**Appendix S1**

**Supplementary methods**

*Interpolating six-hourly NCEP data to hourly*

Six-hourly humidity and pressure data are interpolated to hourly using the native spline function of R (R Development Core Team, 2012). Wind speed and direction are used to derive easterly and northerly wind vectors, which are then spline-interpolated to hourly before back-calculating hourly wind speeds and directions. Hourly incoming solar radiation is derived using microclima’s ‘difprop’ function to partition total incoming shortwave radiation into its direct and diffuse components, determining atmospheric attenuation and then spline-interpolating to hourly. Hourly zenith angles are also computed. Hourly temperature is derived using microclima’s ‘hourlytemp’ function, accounting for hourly variation in emissivity and incoming radiation. Emissivity is derived at six-hourly intervals from incoming and outgoing longwave radiation and then interpolated to hourly.

*Calibrating sub-surface soil temperatures in microclima*

To permit microclima to estimate sub-surface temperatures, we introduce a new parameterisation, whereby the temperature at time *t* (*Tt*) is modelled as a function of temperature in the previous time step (*Tt-1*) and heat exchanges with the soil surface and underlying soil layer, as follows:

Here *So*is the soil surface temperature at time *t* from the linear model mentioned above, *S*2 is the soil temperature at 2 m, assumed to remain constant at mean reference temperature, and *p*1..2 are coefficients estimated by iterative fitting. Throughout, wind is treated as a continuous variable.

*Model tests*

We tested the reliability of our hourly-interpolation method for NCEP data by comparing outputs against data obtained from Culdrose (50.0825°N, 5.2486°W) and Brize Norton (51.750°N, 1.583°W) weather stations in 2010 (the year fewest missing meteorological observations). We tested the terrain and coastal adjustment procedure by calculating daily mean temperatures in January 2015 and comparing our estimates to a one km gridded dataset available from the UK Met Office (Met Office, 2018). January 2015 was selected due to the variety of weather conditions that occurred during this month.

We tested the procedure for predicting time series of soil temperature and soil moisture for the 35 Australian OzNet sites (Smith et al., 2012) used in Kearney and Maino (Kearney & Maino, 2018). We followed the same methods as in the latter study, comparing correlation coefficient *r* and the root mean square deviation *rmsd* for each of the four depths (3-4 cm, 15 cm, 45 cm and 59/75 cm) of results derived with the Australian Water Availability Project (AWAP, aka Australian Gridded Climate Data) forcing data (Jones, Wang, & Fawcett, 2009) used originally, and those created with the microclima-adjusted NCEP data. The code for these analyses is available in Appendix S2.

We tested our procedure for generating high-resolution grids of near-surface temperature against two datasets: (1) a network of iButton temperature loggers deployed across the Lizard Peninsula, UK at a height of 1 m between March 2010 and December 2010 and (2) iButton temperature loggers deployed at 5 cm above the ground in Caerthillian Valley, Lizard Peninsula, UK in May 2010 (see Maclean, Mosedale, & Bennie, 2019 for further details). We also produce side-by-side comparisons of the grids generated using our automated procedures with those generated using methods described in Maclean, Mosedale and Bennie (2019).

The example code below installs the relevant packages and runs an NCEP-based simulation of the site featured in Fig. 2:

install.packages('RNCEP')

install.packages('elevatr')

library(devtools)

install\_github('mrke/NicheMapR')

install\_github('ilyamaclean/microclima')

library(NicheMapR)

loc <- c(146.1103, -34.65478) # (longitude, latitude)

dstart <- "01/01/2007" # start date

dfinish <- "31/12/2010" # end date

DEP <- c(0, 3, 5, 10, 15, 20, 30, 50, 100, 200) # specify depths

save = 1 # save the input data (setting save = 2 uses previously saved input)

micro <- micro\_ncep(loc = loc, dstart = dstart, dfinish = dfinish, DEP = DEP, save = save)

soil <- as.data.frame(micro$soil) # retrieve soil temperatures, minimum shade

soil <- soil[micro$dates >= as.POSIXct("01/01/2008", format = "%d/%m/%Y"), ]

dates <- micro$dates[micro$dates >= as.POSIXct("01/01/2008", format = "%d/%m/%Y")]

# plot the results for 3 and 15 cm soil temperature

par(mfrow = c(2,1))

plot(soil$D3cm ~ dates, xlab = "date", ylab = "3 cm soil temperature (°C)", ylim = c(-10,70), type = "l", main = paste("3 cm soil temperature", sep = ""))

plot(soil$D15cm ~ dates, xlab = "date", ylab = "15 cm soil temperature (°C)", ylim = c(-10,70), type = "l", main = paste("15 cm soil temperature", sep = ""))

Note that setting the parameter ‘save’ to 1 will save the forcing data and setting it to 2 will run it using previously saved forcing data.

