Data in Brief
Open access
Article template

ARTICLE INFORMATION

Article title

Microclimate and canopy effects across forest–grassland ecotones: A dataset from Auckland, New Zealand

Authors

Yuehan Luo*, Peter Yue, Lachlan Reid, Michael Myers.

Affiliations

School of Environment, University of Auckland.

Corresponding author's email address

yluo862@aucklanduni.ac.nz

Keywords

understory conditions; temperature buffering; soil volumetric water content; photosynthetically active radiation (PAR); open environmental data

Abstract

This dataset was generated during the ENVSCI 708 Collaborative Research Project at the University of Auckland. Microclimate and environmental measurements were obtained from 12 fixed sites, comprising six forest and six grassland plots. Variables recorded include soil temperature (°C), atmospheric/surface temperature (°C), canopy cover (%), light intensity (photosynthetically active radiation, µmol m^{-2} s^{-1}), and soil volumetric water content (%). Additional data include soil CO_2 efflux before and after rainfall (g CO_2 m^{-2} h^{-1}), greenhouse gas concentrations (CO_2, CH_4, N_2O; ppm), soil pH, conductivity (µS cm^{-1}), organic matter (%), soil texture class, tea bag decomposition loss (%), and rainwater chemistry (volume, pH, conductivity). Measurements were conducted between 11:00 and 14:00 using handheld thermometers, a Hydrosense II probe (model 4933), a LI-250A PAR sensor, and hemispherical canopy photographs. All data are collated in a single Excel file with standardized variable names and units. R scripts used for data cleaning and figure generation accompany the dataset to support reuse. The dataset is hosted in the ENVSCI 708 SharePoint repository for open access and can be applied in ecological modelling, cross-site comparisons, and educational training.

SPECIFICATIONS TABLE

Subject	Earth & Environmental Sciences
Specific subject area	Ecosystem microclimate buffering and understory stability in response to canopy cover
Type of data	**Type of data:** Tables; Graphs; Figures

	Data format: Raw; Processed; Analysed
Data collection	Microclimate data were collected at 12 fixed sites (six forest, six grassland) on the University of Auckland City Campus. Soil temperature was measured with a digital probe thermometer, soil moisture with a Hydrosense II (model 4933), and photosynthetically active radiation (PAR) with a LI-250A sensor. Canopy cover was quantified from hemispherical photographs. Each variable was measured twice per site between 11:00 and 14:00 under stable weather conditions, and outliers were excluded during quality control.
Data source location	Data were collected at 12 fixed sampling sites (six forest and six grassland) on the University of Auckland City Campus, Sector 100, Auckland, New Zealand (approx. 36.852°S, 174.769°E). The dataset is stored in the ENVSCI 708 SharePoint repository, School of Environment, University of Auckland.
Data accessibility	The dataset supporting this article is publicly available in the University of Auckland SharePoint repository: **ENVSCI 708_Data_Collaborative Research Project.** URL (anonymous access): https://uoa-my.sharepoint.com/:x:/g/personal/yluo862_uoa_auckland_ac_nz/EaVfOEUT4O1Hr8roUzypy0EBjr3zfdnajGurb0IeRcOJXQ?e=saWB7E

VALUE OF THE DATA

- These data provide paired measurements of canopy cover, light intensity (PAR), soil temperature, soil volumetric water content, and atmospheric conditions across 12 fixed forest and grassland sites in Auckland, New Zealand. Such combined measurements are rarely available for forest-grassland ecotones and allow detailed exploration of understory microclimate dynamics.
- The dataset can be reused by researchers examining canopy buffering effects, habitat contrasts, and ecosystem resilience. It is suitable for ecological modelling, cross-site comparisons, and integration with global databases such as SoilTemp [1] to improve understanding of microclimate heterogeneity under climate change.
- The standardized methodology across 12 sites (six forest, six grassland) allows replication and educational use. The dataset is valuable for environmental science teaching, offering students opportunities to practice statistical analysis, visualization, and cross-ecosystem comparisons using real-world field data.
- Supplementary variables measured by other project teams (e.g., soil CO_2 efflux, greenhouse gas concentrations) are also available within the shared repository. While not collected by our team, these data can be linked to our microclimate measurements for cross-disciplinary studies connecting canopy cover and microclimate conditions with soil processes and ecosystem functioning.

