The easyclimate R package: easy access to high-resolution daily climate data for Europe

Verónica Cruz-Alonso (ORCID: 0000-0002-0642-036X) [1,2] §, Christoph Pucher (ORCID: 0000-0002-9269-1907) [3] §, Sophia Ratcliffe (ORCID: 0000-0001-9284-7900) [4], Paloma Ruiz-Benito (ORCID: 0000-0002-2781-5870) [2,5], Julen Astigarraga (ORCID: 0000-0001-9520-3713) [2], Mathias Neumann (ORCID: 0000-0003-2472-943X) [3], Hubert Hasenauer (ORCID: 0000-0003-3469-4031) [3], Francisco Rodríguez-Sánchez (ORCID: 0000-0002-7981-1599) [6,7]

[1] Department of Landscape Architecture, Graduate School of Design, Harvard University, USA

[2] Universidad de Alcalá, Grupo de Ecología Forestal y Restauración (FORECO), Departamento de Ciencias de la Vida, Spain

[3] Institute of Silviculture, Department of Forest and Soil Sciences, University of Natural Resources and Life Sciences, Austria

[4] National Biodiversity Network Trust (NBN Trust), Nottingham, UK

[5] Universidad de Alcalá, Grupo de Investigación en Teledetección Ambiental (GITA), Spain.

[6] Departamento de Biología Vegetal y Ecología, Universidad de Sevilla, Spain.

[7] Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Spain.

§ These authors share the first-author position.

Correspondence:

Verónica Cruz-Alonso, Harvard University, Cambridge, MA, USA. Email: veronica.cral@gmail.com

FUNDING INFORMATION

VCA was supported by the Real Colegio Complutense Postdoctoral Fellowship 2020. CP, MN and HH acknowledge the financial support by the Bio Based Industries Joint Undertaking under the European Union's Horizon 2020 research and innovation program TECH4EFFECT (Techniques and Technologies for Effective Wood Procurement project; Grant Number 720757). PRB was supported by the Community of Madrid Region and the Universidad de Alcalá (Stimulus to Excellence for Permanent University Professors, EPU-INV/2020/010; "Ayudas para la realización de Proyectos para potenciar la Creación y Consolidación de Grupos de Investigación," CG20/CC-005). JA was supported by the FPI fellowship of the Department of Education of the Basque Government. FRS was supported by the VI Plan Propio de Investigación of Universidad de Sevilla (VI PPIT-US) and by

MICINN through European Regional Development Fund [SUMHAL, LIFEWATCH-2019-09-CSIC-13, POPE 2014-2020].

ABSTRACT

In recent decades there has been an increasing demand by ecologists for harmonized climatic data at large spatial scales and spanning long periods. Here we present *easyclimate*, a software package to obtain daily climatic data at high resolution (0.0083º, ~1 km) with R. The package facilitates the downloading and processing of precipitation, minimum and maximum temperatures for Europe from 1950 to 2020. Using *easyclimate* and given a set of coordinates (points or polygons) and dates (days or years), the user can download the climatic information as a tidy table or a raster object. In this package we implemented Cloud-Optimized GeoTIFFs which provide access to daily climate data for thousands of sites/days within minutes, without having to download huge rasters. Daily climate data are not available in many of the current climate databases and are essential for many ecological research questions and applications, including the study of the effects of extreme climatic events related to late-spring frosts, heat waves, or dry periods on plant performance. *easyclimate* taps the potential for climatic data and enables multiple applications in forestry, ecological and vegetation studies across Europe.

Keywords: R package, climate, Europe, extreme climatic events, reproducibility, cloud-optimized geoTIFF, daily data, temperature, precipitation

INTRODUCTION

In recent decades there has been an increasing demand for harmonized daily gridded climatic data at wide spatial scales and spanning long temporal periods. Such data is invaluable for vegetation, wildlife, climatic and hydrological studies and Earth system modelling (Hasenauer et al. 2003; Thornton et al. 2021). Examples are the assesment of climate effects and climate change impacts on European forests (Hlásny et al. 2017; Neumann et al. 2017; Moreno et al. 2018; Archambeau et al. 2020; Ruiz-Benito et al. 2020; George et al. 2021), the initialization of large-scale carbon cycle models (Pietsch & Hasenauer 2006), the spatial-temporal variability of rainfall erosivity (Micić Ponjiger et al. 2021) or the creation of a European net primary production dataset (Neumann et al. 2016).

