DOI: 10.15201/hungeobull.71.1.1 Hungarian Geographical Bulletin **71** 2022 (1) 3–19.

Estimating relative sunshine duration from commonly available meteorological variables for simulating biome distribution in the Carpathian Region

Zoltán SZELEPCSÉNYI[1], Hajnalka BREUER[2] and Nándor FODOR[3]

Abstract

Bright sunshine duration (BSD) data are required for simulating biomes using process-based vegetation models. However, monthly global paleoclimate datasets that can be used in paleo data–model comparisons do not necessarily contain BSD or radiation data. Considering the theoretical and practical aspects, the scheme of Yin, X. (1999) is here recommended to estimate monthly time series of relative BSD using only monthly climate and location data. As a case study for the Carpathian Region, the efficiency of both the original and a variant of that scheme is analysed in this paper. The alternative scheme has high applicability in paleoenvironmental studies. Comparison of the estimated and observed BSD data shows that from May to August, the value of relative root mean squared error in more than 90 percent of the study area does not exceed the threshold of 20 percent, indicating an excellent performance of the original estimation scheme. It is also found that though the magnitude of overestimation for the alternative algorithm is significant in the winter period, the proposed method performs similarly well in the growing season as the original. Furthermore, concerning modelling the distribution of biomes, simulation experiments are performed to assess the effects of modifying some configuration settings: (a) the generation of relative BSD data, and (b) the algorithm used to create quasi-daily weather data from the monthly values. Under both the recent humidity conditions of the study region and the spatial resolution of the climate dataset used, the results can be considered sufficiently robust, regardless of the configuration settings tested. Thus, using monthly temperature and precipitation climatologies, the spatial distribution of biomes can be properly simulated with the configuration settings proposed here.

Keywords: sunshine duration, water balance, biome, plant functional types, data–model comparisons, CarpatClim

Received November 2021, accepted March 2022

Introduction

Today there are a number of tools available to translate climate model outputs into vegetation distribution patterns (see e.g., bioclimatic classification methods: Prentice, K.C. 1990; correlative vs. mechanistic biome models: Yates, D.N. *et al.* 2000; species distribution models: Elith, J. and Leathwick, J.R. 2009). Bioclimatic classification methods (BCMs)

[1] Institute of Archaeology, Research Centre for the Humanities, Eötvös Loránd Research Network, Centre of Excellence of the Hungarian Academy of Sciences, Tóth Kálmán u. 4., H–1097 Budapest, Hungary; Department of Geology and Palaeontology, University of Szeged, Egyetem u. 2–6., H–6722 Szeged, Hungary. Corresponding author's e-mail: szelepcsenyi.zoltan@abtk.hu

[2] Department of Meteorology, ELTE Eötvös Loránd University, Pázmány Péter sétány 1/A, H–1117 Budapest, Hungary. E-mail: bhajni@nimbus.elte.hu

[3] Agricultural Institute, Centre for Agricultural Research, Eötvös Loránd Research Network, Brunszvik u. 2., H–2462 Martonvásár, Hungary. E-mail: fodor.nandor@atk.hu

ORCID: https://orcid.org/0000-0002-9844-4958, https://orcid.org/0000-0002-0271-095X, https://orcid.org/0000-0002-6460-1767

are tools used to transform a set of climate and soil variables into an index-class that can be directly related to biome-level vegetation units (Tapiador, F.J. *et al.* 2019). Correlative models determine a statistical relationship between vegetation distribution and environmental variables (e.g., climate, soil, etc.); then this link is applied to simulate the potential distribution of vegetation under the altered conditions (Yates, D.N. *et al.* 2000; Elith, J. and Leathwick, J.R. 2009). From this point of view, therefore, BCMs (e.g., Köppen, W. 1936) can be considered the simplest correlative biome models. Mechanistic biome models, in contrast, focus on mechanisms determining survival and performance of plants by simulating processes in soil–vegetation–atmosphere systems, such as water, carbon and nutrient cycles (Prentice, I.C. *et al.* 2007).

