open - T max canopy” = β ₁ + (β ₂ *Canopy_index) + (β ₃ *Rural)	OLS	0.051***	0.992	-1.571**	0.026***	0.782*	3.603	2.349
	OLS with robust standard deviation (HC3)	0.051***	0.992	-1.571**	0.026***	0.782*	3.823	2.369
	IRWLS	0.053***	0.993	-1.060*	0.016**	0.937**	2.806	3.295
	Quantile regression	0.033**	0.937	-0.737*	0.012*	0.675*	2.307	2.357
“D (T max open - T min open) - (T max canopy - T min canopy)” = β ₁ + (β ₂ *Canopy_index^2) + (β ₃ *distance_to_thermometer_under_edge_of_canopy)	OLS	0.059***	0.997	0.447	0.0003***	0.127*	4.080	2.399
	OLS with robust standard deviation (HC3)	0.059***	0.997	0.447	0.0003***	0.127**	3.931	2.595
	IRWLS	0.044***	0.982	0.849	0.0002***	0.127*	2.741	2.923
	Quantile regression	0.056**	0.996	0.651	0.0002*	0.142*	2.445	2.197

Obs: for (p) *<5%, **<1%, ***<0.1%. IRWLS reported as robust R². Quantile regression reported as pseudo-R².

Table 6 - Independent sample t-tests

Compared variables	Group	Average difference (°C)	Levene test (p)	Welch two sample t-test (p)	Cohen d effect size	Power (α = 0.05)	Yuen’s test on trimmed means for dependent samples (p) Tr = 0.2 with bootstrap	Wilcoxon signed rank test with continuity correction (p)
T min open - T min canopy	+ 20 metres to light source	-0.78	0.566	0.015	0.30	0.981	0.027	2.2E-16
D = (T max open - T min open) - (T max canopy - T min canopy)	+ 20 metres to light source	1.17	0.203	0.013	-0.31	0.986	0.000	2.2E-16
	Rural	0.87	0.697	0.040	-0.22	0.841	0.072	2.2E-16

Regarding the other multiple regression model, the tests with non-linear models indicated that the squared variable for canopy cover was more strongly correlated with the difference between the daily temperature ranges in the open area and the area under the canopy (D). Therefore, canopy cover was transformed in this way before running the linear regression. Versions including the linear and quadratic transformations of the canopy cover in the same equation did not improve the regression, and thus, the quadratic transformation was used alone. According to the model, a tree with 5% canopy cover would have a daily temperature range that was 0.0075°C (OLS) or 0.0050°C (IRWLS and quantile regression) narrower under the canopy than in the open; furthermore, these values changed to 0.75°C (OLS) or 0.5°C (IRWLS and quantile regression) for a tree with 50% canopy cover and 2.71°C (OLS) or 1.80°C (IRWLS and quantile regression) for a tree with 95% canopy cover. For each metre between the thermometer and the edge of the canopy, the daily temperature range was 0.127°C (OLS and IRWLS) or 0.142°C (quantile regression) narrower than that in the open. Comparing the standardized coefficients, the canopy cover had more influence than did the distance to the edge of the canopy on the difference between the daily temperature range in the open area and the area under the canopy in OLS (even including correction for robust standard deviation); however, this effect gradually decreased in IRWLS and decreased even more in quantile regression. In OLS, no outlier was detected using Cook's distance; however, four outliers were detected using leverage values, influencing 12.5% of the model results. Theoretically, IRWLS and quantile regression would decrease the influence of these outliers in the model and compensate for this bias.

Table 6 presents the results for the t-tests with independent samples. Although Levene's tests indicated that the variance should be homogeneous for the three tests, the Welch adjustment was still applied to all tests as an extra precaution. The three tests had small effect sizes (Cohen "d"); however, they were significant at the 95% confidence level for the parametric and non-parametric versions of the tests. However, the influence of urbanization on D was not significant in the robust version of the test, demanding higher caution when making inferences from this statistical effect. In measurements that were at least 20 metres from light sources, the daily minimum temperatures under the canopy were 0.78°C higher and the temperature ranges under the canopy were 1.17°C narrower than the respective values in the open area.

The difference between the daily temperature range in the open and under the canopy was 0.87°C higher in rural areas than the values in urban areas.

Therefore, all the null hypotheses presented in Table 2 can be rejected with at least 95% confidence using the parametric and non-parametric models. The robust models also supported the same patterns with 95% confidence, except for the influence of urbanization on the temperature range (D). Over a broad spectrum, the modelling results corroborated the respective alternative hypotheses from Table 2. All tests presented a power higher than 0.8, which was advised by Hair et al. (2018) as a cautious threshold that could be used to avoid type II errors. The high power of the tests also indicated a good balance in the experimental design regarding the sample size, effect size, and control of type I and II errors.