Here we present *easyclimate* (Cruz-Alonso et al. 2021), a software package (available from GitHub: https://github.com/VeruGHub/easyclimate) to download and process climate data with R (R Core Team 2021). *easyclimate* has been developed to facilitate the use of high-resolution (0.0083º × 0.0083º, ~1 km^2) daily climate for Europe (longitude: -24.5º - 45.25º; latitude: 25.25º - 75.5º; Figure 1). Daily precipitation and minimum and maximum temperature data are currently available from 1950 to 2020 and hosted at University of Natural Resources and Life Sciences, Vienna, Austria.

The climatic dataset was originally produced by Moreno and Hasenauer (2016). For the production, the coarse daily E-Obs climate data (Cornes et al. 2018) was downscaled by using the finer-resolution WorldClim data (Fick & Hijmans 2017). WorldClim provides global long-term monthly averages of several climatic variables at 0.0083° resolution (approximately 1 km). E-Obs provides gridded daily climate data for Europe at 0.25° resolution (approximately 30 km) by interpolating around 3700 weather stations for temperature and around 9000 stations for precipitation. Downscaling was performed by applying a spatial delta method with a monotone cubic interpolation of anomalies (Mosier et al. 2014; Moreno & Hasenauer 2016). Since its original release, the downscaled climate data has been further developed and updated, and two further releases (v2 and v3) have been published (see the main additions in the releases in the associated documentation in Rammer et al. (2022) and Pucher & Neumann (2022)). The *easyclimate* R package enables easy and fast access to the latest version of the downscaled climate data (v3). We achieved this by implementing Cloud-Optimized GeoTIFFs (https://www.cogeo.org) which provide access to daily climate data for thousands of sites and days within minutes, without having to download huge rasters.

FUNCTIONALITY

The main function in *easyclimate* is *get_daily_climate*, which extracts daily climate data for a given set of coordinates (points or polygons) and a given period of days or years (see examples in *get_daily_climate* help page, and the vignettes Analysing the climate at spatial points for a given period and Analysing the climate of an area for a given period). The output can be either a `data.frame` (Table 1) or a (multilayer) `SpatRaster` object (see Figure 2) with daily climatic values for each point or polygon.

For example, to obtain precipitation data for a single site between 1st and 3rd of January 2001 we can run the next script:

```
library(easyclimate)

coords <- data.frame(lon = -5.36, lat = 37.40)

precip <- get_daily_climate(coords,
                            period = "2001-01-01:2001-01-03",
                            climatic_var = "Prcp")
```

To obtain a multilayer raster with values of maximum temperature for two days (1st January and 7th August 2012) for a region delimited by a polygon, we can do:

```
library(terra)

coords.poly <- vect("POLYGON ((-5.37 40.30, -5.17 40.30, -5.17 40.15, -
```

```
5.37 40.15))")

ras_tmax <- get_daily_climate(
  coords.poly,
  period = c("2012-01-01", "2012-08-07"),
  climatic_var = "Tmax",
  output = "raster"
  )
```

By design, *easyclimate* yields tidy datasets (Wickham 2014) that facilitate calculation of alternative climatic variables and indices following the tidyverse philosophy. In the next example we download daily climatic data (precipitation, minimum and maximum temperature) for a five-year period for a specific location and store them in a `data.frame`. Then we calculate the mean temperature.

```
library(dplyr)

coords <- data.frame(lon = -4.88, lat = 40.82)

daily_output <- get_daily_climate(coords,
                                  period = 2010:2015,
                                  climatic_var = c("Prcp", "Tmin",
"Tmax"))

daily <- daily_output %>%
  mutate(Tmean = (Tmin + Tmax) / 2)
```

To calculate average temperatures and aggregated precipitation by site or time period (Table 2) we can use `group_by` and `summarise` from `dplyr`, or by and `aggregate` from base R:

```
yearclimate <- daily %>%
  mutate(date = as.Date(date),
         year = as.factor(format(date, format = "%Y"))) %>%
  group_by(year) %>%
  summarise(Tmin.year = mean(Tmin),
            Tmean.year = mean(Tmean),
            Tmax.year = mean(Tmax),
            Prcp.year = sum(Prcp))
```