The example code below produces the NCEP-based microclimate grid featured in Fig. 3b:

require(raster)

require(microclima)

require(NicheMapR)

# Calculate leaf area index

l <- lai(aerial\_image[,,3], aerial\_image[,,4]) # leaf area

l <- lai\_adjust(l, veg\_hgt, hgt = 0.05)

x <- leaf\_geometry(veg\_hgt) # leaf angle

fr <- canopy(l, x) # canopy cover

# Calculate ground and canopy albedo

alb <- albedo(aerial\_image[,,1], aerial\_image[,,2], aerial\_image[,,3], aerial\_image[,,4])

albg <- albedo2(alb, fr)

albc <- albedo2(alb, fr, ground = FALSE)

# crop datasets

e <- extent(169400, 169700, 12400, 12700)

l <- crop(if\_raster(l, dtm1m), e)

x <- crop(x, e)

fr <- crop(if\_raster(fr, dtm1m), e)

albg <- crop(if\_raster(albg, dtm1m), e)

albc <- crop(if\_raster(albc, dtm1m), e)

dem <- crop(dtm1m, e)

# Run model over 24 hours

tout <- runauto(dem, "27/05/2010", "28/05/2010", hgt = 0.05,

l = l, x = x, albg = albg, albc = albc)

# Extract data for 13:00 hrs and convert to raster

temp13 <- if\_raster(tout$temps[,,14], dem)

e <- extent(169400, 169700, 12400, 12700)

temp13 <- crop(temp13, e)

mypal <- colorRampPalette(c("darkblue", "blue", "green",

"yellow", "orange", "red"))(255)

par(mfrow=c(1, 1))

plot(temp13, col = mypal, cex.axis = 2, cex.lab = 2, legend.width = 2, axis.args = list(cex.axis = 2))

**Results**

*Mesoclimate*

The unadjusted 2010 temperature data derived from the NOAA-NCEP programme was strongly correlated with hourly data from Culdrose and Brize Norton weather stations (mean *r2* =0.921, mean error = 1.12 °C, RMS error = 1.32 °C, p < 0.001). However, at the Culdrose site, which is classified as a marine pixel in the NCEP dataset, temperature extremes were not fully captured (Fig. S1). Nonetheless, our method of applying elevational adjustments to temperature in this coastal region, and accounting for coastal and cold-air drainage effects, reproduces the patterns in one km gridded data available from the UK Met Office and results in significantly improved estimates (with adjustment: mean error 1.21 °C, RMS error 0.84 °C, without adjustment: mean error 2.09 °C, RMS error 1.50 °C; Fig S2a, b).

*Time-series*

Results for temperature are provided in the main text. For soil moisture (Table S2, Fig. S4), *r* was overall 0.15 worse in absolute terms for NCEP-based predictions and RMS error was 0.17 higher, when compared to AWAP-based predictions, with the latter performing significantly better at all depths for *r* and significantly better at 75 cm for RMS error.

Why was the NCEP data able to provide similarly accurate, and even slightly better, predictions of soil temperature to those made with the Australia-specific data? The AWAP data, despite being of higher spatial resolution, provides only minimum and maximum air temperature values, 9am and 3pm vapour pressure (which were used to calculate an average daily value) and total daily solar radiation (which was used to estimate mean daily cloud cover). In addition, the fine spatial resolution wind speed data we used was of daily temporal resolution (McVicar et al., 2008). The internal temporal disaggregation routines within the NicheMapR microclimate model are based on the assumption that minima and maxima occur at fixed times relative to dawn and solar noon (depending on the variable) (Kearney & Porter, 2017). Thus, the finer temporal resolution (including the appropriate covariation of the weather variables) of the NCEP data, in addition to our terrain-based lapse-rate corrections, may have contributed to the strong performance of the NCEP-based soil temperature predictions.