Discussion

Underlying processes

Table 7 synthesizes the underlying processes that could have caused the effects described in the models. These processes can be discussed based on how Carlson and Boland (1978) developed a surface heat flux/temperature model to compare the canopy effects in urban and rural environments. They proposed that the main factor controlling the contrasting results between urban and rural environments is the ground moisture availability and ground conductance (thermal inertia). According to these authors, the ground moisture availability in cities is usually lower because of reduced areas in evapotranspiration, intensifying the daytime urban heat island effect. Ground conductance tends to be higher in urban areas because structures and artificial surfaces retain heat during the day and release it during the night, increasing night-time minimum temperatures. These processes show the relevance of evapotranspiration and lateral heat emissions (including horizontal radiation and wind advection), which were used for the explanation of many of the patterns described in Table 7. An atmospheric heat flow model proposed by Shashua-Bar and Hoffman (2003) demonstrated how the tree canopy changes the microclimate of urban areas through heat-pump and windbreak effects. The model resulted in general cooling and a decrease in temperature variation (consistent with Table 7), especially during daytime. Heat pumping through tree evapotranspiration (evaporative cooling) was also emphasized by Lewis (1998) as a relevant factor associated with decreasing ground

temperatures, measured in forested and deforested areas.

Table 7 - Patterns described in the models and possible explanations of underlying causes

Phenomenon	Explanation
Lower minimum temperatures under canopy	Night-time heat pump through evapotranspiration; microscale protection from heat island effects in urban areas.
Lower maximum temperatures under canopy; proportionally lower as the canopy cover increases	Shade and daytime heat pump through evapotranspiration
Lower maximum temperatures under canopy in rural areas than in open areas	Less lateral heat emission and light reflection from buildings and other sources during daytime in rural areas
A smaller temperature range under canopy; proportionally smaller as the canopy cover and width increase	Canopy and litter insulating (buffering) the surface from temperature extremes
A smaller temperature range under canopy in rural areas	Less lateral heat emission and light reflection from buildings and other sources during daytime in rural areas
Proximity to a light source increases the minimum daily temperature under canopy	The light source may provide a lateral heat flux to places under the canopy
Proximity to a light source decreases the buffering effect of canopy on the temperature range	The light source may provide a lateral heat flux to places under the canopy

Lower temperatures under tree canopies than on unshaded ground in urban environments were previously reported by Mascaró and Mascaró (2009). Robinette (1972) reported that the tree canopy controls solar radiation and increases air humidity, thus decreasing temperature variation and reducing the thermal amplitude under the vegetation, consistent with the patterns in Table 7. Both Robinette (1972) and Mascaró and Mascaró (2009) also reported that the canopy effect was more effective during the summer because the leaf density and tree evapotranspiration were more intense. Both also found that the canopy effect of groups of trees was cumulatively more intense than that of isolated trees, especially when the various layers of the canopy amplified the absorption of solar radiation and the stratification of air temperature under the vegetation.

Various studies using remote sensing data supported the phenomena that tree removal increases average and maximum temperatures¹ in

both urban (Nichol, 1996; Elmes et al., 2017; Geene and Milward, 2017) and non-urban areas (Godinho et al., 2016) (the latter study also focused on oaks, as this study did), and the findings are consistent with the patterns observed in this study. An exception would be the trees in snow-covered areas, where the trees have a lower albedo and thus absorb more light energy than the bare snow (Bonan et al., 1992). Li et al. (2015), based on remote sensing analysis, proposed that the balance between evaporative cooling and albedo warming of trees would result in higher net cooling at lower latitudes that would decrease towards higher latitudes until reaching negative net values in boreal forests of snow-covered areas. In temperate zones, where most of the measurements of the current study were taken, Li et al. (2015) and Alkama and Cescatti (2016) proposed that the net cooling would be higher in summer and negative during winter (because of snow). However, as the measurements of this experiment were taken in autumn (Northern

¹ It is necessary, nevertheless, to keep in mind that remote sensing studies would mainly reflect surface temperatures, not air temperatures, and would mainly reflect the top of canopy temperatures, not those under canopy. For example, the influence of canopy on the land cover temperature could be very significant in a

remote sensing image, yet evapotranspirational effects would be weakly observable because they are mixed through the air volume.

