

bioclim: An R package for bioclimatic classifications via adaptive water balance

Roberto Serrano-Notivoli^{a,*}, Luis Alberto Longares^b, Rafael Cámara^c

^a Department of Geography, Universidad Autónoma de Madrid, Madrid, Spain

^b Department of Geography and Regional Planning, Universidad de Zaragoza, Zaragoza, Spain

^c Department of Geography, Universidad de Sevilla, Seville, Spain

ARTICLE INFO

Keywords:

Bioclimatology
Water balance
Geographic information systems
R package
Climate
Soil

ABSTRACT

bioclim is a software package in R language for bioclimatic classification based on the Type of Bioclimatic Regime approach, which combines climatic and soil properties to classify a region according to its suitability for plant vegetative activity. We present the software's operating modes, capabilities and limitations, including real-world usage examples. Using monthly temperature, precipitation, and field capacity as inputs, *bioclim* follows a straightforward workflow using three functions to compute: i) a comprehensive water balance describing water resource dynamics throughout the year; ii) a bioclimatic balance to estimate plant vegetative activity; and iii) a collection of bioclimatic intensities quantifying vegetative activity changes. The program uses the results of these functions to classify bioclimatic type at zonal, regional and local scales. The three functions' outputs can be calculated independently, strengthening the software's cross-disciplinary application potential, such as climatology and hydrology. *bioclim* uses numeric and raster formats as input data and contains highly flexible options for a wide range of purposes, from scientific research to end users' representations. The water and bioclimatic balance results can be presented in numerical, graphical, or cartographic forms.

1. Introduction

Bioclimatic classifications help to model the spatial distribution of vegetation and animals worldwide through climate data, due to their ability to synthesise the main climatic characteristics that are important for the presence of different vegetation types. These models have been applied on global (e.g., Kottek et al., 2006; Metzger et al., 2012) and national scales (e.g., Amigo and Ramírez, 1998; Djamali et al., 2011; Pesaresi et al., 2014) for purposes such as modelling vegetation response to climate change (Andrade and Contente, 2020; Kirilenko and Solomon, 1998) or defining climatic suitability for crops (Honorio et al., 2018; Serrano-Notivoli et al., 2020).

Among the most popular classification methods, Köppen developed his first version in the 19th century (Köppen, 1884), which has since been extensively used with few modifications until recently. This method uses a basic summary of precipitation and temperature to categorise areas' climates in terms of three parameters (i.e. climate zone, precipitation regime, and heat level). This approach has achieved good success in terms of biogeographical applications, however, it is too general for modelling vegetation habitats. In the second half of the 20th

century, more specific classifications were developed based on vegetation requirements, such as the Rivas-Martínez et al. (2002) model, which uses a combination of temperature and precipitation indices fitted to plant requirements, or the Holdridge (1967) model, based on the required temperatures for different plant development stages. While these studies establish major ecological classification units, they do not consider the influence of soil parameters as a basic condition of plant development in different climates – this factor is important since soil–climate interactions can be monitored through water balance, an aspect that was not fully considered by the above references. These studies instead used temperature–precipitation indices, thereby simplifying the evaporative losses. Evapotranspiration is a basic parameter derived from water balance to monitor water availability for plants and, while it can be derived only from temperature and precipitation, modern methods can improve its estimation accuracy using a few additional variables such as wind speed or humidity (El-kenawy et al., 2022; Kisi et al., 2022). The evapotranspiration calculation method should be chosen based on the type of climate (Vishwakarma et al., 2022) and using as many climatic variables as possible (Tomas-Burguera et al., 2019).

The Type of Bioclimatic Regime (TBR) method proposed by Cámara

* Corresponding author.

E-mail address: roberto.serrano@uam.es (R. Serrano-Notivoli).

et al. (2020) combines climatic and edaphic properties through a comprehensive water balance that reflects the local behaviour of climate–soil interactions throughout the year. The TBR workflow includes calculating intermediate variables such as: i) a water balance describing water resources throughout the year; ii) a bioclimatic balance to estimate the vegetative activity of plants; and iii) bioclimatic intensities as indicators of vegetative activity changes. These variables are of paramount importance for studies on a broad range of topics including hydrology, water resources, and ecosystem dynamics. Furthermore, vegetation growth is influenced by various climatic requirements that differ depending on the type of ecosystem and plant species. Thus, effectively Modelling the most suitable distribution areas for the growth of different plants requires the use of not only precipitation and temperature but also other derived variables that must be calculated.

The aims of our study are as follows: 1) translate the TBR method into an open-access, user-friendly package to create bioclimatic classifications; 2) improve the spatial analysis of TBR results by integrating geospatial information; and 3) make the approach useful to a wide audience through flexible inputs in terms of water balance calculation and operating options.

Some existing packages of functions are available to compute climatic or bioclimatic classifications, however, Most of them are devoted to calculating a wide variety of climatic, bioclimatic, or agroclimatic indicators derived from precipitation and temperature. Although useful, these packages do not perform bioclimatic classifications because they focus only on climatic aspects from the perspective of vegetation and not as part of an integrated approach, as described by Cámara et al. (2020). Other approaches include ANUCLIM (Xu and Hutchinson, 2013), which computes several climate-based indices, similar to ENVIREM (Title and Bemmels, 2018); ClimClass (Eccel et al., 2016) and kgc (Bryant et al., 2017) can also accomplish Köppen classifications. Furthermore, WatBal (Srinivas et al., 2016) fits a water balance for bioclimatic classifications.

In this paper, we describe for the first time the operation and possibilities of the *bioclim* software, a collection of functions distributed as an R package that can classify data series or raster collections into a bioclimatic classification fitted to vegetation requirements through environmental parameters. This method requires only monthly temperature and precipitation data and a field capacity value representing the soil characteristics. The *bioclim* package is highly flexible and accepts input data in raster format; In this case, the output will be a multi-layer raster. The novelty of this package lies in the software's capability to compute bioclimatic classifications in a geospatial format at any spatial resolution. One of the advantages of *bioclim* over conventional GIS software is the automatic computation of the classifications, without the

need to use multiple layers. Furthermore, the results in both numerical or geospatial formats allow a comprehensive and fully adaptive water balance to be independently computed for additional analyses.

The *bioclim* R package is free and open-source software available through the CRAN repository (<https://cran.r-project.org/package=bioclim>). The package was released under the General Public Licence v3.0 (GPL-3.0) and runs on all the operating systems supported by R (>4.1). *bioclim* was written in R and the program files are <1 MB in size. This is the first release of the software, and its design and development has had no previous versions. Its open-access nature allows for continuous improvement and evolution based on user feedback and reports. The package's functions are based on the theoretical work of Cámara et al. (2020) to address a gap in existing automatic bioclimatic classification methods.

2. Software design and characteristics

The package performs bioclimatic classification through the computation of: 1) a complete water balance to specify the theoretical monthly water resource dynamics and 2) a bioclimatic balance to relate climate and plant development through calculating vegetative activity (Fig. 1). Depending on the input type, which can be numeric or geospatial (i.e. raster) data, the functions can yield outputs in alphanumeric format (table or text) or as a collection of georeferenced layers with the same geographic characteristics (e.g. extent, resolution, projection system, etc.) as the input data. The bioclimatic balance, intensities, and classification are based on the water balance, which forms the core of this approach's methodology. While most of the parameters (e.g. inputs, evapotranspiration method) are user-defined, the main limitations of *bioclim* are from the input data, which condition the final results depending on the data quality, spatial resolution, and representativeness.