- Land managers, restoration practitioners, and conservation planners can use the combined dataset to evaluate how canopy and habitat structure influence microclimate stability and soil water balance, informing strategies for biodiversity protection and climate adaptation.

BACKGROUND

Microclimate is increasingly recognized as a key driver of ecological processes because it mediates species performance, ecosystem functioning, and responses to climate change [2,3]. Forest canopies buffer understory environments by reducing temperature extremes, regulating humidity, and altering light conditions compared with open habitats [4,5]. Edge environments in forest–grassland ecotones also exhibit pronounced microclimate shifts, with canopy structure strongly influencing soil moisture and temperature gradients [6]. Quantifying canopy–microclimate relationships is widely applied using standardized measurements to provide information on biodiversity responses and ecosystem management.

At the methodological level, standardized measurements of soil temperature, humidity, soil moisture, and photosynthetically active radiation (PAR) are widely used to capture fine-scale microclimate variation [1,7]. Pairing these with canopy cover estimates provides insights into how structural habitat differences shape near-surface environments.

In addition to these core microclimate variables, supplementary datasets from other project teams (e.g., greenhouse gas emissions and soil chemistry) are available within the shared repository. These provide potential opportunities for reuse in cross-disciplinary studies that connect canopy cover and microclimate conditions with soil processes and ecosystem functioning. This article reports only the five microclimate variables measured by our group (soil temperature, surface temperature, canopy cover, PAR, and soil volumetric water content).

DATA DESCRIPTION

The dataset is available in the repository [ENVSCI 708_Data_Collaborative Research Project] (SharePoint). It is provided as a single Excel file that contains raw and processed measurements from 12 fixed sampling sites (six forest and six grassland).

The dataset is structured into the following categories:

- Site information: Site ID (forest or grassland/lawn), GPS coordinates, and sampling date.
- Microclimate variables (Sensor 1): Soil temperature (°C), atmospheric/surface temperature (°C), canopy cover (%), light intensity ($\mu mol/m^2/s$, PAR), and soil volumetric water content (%).
- Supplementary environmental variables: Soil CO_2 efflux before and after rain (g $CO_2/m^2/hour$), greenhouse gas concentrations (CO_2, CH_4, N_2O in ppm), soil conductivity (µS/cm), soil pH (mean of three replicates), organic matter percentage, and soil texture classification.
- Decomposition variables: Rooibos mass lost (%) and green mass lost (%) from litter bag experiments.
- Rainwater chemistry: Rain volume (ml), pH, conductivity (µS/cm), and associated metadata.

Table 1 summarises all variables, units, and methods. A study area map with sampling locations is provided (Figure 1) and GPS coordinate records are visualized (Figure 2). Microclimate measurements collected by Sensor 1 are compared between forest and grassland habitats (Figure 3). Canopy cover is presented together with microclimate variables (Figures 4-5). A cross-team linkage between canopy cover and greenhouse gas concentrations is also included (Figure 6).

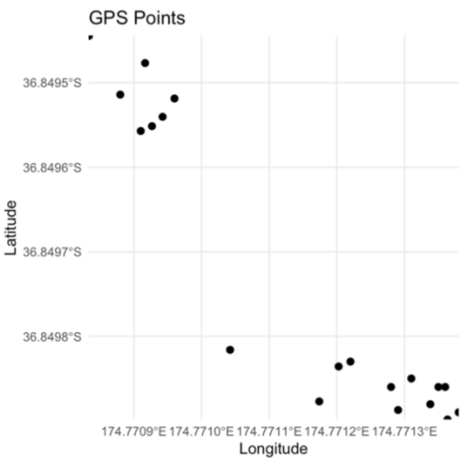
Table 1 Description of variables, units, and measurement methods

Variable name	Unit	Data type	Description
Week of data collection	–	Categorical	Week 4 or Week 5
Name of location	–	Categorical	Sampling site code with identified lawn (grassland) or forest
GPS location	Degrees (°)	Text	Location recorded at each sampling site
Date	DD/MM/YYYY	Date	Sampling date
Soil temperature	°C	Numeric	Measured with digital soil thermometer
Atmospheric/Surface temperature	°C	Numeric	Measured at ground surface
Canopy cover	%	Numeric	Estimated from hemispherical photographs
Light intensity (PAR)	µmol/m²/s	Numeric	Recorded with LI-250A quantum sensor
Soil volumetric water content	%	Numeric	Measured with Hydrosense II (model 4933)
Soil CO_2 efflux (before rain)	g CO_2/m²/hour	Numeric	Measurement by other team
Soil CO_2 efflux (after rain)	g CO_2/m²/hour	Numeric	Measurement by other team
Greenhouse gas concentrations	ppm	Numeric	CO_2, CH_4, N_2O concentrations (gas analyser) by other team
Soil conductivity	µS/cm	Numeric	Mean of three replicate measurements by other team
Soil pH	–	Numeric	Mean of three replicate measurements by other team
Organic matter	%	Numeric	Estimated from soil colour method by other team