Hemisphere) and spring (Southern Hemisphere), these snow-covered areas would have been restricted to high latitudes; therefore, the positive net cooling is consistent with these remote sensing analyses. Alkama and Cescatti (2016) also found that removing tree cover in areas/periods without snow increased maximum air temperatures and amplified diurnal temperature variation, consistent with the patterns in Table 7.

Model reliability

The reliability of these statistical models may be higher in the regions that had higher concentrations of sampling sites, i.e., primarily in the United Kingdom and secondarily in Europe. Therefore, further studies should include a more balanced distribution of measurements around the globe to verify the extent at which these patterns are uniform or differ within distinct regions. The same models as presented in this article were tested using geographical subsets (i.e. only samples from the Northern Hemisphere, from Europe, or from the United Kingdom); however, the changes in the effect sizes, R^2 values, and p-values were so small that they did not justify the removal of the other measurements, especially considering the trade-off with the reduction in spatial coverage. In the same way, model runs that used subsets that excluded shrubs, as well as excluding measurements on artificial ground surfaces, resulted in very small changes that did not justify their exclusion.

Overall, the leverage values of distant measurements in this study were not higher than those in the United Kingdom. Additionally, the outliers detected by the leverage values in the multiple regressions did not include shrubs or samples measured on artificial ground surfaces. Four of these outliers were in the United Kingdom, two were in continental Europe, and none were outside Europe. This result may indicate that the patterns and processes found in this study are robust in many environmental situations; however, there is still a clear need for further research.

Conclusions

The paired t-tests showed a significant effect of tree canopy; specifically, the tree canopy decreased the lower maximum and minimum daily temperature measurements, and it shortened the daily temperature range. The multiple regressions showed that the percentage of canopy cover significantly affected the difference between the maximum temperatures in the open areas and the areas under the canopy, as well as the difference between the temperature ranges. The regressions also showed that urbanization decreased the effect of the canopy on the maximum daily temperatures;

additionally, the distance to the edge of the canopy was positively correlated with the difference between the daily temperature ranges in the open areas and the areas under the canopy. However, the low R^2 values for both regression models indicated that there may be a large effect of other environmental variables that influence temperature. The independent sample t-tests showed that the distance from the light source was significant in decreasing the difference between the minimum daily temperatures in the open areas and the areas under the canopy; however, the distance from the light source increased the difference between the daily temperature ranges. The independent sample t-tests also indicated that urbanized areas have lower differences between daily temperature ranges in the open areas and the areas under the canopy according to the parametric and non-parametric tests but not according to the robust test.

The main underlying processes that could explain these results include canopy shade and wind insulation, litter insulation of the ground surface, and tree heat pumps through evapotranspiration. Light sources and other heat sources, especially in urban environments, may add lateral heat transfer, which would influence the results. The results from the models and the discussion illustrate the relevant influence of trees on the microclimate, and these influences have consequences on ecosystem processes as well as on the thermal comfort in areas inhabited by humans.

This study is also a great example showing how collaborative data collection can be an effective approach to gathering information on environmental processes in a way that integrates local contexts from many parts of the globe. Most previous studies that evaluated the effects of the canopy cover on a large scale were based on remote sensing (Nichol, 1996; Elmes et al., 2017; Greene and Milward, 2017; Godinho et al., 2016; Bonan et al., 1992; Li et al., 2015; Alkama and Cescatti, 2016), and the ones based on field measurements were usually of a limited scale. The novelty of this study, which was based on a collaborative framework, was the performance of field measurements to a larger spatial extent. Nevertheless, the design also adds some uncertainty regarding the uniformity of the data collection procedures among all collaborators. To increase the uniformity, reduce the uncertainty, and allow a larger number of participants, the experimental design needed to be simplified compared with that of previous similar studies, which focused on a limited spatial scale but with a more detailed collection of weather and environmental variables

(Carlson and Boland, 1978; Lewis, 1998; Mascaró and Mascaró, 2009; Robinette, 1972). As discussed in the previous subsection, this study helped corroborate previous hypotheses on processes that have been proposed both from detailed small-scale experiments that lacked confirmation on broader spatial scales and from large-scale remote sensing studies that lacked detailed field measurement validation. Therefore, this study contributes to filling this gap as an interface between existing large-scale and small-scale experiments regarding the effect of the canopy cover on temperature.

Further modelling studies, with appropriate fieldwork designs, could investigate the joint effects of other variables on temperature together with the variables investigated in this article. These variables could include longitude, type of ground surface, tree species, tree high, distance to other tree canopies, distance to buildings (or walls) and distance to coasts.

