Received: 26 August 2021 | Accepted: 19 January 2022

DOI: 10.1111/2041-210X.13877

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

APPLICATION

MCERA5: Driving microclimate models with ERA5 global gridded climate data

David H. Klinges[1] | **James P. Duffy**[2] | **Michael R. Kearney**[3] | **Ilya M. D. Maclean**[2]

[1]School of Natural Resources and Environment, University of Florida, Gainesville, FL, USA

[2]Environment and Sustainability Institute, University of Exeter, Cornwall, UK

[3]School of BioSciences, The University of Melbourne, Vic., Australia

Correspondence
David H. Klinges
Email: dklinges9@gmail.com

Funding information
Met Office Hadley Centre Climate Programme; National Science Foundation, Grant/Award Number: DGE-1842473

Handling Editor: Will Pearse

Abstract

1. Microclimate models predict temperature and other meteorological variables at scales relevant to individual organisms. The broad application of microclimate models requires gridded macroclimatic variables as input. However, the spatial and temporal resolution of such inputs can be a limiting factor on the accuracy of microclimate predictions. Due to its fine resolution and accuracy, the ERA5 reanalysis dataset is emerging as the favoured resource for global historical weather and climate data and has great potential for aiding microclimate modelling.
2. Here we describe MCERA5, an R language package that provides convenient access to, and wrangling of, the ERA5 climate datasets for use in microclimate models. Through this package, we provide functions to query ERA5 data for desired spatial and temporal extents, to correct for spatial biases and process outputs for easy interpretation by ecologists, thereby allowing faster and more accurate microclimate predictions.
3. By validating with empirical observations from multiple biomes globally, we demonstrate that the use of ERA5 climate forcing via MCERA5 improves the prediction accuracy of soil moisture, air temperature and relative humidity as compared to forcing with other globally available data and offers comparable performance when predicting soil temperatures.
4. Through the provision of fine-resolution ERA5 data, the MCERA5 package fits into an ecosystem of tools for modelling microclimate in a spatio-temporally explicit fashion, advancing our ability to efficiently predict microclimate for any place on Earth for the past, present or future. The package also provides convenient access to ERA5 datasets for a range of other applications.

KEYWORDS
biophysical, downscaling, microclimate, modelling, precipitation, R package, soil temperature, spatial

David H. Klinges and James P. Duffy contributed equally to this work.

1 | INTRODUCTION

Numerous studies seek to understand relationships between organisms and climate, but readily available climate datasets are often a poor surrogate for the climatic conditions experienced by plants and animals in nature (Lembrechts et al., 2019). For the systems and processes of interest to ecologists, environmental data on radiation, wind, air temperature or other climatic variables are often not only at inappropriately coarse spatial and temporal resolution, but also provided in formats and units that are not of immediate utility to non-meteorologists. Although mechanistic and statistical approaches exist for downscaling coarse resolution environmental data to predict temperatures at biologically relevant spatial and temporal scales (Bramer et al., 2018; Potter et al., 2013), such methods are limited by the quality and extent of gridded climate or weather data provided as input.

To date, gridded weather data such as the National Centres for Environmental Prediction (NCEP) dataset produced by the National Oceanic and Atmospheric Administration have proven an easily accessible resource (Kalnay et al., 1996). However, most NCEP products are provided at a temporal resolution of 6 hr and a spatial resolution of 2.5 decimal degrees and thus offer only a coarse description of a region or temporal window. Methods exist to downscale the NCEP data (Kearney et al., 2020), but with the launch of ERA5, this new climate product suite will likely supersede NCEP as the primary data source for researchers wishing to calculate microclimate temperatures anywhere on Earth (Figure 1).

ERA5 (ECMWF Re-Analysis 5, Hersbach et al., 2020) is a freely accessible comprehensive climate dataset provided by the European Centre for Medium-Range Weather Forecasts. It is a reanalysis product stemming from the assimilation of numerical weather model forecasts with empirical station observations, thereby offering an appealing balance of low data latency and high accuracy based on near real-time model improvement. ERA5 data are available at a temporal resolution of 1 hr, a spatial resolution of 0.25° and span from 1950 to 5 days from present (note: 1950–1978 data are currently classed as preliminary). Vertically, the data encompass predictions at several soil depths as well as from Earth's 'surface' (2 m above the ground) through 37 pressure levels up into the atmosphere. Uncertainty estimates for climate variables are provided in 3-hr intervals, drawn from variation within a 10-member ensemble of predictions. Monthly averages are also available, yet it is the hourly predictions that provide unprecedented climate forcing data for biological applications.

ERA5 data can be acquired manually from the Climate Data Store (CDS; https://cds.climate.copernicus.eu/) or programmatically using the corresponding application programming interface (API). This API is accessed via Python (van Rossum & Drake, 2002), which may pose a challenge for some in the ecological research community, for whom the R programming environment (R Core Team, 2021) has rapidly established itself as the primary tool used for data analysis (Lai et al., 2019). The ECMWFR package (Hufkens et al., 2019) provides tools in R to securely log in to the CDS using user credentials and then build and execute requests for downloading data to one's machine locally. However, navigating the myriad of available datasets, and the netCDF format of downloaded files, may act as a barrier to researchers less familiar with meteorological data or accustomed to tabular formats. For climate downscaling, it is also necessary to coerce variables to desired spatial and temporal units, apply correction biases and combine variables to compute those needed for microclimate modelling. Thus, there is a need for automated tools that perform the necessary selection and pre-processing of ERA5 data for use in existing microclimate models or other ecological applications. Here we describe the MCERA5 package as a toolkit for acquiring, transforming and formatting the ERA5 forcing data pertinent to microclimate modelling, to predict accurate microclimate almost