Soil texture	Class	Categorical	Field-based soil classification by other team
Rooibos mass lost	%	Numeric	Decomposition bag experiment (rooibos tea) by other team
Green mass lost	%	Numeric	Decomposition bag experiment (green tea) by other team
Rain volume	ml	Numeric	Collected rainfall sample volume by other team
Rain pH	–	Numeric	pH of rainwater sample by other team
Rain conductivity	µS/cm	Numeric	Conductivity of rainwater sample by other team

Figure 1 Study area map showing sampling locations within the study site. Red dots indicate grassland (open lawn) samples and orange dots indicate tree-covered (forest) samples. The base map is an aerial image with north arrow and scale bar included for orientation.

Figure 2 GPS coordinates of sampling sites recorded during fieldwork.



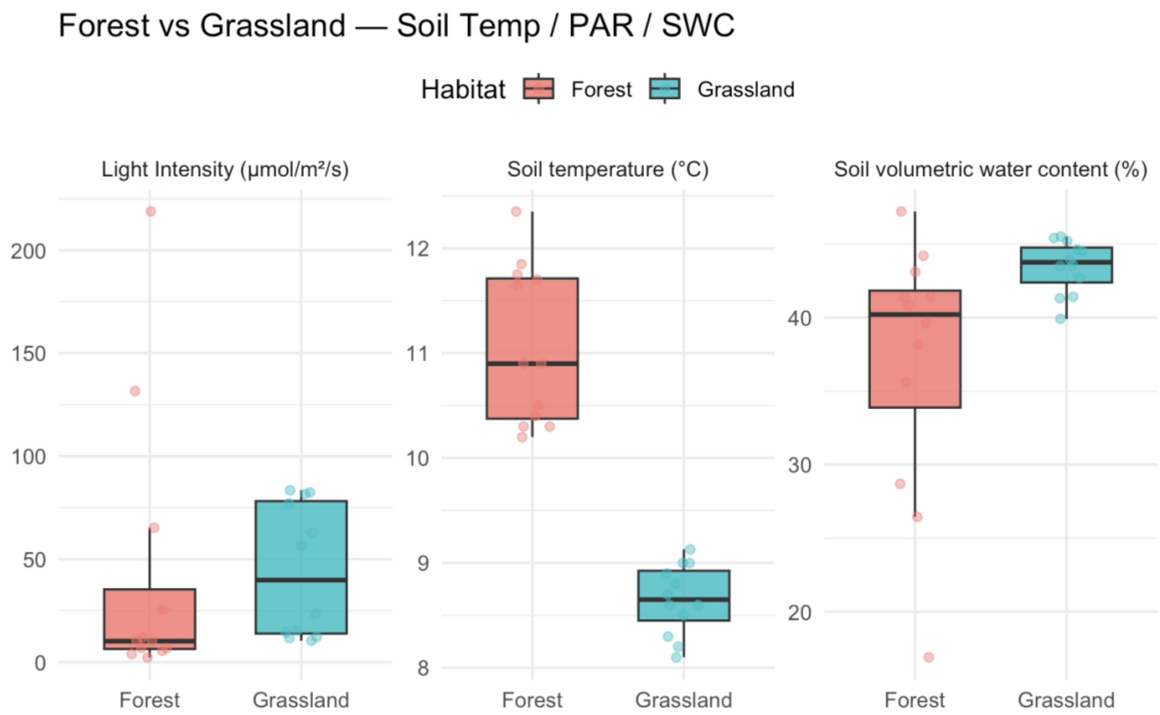


Figure 3 Boxplots comparing key microclimate variables between forest and grassland habitats. Light intensity (PAR), soil temperature (°C), and soil volumetric water content (%). Boxes show interquartile ranges, thick lines represent medians, and jittered points show individual observations.