Paleoclimatological and paleoenvironmental studies often use both outputs from climate model simulations and proxy archives such as fossil pollen records in a common framework. These studies represent a special area of called paleo data–model comparisons (Harrison, S.P. 2013) that have essentially two distinct purposes: (a) to understand the mechanisms of past climate and environmental shifts (e.g., Miller, P.A. *et al.* 2008; Mauri, A. *et al.* 2014), and (b) to provide feedback on the performance of climate/environmental reconstruction approaches (e.g., Prentice, I.C. *et al.* 1998; Webb, III T. *et al.* 1998). These comparisons can be made in two ways, either by collating pollen-inferred climate with that simulated by climate models (e.g., Webb, III T. *et al.* 1998; Mauri, A. *et al.* 2014), or by comparing biome/species distribution estimated using paleoclimate model outputs with vegetation reconstructed from pollen assemblages (e.g., Prentice, I.C. *et al.* 1998; Miller, P.A. *et al.* 2008).

Currently, more and more global datasets are made publicly available that provide bias-corrected monthly climatologies (multi-year averages) from paleoclimate simulations in order to support distribution modelling experiments in various research areas (e.g., paleobiogeography, archaeology). However, these datasets differ significantly in terms of spatial resolution, temporal coverage and time step, and in terms of which climate variables they contain information on. Beyer, R.M. *et al.* (2020), for example, published a monthly global dataset (hereinafter HadCM3-120k), with a horizontal resolution of 0.5°, for temperature, precipitation, cloud cover, relative humidity, and wind speed, and additional parameters related to bioclimatic and biogeochemical conditions, covering the last 120,000 years at a temporal resolution of 1,000–2,000 years. For this, medium-resolution simulations generated by the HadCM3 general circulation model (GCM) for the last 120,000 years were combined with high-resolution simulations prepared by the HadAM3H GCM for the last 21,000 years and a recent observational dataset. (For details of the above-mentioned GCMs, see Valdes, P.J. *et al.* 2017.) Then, Krapp, M. *et al.* (2021) extended access to the climate variables in question for the last 800,000 years by performing a statistical-based reconstruction using the above-mentioned simulations. And recently, Karger, D.N. *et al.* (in review) shared with the scientific community their dataset, called CHELSA-TraCE21k v1.0, that includes monthly climatologies for both temperature and precipitation, and other bioclimatic variables, with a spatial resolution of 30 arc-sec at a 100-year time step for the last 21,000 years. For this, a transient simulation generated by the CCSM3 GCM (He, F. 2011) was downscaled considering the temporal change of orography.

The HadCM3-120k was specifically generated to feed mechanistic biome models, while the CHELSA-TraCE21k v1.0 was clearly developed to support paleoecological studies using correlative species distribution models (SDMs). In SDMs (e.g., MaxEnt: Phillips, S.J. *et al.* 2006), bioclimatic variables, i.e., annual and seasonal measures derived from monthly values of temperature and precipitation, are generally used as environmental predictors, besides topographic variables. In contrast, due to the simulation of energy and water fluxes, for applying mechanistic biome models, a meteorological variable directly related

to radiation (e.g., cloud cover, sunshine duration) is also required, besides temperature and precipitation data. Beyer, R.M. *et al.* (2020) and Krapp, M. *et al.* (2021), relying on their own datasets, also estimated the evolution of the global biome distribution using one of the best-known mechanistic biome models, called BIOME4 (Kaplan, J.O. 2001). The BIOME*n* models (e.g., BIOME: Prentice, I.C. *et al.* 1992; BIOME4) estimate the net radiation as a function of latitude, temperature, and bright sunshine duration (BSD). Thus, to apply the above-mentioned vegetation model to their own datasets, the authors used different empirical linear relationships (Doorenbos, J. and Pruitt, W.O. 1977; Hoyt, D.V. 1977) to convert cloud cover data to the percentage of possible sunshine hours. Unfortunately, the CHELSA-TraCE21k v1.0 does not contain cloudiness or BSD data, so it is not directly suitable for feeding the above-mentioned BIOME*n* models. However, estimating BSD data from commonly available meteorological variables may be a solution to overcome the lack of data.

Although there are several estimation methods for calculating monthly values of BSD (Kandirmaz, H.M. *et al.* 2014), in this study, the use of a method developed by Yin, X. (1999) is recommended. Yin, X. (1999) used monthly data of 729 worldwide stations for finding a generic algorithm that captures global variability of BSD data in relation to temperature, precipitation, and geographic location. Regression models for estimating monthly mean daily values of BSD are usually set only for smaller regions due to limited access to reliable station data (e.g., Italy: Stanghellini, C. 1981), and/or use parameters that are not readily available from global gridded climate datasets (e.g., the number of wet days per month: Castellvi, F. 2001). Thus, what makes the method proposed by Yin, X. (1999) special is that it is globally parameterized and uses only monthly climate and location data that are widely available.