*Model calibration for gridded temperature*

The simpler model calibration procedure implemented within microclima compares well with the more complex, first-principles method implemented within NicheMapR, thereby justifying the use of NicheMapR to derive calibration coefficients that would normally be obtained from experimental data (Fig S5). While temperature extremes are not always captured, particularly when soil moisture conditions are dry, the simpler fitting procedure reproduces above ground temperatures derived using NicheMapR with a mean error of 1.65 °C and an RMS error of 2.49 °C, and below ground temperatures with a mean error of 0.61 °C and an RMS error of 0.83 °C.

*Temperature grids*

Spatial patterns in temperatures at 1 m above the ground are well-reproduced by our automated procedure, in comparison to estimates generated using models calibrated with experimental data (Fig S2c, d). However, they are typically less variable than those derived from models calibrated using experimental data with coefficient estimates, particularly for radiation, which is lower when estimated using NicheMapR than when estimated using temperature logger data (Table S3). This contrasts with temperatures estimated at 5 cm above the ground where coefficient estimates, particularly for radiation, higher when estimated using NicheMapR than when estimated using temperature logger data (Table S4), though the radiation estimates themselves are less variable than when locally sourced data are used.

**Table S1.** Summary statistics for the correlation coefficient *r* and Root Mean Square (RMS) error for observed and predicted soil temperature (°C) across the 35 Oznet soil temperature and moisture monitoring sites when using either a) the Australian Water Availability or b) the National Centers for Environmental Prediction (NCEP) gridded weather data as forcing variables. Also shown c) are P-values and respective differences in mean values (AWAP - NCEP) of *r* and RMS error between simulations using the two different forcing data sets.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **r** |  | **RMS error** |  |
| a) AWAP | **depth (cm)** | **mean** | **stdev** | **mean** | **stdev** |
|  | 4 | 0.85 | 0.03 | 5.29 | 0.78 |
|  | 15 | 0.95 | 0.06 | 2.79 | 0.84 |
|  | 45 | 0.98 | 0.01 | 2.70 | 0.83 |
|  | 75 | 0.98 | 0.02 | 2.58 | 1.25 |
|  | overall | 0.93 | 0.07 | 3.55 | 1.47 |
| b) NCEP | **depth (cm)** | **mean** | **stdev** | **mean** | **stdev** |
|  | 4 | 0.94 | 0.01 | 3.64 | 0.63 |
|  | 15 | 0.95 | 0.04 | 2.77 | 1.00 |
|  | 45 | 0.98 | 0.02 | 2.40 | 1.51 |
|  | 75 | 0.97 | 0.02 | 1.95 | 0.58 |
|  | overall | 0.96 | 0.03 | 2.96 | 1.12 |
| c) comparison | **depth (cm)** | **P** | **Delta** | **P** | **Delta** |
|  | 4 | <0.0001 | -0.0941 | <0.0001 | 1.6353 |
|  | 15 | 0.6167 | 0.0025 | 0.8928 | 0.0247 |
|  | 45 | 0.0353 | 0.0073 | 0.3725 | 0.2961 |
|  | 75 | 0.5488 | 0.0021 | 0.1590 | 0.6242 |

**Table S2.** Summary statistics for the correlation coefficient *r* and Root Mean Square (RMS) error for observed and predicted volumetric soil moisture (m3/m3) across the 35 Oznet soil temperature and moisture monitoring sites when using either a) the Australian Water Availability or b) the National Centers for Environmental Prediction (NCEP) gridded weather data as forcing variables. Also shown c) are P-values and respective differences in mean values (negative values mean that NCEP is higher) of *r* and RMS error between simulations using the two different forcing data sets.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **r** |  | **RMS error** |  |
| a) awap | **depth (cm)** | **mean** | **stdev** | **mean** | **Stdev** |
|  | 4 | 0.79 | 0.07 | 8.42 | 3.14 |
|  | 15 | 0.76 | 0.09 | 6.84 | 2.11 |
|  | 45 | 0.52 | 0.40 | 7.37 | 4.05 |
|  | 75 | 0.53 | 0.29 | 7.07 | 4.14 |
|  | overall | 0.65 | 0.28 | 7.42 | 3.46 |
| b) NECP | **depth (cm)** | **Mean** | **stdev** | **mean** | **stdev** |
|  | 4 | 0.60 | 0.12 | 8.04 | 2.01 |
|  | 15 | 0.56 | 0.13 | 6.88 | 2.08 |
|  | 45 | 0.40 | 0.28 | 7.75 | 3.76 |
|  | 75 | 0.43 | 0.24 | 7.73 | 4.30 |
|  | overall | 0.50 | 0.22 | 7.59 | 3.19 |
| c) comparison | **depth (cm)** | **P** | **delta** | **P** | **delta** |
|  | 4 | <0.0001 | 0.1922 | 0.3040 | 0.3845 |
|  | 15 | <0.0001 | 0.1932 | 0.8692 | -0.0345 |
|  | 45 | 0.0144 | 0.1206 | 0.2132 | -0.3830 |
|  | 75 | 0.0385 | 0.0962 | 0.0305 | -0.6624 |