The results of the package *easyclimate* can be used directly or serve as input to calculate climatic indices with other packages, such as ClimInd (Reig-Gracia et al. 2021) or SPEI (Beguería & Vicente-

Serrano 2017) (see some examples in the vignette [Calculating basic climatic indices with data from easyclimate](#)).

easyclimate ADVANTAGES

Although the entire downscaled climatic data is available for download as GeoTIFF raster layers in a public FTP server (ftp://palantir.boku.ac.at/Public/ClimateData/), for small to moderately-sized areas (e.g. less than 10000 sites or 10000 km^2), the Cloud-Optimised GeoTIFF technology implemented in *easyclimate* allows to efficiently extract the data and can save significant time. Furthermore, with *easyclimate* we avoid downloading large rasters (several GB for each year) requiring storage space on local or remote servers, energy and resources (Hischier et al. 2015; Hilty & Aebischer 2015). In this sense, *easyclimate* becomes even more efficient if we are interested in climate data for multiple years and a small number of sites. For querying climate data from larger areas, it is recommended to download the raster layers and extract the data to local storage (e.g. using the `extract` function from `terra` R package, Hijmans (2021)), to avoid overloading the FTP server.

As a test comparing the two methodologies (i.e. raster downloading and local extraction and the *easyclimate* method), we downloaded daily precipitation data for one year in an area of *ca*. 100 km^2. While the local download and extraction took 9-10 minutes in a laptop with good internet connection and stored ~5960 Mb, *easyclimate* took ~17 seconds to obtain the same data storing only the final dataset (2.3 Mb).

```r
library(terra)

# Method 1: Raster downloading and local data extraction

coords.poly <- terra::vect("POLYGON ((-5.039 40.913, -4.919 40.913, -4.918 40.825, -5.039 40.825))")

raster.url <- "ftp://palantir.boku.ac.at/Public/ClimateData/v3/AllDataRasters/prec/DownscaledPrcp2010.tif"

options(timeout = max(10000, getOption("timeout")))

system.time({
   download.file(raster.url, destfile = "prcp2010.tif", mode = "wb")
   prcp2010.ras <- terra::rast("prcp2010.tif")
```

```
  prcp2010.data <- terra::extract(prcp2010.ras, coords.poly, xy = TRUE)
})

   user  system elapsed 
  18.55   76.25  562.56 

# Method 2: Obtain the same data using easyclimate

system.time(
  prcp2010.data_2 <- get_daily_climate(
    coords.poly,
    period = 2010,
    climatic_var = "Prcp"
  )
)

   user  system elapsed 
   2.14    2.76   17.21 
```

APPLICATIONS IN PLANT SCIENCES

Plant distribution as well as plant growth, phenology, respiration and mortality are strongly driven by weather conditions (e.g. Kunstler et al. 2021). Any aggregation of climate data to average monthly or annual numbers may hide important climate effects on plants specifically if we expect changing environmental conditions. In this sense, daily climate data are of interest for many ecological research questions and applications, including the study of the effects of late-spring frosts (Zohner et al. 2020), heat waves or dry periods on plant performance (Cruz-Alonso et al. 2020). However, accessing and processing such daily climate data is often cumbersome, even more if harmonized data are required at large spatial scales, and instead researchers use monthly or annual climate data.

In the next example we show how we can easily calculate the number of spring days with freezing temperatures (below zero), and the mean minimum temperature reached over several years (Table 3; see “Calculating basic climatic indices with data from easyclimate” vignette for other examples).

```
spring.months <- c("03","04","05") # March to May

springfrost_event <- daily %>%
  mutate(date = as.Date(date),
         month = format(date, format = "%m"),
         year = as.factor(format(date, format = "%Y"))) %>%
```

```
    filter(month %in% spring.months) %>%
    mutate(event = ifelse(Tmin < 0, 1, 0))

springfrost_peryear <- springfrost_event %>%
    group_by(year) %>%
    mutate(n.frost = sum(event))

springfrost <- springfrost_peryear %>%
    filter(event == 1) %>%
    group_by(ID_coords, year, n.frost) %>%
    summarise(Tmin.frost.avg = mean(Tmin))
```

By providing data with large temporal and spatial extent, *easyclimate* is a valuable tool with multiple applications in forestry, ecological and vegetation studies across Europe. For example, the harmonization and use of European forest inventories has been a priority for scientist and governments in the last decade (http://enfin.info, https://more.bham.ac.uk/treemort/datasets/, https://efi.int/knowledge/models/efiscen/inventory, http://project.fundiveurope.eu/). Often these repeat inventories are at a spatial resolution of few kilometers (Mauri et al. 2017), making *easyclimate* a powerful tool at extracting climatic information that can be used to support a better understanding of the effect of climate on forest dynamics at large spatial extents.