To our knowledge, by means of gridded climate datasets, the performance of the estimation scheme proposed by Yin, X. (1999) has not yet been evaluated. Our current level of knowledge would indicate that there is currently no global gridded observational dataset to which the following three statements are true without exception: (i) it contains monthly values for BSD, temperature, and precipitation; (ii) it contains monthly meteorological data for a long period without time averaging; and (iii) it was developed based on station observations. Currently, there is only access to global datasets that also use remotely sensed and reanalysis data to produce gridded climate information, and for which values of BSD can only be derived from another meteorological variable (e.g., incoming solar radiation: TerraClimate, Abatzoglou, J.T. *et al.* 2018; cloud cover: CRU TS v4, Harris, I. *et al.* 2020). However, two regional climate databases are known that provide station-based meteorological fields for the above-mentioned three variables for continuous periods: CarpatClim (Spinoni, J. *et al.* 2015) and HadUK-Grid (Hollis, D. *et al.* 2019).

In consideration of the literature discussed above, the first objective of this study is to assess the accuracy of the scheme under discussion using the monthly climate data provided by the CarpatClim dataset. The second goal of this paper is to test how the quality of estimates changes as a result of proposed modifications of the approach, which are justified by its applicability in paleoenvironmental studies. The amount of incoming and outgoing radiation influences the growing conditions of plants, so the error in estimating the relative BSD can cause problems in the modelling of the biome distribution. Therefore, as a case study for the Carpathian Region, evaluation of the impact of the estimated relative BSD on simulation of energy and water fluxes and biome designation is the third aim of this study.

Materials and methods

Estimation of monthly mean relative sunshine duration

The monthly mean daily values of relative BSD (*RSD*, dimensionless) for a given month is estimated using the following parametric regression model as recommended by Yin, X. (1999):

$$RSD = e^{\left[-1.65 \cdot p \cdot f_o \cdot \left(1+\sum f_i\right)\right]}, \quad (1)$$

with

$$f_o = f(R_E) \cdot f(P) \cdot f(T, E_P) \cdot f(\phi), \quad (2)$$

where

$$f(R_E) = [1 + 0.756 \cdot R_E \cdot (3 - R_E)]^{-1}, \quad (3)$$

$$f(P) = \frac{1 + 0.785 \cdot P}{1 + 0.222 \cdot P}, \quad (4)$$

$$f(T, E_P) = 1 + \frac{(7.66 \cdot I_{T<0} - 4.98) \cdot T + E_P^2}{184}, \quad (5)$$

$$f(\phi) = 1 + 0.512 \cdot \sin\left(\frac{(\phi + 15) \cdot 2\pi}{77}\right), \quad (6)$$

and

$$f_i = \begin{cases} \dfrac{15.6 - T_{am}}{46.3}, & \text{if island or Australia} \\ 0.331 - 0.213 \cdot E_{Pam} - \dfrac{|T_{am}|}{72.5} + \dfrac{\max(T_{min}, 0)}{30}, & \text{if North America} \\ 0, & \text{if South America} \\ \dfrac{1}{10.2} + \dfrac{P_7 - P_1}{36.8 \cdot P_{am}} \cdot \left(2.3 - \dfrac{P_7 - P_1}{P_{am}}\right), & \text{if Eurasia} \\ -0.643 + 0.314 \cdot R_E + (31.8 - T_{am}) \cdot \left(\dfrac{T_{am}}{140}\right)^2, & \text{if monsoonal Asia} \\ \dfrac{7.26 - T_{ar}}{32.3}, & \text{if Africa} \end{cases}, \quad (7)$$