**Table S3.** Comparison of coefficient estimates obtained using NicheMapR with those obtained using iButton temperature loggers at one metre above ground across the Lizard Peninsula, Cornwall, UK in 2010.

|  |  |  |
| --- | --- | --- |
| **Coefficient** | **Estimate (logger data)** | **Estimate (NicheMapR)** |
| Intercept | -0.018 | 0.124 |
| Log (wind) | 0.289 | -0.090 |
| Net radiation | 3.112 | 1.254 |
| Log (wind) x net radiation | -0.760 | -0.598 |

**Table S4.** Comparison of coefficient estimates obtained using NicheMapR with those obtained using iButton temperature loggers at 5 cm above ground at Caerthillian Valley, Lizard Peninsula, UK in May 2010.

|  |  |  |
| --- | --- | --- |
| **Coefficient** | **Estimate (logger data)** | **Estimate (NicheMapR)** |
| Intercept | -0.781 | 0.279 |
| Log (wind) | 0.441 | -0.211 |
| Net radiation | 3.277 | 5.506 |
| Log (wind) x net radiation | -0.954 | -2.174 |

Figures

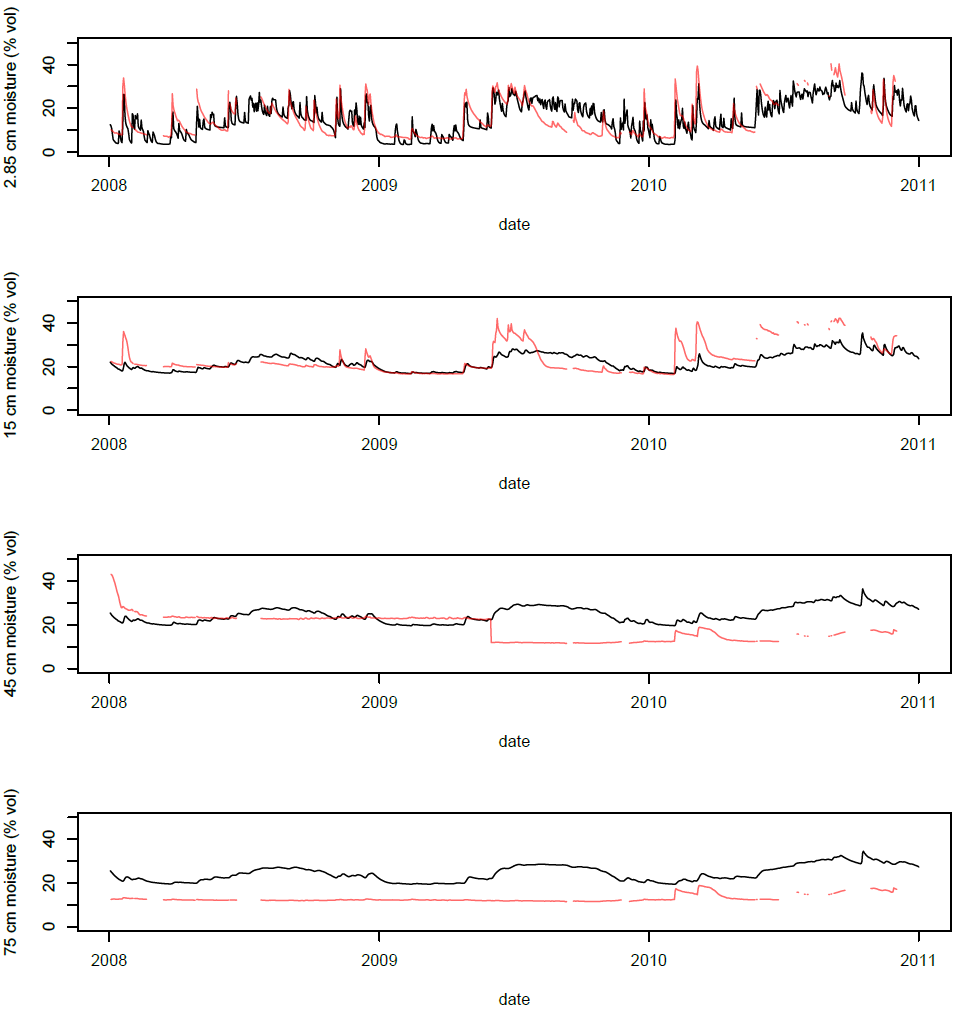
|  |  |  |
| --- | --- | --- |
| **(a)** | **(b)** | **(c)** |
|  |  | https://lh6.googleusercontent.com/M9DvQ-kO8PRMZbk8XZV7-Mxy4hiCeAzIFCIht2l9gxPrdyrc_f30mLxGbADUWLZB1GW0Zrtl8x0MK1ou5lcfFjBCDxP3qu3oEnFts5mDKAozjxhlK6g9YI2-nYdkWNxYFCQu-K8d |