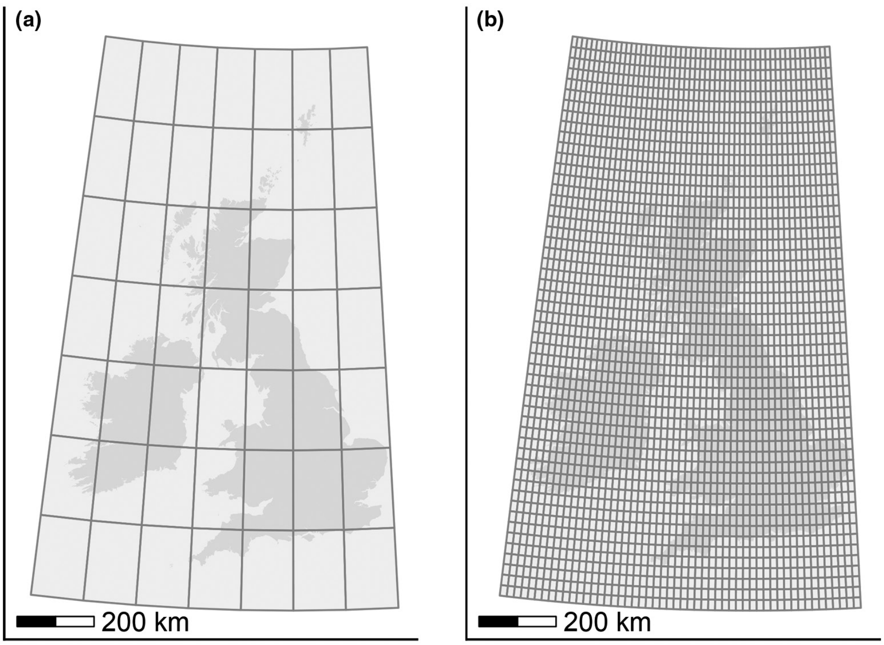

FIGURE 1 The spatial resolution of NOAA NCEP (a) and ERA5 (b) gridded climate data across the extent of the British Isles. *Note*: both datasets are produced with global extents.

automatically. We then integrate it with a set of existing microclimate models and demonstrate corresponding improvements in prediction accuracy.

2 | THE MCERA5 PACKAGE

The aim of the MCERA5 package is to provide a set of tools to download variables of relevance to microclimate modelling, then process and format these data to provide users with fine-resolution climate forcing data (Figure 2). The package is designed to be interoperable with a series of other R packages including NICHEMAPR (Kearney & Porter, 2017), MICROCLIMA (Maclean et al., 2019) and MICROCLIMC (Maclean & Klinges, 2021), which all feature models for mechanistically predicting local microclimate when provided data on climate forcing, vegetation and soil characteristics of a site or region. NICHEMAPR provides a vertical-flux air and soil microclimate model paired with energy and mass exchange models for ectothermic and endothermic organisms; MICROCLIMA calculates the effect of physical forces on near-ground temperature at a spatial point, and when provided empirical parameters, then applies these relationships across a grid to spatially map microclimate; MICROCLIMC draws upon first principles of physics to represent heat exchange above, below and within canopies of a variety of vegetated habitats. Collectively, these packages form complementary approaches for predicting spatio-temporally explicit microclimate variables such as temperature, relative humidity and soil moisture. The integration of MCERA5 with these packages provides convenient use of ERA5 data for predicting local microclimate, but MCERA5 also functions as a stand-alone portal to curated ERA5 climate and weather data for the past or near-present and across wide spatial extents. The code and vignette for the MCERA5 package is hosted on GitHub at https://github.com/dklinges9/mcera5 and can be installed in R using *remotes::install_github("dklinges9/mcera5")*. Some functions in the MCERA5 package rely on several other underlying R packages: PLYR (Wickham, 2020), DPLYR (Wickham et al., 2021), MAGRITTR (Bache et al., 2020), ECMWFR (Hufkens et al., 2019), NCDF4 (Pierce, 2021), ABIND (Heiberger, 2016), LUBRIDATE (Spinu et al., 2018), TIDYNC (Sumner, 2021) and MICROCLIMA (Maclean et al., 2019). These packages are thus specified as dependencies and automatically installed along with MCERA5 if not already present on a user's machine. Here, we describe the workflow in R and provide a worked example as an online supplemental file (also available as a vignette in the R package).

3 | WORKFLOW

3.1 | Building a request and downloading data

Note: at times, when the CDS server is busy, there can be an overhead time of several hours for downloading data from the CDS, even for small amounts of data (<10 MB). Such overhead time is independent of MCERA5, yet can be reduced according to the spatial/temporal dimensions of the query. Generally, querying temporal durations >1 year causes a delay, while query of wide spatial extents can occur rapidly. It is therefore most efficient to download time-series data in temporal chunks (e.g. monthly basis) each specifying a region encompassing multiple points of interest. We also recommend users track current usage of the CDS at https://cds.climate.copernicus.eu/live/queue and ECMWF news at https://confluence.ecmwf.int/#all-updates.