2.1. Input data

The only three required inputs for *bioclim* are the temperature (T), precipitation (P), and field capacity (CC). The first two parameters can be numeric vectors of length 12 containing average monthly values of T and P at a single location; in this case, the latitude is also required for internal theoretical calculations of evapotranspiration. The other option is to provide 24 raster layers (i.e. one for each month and variable) with identical spatial domain, resolution, and coordinate system. When T and P are provided as raster inputs, the outputs are always in raster format.

Field capacity is a theoretical calculation of the soil's water retention

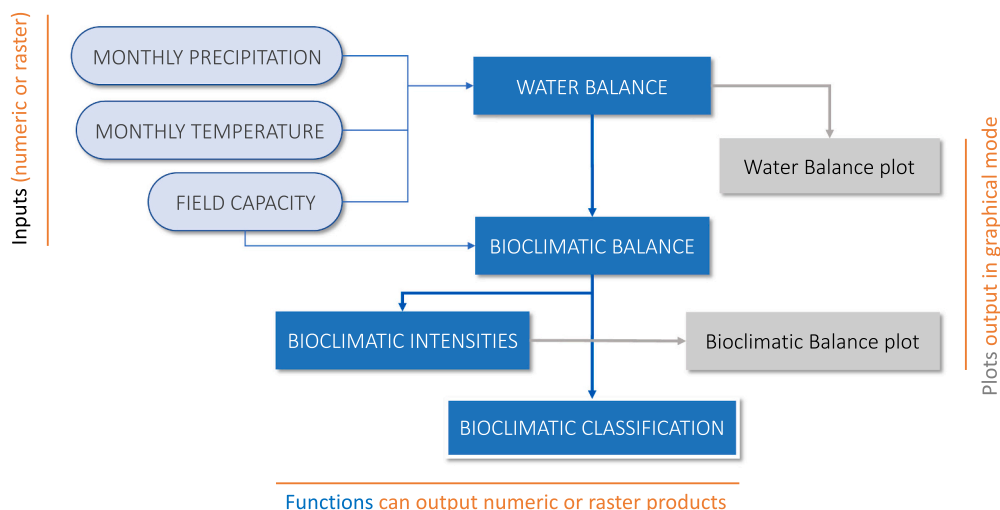


Fig. 1. Workflow of *bioclim* package.

capacity. The depth of the soil where water loss occurs by evapotranspiration is defined by the depth of the root system of the existing vegetation. The proportions of sand-, silt-, and clay-sized particles define the porosity of a soil and, therefore, its water retention capacity; However, this capacity is influenced by the type of vegetation present above the soil which, depending on the depth of the roots, may decrease or favour water accumulation. On this premise, López Cadenas et al. (1986) mathematically established CC as:

$$CC = RC \bullet Rd$$

where RC is the retention capacity of the water (in millimetres) in soil texture types, and Rd is the average root depth (in metres) of different types of vegetation and crops. The authors provided guiding values for both variables (see Table S1). CC can be provided as an input in all functions of *bioclim*, either as a single numerical value or as a raster layer.

2.2. Water balance

The Water balance calculation is the first step used in assessing monthly water availability; in *bioclim*, this calculation is based on the application of the Thornthwaite and Mather (1955) method. If provided, this function accepts a pre-computed potential evapotranspiration set of monthly layers (only in raster mode) as an input instead of using the Thornthwaite–Mather method.

The function outputs 12 variables (Table 1) at a monthly scale that can be grouped into three categories, i.e. climatic (*T*, *P*, *PET*, *TEAW*), soil (*PAWL*, *ST*, *i_{ST}*, *RET*, *MD*, *ME*), and hydrological (*r*, *rP*) variables. When working with numerical inputs, the *watbal()* function must be used, with the results provided as a table, whereas raster inputs require the use of the *watbalRaster()* function, which produces 144 (.tif) files (i.e. 12 monthly layers for each of the 12 variables).

2.3. Bioclimatic balance

The bioclimatic balance (BB) calculation in *bioclim* is adapted from Montero de Burgos and González (1974) and aims to calculate the production based on temperature and rainfall patterns. The BB calculation estimates, throughout the year, periods of positive vegetative activity (i.e. with production of biomass) and periods of vegetative activity stagnation (i.e. without production of biomass), relating climate and plant development. The variables derived from the water balance step help to calculate the indicators of these changes, called bioclimatic intensities (BIs). These BIs provide insights into plant vegetative activity or stagnation and, therefore, their productivity (biomass productivity) in a specific climate. Four BIs are computed by the BB: potential (*PBI*), real (*RBI*), conditioned (*CBI*), and free (*FBI*), representing the vegetative activity based on the water balance provided as an input.

The relevant bioclimatic balance function is *biobal()*; when using numerical inputs, this function yields eight variables (Table 2) at a monthly scale that are required to compute the four BIs (Table 3). In contrast, The raster mode requires the *biobalRaster()* function, which only yields the four BIs on a monthly scale.

BIs are measured in bioclimatic units (*bcu*), which indicate the climate productivity assuming there are no deficiency factors. A single *bcu* is equivalent to the number of plant material cubic metres per hectare in one year, thus the BIs are a convenient indicator of biomass production.

The BB output is a table containing monthly values; assuming the function input is numerical and not raster, the results can be plotted to create a graphical visualisation of the BIs. However, this graphic can be further improved with the calculation, through the *bioint()* function, of two variants of the original four BIs (i.e. *PBI*, *RBI*, *CBI* and *FBI*): warm (w) and cold (c). The warm BIs measure positive vegetative activity, whereas the cold BIs measure the degree of vegetative stagnation. In

Table 1

Output variables from water balance calculation.

Variable	Description	Mode of calculation
<i>T</i>	Monthly temperature	Same values as input
<i>P</i>	Monthly precipitation	Same values as input
<i>PET</i>	Potential evapotranspiration	Adapted from Beguerfa and Vicente-Serrano (2017).
<i>TEAW</i>	Theoretical exceedance of available water	Difference between monthly precipitation and potential evapotranspiration (m) $TEAW_m = P_m - PET_m$
<i>PAWL</i>	Potential accumulated losses of water in each month	Represents the accumulated negative differences between precipitation and potential evapotranspiration. Calculated using a sequential process starting in the first month where $TEAW < 0$. $PAWL_m \begin{cases} 0 & TEAW_m \geq 0 \\ PALW_{m-1} + TEAW_m & TEAW_m < 0 \end{cases}$
<i>ST</i>	Water (moisture) stored in soil	Represents the accumulated water in the soil, with the field capacity as its upper limit. When $TEAW$ is positive and greater than CC , then $ST_m = CC$. Otherwise, the monthly ST is calculated as follows: $ST_m \begin{cases} ST_{m-1} + TEAW_m & TEAW_m \geq 0 \\ CC \bullet \exp(PALW_m/CC) & TEAW_m < 0 \end{cases}$
<i>i_{ST}</i>	Soil moisture change compared to previous month	$i_{ST}_m = ST_m - ST_{m-1}$
<i>RET</i>	Real evapotranspiration	When P is greater than $TEAW$, then $RET = PET$; otherwise, RET depends on the monthly changes in soil moisture. $RET_m \begin{cases} PET_m & P_m \geq PET_m \\ P_m + i_{ST}_m & P_m < PET_m \end{cases}$
<i>MD</i>	Moisture deficit	Difference between RET and PET . $MD_m = RET_m - PET_m$
<i>ME</i>	Moisture excess	When $TEAW$ is positive, the moisture excess is set as the difference between the precipitation and the sum of RET and the monthly change in soil moisture. $ME_m \begin{cases} P_m - (RET_m + i_{ST}_m) & TEAW_m > 0 \\ 0 & TEAW_m \leq 0 \end{cases}$
<i>r</i>	Surface runoff	This variable is conditioned to the months in which $TEAW$ is positive and ST is greater than CC . $r_m \begin{cases} 0.5(ME_m + r_{m-1}) & TEAW_m > 0 \text{ and } ST_m > CC \\ 0 & \text{otherwise} \end{cases}$
<i>rP</i>	Percentage of precipitation occurring as runoff	$rP_m = r_m \bullet 100/P_m$