ACKNOWLEDGEMENTS

We acknowledge the support by the University of Natural Resources and Life Sciences, Vienna, Austria.

AUTHOR CONTRIBUTIONS

Conceptualization: V.C.-A., P.R.-B., and F.R.-S.
Data curation: C.P. and M.N.
Formal analysis: V.C.-A. and F.R.-S.
Funding acquisition: M.N. and H.H.
Investigation: C.P. and M.N.
Methodology: V.C.-A., S.R., P.R.-B., J.A., and F.R.-S.
Project administration: V.C.-A.
Resources: C.P., M.N., and H.H.
Software: V.C.-A., S.R., and F.R.-S.
Supervision: V.C.-A., P.R.-B., and F.R.-S.
Validation: V.C.-A., P.R.-B., J.A., and F.R.-S.

Visualization: V.C.-A., P.R.-B., and J.A.
Writing - original draft: V.C.-A. and F.R.-S.
Writing - review & editing: V.C.-A., C.P., S.R., P.R.-B., J.A., M.N., H.H., and F.R.-S.

DATA AVAILABILITY STATEMENT

The climate datasets are publicly available at ftp://palantir.boku.ac.at/Public/ClimateData/. The easyclimate R package is available on GitHub (https://github.com/VeruGHub/easyclimate).

REFERENCES

Archambeau, J., Ruiz-Benito, P., Ratcliffe, S., Fréjaville, T., Changenet, A., Muñoz Castañeda, J.M., Lehtonen, A., Dahlgren, J., Zavala, M.A., & Benito Garzón, M. 2020. Similar patterns of background mortality across Europe are mostly driven by drought in European beech and a combination of drought and competition in Scots pine. *Agricultural and Forest Meteorology* 280: 107772.

Beguería, S., & Vicente-Serrano, S.M. 2017. *SPEI: Calculation of the standardised precipitation-evapotranspiration index*.

Cornes, R.C., Schrier, G. van der, Besselaar, E.J.M. van den, & Jones, P.D. 2018. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres* 123: 9391–9409.

Cruz-Alonso, V., Rodríguez-Sánchez, F., Pucher, C., Ruiz-Benito, P., Astigarraga, J., Neumann, M., & Ratcliffe, S. 2021. *Easyclimate: Easy access to high-resolution daily climate data for europe*.

Cruz-Alonso, V., Villar-Salvador, P., Ruiz-Benito, P., Ibáñez, I., & Rey-Benayas, J.M. 2020. Long-term dynamics of shrub facilitation shape the mixing of evergreen and deciduous oaks in Mediterranean abandoned fields (C. García, Ed.). *Journal of Ecology* 108: 1125–1137.

Fick, S.E., & Hijmans, R.J. 2017. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302–4315.

George, J.-P., Neumann, M., Vogt, J., Cammalleri, C., & Lang, M. 2021. Assessing effects of drought on tree mortality and productivity in European forests across two decades: a conceptual framework and preliminary results. *IOP Conference Series: Earth and Environmental Science* 932: 012009.

Hasenauer, H., Merganicova, K., Petritsch, R., Pietsch, S.A., & Thornton, P.E. 2003. Validating daily climate interpolations over complex terrain in Austria. *Agricultural and Forest Meteorology* 119: 87–107.

Hijmans, R.J. 2021. *Terra: Spatial data analysis*.

Hilty, L.M., & Aebischer, B. (Eds.). 2015. *ICT Innovations for Sustainability. Advances in intelligent systems and computing.* Springer International Publishing, Cham.

Hischier, R., Coroama, V.C., Schien, D., & Ahmadi Achachlouei, M. 2015. Grey Energy and Environmental Impacts of ICT Hardware. In Hilty, L.M. & Aebischer, B. (eds.), pp. 171–189. Springer International Publishing, Cham.