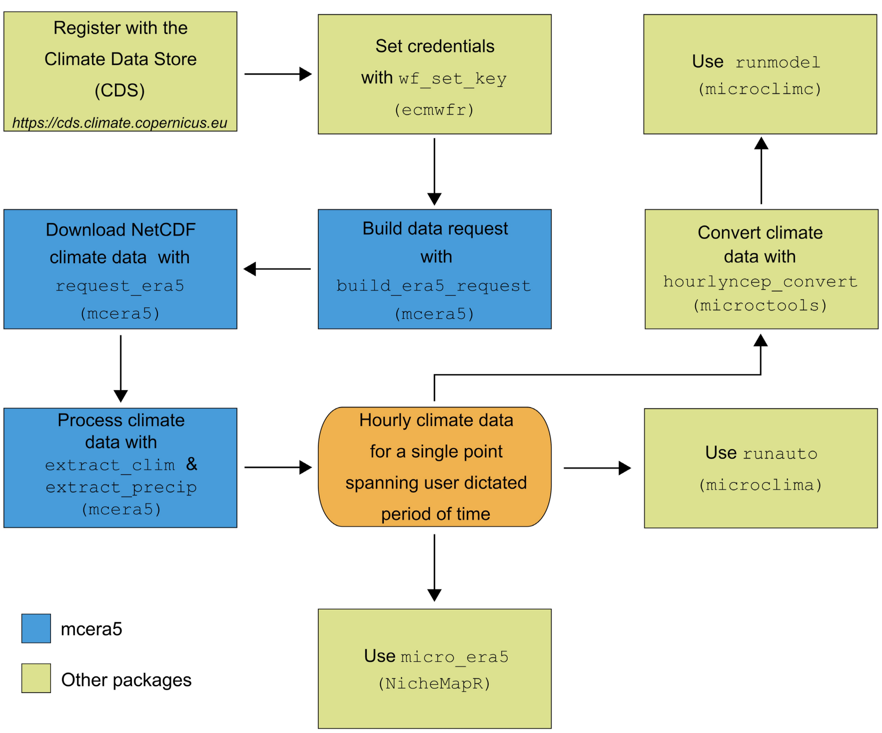

FIGURE 2 Workflow showing how MCERA5 fits into a wider process of data acquisition, processing and microclimate modelling. ERA5 data are stored on the CDS server, and after a user has registered their credentials via *wf_set_key*, they can efficiently structure a data query using the MCERA5 function *build_era5_request* and then submit that query using *request_era5*. Upon successful download of the corresponding file(s), meteorological predictions for a given spatial/temporal domain within the data can be extracted as a data frame using *extract_clim* and *extract_precip*. These data frames are immediately ready for use with several microclimate modelling R packages, including NICHEMAPR, MICROCLIMA and MICROCLIMC.

Due to the size of datasets and the speed at which they can be accessed through the CDS, acquisition and processing cannot be performed in tandem on-the-fly. Data must first be acquired and then, once downloaded, further processing steps can be performed. Once a user has registered with the CDS (https://cds.climate.copernicus.eu/user/register), they are able to access their UID and API keys (currently provided at the bottom of a user's profile page), which are a set of credentials that allow API access. A user can then add these to their machine's local keychain via the ECMWFR::*wf_set_key* function, which will allow queries to be submitted to the CDS from that machine. The next step is to build a data request, which is achieved with the *build_era5_request* function. This function takes, as inputs, four values defining a bounding box (in decimal degree longitude/latitude), start and end times and a user-defined prefix for file names once they are downloaded. Once executed, a request in list format is produced containing query metadata and a string of the variables to be downloaded and used in data processing further down the line. Next, *request_era5* takes the list created in the previous step and sends a request for a netCDF file containing the desired data to the CDS. User credentials and a local destination for the downloaded files are also required inputs. These steps can be implemented in code as follows:

```
# Store your UID and API key credentials as R ob-
jects
uid <- "*****"
cds_api_key <- "********-****-****-****-*********
***"
# Use `ecmwfr` package to register your machine
with your credentials
ecmwfr::wf_set_key(user = uid,
key = cds_api_key,
service = "cds")
# Designate your desired bounding coordinates (in
WGS84 / EPSG:4326)
xmn <- -4
xmx <- -2
ymn <- 49
ymx <- 51
# Designate your desired temporal extent
st_time <- lubridate::ymd("2010:02:26")
en_time <- lubridate::ymd("2010:03:01")
# Set a unique prefix for the filename (here based
on spatial
# coordinates), and the file path for downloaded .
nc files (here,
# the user's working directory)
file_prefix <- "era5_-4_-2_49_51"
file_path <- getwd()
# Build a request
req <- build_era5_request(xmin = xmn, xmax = xmx,
ymin = ymn, ymax = ymx,
start_time = st_time,
end_time = en_time,
outfile_name = file_prefix)
# Submit your request
request_era5(request = req, uid = uid, out_path =
file_path)
```

By default, MCERA5 queries the 'ERA5 hourly data on single levels from 1979 to present' dataset, which provides the most appropriate vertical and temporal resolutions for microclimate modelling. If users specify a duration longer than 1 year, the query is first downloaded as a separate netCDF file for each year to increase download speed, and then these files are combined (users can turn off file combination by specifying the parameter combine = FALSE in *request_era5*). At this stage, the user then waits for the netCDF to be downloaded to their machine and will receive a confirmation message ('ERA5 netCDF file successfully downloaded') from the R console upon completion.

3.2 | Processing data

Once the netCDF file(s) have been downloaded, data for a single point can be extracted. This is achieved using *extract_clim*, which creates a data frame of hourly climate variables (Table 1) and *extract_precip* which creates a vector of daily or hourly precipitation values. Both functions require a path to a downloaded ERA5 netCDF file as well as longitude, latitude, start and end times as inputs: # List the path of an .nc file that was downloaded via# `request_era5()` my_nc <- paste0(getwd(), "/era5_-4_-2_49_51_2010.nc")# Specify desired single point (within the bounds of your .nc file) x <- -3.654600 y <- 50.640369# Gather all hourly variables, with spatial and temporal dimensions# matching the extent, or a subset, of data in one downloaded file point_out <- extract_clim(nc = my_nc, long = x, lat = y, start_time = st_time, end_time = en_time) # You can then inspect the data frame head(point_out)# Gather daily precipitation point_out_precip <- extract_precip(nc = my_nc, long = x, lat = y, start_time = st_time, end_time = en_time, convert_daily = TRUE)

By default, *extract_clim* and *extract_precip* also apply an inverse distance weighting calculation. This means that if the user requests data for a point that does not match the regular grid found in the ERA5 dataset (i.e. the centre point of each ERA5 grid cell), the four nearest neighbouring data points to the requested location will be used to create a weighted average of each climate variable, thereby providing a better estimate of location conditions. Furthermore, *extract_clim* allows the user to apply a diurnal temperature range correction to the data. The diurnal temperature ranges of ERA5 are artificially lower in grid cells classed as sea as opposed to land. It may thus be useful to apply a correction if estimates are required for a terrestrial location in predominantly marine grid cells. If applied, an internal function is evoked that uses the land/sea value in the downloaded netCDF file to adjust temperature values by the factor provided