All of the above variables can be plotted with the *plotWatbal()* function, which requires the result of the water balance function as an input. The resulting plots show the variables' values on a monthly scale.

addition, the function calculates a fifth BI, the Dry Bioclimatic Intensity (DBI), which quantifies the drought intensity during the dry season (if it exists), thus measuring biomass productivity stagnation due to drought. In summary, the *bioint()* function yields 10 variables:

- *PBIw*, *RBIw*, *CBIw*, *FBIw*, *DBIw*: Potential, Real, Conditioned, Free, and Dry bioclimatic intensities, respectively, during positive plant vegetative activity.
- *PBIc*, *RBIc*, *CBIc*, *FBIc*, *DBIc*: Potential, Real, Conditioned, Free and Dry bioclimatic intensities, respectively during the absence of vegetative activity.

The result of the *bioint()* function is required to run the *plotBiobal()* function, which produces a plot of the monthly values of the BIs to evaluate vegetative activity throughout the year, including showing the months in which hydric/thermal vegetative stagnation occur.

Table 2
Output variables from bioclimatic balance calculation.

Variable	Description	Mode of calculation
<i>AIP</i>	Available infiltrated precipitation	The result of weighting monthly precipitation with that month's surface runoff. <i>rP</i> from water balance is used to compute this parameter. $AIP_m = C_{em} \cdot P_m$, where $C_{em} = 1 - (rP_m / 100)$
<i>T</i>	Monthly temperature	Same values as <i>T</i> from the water balance.
<i>PET</i>	Potential evapotranspiration (Thornthwaite, 1948)	This parameter helps to identify the water requirements of vegetation. Same values as <i>PET</i> from water balance.
<i>RE</i>	Residual evapotranspiration	When the available water is less than <i>PET</i> , vegetative activity progressively decreases and plants activate their defence mechanisms (closing of stomas, defoliation, etc.). At this point, plant evapotranspiration continues but at its minimum, considered to be 20% of <i>PET</i> by Montero de Burgos and González (1974). $RE_m = PET_m / 5$
<i>AW</i>	Available water	The sum of monthly precipitation and the surplus (<i>S</i>) from the previous month. $AW_m = P_m \cdot S_{m-1}$
<i>S</i>	Surplus	Represents the excess water in the month, calculated as the difference between <i>AW</i> and <i>PET</i> . $S_m = AW_m \cdot PET_m$ If <i>S</i> is negative, it will take a value determined by the relationship between the available water (<i>AW</i>) and the residual evapotranspiration (<i>RE</i>): $S_m \begin{cases} 0 & AW_m > RE_m \\ RE_m - AW_m & otherwise \end{cases}$
<i>CWA</i>	Coefficient of water availability	<i>CWA</i> is expressed as a ratio between the difference in available water and residual evapotranspiration and the difference between <i>PET</i> and residual evapotranspiration. $CWA_m = (AW_m - RE_m) / (PET_m - RE_m)$
<i>T75</i>	Temperature minus 7.5 °C	The result of subtracting 7.5 from each monthly temperature. Positive results are directly related to an increased plant growth velocity. The minimum temperature of 7.5 °C is the level at which plant vegetative activity begins, assuming no other soil or climatic limitations.

2.4. Bioclimatic classification

Bioclimatic classification is performed through the function *biotype()* when working with numeric inputs and with function *biotypeRaster()* when the inputs are in raster format. Both functions accept multiple data types, from raw precipitation and temperature variables to pre-computed water and bioclimatic balances.

Three classification types are performed depending on the spatial scale (Fig. 2): 1) zonal (macro-scale), based on large, latitude-based climatic groups; 2) regional (meso-scale), also called TBRs, based on thermal and humidity requirements of vegetation; and 3) local (micro-scale), where the TBRs are fine-tuned depending on their humidity character.

Using climatic data (from stations or raster products), Zonal classification defines the region in terms of major climatic units. The results from water and bioclimatic balances then provide the TBR regional classification, which is based on the limits of temperature and precipitation for plant development. Finally, the Thornthwaite index defines the local effect of climate on soil moisture (Thornthwaite, 1948).

The TBR regional classification translates climatic data into biological restrictions. For instance, temperature limitations can determine five different scenarios for plant development:

Table 3
Bioclimatic intensities calculated from bioclimatic balance output.

Bioclimatic intensity	Description	Mode of calculation/Interpretation
<i>PBI</i>	Potential bioclimatic intensity	The theoretical BI when there are no water restrictions, representing a measure of the maximum vegetative activity. <i>PBI</i> can be assumed as the vegetative activity of irrigated land.
<i>RBI</i>	Real bioclimatic intensity	<i>RBI</i> is calculated through the water availability based on the observed climatic values. As vegetative activity depends on water availability throughout the year, this parameter can proportionally decrease depending on the relation between <i>PET</i> and <i>AW</i> .
<i>CBI</i>	Conditioned bioclimatic intensity	<i>CBI</i> represents the period of water recovery after a drought until the water input compensates for the imbalance, i.e. until plants have sufficient water to recover their vegetative activity. Different species are variably adapted to <i>CBI</i> periods depending on their water requirements; for instance, herbaceous plants require shorter <i>CBI</i> periods to compensate for drought than larger species.
<i>FBI</i>	Free bioclimatic intensity	<i>FBI</i> occurs in periods without drought, after <i>CBI</i> , and represents the capacity of the location (based on climatic data) to produce biomass, considering the limitations of drought and cold. It is equivalent to the <i>RBI</i> .