Hlásny, T., Trombik, J., Bošeľa, M., Merganič, J., Marušák, R., Šebeň, V., Štěpánek, P., Kubišta, J., & Trnka, M. 2017. Climatic drivers of forest productivity in Central Europe. *Agricultural and Forest Meteorology* 234-235: 258–273.

Kunstler, G., Guyennon, A., Ratcliffe, S., Rüger, N., Ruiz-Benito, P., Childs, D.Z., Dahlgren, J., Lehtonen, A., Thuiller, W., Wirth, C., Zavala, M.A., & Salguero-Gomez, R. 2021. Demographic performance of European tree species at their hot and cold climatic edges (B. Leys, Ed.). *Journal of Ecology* 109: 1041–1054.

Mauri, A., Strona, G., & San-Miguel-Ayanz, J. 2017. EU-Forest, a high-resolution tree occurrence dataset for Europe. *Scientific Data* 4: 160123.

Micić Ponjiger, T., Lukić, T., Basarin, B., Jokić, M., Wilby, R.L., Pavić, D., Mesaroš, M., Valjarević, A., Milanović, M.M., & Morar, C. 2021. Detailed Analysis of Spatial–Temporal Variability of Rainfall Erosivity and Erosivity Density in the Central and Southern Pannonian Basin. *Sustainability* 13: 13355.

Moreno, A., & Hasenauer, H. 2016. Spatial downscaling of european climate data. *International Journal of Climatology* 36: 1444–1458.

Moreno, A., Neumann, M., & Hasenauer, H. 2018. Climate limits on European forest structure across space and time. *Global and Planetary Change* 169: 168–178.

Mosier, T.M., Hill, D.F., & Sharp, K.V. 2014. 30-Arcsecond monthly climate surfaces with global land coverage. *International Journal of Climatology* 34: 2175–2188.

Neumann, M., Moreno, A., Thurnher, C., Mues, V., Härkönen, S., Mura, M., Bouriaud, O., Lang, M., Cardellini, G., Thivolle-Cazat, A., Bronisz, K., Merganic, J., Alberdi, I., Astrup, R., Mohren, F., Zhao, M., & Hasenauer, H. 2016. Creating a Regional MODIS Satellite-Driven Net Primary Production Dataset for European Forests. *Remote Sensing* 8: 554.

Neumann, M., Mues, V., Moreno, A., Hasenauer, H., & Seidl, R. 2017. Climate variability drives recent tree mortality in Europe. *Global Change Biology* 23: 4788–4797.

Pietsch, S.A., & Hasenauer, H. 2006. Evaluating the self-initialization procedure for large-scale ecosystem models. *Global Change Biology* 12: 1658–1669.

Pucher, C., & Neumann, M. 2022. *Description and Evaluation of Downscaled Daily Climate Data Version 3. figshare. Online resource.*

R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rammer, W., Pucher, C., & Neumann, M. 2022. *Description, Evaluation and Validation of Downscaled Daily Climate Data Version 2. figshare. Online resource.*

Reig-Gracia, F., Vicente-Serrano, S.M., Dominguez-Castro, F., & Bedia-Jiménez, J. 2021. *ClimInd: Climate indices*.

Ruiz-Benito, P., Vacchiano, G., Lines, E.R., Reyer, C.P.O., Ratcliffe, S., Morin, X., Hartig, F., Mäkelä, A., Yousefpour, R., Chaves, J.E., Palacios-Orueta, A., Benito-Garzón, M., Morales-Molino, C., Camarero, J.J., Jump, A.S., Kattge, J., Lehtonen, A., Ibrom, A., Owen, H.J.F., & Zavala, M.A. 2020. Available and missing data to model impact of climate change on European forests. *Ecological Modelling* 416: 108870.

Thornton, P.E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., & Wilson, B.E. 2021. Gridded daily weather data for North America with comprehensive uncertainty quantification. *Scientific Data* 8: 190.