**Figure S1.** Hourly temperature data recorded at (a) Culdrose (50.083°N, 5.249°W) and (b) Brize Norton (51.750°N, 1.583°W) weather stations in 2010 (red) versus temperature data derived from the NOAA-NCEP programme (dark grey / black). In (c) representation of the sea and land by NCEP-NOAA data over the United Kingdom and NW France, and the locations of the two weather stations, are shown. While the overall trends in temperature are similar, temperature extremes are not fully captured in the NOAA-NCEP datasets, particularly at Culdrose, which is assumed by NCEP-NOAA data to lie within the sea. This failure to capture extremes is partially corrected by the calculation and application of coastal effects during the downscaling process (Fig. 2a, b).

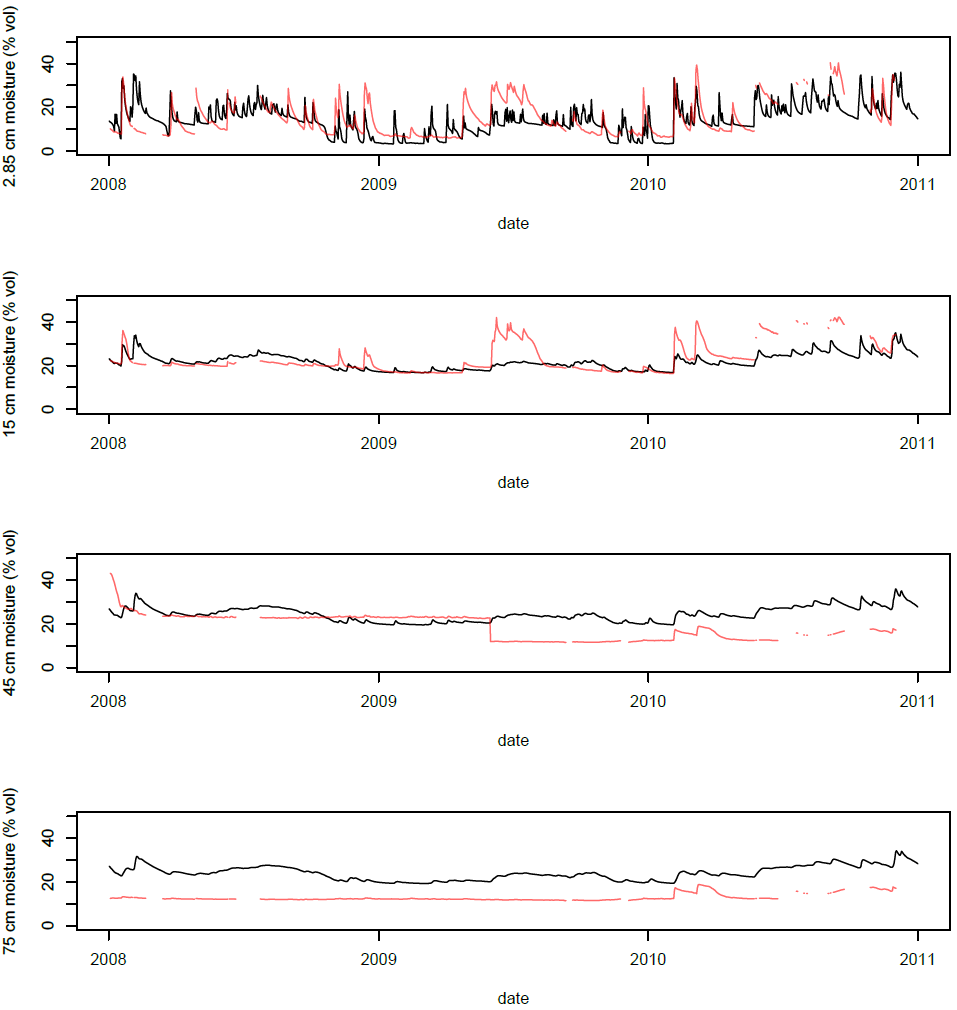
|  |  |
| --- | --- |
| **(a)** | **(b)** |
|  | |
| **(c)** | **(d)** |
|  | |