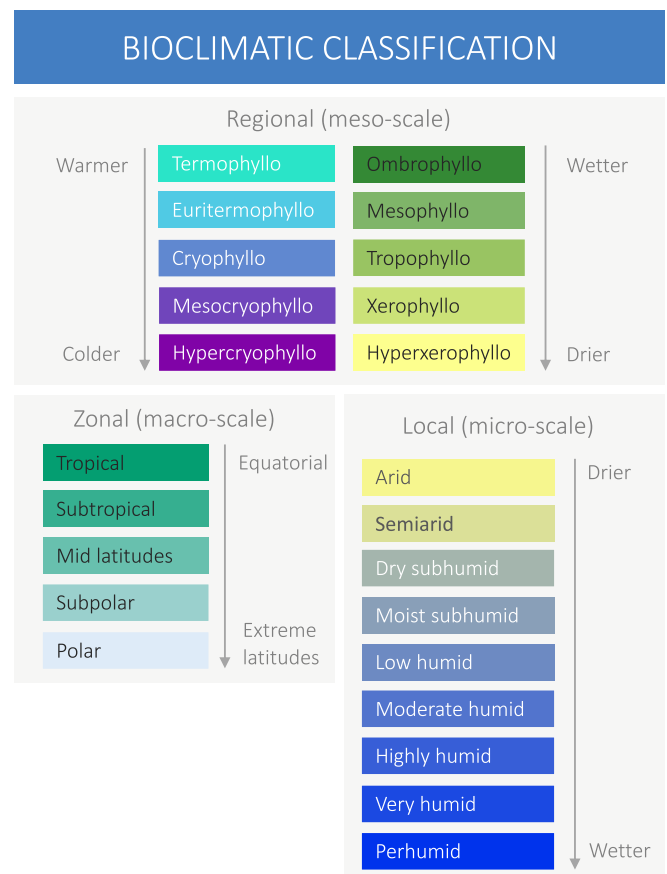


Fig. 2. Bioclimatic types resulting from classification at different spatial scales.

- *Termophyllia*: absence of thermal restrictions and low annual thermal oscillation.
- *Euritermophyllia*: significant temperature variation throughout the year but without vegetative activity stagnation due to thermal causes.
- *Cryophyllia*: cold temperatures produce one to five months of vegetative activity stagnation.

- *Mesocryophyllia*: cold temperatures produce six to nine months of vegetative activity stagnation, conditioning the presence of broad-leaved trees.
- *Hypercryophyllia*: cold temperatures produce six to nine months of vegetative activity stagnation, conditioning all woody plants.

In parallel, water scarcity also determines various limitations

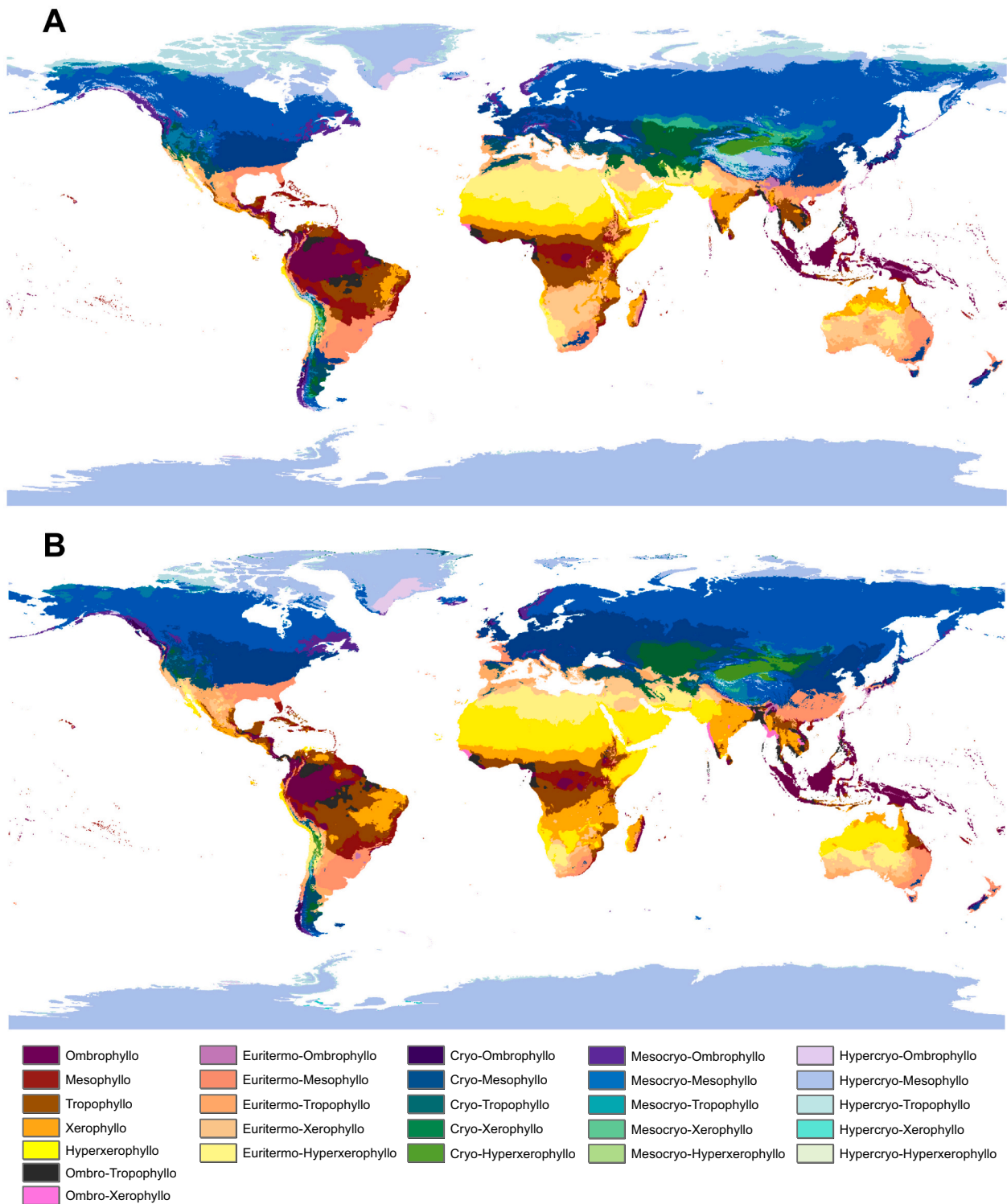


Fig. 3. Bioclimatic classification based on TBRs under (A) historical climatic conditions (1970–2000) and (B) a severe future climate change scenario (SSP585, 2061–2080).

depending on its intensity:

- *Ombrophyllia*: absence of water restrictions (i.e. precipitation in all months exceeds 60 mm).
- *Mesophyllia*: moderate water scarcity without vegetative activity stagnation. Some months with water deficit in soil.
- *Tropophyllia*: water scarcity with vegetative activity stagnation from one to four months.
- *Xerophyllia*: water scarcity with prolonged vegetative activity stagnation from five to eight months.
- *Hyperxerophyllia*: water scarcity with very long vegetative activity stagnation from nine to 12 months.

Using this classification on worldwide climate data, [Cámara et al. \(2020\)](#) identified five macro-scale zonal categories, 27 meso-scale TBRs, and 157 micro-scale subtypes.

3. Examples of potential applications

To show the potential applications of the package, we present examples at three spatial scales: 1) global, illustrating the bioclimatic classification of the entire world's land area at both the present time and under a severe climate change scenario; 2) regional, showing a collection of output variables derived from a water balance calculation of the Mediterranean basin; and 3) local, performing a complete bioclimatic classification of a mid-latitude city.

We used present (1970–2000) and future (2061–2080) climatic data (temperature and precipitation) from WorldClim (version 2.1) ([Fick and Hijmans, 2017](#)) and a constant value of 400 as the field capacity. The Potential evapotranspiration was calculated using the default Thornthwaite's method in *bioclim*. The complete code for reproducing the examples is included as supplementary material.