Wickham, H. 2014. Tidy Data. *Journal of Statistical Software* 59:

Zohner, C.M., Mo, L., Renner, S.S., Svenning, J.-C., Vitasse, Y., Benito, B.M., Ordonez, A., Baumgarten, F., Bastin, J.-F., Sebald, V., Reich, P.B., Liang, J., Nabuurs, G.-J., De-Miguel, S., Alberti, G., Antón-Fernández, C., Balazy, R., Brändli, U.-B., Chen, H.Y.H., Chisholm, C., Cienciala, E., Dayanandan, S., Fayle, T.M., Frizzera, L., Gianelle, D., Jagodzinski, A.M., Jaroszewicz, B., Jucker, T., Kepfer-Rojas, S., Khan, M.L., Kim, H.S., Korjus, H., Johannsen, V.K., Laarmann, D., Lang, M., Zawila-Niedzwiecki, T., Niklaus, P.A., Paquette, A., Pretzsch, H., Saikia, P., Schall, P., Šebeň, V., Svoboda, M., Tikhonova, E., Viana, H., Zhang, C., Zhao, X., & Crowther, T.W. 2020. Late-spring frost risk between 1959 and 2017 decreased in North America but increased in Europe and Asia. *Proceedings of the National Academy of Sciences* 117: 12192–12200.

Figure 1

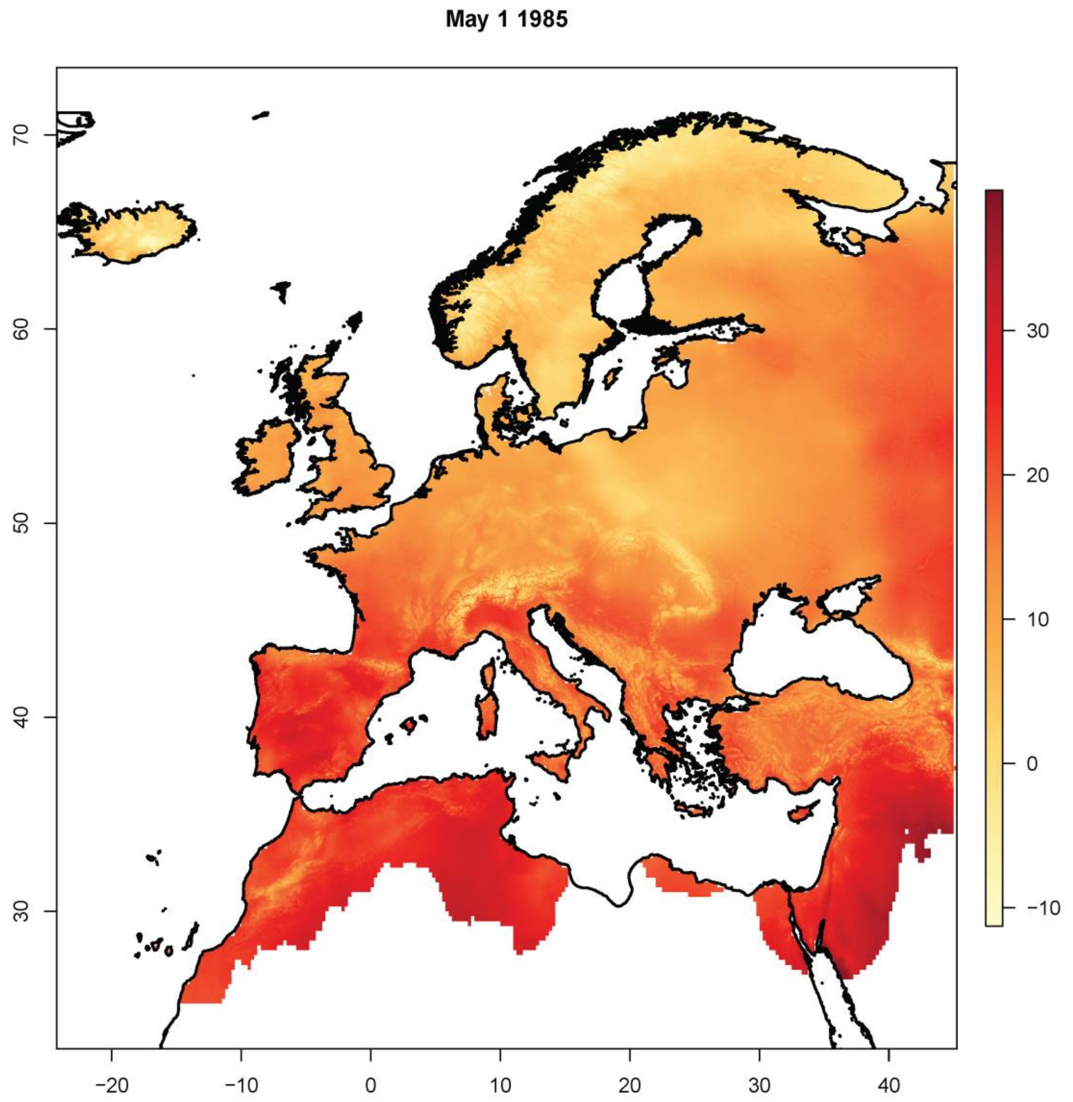

Figure 1: Example raster of maximum temperature (°C) to show the spatial coverage of the daily downscaled climate data.