Acknowledgements

The authors thank all the collaborators involved in data collection. Special thanks to the research team of the School of Environment, Earth and Ecosystems Sciences of The Open University of the United Kingdom for designing and managing the collaborative experiment and for the preliminary consistency analysis of the database. Thanks to Apinyapon Seingyai for the suggestions regarding robust statistical modelling.

References

- Alkama, R., Cescatti, A. 2016. Biophysical climate impacts of recent changes in global forest cover. *Science*, 351(6273): 600-604. <https://doi.org/10.1126/science.aac8083>
- Bonan, G. B., Pollard, D., Thompson, S. L. 1992. Effects of boreal forest vegetation on global climate. *Nature*, 359(6397):716. <https://doi.org/10.1038/359716a0>
- Breusch, T. S., Pagan, A. R. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47:1287-1294. <https://doi.org/10.2307/1911963>
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., West, G. B. 2004. Toward a metabolic theory of ecology. *Ecology*, 85(7):1771-1789. <https://doi.org/10.1890/03-9000>
- Carlson, T. N., Boland, F.E. 1978. Analysis of Urban-Rural Canopy Using a Surface Heat Flux/Temperature Model. *J Appl Meteorology*, 17(7):998-1013. [https://doi.org/10.1175/1520-0450\(1978\)017<0998:AOURCU>2.0.CO;2](https://doi.org/10.1175/1520-0450(1978)017<0998:AOURCU>2.0.CO;2)
- Cook, R. D., Weisber, S. 1982. Residuals and influence in regression. Chapman & Hall, New York.
- Cragg, J. G., Uhler, R. S. 1970. The demand for automobiles. *The Canadian J Economics*, 3:386-406. <https://doi.org/10.2307/133656>
- Davies, S., Gowing, D. 2016. S369 Practical Task Guide. The Open University, Milton Keynes.
- Durbin, J., Watson, G. S. 1971. Testing for serial correlation in least squares regression. III. *Biometrika*, 58(1):1-19. <https://doi.org/10.2307/2334313>
- Elmes, A., Rogan, J., Williams, C., Ratick, S., Nowak, D., Martin, D. 2017. Effects of urban tree canopy loss on land surface temperature magnitude and timing. *ISPRS J Photogrammetry Remote Sens*, 128:338-353. <https://doi.org/10.1016/j.isprsjprs.2017.04.011>
- Godinho, S., Gil, A., Guiomar, N., Costa, M. J., Neves, N., 2016. Assessing the role of Mediterranean evergreen oaks canopy cover in land surface albedo and temperature using a remote sensing-based approach. *Appl Geography*, 74:84-94. <https://doi.org/10.1016/j.apgeog.2016.07.004>
- Gowing, D. J. G., Davies, S. J. M., Denny, H., Edwards, N. R., Gauci, V., Gillman, M. P., Halliday, T. R., Stevens, C. J., 2008. Ecosystems. The Open University, Milton Keynes.
- Greene, C. S., Millward, A. A. 2017. Getting closure: The role of urban forest canopy density in moderating summer surface temperatures in a large city. *Urban Ecosyst* 20(1):141-156. <https://doi.org/10.1007/s11252-016-0586-5>
- Hair, J. F., Anderson, R. E., Tatham, R. L., Black, W. C. 2018. Multivariate Data Analysis (8th ed.). Cengage Learning EMEA, Hampshire.
- Henson, R. K., Smith, A. D. 2000. State of the art in statistical significance and effect size reporting: A review of the APA Task Force report and current trends. *J Res Development Education*, 33(4):285-296.
- Holland, P., Welsch, R. E. 1977. Robust regression using iteratively reweighted least-squares. *Comm Statist Theory Meth*, 6:813-888. <https://doi.org/10.1080/03610927708827533>
- King, G., Roberts, M. E. 2014. How robust standard errors expose methodological problems they do not fix, and what to do about it. *Political Analysis*, 23(2):159-179. <https://doi.org/10.1093/pan/mpu015>
- Koenker, R., Bassett, G., 1978. Regression Quantiles. *Econometrica*, 46(1):33-50. <https://doi.org/10.2307/1913643>
- Koller, M., Stahel, W.A., 2011. Sharpening Wald-type inference in robust regression for small samples. *Computational Statistics Data* 3285