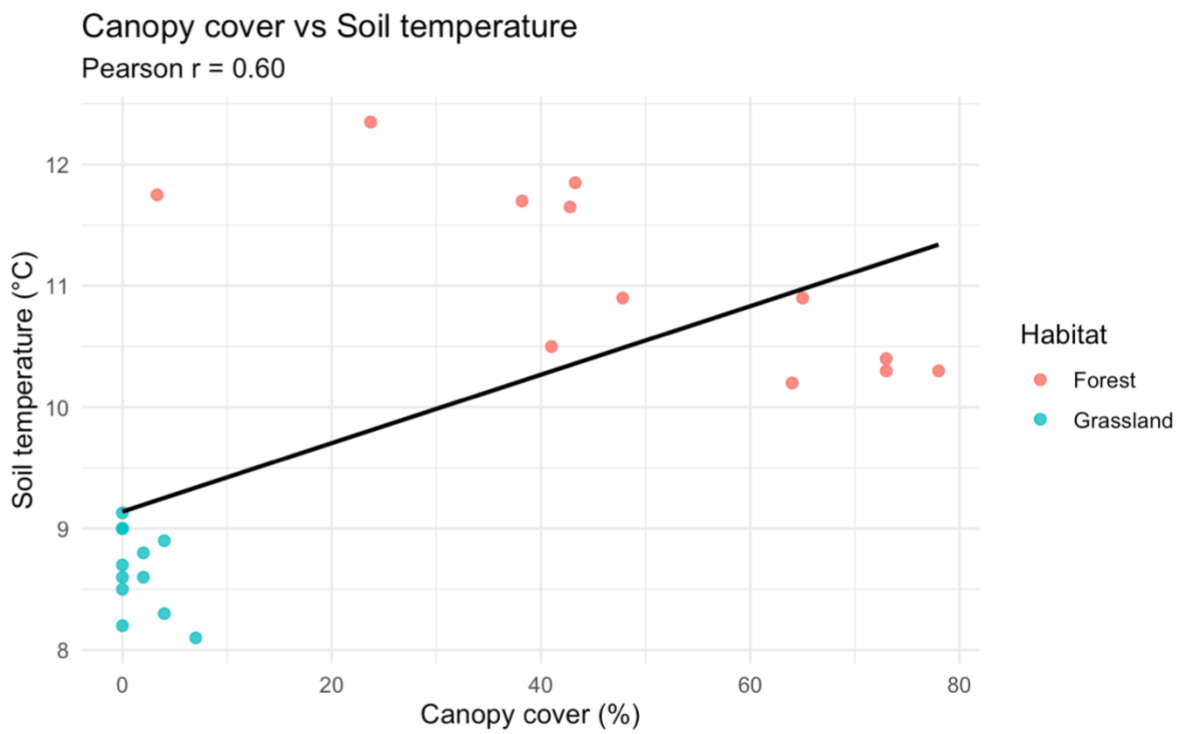


Figure 4 Relationship between canopy cover (%) and soil temperature (°C) across forest and grassland sites. Points are coloured by habitat (red = forest, blue = grassland). The solid line shows a fitted linear regression, with Pearson correlation coefficient r = 0.60.