3.1. Global scale: variations of bioclimates under climate change scenarios

In this example, we compute the complete bioclimatic classification ([Cámara et al., 2020](#)) for the whole Earth's land area both at present and under a severe climate change scenario ([Fig. 3](#)). While the calculation is relatively time-consuming due to the spatial resolution of the raster layers used, the coding is simple:

```
waterbalance <- watbalRaster(temp, prec, CC = 400)
biotypes <- biotypeRaster(bh = waterbalance)
```

Several options can be easily specified, such as defining the number of CPU cores used to perform the computation or the output folder if files should be saved to disk. For this example, both functions were run twice in raster format, with *temp* and *prec* changed to the historical or future climate information as required. In both cases, the *biotypeRaster()* function yielded three layers with classifications at different spatial scales; here, We show only the meso-scale TBR classification.

3.2. Regional scale: water balance of the Mediterranean Basin

The *bioclim* package yields not only the bioclimatic classifications as a final product but also the results of intermediate functions. In this example at a regional scale, the *watbalRaster()* function is used to create a collection of maps of various water balance variables in the Mediterranean basin on a monthly scale. The results of PET and RET for May to August ([Fig. 4](#)) show a latitudinal gradient from humid to dry conditions from north to south in all months in both variables, except for in areas of mountain ranges. While PET ranges from <50 to >250 mm in this gradient, RET shows significantly lower values (maxima below 200 mm) with an inverted spatial pattern, i.e. lower RET in the south compared to the north. The RET reflects more realistic evapotranspiration due to the

accumulated water loss in previous months, meaning that, although temperature conditions would favour high theoretical evapotranspiration rates in July and August, there is no remaining available water to evaporate and the incident precipitation is insufficient to increase the evapotranspiration rates.

3.3. Local scale: complete bioclimatic classification of a single location

The functions from the previous examples can be also applied to numerical values for a single location. In this example, a complete bioclimatic balance is demonstrated for the city of Seville (Spain).

This analysis commences with the water balance step by manually introducing monthly temperature and precipitation values, which produces a table containing all the variables on a monthly scale ([Table 4](#)).

```
wb <- watbal(t = c(10,11.5,14,16.5,20,24.5,27.5,28,24.5,19.5,
14.5,11),
p = c(55,73,84,58,33,23,2,2,28,66,94,71),
lat = 37.38,
CC = 400)
```

All the variables of the water balance are simplified in the resulting plot ([Fig. 5](#)) through four coloured categories (plus one transparent), representing the soil's water management during every month of the year ([Table 5](#)). (See [Table 6](#).)

The bioclimatic balance, through the *biobal()* function, produces a monthly table with eight variables and four bioclimatic intensities ([Table 2](#)) that must be refined to compute warm and cold variants. These variants illustrate periods of vegetative activity stagnation and are calculated through the *bioint()* function; this function produces monthly values that are used in the *plotBiobal()* function to plot the complete bioclimatic balance throughout the year. For more details on the calculation of bioclimatic intensities, see [Fidalgo and Cámara \(2022\)](#).

```
bb <- biobal(wb, CC = 400)
intens <- bioint(bb)
plotBiobal(intens)
```

The single-location bioclimatic balance example (Seville, [Fig. 6](#)) shows a climatic regime with mild temperatures and low rainfall, leading to an absence of cold-induced vegetative activity stagnation. In this location, plant development is not temperature-limited as all months are above 7.5 °C, consequently, there are no cold bioclimatic intensities. The only pause in plant development is in August due to dry conditions, which is the result of four consecutive months of water deficit and intense water demand by the soil, as shown in the water balance ([Fig. 5](#)).

The bioclimatic balance calculation provides the required information to classify the bioclimate at different spatial scales. The *biotype()* function assists in this process through the *mode* argument, which accepts a value of 'zonal' for macro-scale classification, 'TBR' for meso-scale classification, and 'sub' for micro-scale classification. The last option requires the water balance and field capacity because this approach is based on the calculation of Thornthwaite's index.

```
biotype(bb = bb, mode = 'zonal')
biotype(bb = bb, mode = 'TBR')
biotype(wb = wb, CC = 400, mode = 'sub')
```

In the above example, the city of Seville is classified as a 'Subtropical' climate at a large scale, the 'Euritermo-Mesophyllo' with the TBR option, and 'Semiarid Euritermo-Tropophyllo' using the complete classification method.

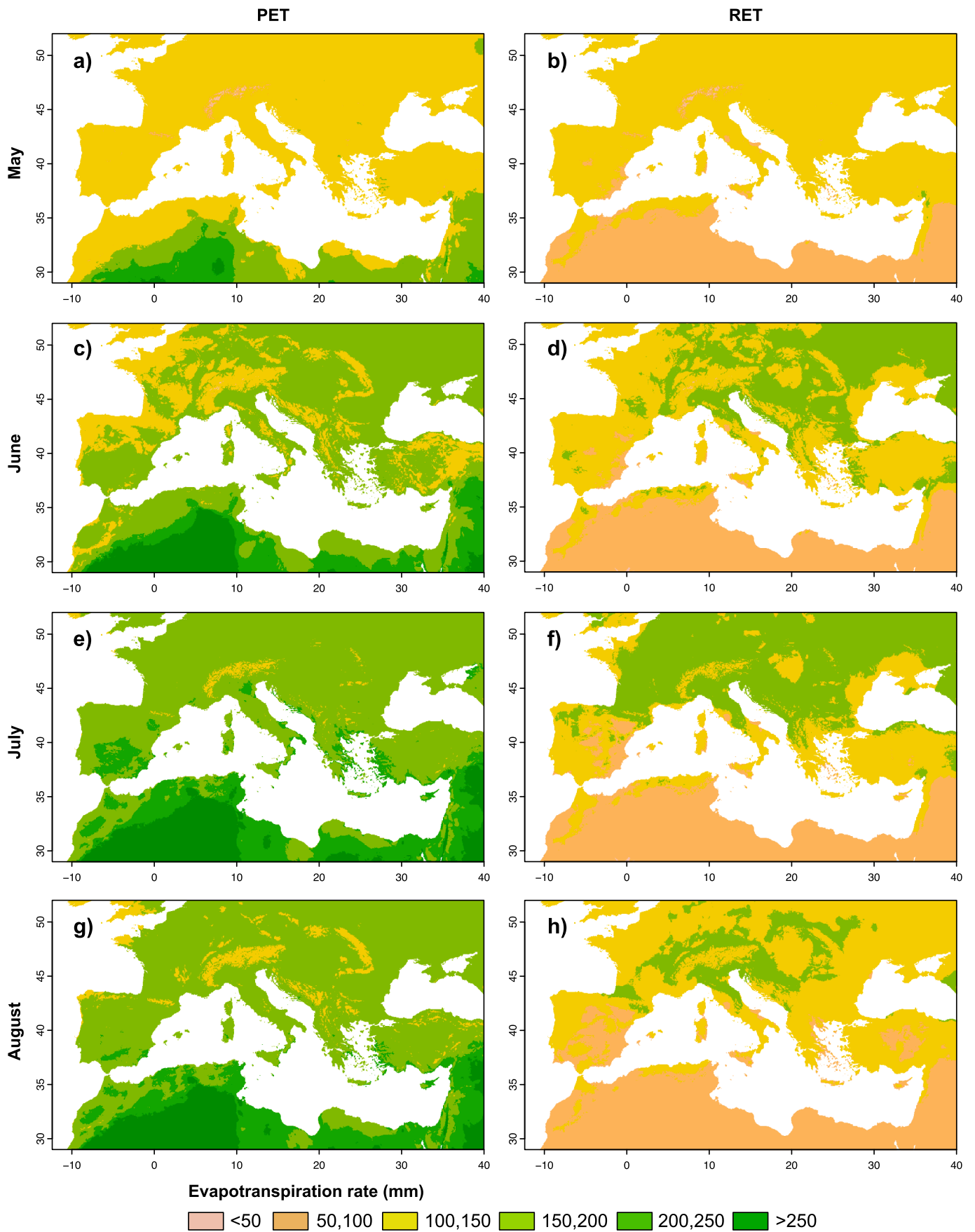


Fig. 4. Potential (left column) and real (right column) evapotranspiration in the Mediterranean basin calculated for May, June, July, and August (rows).