TABLE 1 The names of the processed ERA5 variables used for microclimate modelling, a summary of processing conducted in mcera5 to derive these variables and the corresponding original climate products acquired from the ERA5 single-level dataset

mcera5 variable	Processing	ERA5 variables used
Temperature (°C)	Converted from Kelvin to Centigrade	2 m air temperature
Pressure (Pa)	NA	Surface Pressure
Specific humidity (kg/kg)	Derived from dewpoint temperature, air temperature and surface pressure based on the calculations of Bolton (1980)	2 m dewpoint temperature, 2 m air temperature and surface pressure
Windspeed (m/s)	Derived from 10 m zonal and 10 m meridional orthogonal wind velocities, adjusted to 2 m height using a standard logarithmic height profile	10 m zonal (u, towards east) wind velocity and 10 m meridional (v, towards north) wind velocity
Wind direction, azimuth (degrees from N)	Derived from 10 m zonal and 10 m meridional orthogonal wind velocities	10 m u component of wind and 10 m v component of wind
Emissivity (0–1)	Derived as downward longwave radiation flux divided by the sum of net longwave radiation flux and downward longwave radiation flux (downward flux negative)	Mean surface downward longwave radiation flux and mean surface net longwave radiation flux
Net longwave radiation (MJ m^{-2} hr^{-1})	Converted from W m^{-2} to MJ m^{-2} hr^{-1}	Mean surface net longwave radiation flux
Upward longwave radiation (MJ m^{-2} hr^{-1})	Derived as sum of net longwave radiation flux and downward longwave radiation flux (downward flux negative) and converted to MJ m^{-2} hr^{-1}	Mean surface net longwave radiation flux and mean surface downward longwave radiation flux
Downward longwave radiation (MJ m^{-2} hr^{-1})	Converted to MJ m^{-2} hr^{-1}	Mean surface downward longwave radiation flux
Direct normal irradiance (MJ m^{-2} hr^{-1})	Converted from continuous measure to instantaneous measure for the hour, adjusted from horizontal to normal and converted to MJ m^{-2} hr^{-1}	Total sky direct solar radiation at surface
Diffuse radiation (MJ m^{-2} hr^{-1})	Total downward surface solar radiation minus direct solar radiation (for a flat surface), corrected to provide an instantaneous measure for the hour and converted to MJ m^{-2} hr^{-1}	Surface solar radiation downwards and total sky direct solar radiation at surface
Solar zenith angle (degrees)	Calculated from location, date and time using the equation of time applied with Julian dates and computation of solar declination	Longitude, Latitude, Time, Date
Cloud cover (%)	Multiplied by 100	Total cloud cover

using the formula $DTR_C = DTR[(1 - p_l)C_f + 1]$, where DTR_C is the corrected diurnal temperature range, DTR is the diurnal temperature range in the ERA5 dataset, p_l is the proportion of the grid cell that is land and C_f is the correction factor. The default function input value is a correction based on calibration against the UK Met Office 1-km^2 gridded dataset of daily maximum and minimum temperatures (Hollis et al., 2019), itself calibrated and validated against a network of (on average) 1,203 weather stations distributed across the United Kingdom. Further details of appropriate corrections to apply are documented in the *extract_clim* help file. Beyond these corrections, *extract_clim* also converts and synthesizes ERA5 variables into forms more familiar and of greater utility to ecologists and evolutionary biologists (Table 1). For example, wind velocities expressed as vector components (u and v) are converted to wind speed and direction and direct normal irradiance is converted from a continuous average over the preceding hour to an instantaneous measure for the hour. By default, *extract_precip* sums up hourly ERA5 precipitation to the daily level, which is required for the aforementioned microclimate models. However, users can instead receive hourly values by setting convert_daily = FALSE. Processing of climate data employs tidyverse-style pipelines levering the syntax provided via the tidync package (Sumner, 2021).

Once complete, the user will have a data frame of climate variables and vector of precipitation, ready for use with multiple microclimate modelling tools in R (Figure 2) or for use in other ecological applications.

4 | CASE STUDIES

4.1 | Using microclima to calculate hourly temperatures on the Roseland Peninsula, Cornwall, UK

Here we show how mcera5 can be used in conjunction with the microclima package to construct hourly microclimate temperature estimates for 2019 for a grassy region in Cornwall, UK (50.207°N, 4.918°W). The *runauto* function from microclima, which parameterizes and runs a microclimate model, was used twice—first automatically acquiring NCEP forcing data, and second with ERA5 forcing data that had been acquired and processed using mcera5. The microclimate model was run