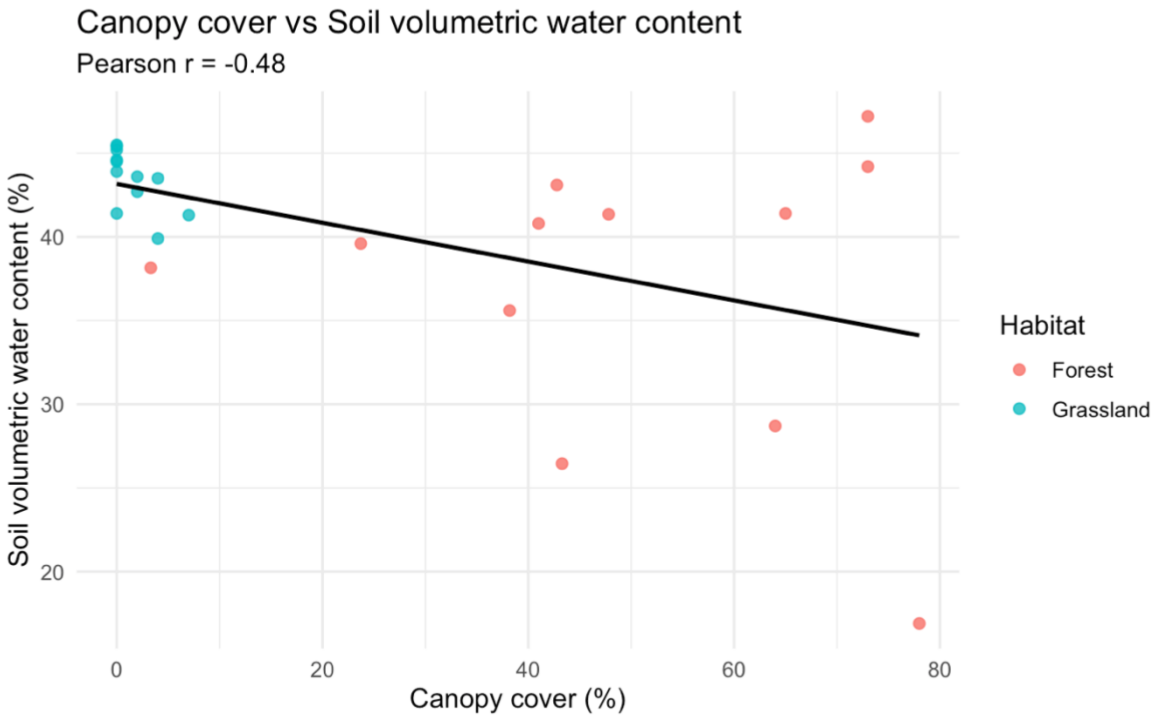


Figure 5 Relationship between canopy cover (%) and soil volumetric water content (%) across forest and grassland sites. Points are coloured by habitat (red = forest, blue = grassland). The solid line represents a fitted linear regression, with Pearson correlation coefficient r = –0.48.

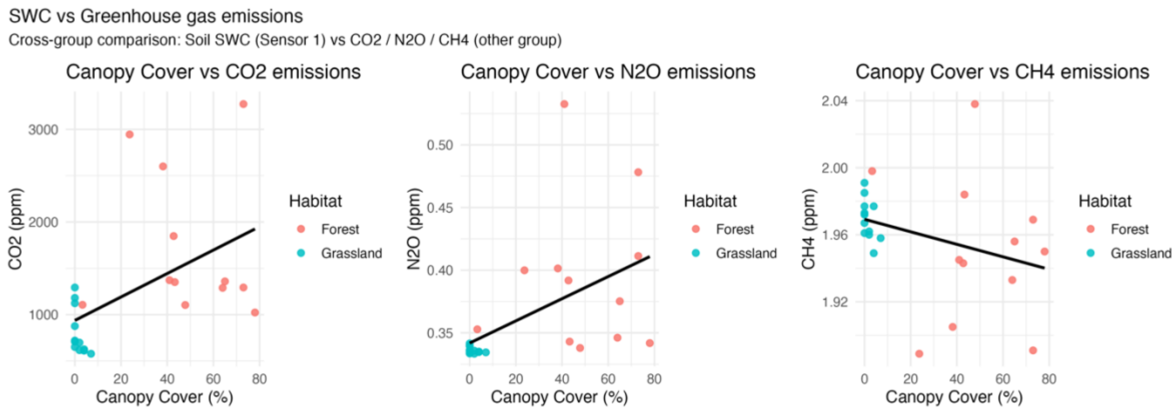


Figure 6 Relationships between canopy cover (%) and greenhouse gas concentrations (CO_2, N_2O, CH_4; ppm) obtained from cross-team measurements. Points denote site-level observations coloured by habitat (red = forest, blue = grassland). Solid lines represent fitted linear regressions for visual interpretation.

EXPERIMENTAL DESIGN, MATERIALS AND METHODS

Study location and design

Sampling was carried out during Weeks 4-5 of the ENVSCI 708 Collaborative Research Project at the University of Auckland campus grounds. Two contrasting habitats were selected: tree-covered forest and open grassland (lawn). At each habitat, multiple sampling points were established. GPS coordinates were recorded for all sites. Sampling points were georeferenced using handheld GPS devices. As site markers could not be retained between weeks, locations were re-identified by coordinates, introducing small positional deviations. All measurements were collected between 11:00 and 14:00 to minimise diurnal variability.