where p is the station atmospheric pressure in relation to the pressure at sea level, f_o represents global trends, f_i gives regional modifications, $\sum f_i$ is the summation of values of any regional functions that are applicable to a particular location, R_E is the monthly average of hourly solar irradiance for cloudless-sky conditions (in MJ m^{-2} hr^{-1}), P is the monthly mean precipitation intensity (in mm dy^{-1}), T is the monthly mean air temperature (in °C), E_p is the monthly average of daily potential evapotranspiration (in mm dy^{-1}), ϕ is the latitude (in decimal degrees), $I_{T<0}$ is a temperature indicator (1 if $T \leq 0$ °C, and 0 if $T > 0$ °C), T_{am} is the annual mean temperature (in °C), E_{Pam} is the annual average of daily potential evapotranspiration (in mm dy^{-1}), T_{min} is the lowest monthly mean air temperature (in °C), P_7 is the monthly mean precipitation intensity (in mm dy^{-1}) in the warmest month (fixed at July for the Northern Hemisphere, and January for the Southern Hemisphere), P_1 is the monthly mean precipitation intensity (in mm dy^{-1}) in the coldest month (fixed at January for the Northern Hemisphere, and July for the Southern Hemisphere), P_{am} is the annual mean precipitation intensity (in mm dy^{-1}), T_{ar} is the annual range of monthly mean air temperature (annual diurnal range) (in °C).

The values of R_E are calculated as proposed by Yin, X. (1997a), with a minor modification, using the daytime means of optical air mass and cosine zenith. The former is computed as recommended by Yin, X. (1997a), while the latter is estimated by using Eq. 5 of Yin, X. (1997b). Furthermore, in contrast to the original approach, where the solar constant was fixed at 4.9212 MJ m^{-2} hr^{-1}, its value is corrected, according to Yin, X. (1999), by calendar day for the variable ellipticity of the Earth's orbit using the scheme of Brock, T.D. (1981). In these calculations, the values of solar declination and daylength are derived by using the approach of Brock, T.D. (1981). The values of E_p are computed using Eq. A10 of Yin, X. (1998). The value of E_{Pam} is calculated as a weighted mean of the E_p values using the number of days in months as weights.

Simulation of biome distribution

In this study, the BIOME model (Prentice, I.C. *et al.* 1992) is applied to simulate the spatial distribution of biome-level vegetation units. First, the presence of each plant functional type (PFT) that is a group of plant types with similar ecophysiological behaviour is estimated under given climatic conditions. To do this, it is necessary to check which of the 14 PFTs defined can occur considering the environmental constraints associated with their climatic tolerances and requirements *(Table 1).* After this, the dominance class value (D) of each PFT is examined and only those in the highest class (with lowest D) present are retained. Finally, to infer the biome type, retained PFTs are combined with each other by taking into account rules formalized in *Table 2.*

Table 1. Dominance class (D) and environmental constraints for each plant functional type used in the model*

Abbreviation	Plant functional type	D	T_C		GDD_0	GDD_5	T_W	α	
			min	max	min	min	min	min	max
tr.e.t	Tropical evergreen tree	1	15.5	–	–	–	–	0.80	–
tr.r.t	Tropical rain-green tree	1	15.5	–	–	–	–	0.45	0.95
w-te.e.t	Warm temperate evergreen tree	2	5.0	–	–	–	–	0.65	–
te.s.t	Temperate summer-green tree	3	−15.0	15.5	–	1,200	–	0.65	–
c-te.c.t	Cool temperate conifer tree	3	−19.0	5.0	–	900	–	0.65	–
bo.e.t	Boreal evergreen conifer tree	3	−35.0	−2.0	–	350	–	0.75	–
bo.s.t	Boreal summer-green tree	3	–	5.0	–	350	–	0.65	–
sb.suc	Sclerophyll/succulent	4	5.0	–	–	–	22	0.28	–
wa.g.s	Warm grass/shrub	5	–	–	–	–	–	0.18	–
cl.g.s	Cool grass/shrub	6	–	–	–	500	–	0.33	–
cd.g.s	Cold grass/shrub	6	–	–	100	–	22	0.33	–
h.d.s	Hot desert shrub	7	–	–	–	–	–	–	–
c.d.s	Cold desert shrub	8	–	–	100	–	–	–	–
p.d	Polar desert	9	–	–	–	–	–	–	–

*T_C = mean temperature of the coldest month (in °C); GDD_0 = growing degree-days above a 0 °C base (in °C day); GDD_5 = growing degree-days above a 5 °C base (in °C day); T_W = mean temperature of the warmest month (in °C); α = Priestley–Taylor coefficient at an annual time scale (dimensionless).