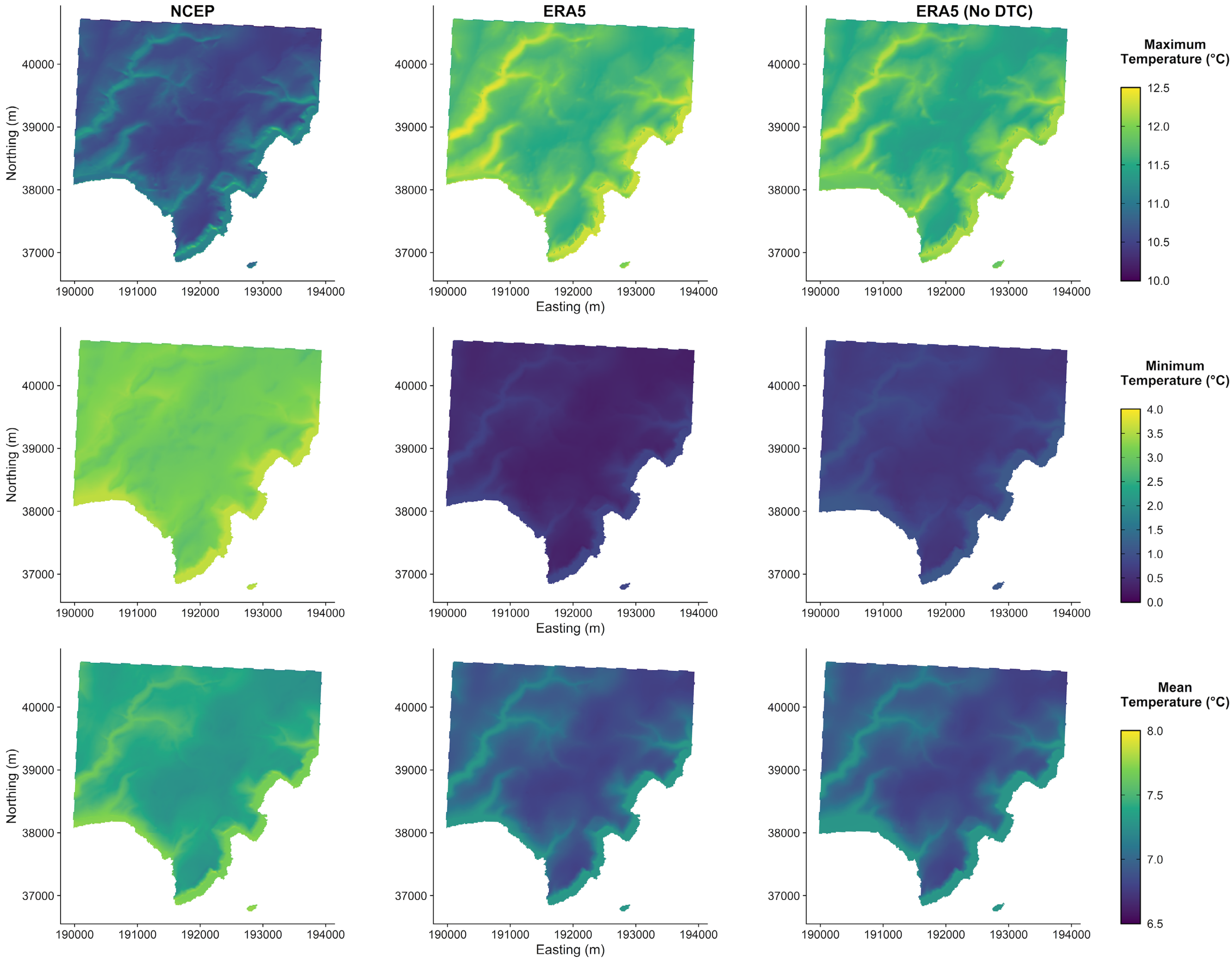

FIGURE 3 Outputs from microclimate model in the MICROCLIMA package applied to the Roseland Peninsula, Cornwall, UK (50.207°N, 4.918°W). Maps in the left-hand column show temperatures using NCEP forcing data, those in the middle column show temperatures using ERA5 forcing data accessed via MCERA5 with a diurnal temperature correction value of 1.285 and those in the right-hand column show temperatures using ERA5 forcing data accessed via MCERA5 without diurnal temperature correction (no DTC).

for the duration of January 2019, at a height of 0.1 m using the habitat parameters built into MICROCLIMA for 'Short grasslands'. Maximum, minimum and mean temperatures all differed between estimates created with NCEP and ERA5 forcing data (Figure 3), the greatest of which was that NCEP-derived minimum temperatures were ~3°C warmer. ERA5-derived maximum temperatures, both with and without diurnal temperature correction, were greater than NCEP-forced data, and temporal variance was lower within predictions forced by NCEP than forced by ERA5. These differences can be likely attributed to how well each climate product captures the presence of land or sea: the NCEP grid cell covering the Roseland Peninsula contains more sea than land (Figure S1), whereas the fine-grain ERA5 cells considered in the distance weighting were a mixture of land and sea, better reflecting the shape of the coastline (Figure S1). In tandem, increasing spatial resolution of input data and integrating the estimates of coastal effects on local climate, as is done with MCERA5, can result in significantly different microclimate predictions.

4.2 | Using NicheMapR to generate point predictions of soil temperature and moisture

We also demonstrate the interoperability of MCERA5 with the NicheMapR package to generate point predictions of microclimate for a specified temporal window. For this case study, we used the newly implemented *micro_era5* function in NicheMapR, which specifically uses MCERA5 to access ERA5 forcing data. We also used NicheMapR's *micro_ncep* function to generate a comparison against NCEP forcing data.

Using each function, we replicated a test against empirical soil temperature and soil moisture observations in an open paddock in eastern Australia for two depths (Yanco Site 2, Smith et al., 2012, Kearney et al., 2020; Kearney & Maino, 2018). For reference, we compare these predictions to those generated with the *micro_aust* function which uses the AWAP ('Australian Water Availability Project') 5-km resolution daily gridded weather product for Australia, to compare ERA5-forced predictions against those forced by a finer-resolution regional product (Jones et al., 2009; Kearney & Maino, 2018). Predictions for soil temperature were similarly accurate for all forcing datasets but soil moisture predictions from ERA5 forcing were considerably more accurate than those from NCEP (14.2% decrease in root-mean-square deviation from empirical data) and approached the accuracy and precision of predictions made with the finer-resolution AWAP data (Figure 4, Table 2).

4.3 | Using MICROCLIMC to generate point predictions of shaded microclimates

We also made additional predictions of air temperature and relative humidity at 1.5 m height from 1 January 2017 to 31 December 2017 within a deciduous broadleaved forest in Massachusetts, USA at the site of an ecosystem flux tower supplying long-term measurements by fine-wire thermistors (42°32′N, 72°11′W, Munger, 2021, Urbanski et al., 2007). Here we used the package MICROCLIMC (Maclean & Klinges, 2021), which is especially equipped to model heat transfer and storage within and below vegetation canopies, and validated predictions against empirical measurements. Vegetation parameter estimates of the site were derived using the parameterization for 'Deciduous broadleaved forest' provided by *PAIfromhabitat* in the MICROCTOOLS package (Maclean & Klinges, 2021), a companion to MICROCLIMC. We modelled below-canopy air temperature and relative humidity assuming steady-state conditions using the function *runmodelS* and performed two model runs: once with NCEP forcing and again with ERA5 forcing (acquired via MCERA5). Predictions forced by ERA5 had better fit to empirical temperature and relative humidity data than NCEP-driven predictions throughout the year: root-mean-square deviation decreased 30.6% and 33.1% for air temperature and relative humidity, respectively, and improvements held at both coarse and fine-temporal resolutions (Figure 5; Figure S2; Table 2).