- Analysis, 55(8):2504–2515.
<https://doi.org/10.1016/j.csda.2011.02.014>
- Levene, H., 1960. Robust test for equality of variances. In: Olkin I, Ghurye SG, Hoëffding W, Madow WG, Mann HB (eds.) Contributions to Probability and Statistics: Essays in Honour of Harold Hotelling. Stanford University Press, Stanford, pp 278–292.
- Lewis, T., 1998. The effect of deforestation on ground surface temperatures. *Glob Planet Change*, 18(1-2):1-13.
[https://doi.org/10.1016/S0921-8181\(97\)00011-8](https://doi.org/10.1016/S0921-8181(97)00011-8)
- Li, Y., Zhao, M., Motesharrei, S., Mu, Q., Kalnay, E., Li, S., 2015. Local cooling and warming effects of forests based on satellite observations. *Nat communications*, 6:6603.
<https://doi.org/10.1038/ncomms7603>
- Mascaró, L., Mascaró, J. J. 2009. *Ambiência Urbana – Urban environment*. Masquatro Editora, Porto Alegre.
- Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determination. *Biometrika*, 78:691-692.
<https://doi.org/10.1093/biomet/78.3.691>
- Nichol, J. E. 1996. High-resolution surface temperature patterns related to urban morphology in a tropical city: A satellite-based study. *J Appl Meteorology*, 35(1):135-146.
[https://doi.org/10.1175/1520-0450\(1996\)035<0135:HRSTPR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1996)035<0135:HRSTPR>2.0.CO;2)
- Pituch, K. A., Stevens, J., 2015. *Applied multivariate statistics for the social sciences* (6th ed.). Routledge.
- Renaud, O., Victoria-Feser, M. P. 2010. A robust coefficient of determination for regression, *J Statistical Plan Inference*, 140, 1852-1862.
<https://doi.org/10.1016/j.jspi.2010.01.008>
- Ricklefs, R. E., Relyea, R. 2019. *Ecology: The economy of nature* (8th ed.). Freeman, W. H. & Company.
- Robinette, G. O., 1972. *Plants, people, and environmental quality: A study of plants and their environmental functions* (1st ed.). U.S. Dept. of the Interior, National Park Service, Washington DC.
- Salmund, J. A., Tadaki, M., Vardoulakis, S., Arbuthnott, K., Coutts, A, Demuzere, M., Dirks, K. M., McInnes, R. N., Wheeler, B. W., 2016. Health and climate related ecosystem services provided by street trees in the urban environment. *Environmental Health*, 15(1)S36:95-107.
<https://doi.org/10.1186/s12940-016-0103-6>
- Shapiro, S. S., Wilk, M. B., 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591-611.
<https://doi.org/10.2307/2333709>
- Shashua-Bar, L., Hoffman, M. E. 2003. Geometry and orientation aspects in passive cooling of canyon streets with trees. *Energy and Buildings*, 35(1):61–68. [https://doi.org/10.1016/s0378-7788\(02\)00080-4](https://doi.org/10.1016/s0378-7788(02)00080-4)
- Smith, D. M., Jarvis, P. G., 1998. Physiological and environmental control of transpiration by trees in windbreaks. *For Ecology Management*, 105(1):159-173. [https://doi.org/10.1016/S0378-1127\(97\)00292-2](https://doi.org/10.1016/S0378-1127(97)00292-2)
- Tyrväinen, L., Pauleit, S., Seeland, K., Vries, S., 2005. Benefits and uses of urban forests and trees. In Schipperijn J (ed), *Urban Forests and Trees*. Springer-Verlag, Netherlands, Chapter 4, pp 81-114.
- Welch, B. L., 1938. The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29, 350-362. <https://doi.org/10.2307/2332010>
- White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838. <https://doi.org/10.2307/1912934>
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics*, 1:80-83.
<https://doi.org/10.2307/3001968>
- Yohai, V. J., Zamar, R. H., 1988. High breakdown-point and high efficiency estimates for regression. *Journal of the American Statistical Association*, 83(402):406-413.
<https://doi.org/10.2307/2288856>
- Yuen, K. K., 1974. The two-sample trimmed t for unequal population variances. *Biometrika*, (1):165-170.
<https://doi.org/10.2307/2334299>
- Vasconcelos, V. V. 2021a. Canopy cover reference sheet. Feb.
<http://dx.doi.org/10.13140/RG.2.2.25874.58566>
- Vasconcelos, V. V. 2018. Influence-of-canopy-on-temperature, r-code.
<https://github.com/influence-of-canopy-on-temperature/r-code>
- Vasconcelos, V. V. 2021b. Dataset and code for Influence of canopy cover on surface temperature. <https://doi.org/10.13140/RG.2.2.11613.95208>