Table 2. A list of biome types used in the model and their generation rules

Abbreviation	Biome type	Plant functional types
TRRA	Tropical rain forest	tr.e.t
TRSE	Tropical seasonal forest	tr.e.t + tr.r.t
TRDR	Tropical dry forest/savannah	tr.r.t
WAMX	Broad-leaved evergreen/warm mixed forest	w-te.e.t
TEDE	Temperate deciduous forest	te.s + c-te.c.t + bo.s.t
COMX	Cool mixed forest	te.s.t + c-te.c.t + bo.e.t + bo.s.t
COCO	Cool conifer forest	c-te.c.t + bo.e.t + bo.s.t
TAIG	Taiga	bo.e.t + bo.s.t
CLMX	Cold mixed forest	c-te.c.t + bo.s.t
CLDE	Cold deciduous forest	bo.s.t
XERO	Xerophytic woods/scrub	sb.suc
WAST	Warm grass/shrub	wa.g.s
COST	Cool grass/shrub	cl.g.s + cd.g.s
TUND	Tundra	cd.g.s
HODE	Hot desert	h.d.s
SEDE	Semi-desert	c.d.s
PODE	Polar desert	p.d

Note: Each biome type is arising as a combination of dominant plant functional types.

In the BIOME model, the plant-available moisture is characterized by the Priestley–Taylor coefficient (α, dimensionless). Here, the values of α at an annual time scale are computed by using the SPLASH v.1.0 model (Davis, T.W. *et al.* 2017), through the simulation of seasonal changes in both surface energy fluxes and climatic water balance. In the BIOME model, in order to quantify heat requirement, the growing degree-days (*GDD*, in °C day) is used, which can be obtained by summing the values of daily temperature above a certain base temperature. To calculate values of *GDD*, the values of daily mean temperature are required; furthermore, besides temperature and precipitation data, the relative BSD must also be used in the SPLASH v.1.0 model, on a daily basis. The methods for generating these daily values are described in more detail below.

Evaluation methodology

Monthly time series of the temperature, precipitation and sunshine duration, along with location data, are required for evaluating the performance of the procedure proposed by Yin, X. (1999). The CarpatClim dataset provides access to the three meteorological variables relevant to the assessment for the time period 1960–2010, with a horizontal resolution of 0.1°, covering nine countries with 5,895 grid cells *(Figure 1).* For this reason, using data derived from this dataset, *RSD* values for each year in the period 1961–2010 are estimated using the scheme developed by Yin, X. (1999), and the estimates are compared to the observed data. Observed values of *RSD* are determined by a two-step procedure, following Spinoni, J. *et al.* (2015): (i) the monthly amount of BSD to which the CarpatClim dataset provides access is divided by the number of days in a given month to calculate the monthly mean for BSD, and then (ii) this value is divided by the monthly mean for daylength that is calculated using Eq. 8.5.3 of Iqbal, M. (1983). Finally, in each grid cell, values of the root mean square error normalized by the mean value of observed data (*RRMSE*, in percentages) are computed between the observed and estimated 50-year time series of *RSD*, separately for each month.

To apply the parametric regression model proposed by Yin, X. (1999) to monthly global paleoclimate datasets already described in the introduction, two modifications are needed to use. A feature of such paleoclimate datasets is that they represent climatic conditions averaged over a longer period (typically 30 or 50 years) at each time step. For this reason, it is considered necessary to in-