Soil and microclimate measurements

- Soil volumetric water content (SWC, %): measured with a Hydrosense II probe (model 4933). The probe rods were fully inserted into the soil (~ 5 cm depth) until readings stabilised.
- Light intensity (PAR, $\mu mol\ m^{-2}\ s^{-1}$): measured at ground level using a LI-250A quantum sensor (LI-COR Biosciences) [8]. The sensor was placed directly on the soil surface, and the 15-second average was recorded.
- Canopy cover (%): quantified using the Canopy Cover smartphone application, with images taken vertically upward at each sampling location. The app algorithm automatically calculated canopy cover percentage.
- Soil temperature (°C): recorded at ~ 5 cm depth using digital soil thermometers. The probe was fully inserted into the soil, left in place until the reading stabilised, and maintained in position during measurement to avoid artefacts.
- Atmospheric/Surface temperature (°C): measured with handheld thermometers at 1 m above the soil surface.

Data handling and quality control

All measurements were recorded in the shared project Excel repository (ENVSCI 708_Data_Collaborative Research Project.xlsx). Data were screened for consistency, unit conformity, and missing values. Column names containing irregular spaces or symbols were standardised before analysis. At each sampling point, two replicate measurements were taken and averaged to reduce instrument noise. Outliers were identified using a 1.5 × interquartile range (IQR) rule and excluded prior to analysis.

Integration of supplementary data

In addition to the five core microclimate variables measured by our team (soil temperature, atmospheric/surface temperature, canopy cover, PAR, and soil volumetric water content), supplementary variables (e.g., soil CO_2 efflux, greenhouse gas concentrations) were contributed by other project groups. These were incorporated into the shared Excel dataset with standardized variable names, units, and metadata. Cross-team integration allowed the generation of comparative figures (e.g., canopy cover vs. greenhouse gas emissions) while ensuring that each dataset could still be independently identified by variable name and measurement source.

Software and code

Data analysis and figure generation were performed in R (version 4.4.3) [9]. The following packages were used:

- tidyverse for data cleaning and wrangling
- readxl for Excel import
- ggplot2 for figure generation
- patchwork for multi-panel figure assembly
- ggdist for raincloud plots
- ggrepel for point labelling

All scripts are provided in /code/ for reproducibility.

LIMITATIONS

Several limitations should be noted regarding the dataset. First, the number of sampling points per habitat (forest and grassland) was restricted due to time constraints during fieldwork, which limits spatial representativeness [10]. Second, all measurements were collected during daytime (11:00–14:00), meaning that diurnal variability in soil temperature, light intensity, and soil water content was not captured. Third, environmental conditions such as recent rainfall and cloud cover may have influenced individual measurements, introducing short-term variability. Fourth, the use of handheld instruments and a smartphone application, while practical, may be subject to small measurement inaccuracies compared with laboratory-grade equipment [11]. Fifth, ecotone environments are inherently heterogeneous at fine spatial scales, with steep gradients in light, soil moisture, and temperature across short distances. This small-scale variability may not be fully captured by the limited number of sites in this dataset, which is a common challenge in ecotone research [12]. Sixth, small positional deviations occurred because handheld GPS accuracy was limited and site markers could not be retained between weeks, requiring relocation of sampling points. Finally, minor inconsistencies in data entry were observed in the shared Excel file (e.g., trailing spaces or unit formatting), which required standardisation during data curation.

CRediT AUTHOR STATEMENT

Yuehan Luo: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing – Original draft, Writing – Review & Editing, Project administration.

Peter Yue: Data curation, Investigation, Methodology, Conceptualization, Resources.

Lachlan Reid: Conceptualization, Investigation, Methodology.

Michael Myers: Conceptualization, Investigation, Methodology.

ACKNOWLEDGEMENTS

The authors thank Luitgard Schwendenmann for supervision and guidance throughout the ENVSCI 708 Collaborative Research Project. We also acknowledge the support of the Waipapa Taumata Rau, University of Auckland for providing field equipment and learning resources.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] J.J. Lembrechts, J. Aalto, M.B. Ashcroft, P. De Frenne, M. Kopecký, J. Lenoir, et al., SoilTemp: A global database of near-surface temperature. *Global Change Biology, 26*(11), 6616–6629. https://doi.org/10.1111/gcb.15123

[2] F. Zellweger, P. De Frenne, J. Lenoir, P. Vangansbeke, K. Verheyen, M. Bernhardt-Römermann, et al., Forest microclimate dynamics drive plant responses to warming. *Science* 368 (2020) 772-775. https://doi.org/10.1126/science.aba6880