Table 1

Table 1: Daily precipitation (Prcp; mm) for a given site obtained with easyclimate. Longitude (lon, x) and latitude (lat, y; WGS84) and date (YYYY-MM-DD) are also shown.

lon	lat	date	Prcp
-5.36	37.4	2001-01-01	8.64
-5.36	37.4	2001-01-02	0.00
-5.36	37.4	2001-01-03	2.93

Figure 2

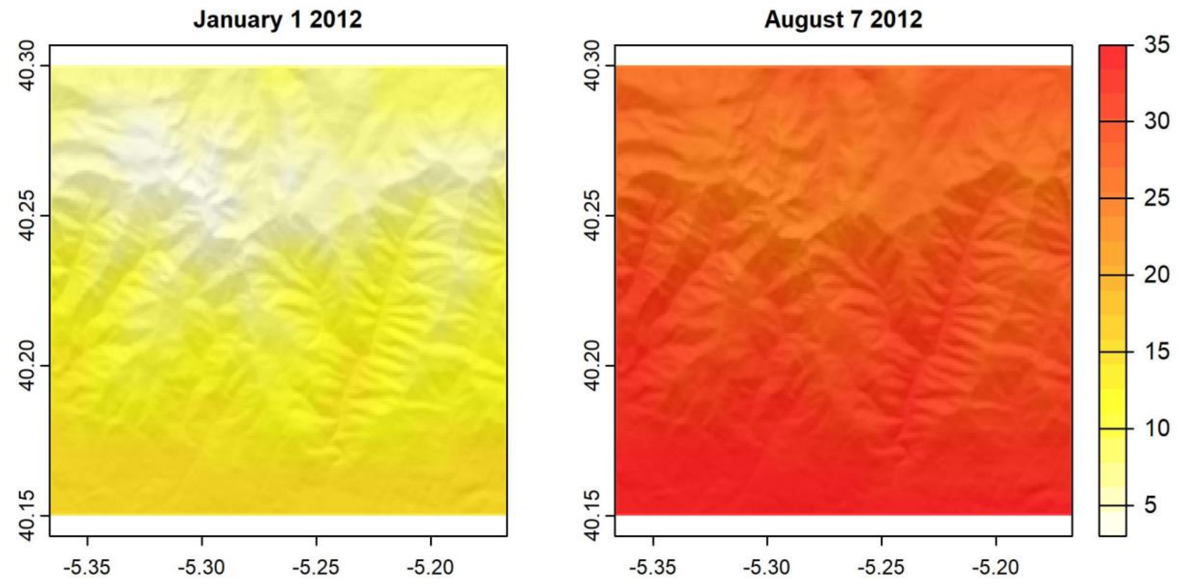

Figure 2: A multilayer raster of maximum temperature values for a given polygon in two different days of the year.

Table 2

Table 2: Yearly climatic values for a given site extracted with easyclimate. Tmin.year = Minimum temperature (ºC), Tmean.year = Mean temperature (ºC), Tmax.year = Maximum temperature (ºC), Prcp.year = Precipitation (mm)

year	Tmin.year	Tmean.year	Tmax.year	Prcp.year
2010	6.2	12.2	18.2	385.3
2011	6.9	13.3	19.7	281.9
2012	5.7	12.1	18.6	314.8
2013	5.7	11.8	18.0	354.9
2014	6.9	13.0	19.1	341.5
2015	6.7	13.3	19.8	221.3

Table 3

Table 3: No. of days with temperature below zero in spring (n.frost) and mean minimum temperature reached (Tmin.frost.avg; ºC) extracted with easyclimate

year	n.frost	Tmin.frost.avg
2010	10	-2.7790000
2011	4	-2.0325000
2012	11	-1.6290909
2013	11	-1.9190909
2014	3	-0.8733333
2015	7	-2.4728571