**Figure S2.** Side by side comparisons of gridded datasets derived using methods described in Maclean, Mosedale and Bennie(2019) (a, c) with those obtained using the automated procedures described in the present paper (b, d). Top: a one km gridded dataset of daily temperature Met Office (2018) (a) is compared to our method for adjusting 250 km NCEP data for coastal exposure elevation and cold-air drainage (b) on 7th January 2017 across Cornwall, UK. Bottom: a 100 m resolution dataset of temperatures at one m height generated using methods described in Maclean, Mosedale and Bennie (2019) (c) is compared to estimates at the same height using automated procedures for adjusting ~200 km resolution NCEP data (d) on 4th May 2010 11:00 across the Lizard Peninsula, Cornwall, UK. Here canopy cover is assumed to be zero and surface albedo is held constant at 0.15.

1. AWAP



1. NCEP



**Figure S3.** Observed (red) and predicted (black) volumetric soil moisture for one of the Yanco OzNet sites for the years 2008-10 using a) the Australian Water Availability Project (AWAP) daily weather grids or b) down-scaled and disaggregated National Centers for Environmental Prediction (NCEP) daily weather as forcing data.

|  |  |
| --- | --- |
| **(a)** | **(d)** |
|  |  |
| **(b)** |
|  |
| **(e)** |
|  |
| **(c)** |
|  |

**Figure S4.** Comparison of hourly temperature estimates derived using NicheMapR (red) with those derived using the simpler model calibration procedures available with microclima (black). Reference air temperature (grey) is also shown. In a-c, time series of temperatures 5cm above ground (a, b) and 5 cm below ground (c) are shown. In (d, 5 cm above ground) and (e, 5 cm below ground) the two sets of temperature data for the whole of 2010 are plotted against each other. The red lines represent 1:1 relationships. All data were generated for 2010 at Caerthillian Valley, Lizard Peninsula, UK (the location 49.968°N, 5.2157°W).



**Fig. S5.** Observed and predicted temperatures (Caerthillian Valley, 49.969ºN, 5.215ºW, during May 2010). Temperatures recorded by iButtons placed 5 cm above the ground are compared to outputs obtained from the automated microclimate model. Ground and canopy albedo are fixed at 0.15 and 0.23 respectively and canopy shading is estimated from habitat type.

**References**

Jones, D., Wang, W., & Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, *58*(04), 233–248. doi:10.22499/2.5804.003

Kearney, M. R., & Maino, J. L. (2018). Can next-generation soil data products improve soil moisture modelling at the continental scale? An assessment using a new microclimate package for the R programming environment. *Journal of Hydrology*, *561*, 662–673. doi:10.1016/j.jhydrol.2018.04.040

Kearney, M. R., & Porter, W. P. (2017). NicheMapR - an R package for biophysical modelling: the microclimate model. *Ecography*, *40*(5), 664–674. doi:10.1111/ecog.02360

Maclean, I. M. D., Mosedale, J. R., & Bennie, J. J. (2019). Microclima: an R package for modelling meso- and microclimate. *Methods in Ecology and Evolution*, *10*(2), 280–290. doi:10.1111/2041-210X.13093

McVicar, T. R., Van Niel, T. G., Li, L. T., Roderick, M. L., Rayner, D. P., Ricciardulli, L., & Donohue, R. J. (2008). Wind speed climatology and trends for Australia, 1975-2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophysical Research Letters*, *35*(20), L20403. doi:10.1029/2008GL035627

R Development Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., … Richter, H. (2012). The Murrumbidgee soil moisture monitoring network data set. *Water Resources Research*, *48*(7), n/a-n/a. doi:10.1029/2012WR011976