5 | CONCLUSION

The finer spatial and temporal resolution of the ERA5 dataset should generally produce more accurate predictions of microclimate than those based on the NCEP data or other global climate products that are provided at coarser scales. We found this to be true for barren, sparsely vegetated and canopied systems in our case studies, with on average a 13.4% improvement in root-mean-square deviation of microclimate predictions when models were enabled by MCERA5 as opposed to driven by NCEP (Table 2). Improvements over other forcing datasets were in some cases small when predicting soil and air temperature as compared to the more noticeable improvement in humidity and soil moisture predictions. The MICROCLIMA and MICROCLIMC mesoclimatic temperature corrections, including lapse-rate adjustments, can moderately compensate for the low spatial resolution of NCEP air temperature, but it is not possible to similarly adjust precipitation and humidity data. Thus, soil moisture, humidity and snow predictions should be especially more accurate when the ERA5 dataset is used for microclimate prediction. We also expect the ERA5 dataset to perform better than NCEP in mountainous regions, where the higher spatial resolution of ERA5 better captures topographic complexity. These strengths make the MCERA5 package highly suitable for locations where predicting microclimate is of most need, such as within forests (De Frenne et al., 2021), snowy landscapes (Niittynen et al., 2018) and across elevation gradients (Klinges & Scheffers, 2021). Although the package's design is tailored for

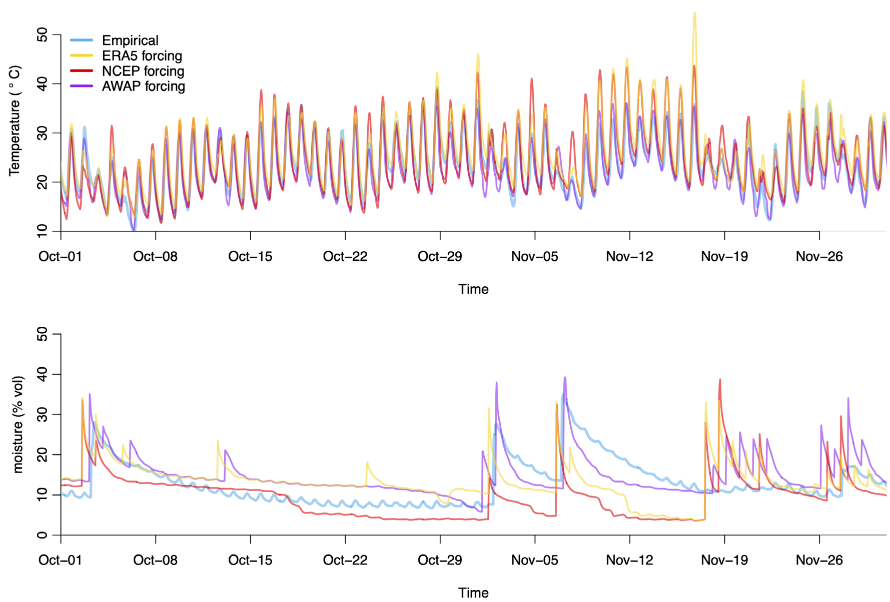

FIGURE 4 Two months of empirical and predicted (a) soil temperature and (b) soil moisture at a depth of 3 cm for an unshaded paddock in eastern Australia during spring 2008 (empirical data from Kearney & Maino, 2018). Predictions generated using the NICHEMAPR package functions *micro_aust* (AWAP), *micro_ncep* (NCEP) and *micro_era5* (ERA5), the latter of which was enabled by MCERA5. ERA5 forcing outperformed NCEP and had comparable performance to AWAP forcing, which is only regionally available.

TABLE 2 Summary statistics from a comparison between empirical observations (Kearney & Maino, 2018; Urbanski et al., 2007) and predicted soil temperature, soil moisture, air temperature and relative humidity values using NCEP, AWAP and ERA5 forcing data, the latter of which was enabled by MCERA5. Validation statistics reported are the Pearson's correlation coefficients (*r*) of empirical data and predictions, and the root-mean-square deviation (RMSD) of predictions from empirical data. The use of MCERA5 improved predictions compared to NCEP forcing and generated prediction accuracy competitive against the finer-resolution (yet only regionally available) AWAP

Site	Variable	Depth/height (cm)	Forcing	*r*	RMSD
Harvard Forest, Massachusetts, USA	Air temperature	150	ERA5	0.98	4.08
			NCEP	0.95	5.88
	Relative humidity	150	ERA5	0.79	23.92
			NCEP	0.48	35.75
Yanco Site 2, New South Wales, AUS	Soil temperature	3	ERA5	0.97	2.8
			NCEP	0.96	2.8
			AWAP	0.94	3.2
		15	ERA5	0.98	1.9
			NCEP	0.98	1.7
			AWAP	0.99	1.7
	Soil moisture	3	ERA5	0.73	5.5
			NCEP	0.62	6.7
			AWAP	0.78	5
		15	ERA5	0.68	5.1
			NCEP	0.61	5.7
			AWAP	0.65	5.2