Table 4
Water balance for Seville (Spain).

Month	T	P	PET	TEAW	PALW	ST	i_ST	RET	MD	ME	r	rP
J	10.0	55.0	16.5	38.5	0	218.1	38.5	16.5	0	0	0	0
F	11.5	73.0	21.2	51.8	0	269.9	51.8	21.2	0	0	0	0
M	14.0	84.0	38.4	45.6	0	315.5	45.6	38.4	0	0	0	0
A	16.5	58.0	56.8	1.2	-93.4	316.7	1.2	56.8	0	0	0	0
M	20.0	33.0	92.6	-59.6	-153	272.9	-43.8	76.8	-15.7	0	0	0
J	24.5	23.0	138.5	-115.5	-268.4	204.5	-68.4	91.4	-47.1	0	0	0
J	27.5	2.0	177.5	-175.5	-443.9	131.8	-72.6	74.6	-102.9	0	0	0
A	28.0	2.0	173.4	-171.4	-615.4	85.9	-46.0	48.0	-125.5	0	0	0
S	24.5	28.0	117.6	-89.6	-704.9	68.7	-17.2	45.2	-72.3	0	0	0
O	19.5	66.0	69.5	-3.5	-708.4	68.1	-0.6	66.6	-2.9	0	0	0
N	14.5	94.0	34.1	59.9	0	128	59.9	34.1	0	0	0	0
D	11.0	71.0	19.4	51.6	0	179.6	51.6	19.4	0	0	0	0

The values of the table are used to produce the graphical representation that can be built with the *plotWatbal()* function.

plotWatbal(wb)

T: temperature (°C), P: precipitation (mm), PET: potential evapotranspiration (mm), TEAW: theoretical exceedance of available water (mm), PALW: potential accumulated loss of water (mm), ST: soil moisture (mm), i_ST: change in soil moisture (mm), RET: real evapotranspiration (mm), MD: Moisture deficit (mm), ME: Moisture excess (mm), r: surface runoff (mm), and rP: percentage of runoff (%).

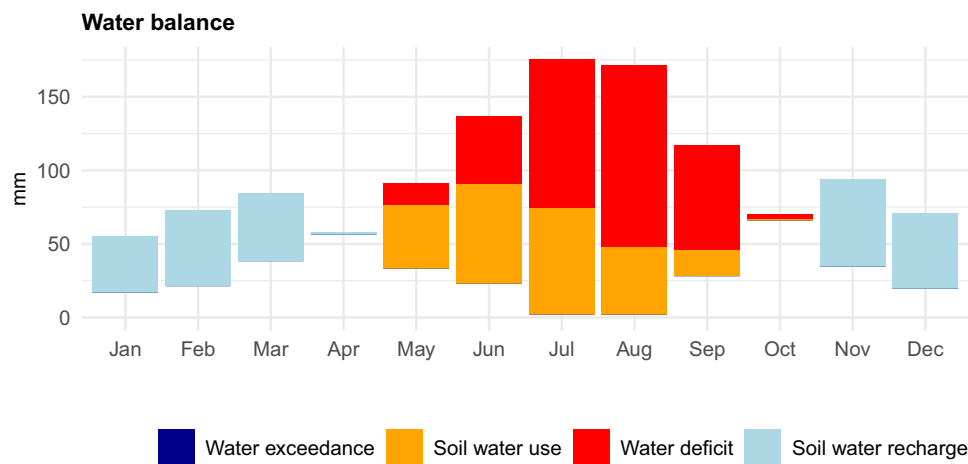


Fig. 5. Water balance graphical representation of Seville (Spain).

Table 5
Variables calculated to build the water balance graph.

Variable	Description	Mode of calculation
<i>BL</i>	Base level	Transparent; represents the starting point of the available water in the month. Its value is the minimum of P and PET. $BL_m = \min \{P_m, PET_m\}$
<i>WE</i>	Water exceedance	Shown in dark blue; occurs when the real and potential evapotranspiration are equivalent in the same month and a moisture excess exists. $WE_m \begin{cases} P_m - BL_m & (RET_m = PET_m) \text{ ME}_m > 0 \\ 0 & \text{otherwise} \end{cases}$
<i>SWU</i>	Soil water use	Shown in orange; represents the months in which PET is higher than P. $SWU_m \begin{cases} RET_m - (BL_m + WE_m) & TEAW_m > 0 \\ 0 & TEAW_m \leq 0 \end{cases}$
<i>WD</i>	Water deficit	Shown in red; indicates the absence of available water in the months when the potential evapotranspiration is higher than the real evapotranspiration. $WD_m \begin{cases} PET_m - (BL_m + WE_m + SWU_m) & PET_m > RET_m \\ 0 & PET_m \leq RET_m \end{cases}$
<i>SWR</i>	Soil water recharge	Shown in light blue; quantifies the water excess based on the difference between precipitation and the sum of the previous factors. $SWR_m \begin{cases} P_m - (BL_m + WE_m + SWU_m + WD_m) & TEAW_m \geq 0 \text{ ((RET}_m = PET_m) \text{ ME}_m = 0) \\ 0 & \text{otherwise} \end{cases}$

4. Discussion and conclusions

While some existing software programs already calculate climatic classifications (e.g. Bryant et al., 2017; Eccel et al., 2016), the *bioclim* package extends far beyond traditional approaches that use a few functions with the aim of performing a complete bioclimatic analysis of a single location or a region of any size using raster data as an input. The

bioclim package provides three simple functions to compute water balance, bioclimatic balance, and bioclimatic classification using only monthly temperature and precipitation as input data and a field capacity value. While these values can be long-term averages or theoretical estimates, as shown in the examples in this paper, field observations could also be used as input to obtain a reliable representation of an area's short-term water and bioclimatic balance. For example, *bioclim* has

Table 6
Bioclimatic balance of Seville (Spain).