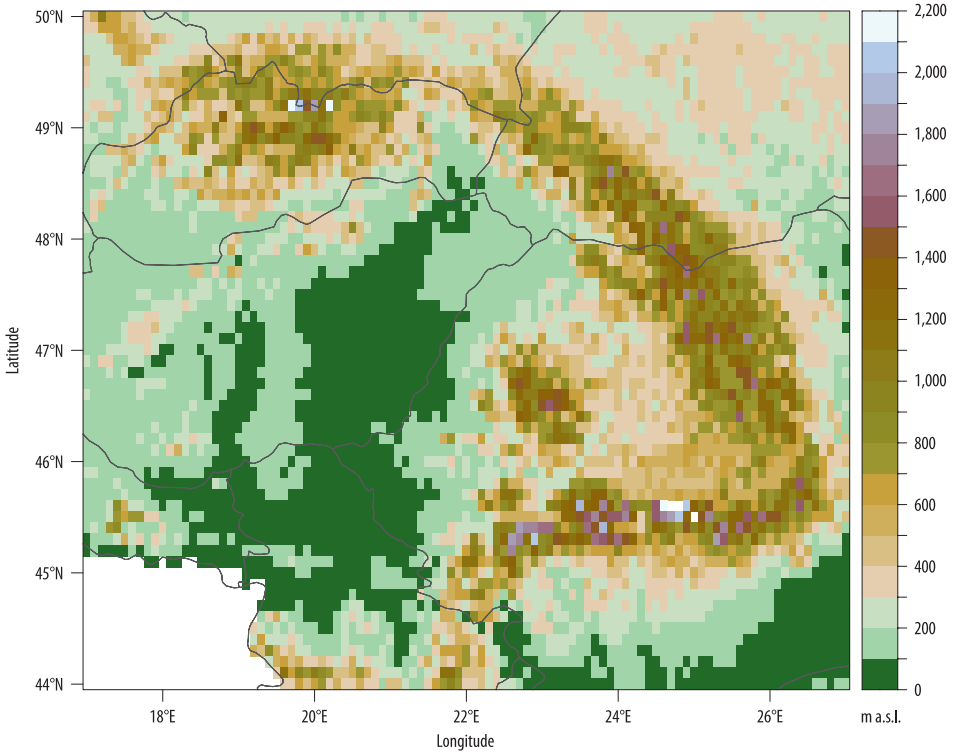


Fig. 1. Topography of the Carpathian Region based on the CarpatClim dataset (Spinoni, J. *et al.* 2015)

vestigate how the scheme performs when applied to multi-year averages instead of single year time series. Furthermore, the regression model of Yin, X. (1999) also uses the monthly mean of hourly solar irradiance to estimate the global trend, but uses an algorithm to estimate the value of R_E that cannot be applied without modification in paleoclimatological studies because it does not consider changes in the Earth's orbital parameters. For this reason, here, it is recommended that the value of R_E be calculated using the algorithm used in the SPLASH v.1.0 model, with the addition that orbital parameters are calculated using the method of Berger, A. and Loutre, M.F. (1991). In this approach, first, the daily solar radiation at the top of the atmosphere is calculated (Eq. 7 in Davis, T.W. *et al.* 2017), and then this value is multiplied by the atmospheric transmissivity to obtain the value of daily surface radiation. In this case, as well, cloudless conditions are assumed, i.e., the transmission coefficient is taken into account with a universal value of 0.75, however, its value is modified as a function of elevation by using the scheme of Allen, R.G. (1996). The daylength is calculated via Eq. 1.6.11 in Duffie, J.A. and Beckman, W.A. (1991), using the sunset hour angle (Eq. 8 in Davis, T.W. *et al.* 2017). Finally, the mean hourly surface radiation is derived as the quotient of the daily surface radiation and the daylength. In this study, using the CarpatClim dataset, values of *RSD* for the period 1981–2010 are computed in two ways: (A) by averaging the time series estimated using the initial scheme for each year, and (B) by applying the scheme to 30-year averages, with the provision that the values of R_E are calculated for year 1995 using the algorithm described above.

When modelling the distribution of biomes, monthly climatologies must be converted to daily values, in order to simulate seasonal changes in both surface energy fluxes and climatic water balance. In the description of the water balance module used in the initial version of the BIOME model, Prentice, I.C. *et al.* (1993) have recommended for this that monthly values are interpolated linearly between mid-month days. However, this approach is unsound because it is not mean-preserving (the monthly means of the interpolated daily values will generally not match the original monthly values). When presenting the SPLASH v.1.0 model, Davis, T.W. *et al.* (2017) simply suggested that monthly mean values are assumed constant over each day of the month. This procedure is suitable in terms of the monthly averages, but it generates unrealistic time series. In the 1990s, several mean-preserving methods (see e.g., Epstein, E.S. 1991; Lüdeke, M.K.B. *et al.* 1994) were developed to address this issue. Here, quasi-daily values are constructed in two ways: (a) monthly averages of temperature and *RSD* are assumed constant, and the monthly precipitation sum is divided equally across each day of the month; and (b) for temperature and *RSD*, the 'harmonic' interpolation technique described by Epstein, E.S. (1991) is used, with a correction of physically impossible values, and in the case of precipitation, the temporal scaling using an iterative interpolation technique described by Lüdeke, M.K.B. *et al.* (1994) is applied, with a damping variable of 0.7 for each month.