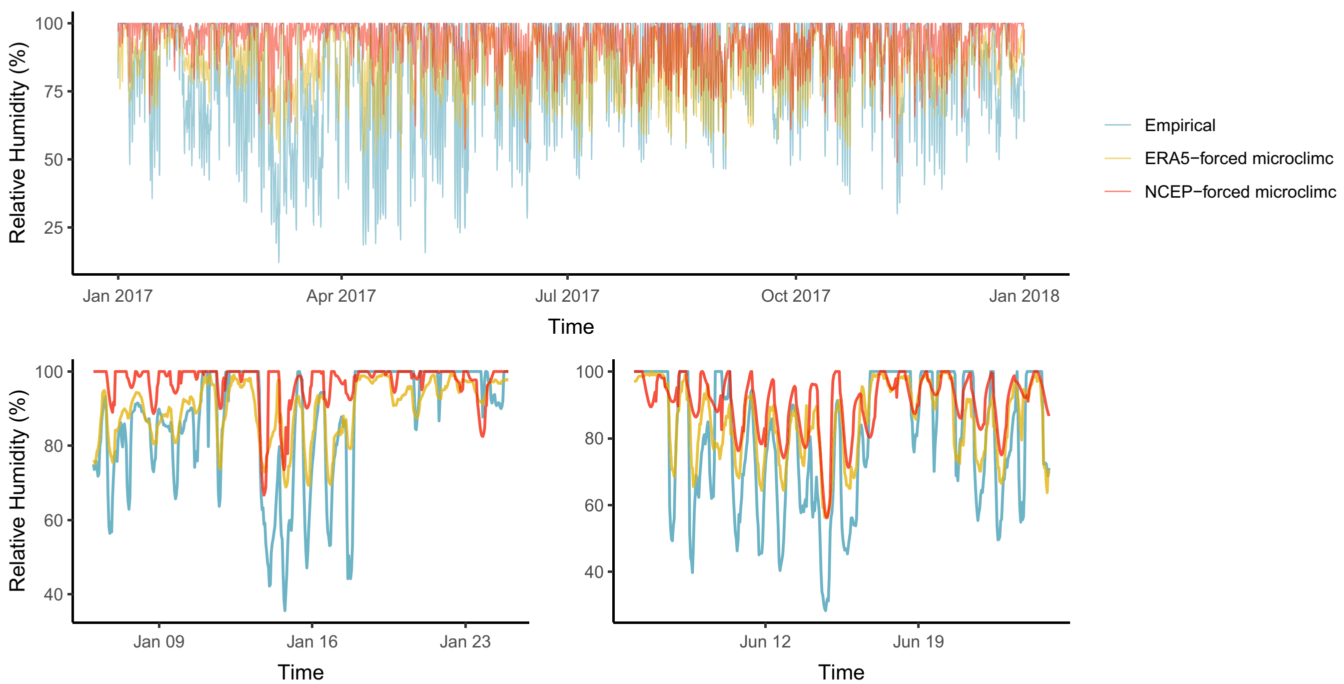

FIGURE 5 Empirical relative humidity underneath the canopy of a deciduous broadleaved forest (42°32′N, 72°11′W, Urbanski et al., 2007) and model predictions from the MICROCLIMC package forced with NCEP and ERA5 data (acquired via MCERA5). The top panel provides a year time series, and bottom panels are subsets in January (northern hemisphere winter) and July (northern hemisphere summer) of the same time series to demonstrate variability at finer temporal resolution.

driving microclimate models, given its open-access licence the code can be re-adapted to query different CDS variables, thereby providing for a broad range of environmental applications. By supplying ecologists and evolutionary biologists with easy access to ERA5 data in table form as well as processing ERA5 to provide as input to popular microclimate models, the MCERA5 package helps advance a holistic workflow of generating high-resolution, accurate and efficient microclimate predictions anywhere on Earth.

ACKNOWLEDGEMENTS

The authors are very grateful for the work by the authors of the ECMWFR package, Koen Hufkens, Reto Stauffer and Elio Campitelli, whose code has provided a strong foundation upon which to create some of the functions in MCERA5. They thank Associate Editor Will Pearse as well as Ofir Levy and two anonymous reviewers for their generous and constructive comments, all of which substantially improved the manuscript and software. J.D.P was supported by a grant to I.M.D.M from the Met Office Hadley Centre Climate Programme (HCCP) funded by BEIS and Defra. D.H.K. was supported by the National Science Foundation Graduate Research Fellowship (DGE-1842473).

CONFLICT OF INTEREST

The authors have no conflict of interest to claim.

AUTHORS' CONTRIBUTIONS

J.P.D. led development of the software, with specific function development by D.H.K., M.R.K. and I.M.D.M.; Subsequent maintenance of the software is provided by D.H.K. Validation was performed by D.H.K., J.P.D. and M.R.K. D.H.K. and J.P.D. wrote the manuscript with critical contributions from all authors.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13877.

DATA AVAILABILITY STATEMENT

No new data were included as part of this article. Version 1.0.0 of the MCERA5 package is available at https://github.com/dklinges9/mcera5 and citable via Zenodo at 10.5281/zenodo.5998263 (Klinges & Duffy, 2022). The NicheMapR release relevant to this paper (v2.0.0) is 10.5281/zenodo.3478635, and the microclima release (v2.0.0) is 10.5281/zenodo.3484589. These packages include all data related to this manuscript with the exception of the Australian soil temperature and moisture observations and the Massachusetts, USA air temperature and relative humidity observations. Australian measurements are freely accessible for direct download via links embedded at https://www.oznet.org.au/mdbdata/mdbdata.html#M1 (Smith et al., 2012), with further documentation available at https://www.oznet.org.au/y2.html. Massachusetts data, offered by the AmeriFlux network, can be downloaded at https://doi.org/10.17190/AMF/1246059 (Munger, 2021; Urbanski et al., 2007), by pressing the 'Download Data' button, logging in at the secure portal, selecting 'AmeriFlux BASE', pressing 'Confirm and Go to Next Step', selecting an applicable licence and again pressing 'Confirm and Go to Next Step' twice, describing the intended use of the data and agreeing to the licence and finally clicking on embedded links now visible on the webpage.