Month	AIP	T	PET	RE	AW	S	CWA	T75	PBI	RBI	FBI	CBI
J	55.0	10.0	16.5	3.3	159.9	143.4	11.9	2.5	0.5	0.5	0.5	0
F	73.0	11.5	21.2	4.2	216.4	195.2	12.5	4.0	0.8	0.8	0.8	0
M	84.0	14.0	38.4	7.7	279.2	240.8	8.8	6.5	1.3	1.3	1.3	0
A	58.0	16.5	56.8	11.4	298.8	242	6.3	9.0	1.8	1.8	1.8	0
M	33.0	20.0	92.6	18.5	275.0	182.4	3.5	12.5	2.5	2.5	2.5	0
J	23.0	24.5	138.5	27.7	205.4	66.9	1.6	17.0	3.4	3.4	3.4	0
J	2.0	27.5	177.5	35.5	68.9	0	0.2	20.0	4	0.9	0.9	0
A	2.0	28.0	173.4	34.7	2.0	0	-0.2	20.5	4.1	-1	0	0
S	28.0	24.5	117.6	23.5	28.0	0	0	17.0	3.4	0.2	0	0.2
O	66.0	19.5	69.5	13.9	66.0	0	0.9	12.0	2.4	2.2	1	1.2
N	94.0	14.5	34.1	6.8	94.0	59.9	3.2	7.0	1.4	1.4	1.4	0
D	71.0	11.0	19.4	3.9	130.9	111.5	8.2	3.5	0.7	0.7	0.7	0

AIP: Available infiltrated precipitation (mm), T: temperature (°C), PET: potential evapotranspiration (mm), RE: residual evapotranspiration (mm), AW: available water (mm), S: surplus (mm), T75: temperature excess over 7.5 °C (°C), PBI: potential bioclimatic intensity, RBI: real bioclimatic intensity, FBI: free bioclimatic intensity, and CBI: conditioned bioclimatic intensity.

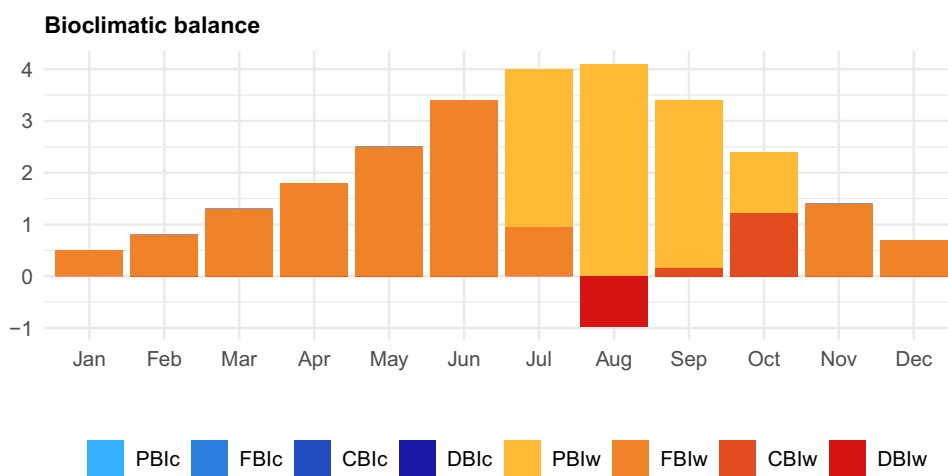


Fig. 6. Bioclimatic balance graphical representation of the bioclimatic balance for Seville (Spain). PBIc: cold Potential BI, FBIC: cold Free BI, CBIC: cold Conditioned BI, DBIC: cold Dry BI, PBIw: warm Potential BI, FBIw: warm Free BI, CBIw: warm Conditioned BI, and DBIw: warm Dry BI.

significant potential for monitoring experimental areas with meteorological and soil information but can also be applied to large regions using GIS-based data. In this regard, some efforts have previously been made to interpolate the results of bioclimatic classifications at point locations to ungauged areas using various techniques (e.g. Garzón-Machado et al., 2014; Mesquita and Sousa, 2009; Passarella et al., 2020). However, direct interpolation of bioclimatic classification results can introduce unknown uncertainties due to the underlying assumption that the spatial variation of all climatic variables participating in the classification is the same without considering factors such as topography or distance from the coast. *bioclim* avoids this issue by using raster datasets of climatic variables as an input, instead relying on the quality of these inputs. Furthermore, many climatic gridded datasets also include uncertainty as a measure of reliability at a pixel level (Serrano-Notivoli and Tejedor, 2021), which can be useful to evaluate the results at any location in the study area.

The potential applications of *bioclim* are diverse: from the simple and robust bioclimatic classification of vegetation communities to the detailed study of any of the components of the water balance at any spatial scale. The package is limited only by the information used to run the functions, as the final precision of the results is determined by the resolution and accuracy of the input data. Although the software's potential uses are wide-ranging, the package in its current form also has certain limitations that can compromise the precision of its results, most of which are related to the input data. In this context, soil information is

typically the most difficult data to obtain, especially for large regions. However, when field data are not available, models such as those developed by Batjes et al. (2020) can be used to provide detailed information. At a methodological level, the calculation process in *bioclim* is robust, however, there is scope for improvement within the potential evapotranspiration (PET) method. Thornthwaite's method has some limitations in terms of correctly reproducing the water balance in specific scenarios due to using only the average temperature and precipitation, as demonstrated in various applications (e.g. Hashemi and Habibian, 1979; Navarro et al., 2022; van der Schrier et al., 2011). While a change of method is possible, this implies the inclusion of new climatic variables as input data. For example, the Penman-Monteith method (Allen et al., 1998) has been demonstrated as the most accurate approach for PET calculations (Tomas-Burguera et al., 2019), however, this technique requires further climatic information such as wind, solar radiation, etc. Although these data can be extracted from different modelled sources such as climatic reanalysis, the necessity for more data types implies greater uncertainty than using observations. At this time, *bioclim* is not able to compute PET using any method other than Thornthwaite's, however, the package can accept an input dataset to replace the calculation (in raster mode).

The *bioclim* package is primarily aimed at scientific use, mostly within the earth sciences; while biogeographers, climatologists, and hydrologists are likely to be the main users, this software can also be useful for specific purposes outside of the climatic contextualisation of

the vegetative activity. For instance, *bioclim* can be applied to the modelling of habitats, first providing a general framework of the potential species distribution based on a climatic division of the study area and then using the individual variables that are yielded by the package's functions.

The open-access code of *bioclim* provides a high level of transparency for data analysis and allows for the possibility of adapting the code to meet users' requirements as there are no limitations in terms of modifying the parameters and functions. *bioclim* addresses a previously un-solved need for a deterministic classification of the Earth based on rigorous climatic and biogeographical criteria. The result is a software package with the capability to generate bioclimatic classifications in a simple, automatic, and reliable way both for single locations and large regions.

Funding

The present research was conducted within the framework of project SI3-PJI-2021-00398 funded by the Universidad Autónoma de Madrid and the Comunidad de Madrid.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

RSN and LAL are supported by the Government of Aragón through the "Program of research groups" (group H09_20R, "Climate, Water, Global Change, and Natural Systems").