[3] P. De Frenne, F. Zellweger, F. Rodríguez-Sánchez, B.R. Scheffers, K. Hylander, M. Luoto, *et al.* Global buffering of temperatures under forest canopies. *Nat Ecol Evol* 3 (2019) 744–749. https://doi.org/10.1038/s41559-019-0842-1

[4] J. Chen, S.C. Saunders, T.R. Crow, R.J. Naiman, K.D. Brosofske, G.D. Mroz, et al., Microclimate in Forest Ecosystem and Landscape Ecology: Variations in local climate can be used to monitor and compare the effects of different management regimes, *BioScience*, Volume 49, Issue 4, April 1999, Pages 288–297, https://doi.org/10.2307/1313612

[5] P. De Frenne, J. Lenoir, M. Luoto, B.R. Scheffers, F. Zellweger, J. Aalto, et al., (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. *Global Change Biology, 27*(11), 2279–2297. https://doi.org/10.1111/gcb.15569

[6] Didham, R. K., & Lawton, J. H. (1999). Edge structure determines the magnitude of changes in microclimate and vegetation structure in tropical forest fragments. *Biotropica, 31*(1), 17–30. https://doi.org/10.1111/j.1744-7429.1999.tb00113.x

[7] Von Arx, G., Graf Pannatier, E., Thimonier, A., & Rebetez, M. (2013). Microclimate in forests with varying leaf area index and soil moisture: Potential implications for seedling establishment in a changing climate. *Journal of Ecology, 101*(5), 1201–1213. https://doi.org/10.1111/1365-2745.12121

[8] LI-COR Biosciences. LI-250A Light Meter Instruction Manual. LI-COR; 2012. https://licor.com (DOI not available; official manual).

[9] R Core Team. (2023). R: A language and environment for statistical computing. The R Project for Statistical Computing. https://www.r-project.org/ (DOI not available; software reference).

[10] Holden, Z. A., Abatzoglou, J. T., Luce, C. H., & Baggett, L. S. (2011). Empirical downscaling of daily minimum air temperature at very fine resolutions in complex terrain. *Agricultural and Forest Meteorology, 151*(8), 1066–1073. https://doi.org/10.1016/j.agrformet.2011.03.011

[11] Gałuszka, A., Migaszewski, Z. M., & Namieśnik, J. (2015). Moving your laboratories to the field—Advantages and limitations of the use of field portable instruments in environmental sample analysis. *Environmental Research, 140,* 593–603. https://doi.org/10.1016/j.envres.2015.05.017

[12] Gosz, J. R. (1993). Ecotone hierarchies. *Ecological Applications, 3*(3), 369–376. https://doi.org/10.2307/1941905

DECLARATION OF USE FOR GENERATIVE AI

I, **Yuehan Luo** by confirm that on **6 September 2025**:

1. The document I have submitted was written by me.
2. I acknowledge awareness of any updates to the generative AI tools used, up to the date of this submission. This includes AI plug-ins or assistants included in existing programs, such as Grammarly. I take responsibility for any fabricated references or factual errors stemming from the use of these tools.
3. I have informed myself of the limitations and implications of using generative AI and related technologies, including the reinforcement of biases and propensity for fabrication.
4. I have used these tools ethically, including not uploading confidential, private, personal, or otherwise sensitive information.
5. When using AI, I have ensured that the work produced is still my own and I understand that submitting output from a generative AI tool as my own is NOT acceptable. I understand that I am expected to build on the output, ensuring any submissions are that of my own ideas and knowledge.
6. To assist with maintaining academic integrity, I have appropriately acknowledged any use of generative AI in my work (list below as applicable).
7. I acknowledge that any undeclared use of generative AI will constitute academic dishonesty and will be dealt with according to relevant University policy.
8. I understand that I will be held accountable and liable for any academic misconduct that arises in breach of any relevant University policy, as well as the consequences of such infringements.

Acknowledgement of Generative AI Tools Used

Please add more rows to the table as needed so that each tool used in the creation of your assessment submission are included in the declaration.

Generative AI Tool Used (Please List Each Separately)	Purpose of Use	Briefly Explain the Extent of Use
ChatGPT (OpenAI, GPT-5)	Grammar and language editing	Used to improve grammar, clarity, and academic tone of draft text. All substantive content and interpretations were produced by the author.