ORCID

David H. Klinges https://orcid.org/0000-0002-7900-9379
James P. Duffy https://orcid.org/0000-0001-7971-3924
Michael R. Kearney https://orcid.org/0000-0002-3349-8744
Ilya M. D. Maclean https://orcid.org/0000-0001-8030-9136

REFERENCES

Bache, S. M., Wickham, H., Henry, L., & RStudio. (2020). *magrittr: A forward-pipe operator for R* (2.0.1) [Computer software]. Retrieved from https://CRAN.R-project.org/package=magrittr

Bolton, D. (1980). The computation of equivalent potential temperature. *Monthly Weather Review*, *108*(7), 1046–1053. https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2

Bramer, I., Anderson, B. J., Bennie, J., Bladon, A. J., De Frenne, P., Hemming, D., Hill, R. A., Kearney, M. R., Körner, C., Korstjens, A. H., Lenoir, J., Maclean, I. M. D., Marsh, C. D., Morecroft, M. D., Ohlemüller, R., Slater, H. D., Suggitt, A. J., Zellweger, F., & Gillingham, P. K. (2018). Chapter three—Advances in monitoring and modelling climate at ecologically relevant scales. In D. A. Bohan, A. J. Dumbrell, G. Woodward, & M. Jackson (Eds.), *Advances in ecological research* (Vol. *58*, pp. 101–161). Academic Press. https://doi.org/10.1016/bs.aecr.2017.12.005

De Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., Ashcroft, M. B., Christiansen, D. M., Decocq, G., Pauw, K. D., Govaert, S., Greiser, C., Gril, E., Hampe, A., Jucker, T., Klinges, D. H., Koelemeijer, I. A., Lembrechts, J. J., Marrec, R., … Hylander, K. (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

Heiberger, T. P. (2016). *abind: Combine multidimensional arrays* (1.4-5) [Computer software]. Retrieved from https://CRAN.R-project.org/package=abind

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hollis, D., McCarthy, M., Kendon, M., Legg, T., & Simpson, I. (2019). HadUK-Grid – A new UK dataset of gridded climate observations. *Geoscience Data Journal*, *6*(2), 151–159. https://doi.org/10.1002/gdj3.78

Hufkens, K., Stauffer, R., & Campitelli, E. (2019). khufkens/ecmwfr: Ecmwfr. *Zenodo*, https://doi.org/10.5281/zenodo.2647541

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

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., … Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–472. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kearney, M. R., Gillingham, P. K., Bramer, I., Duffy, J. P., & Maclean, I. M. D. (2020). A method for computing hourly, historical, terrain-corrected microclimate anywhere on earth. *Methods in Ecology and Evolution*, *11*(1), 38–43. https://doi.org/10.1111/2041-210X.13330

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. https://doi.org/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. https://doi.org/10.1111/ecog.02360

Klinges, D. H., & Duffy, J. P. (2022). dklinges9/mcera5: Driving microclimate models with ERA5 global gridded climate data. *Zenodo*, https://doi.org/10.5281/zenodo.5998263

Klinges, D. H., & Scheffers, B. R. (2021). Microgeography, not just latitude, drives climate overlap on mountains from tropical to polar ecosystems. *The American Naturalist*, *197*(1), 75–92. https://doi.org/10.1086/711873

Lai, J., Lortie, C. J., Muenchen, R. A., Yang, J., & Ma, K. (2019). Evaluating the popularity of R in ecology. *Ecosphere*, *10*(1), e02567. https://doi.org/10.1002/ecs2.2567

Lembrechts, J. J., Lenoir, J., Roth, N., Hattab, T., Milbau, A., Haider, S., Pellissier, L., Pauchard, A., Backes, A. R., Dimarco, R. D., Nuñez, M. A., Aalto, J., & Nijs, I. (2019). Comparing temperature data sources for use in species distribution models: From in-situ logging to remote sensing. *Global Ecology and Biogeography*, *28*(11), 1578–1596. https://doi.org/10.1111/geb.12974

Maclean, I. M. D., & Klinges, D. H. (2021). Microclimc: A mechanistic model of above, below and within-canopy microclimate. *Ecological Modelling*, *451*, 109567. https://doi.org/10.1016/j.ecolmodel.2021.109567

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. https://doi.org/10.1111/2041-210X.13093

Munger, J. W. (2021). *AmeriFlux BASE US-Ha1 Harvard Forest EMS tower (HFR1) (version 17-5) [dataset]*. AmeriFlux AMP. https://doi.org/10.17190/AMF/1246059

Niittynen, P., Heikkinen, R. K., & Luoto, M. (2018). Snow cover is a neglected driver of Arctic biodiversity loss. *Nature Climate Change*, *8*(11), 997–1001. https://doi.org/10.1038/s41558-018-0311-x

Pierce, D. (2021). *ncdf4: Interface to Unidata netCDF (version 4 or earlier) format data files* (1.17.1) [computer software]. Retrieved from https://CRAN.R-project.org/package=ncdf4

Potter, K. A., Woods, H. A., & Pincebourde, S. (2013). Microclimatic challenges in global change biology. *Global Change Biology*, *19*(10), 2932–2939. https://doi.org/10.1111/gcb.12257

R Core Team. (2021). *A language and environment for statistical computing (4.1.0) [computer software]*. R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., & Richter, H. (2012). The Murrumbidgee soil moisture monitoring network data set. *Water Resources Research*, *48*(7), 1–6. https://doi.org/10.1029/2012WR011976

Spinu, V., Grolemund, G., Wickham, H., Lyttle, I., Constigan, I., Law, J., Mitarotonda, D., Larmarange, J., Boiser, J., & Lee, C. H. (2018). *lubridate: Make dealing with dates a little easier* (1.7.4) [computer software]. Retrieved from https://CRAN.R-project.org/package=lubridate

Sumner, M. (2021). *Tidync [R]*. rOpenSci. Retrieved from https://github.com/ropensci/tidync

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., & Munger, J. W. (2007). Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest. *Journal of Geophysical Research: Biogeosciences*, *112*(G2), 1–25. https://doi.org/10.1029/2006JG000293

van Rossum, G., & Drake, F. L. (2002). *Python reference manual (2.2.1) [computer software]*. iUniverse.

Wickham, H. (2020). *plyr: Tools for splitting, applying and combining data (1.8.6) [computer software]*. Retrieved from https://CRAN.R-project.org/package=plyr

Wickham, H., François, R., Henry, L., Müller, K., & RStudio (2021). *dplyr: A grammar of data manipulation (1.0.7) [computer software]*. Retrieved from https://CRAN.R-project.org/package=dplyr

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Klinges, D. H., Duffy, J. P., Kearney, M. R., & Maclean, I. M. (2022). MCERA5: Driving microclimate models with ERA5 global gridded climate data. *Methods in Ecology and Evolution*, *00*, 1–10. https://doi.org/10.1111/2041-210X.13877