References

- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-guidelines for computing crop water requirements-FAO irrigation and drainage paper 56. *Fao, Rome* 300 (9), D05109.
- Amigo, J., Ramirez, C., 1998. A bioclimatic classification of Chile: woodland communities in the temperate zone. *Plant Ecol.* 136, 9–26.
- Andrade, C., Contente, J., 2020. Köppen's climate classification projections for the Iberian Peninsula. *Clim. Res.* 81, 71–89.
- Batjes, N.H., Ribeiro, E., van Oostrum, A., 2020. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). *Earth Syst. Sci. Data* 12, 299–320. <https://doi.org/10.5194/essd-12-299-2020>.
- Beguiería, S., Vicente-Serrano, S.M., 2017. SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index. R package version 1.7. <https://CRAN.R-project.org/package=SPEI>.
- Bryant, C., Wheeler, N.R., Rubel, F., French, R.H., 2017. kgc: Koeppen-Geiger Climatic Zones. R package version 1.0.0.2. <https://cran.r-project.org/package=kgc>.
- Cadenas, López, del Llano, F., Mintegui, J.A., 1986. *Hidrología de superficie: Tomo I*. Fundación Conde del Valle de Salazar, Madrid.
- Cámara, R., Díaz del Olmo, F., Martínez Batlle, J.R., 2020. TBRs, a methodology for the multi-scalar cartographic analysis of the distribution of plant formations. *B. Asoc. Geogr. Esp.* 85 (2915), 1–38. <https://doi.org/10.21138/bage.2915>.
- Djamali, M., Akhiani, H., Hhoshraresh, R., Andrieu-Ponel, V., Ponel, P., Brewer, S., 2011. Application of the global bioclimatic classification to Iran: implications for understanding the modern vegetation and biogeography. *Ecologia Mediterranea* 37 (1), 91–114.
- Eccel, E., Cordano, E., Toller, G., 2016. ClimClass: Climate Classification According to Several Indices. R package version 2.1.0. <https://CRAN.R-project.org/package=ClimClass>.
- El-kenawy, A., Zerouali, B., Bailek, N., Bouchouich, K., Hassan, M.A., Almorox, J., Kuriqi, A., Eid, M., Ibrahim, A., 2022. Improved weighted ensemble learning for predicting the daily reference evapotranspiration under the semi-arid climate conditions. *Environ. Sci. Pollut. Res.* <https://doi.org/10.1007/s11356-022-21410-8>.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37 (12), 4302–4315. <https://doi.org/10.1002/joc.5086>.
- Fidalgo, C., Cámara, R., 2022. Metodología fitoclimática. Los diagramas bioclimáticos de Montero de Burgos y González Rebollar. Asociación Española de Geografía, Madrid, España.
- Garzón-Machado, V., Otto, R., del Arco Aguilar, M.J., 2014. Bioclimatic and vegetation mapping of a topographically complex oceanic island applying different interpolation techniques. *Int. J. Biometeorol.* 58, 887–899.
- Hashemi, F., Habibian, M.T., 1979. Limitations of temperature-based methods in estimating crop evapotranspiration in arid-zone agricultural development projects. *Agric. Meteorol.* 20 (3), 237–247. [https://doi.org/10.1016/0002-1571\(79\)90025-6](https://doi.org/10.1016/0002-1571(79)90025-6).
- Holdridge, L.R., 1967. Life Zone Ecology. Tropical Science Center, San José, Costa Rica.
- Honorio, F., García-Martín, A., Moral, F.J., Paniagua, L.L., Rebollo, F.J., 2018. Spanish vineyard classification according to bioclimatic indexes. *Austr. J. Grape Wine Res.* 24 (3), 335–344.
- Kirilenko, A.P., Solomon, A.M., 1998. Modeling dynamic vegetation response to rapid climate change using bioclimatic classification. *Clim. Chang.* 38, 15–49.
- Kisi, O., Mirboluki, A., Naganna, S.R., Malik, A., Kuriqi, A., Mehraein, M., 2022. Comparative evaluation of deep learning and machine learning in modelling pan evaporation using limited inputs. *Hydro. Sci. J.* 67 (9), 1309–1327.
- Köppen, W., 1884. Die Wärmezonen der Erde, nach der Dauer der heissen, gemässigten und kalten Zeit und nach der Wirkung der Wärme auf die organische Welt betrachtet. *Meteorol. Z.* 1, 215–226.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* 15 (3), 259–263.
- Mesquita, S., Sousa, A.J., 2009. Bioclimatic mapping using geostatistical approaches: application to mainland Portugal. *Int. J. Climatol.* 29 (14), 2156–2170.
- Metzger, M.J., Bince, R.G.H., Jongman, R.H.G., Sayre, R., Trabucco, A., Zomer, R., 2012. A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. *Glob. Ecol. Biogeogr.* 22(5), 630–638.
- Montero de Burgos, J.L., González, J.L., 1974. *Diagramas bioclimáticos*. ICONA. Ministerio de Agricultura, Madrid.
- Navarro, A., Merino, A., Sánchez, J.L., García-Ortega, E., Martín, R., Tapiador, F.J., 2022. Towards better characterization of global warming impacts in the environment through climate classifications with improved global models. *Int. J. Climatol.* <https://doi.org/10.1002/joc.7527>.
- Passarella, G., Bruno, D., Lay-Ekuakille, A., Maggi, S., Masciale, R., Zaccaria, D., 2020. Spatial and temporal classification of coastal regions using bioclimatic indices in a Mediterranean environment. *Sci. Total Environ.* 700, 134415.
- Pesaresi, S., Galdenzi, D., Biondi, E., Casavecchia, S., 2014. Bioclimate of Italy: application of the worldwide bioclimatic classification system. *J. Maps* 10 (4), 538–553.
- Rivas-Martínez, S., Rivas-Saenz, S., Penas, A., 2002. *Worldwide Bioclimatic Classification System*. Backhuys Pub, Kerkwerve, The Netherlands.
- Serrano-Notivoli, R., Tejedor, E., 2021. From rain to data: a review of the creation of monthly and daily station-based gridded precipitation datasets. *WIREs Water* 8 (6), e1555.
- Serrano-Notivoli, R., Tomás-Burguera, M., Martí, A., Beguería, S., 2020. An integrated package to evaluate climatic suitability for agriculture. *Comput. Electron. Agric.* 17, 105473.
- Srinivas, S., Srinivas, C.V., Nair, K.M., Naidu, L.G.K., Sarkar, D., Dingh, S.K., 2016. A climatic water balance model 'WatBal' for bio-climatic classification and agro-climatic analysis. *Eco. Env. & Cons.* 22 (1), 173–180.
- Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. *Geogr. Rev.* 38, 55–94. <https://doi.org/10.2307/2107309>.
- Thornthwaite, C.W., Mather, R.J., 1955. *The Water Balance*. Drexel Institute of Technology, Laboratory of Climatology, Centerton, NJ.
- Title, P.O., Bemmels, J.B., 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*. 41, 291–307. <https://doi.org/10.1111/ecog.02880>.
- Tomas-Burguera, M., Vicente-Serrano, S.M., Beguería, S., Reig, F., Latorre, B., 2019. Reference crop evapotranspiration database in Spain (1961–2014). *Earth Syst. Sci. Data* 11, 1917–1930. <https://doi.org/10.5194/essd-11-1917-2019>.
- van der Schrier, G., Jones, P.D., Briffa, K.R., 2011. The sensitivity of the PDSI to the Thornthwaite and penman-Monteith parameterizations for potential evapotranspiration. *J. Geophys. Res.-Atmos.* 116 (D3), D03106. <https://doi.org/10.1029/2010JD015001>.
- Vishwakarma, D.K., Pandey, K., Kaur, A., Kushwaha, N.L., Kumar, R., Ali, R., Elbeltagi, A., Kuriqi, A., 2022. Methods to estimate evapotranspiration in humid and subtropical climate conditions. *Agric. Water Manag.* 261, 107378.
- Xu, T., Hutchinson, M.F., 2013. New developments and applications in the ANUCLIM spatial climatic and bioclimatic modelling package. *Environ. Model. Softw.* 40, 267–279. <https://doi.org/10.1016/j.envsoft.2012.10.003>.