In this study, we assess the effects of the choice of the method used to generate the quasi-daily values and of the source of the BSD data on the results in terms of the spatial distribution of the bioclimatic variables used in the BIOME model. Finally, biome maps simulated under various model configuration settings are compared using the Kappa statistic (Cohen, J. 1960), which value ranges from 0 to 1, with 0 representing totally different patterns and 1 indicating complete agreement.

Results and discussion

One of the key objectives of this study is to attempt to evaluate the performance of the estimation procedure for *RSD* using data provided by the CarpatClim database. The performance of the scheme proposed by Yin, X. (1999) is assessed based on the root

mean square error normalized by the mean value of observed data (*RRMSE*, in percentages) calculated between the observed and estimated values for the period 1961–2010, separately for each month *(Figure 2).* From May to August, the *RRMSE* value in more than 90 percent of the study area does not exceed the threshold of 20 percent below which the model performance can be considered excellent, according to Bellocchi, G. *et al.* (2002). In the period from April to October, in nearly 99 percent of the grid cells with elevation smaller than 500 m a.s.l., the value of *RRMSE* is less than 40 percent, which is the limit of the model performance still considered acceptable based on the work of Bellocchi, G. *et al.* (2002). In the summer months, the *RRMSE* value in almost 90 percent of the lower regions (elevation < 500 m a.s.l.) does not even exceed the threshold of 15 percent. Interestingly, in the winter months, the estimation scheme performs better in the higher than in the lower elevation areas (66 ± 3% and 45 ± 28% of the regions at elevations above and below 500 m a.s.l., respectively, with a threshold of 40%). Although not within the scope of this study, it should be pointed out that the inconsistency between the measured and estimated values found in the Ukrainian section of the Carpathians suggests (see *Figure 2*) that one or even more of the climate fields used in the assessment may contain significant errors in this region. However, an explanation of this requires a more detailed analysis.

An important objective of this paper is to assess how the accuracy of the estimates changes when the scheme is adapted for applying to paleoclimate datasets. To study this, values of *RSD* for the period 1981–2010 are calculated in two ways *(Figure 3)*: (A) by averaging the time series estimated using the initial scheme for each year, and (B) by applying the scheme to 30-year averages. The estimated results are compared to the averages of the measured values over the period 1981–2010 *(Figure 3,* b*).* For the time window used here, we can see that in the period from March to September, the estimation method proposed here performs even better than the initial algorithm. In these months, i.e., in the most important period in terms of the evapotranspiration processes, with one exception, the value of *RRMSE* calculated for the whole study area does not exceed the threshold of 10 percent (see the second row in *Figure 3,* b), which indicates a very good quality of the estimates. (As previously indicated, the Ukrainian part of the Carpathians is the main contributor to the observed discrepancies.)

In the context of *Figure 3*, it is important to underline that when applying the modified estimation scheme to 30-year averages, the overestimation is very high in the winter months (in January, its value exceeds the value of 0.15 over almost half of the region), which, combined with a low (around 0.254 in January) benchmark, results in very high values of *RRMSE*: 60.8 percent in January and almost 30 percent in February. For both estimation methods, the difference in winter months, which is also highlighted above in relation to the *Figure 2*, is probably related to the formation of conditions for cold-air pool (CAP), which is a typical weather situation in the Carpathian Basin (Szabóné André, K. *et al.* 2021). Namely, the CAP conditions are extremely favourable for the formation of fog which lead to less surface solar radiation. Considering that the model is globally parameterized, it is impractical to expect it to capture such local effects, but fortunately, this model weakness has little relevance in simulating important processes for plants, as it will also be shown later.

The ultimate goal of this study is to examine how sensitive the BIOME model is to change configuration settings. We are interested in how the results change when on the one hand, the measured time series of *RSD* are replaced by estimates produced by different algorithms, and on the other hand, the technique for generating daily weather data is made more sophisticated. For the latter aspect, the simulations are performed in two ways: (a) monthly means are assumed constant over each day of the month, and (b) different mean-preserving interpolation techniques are applied (for details, see evaluation methods).

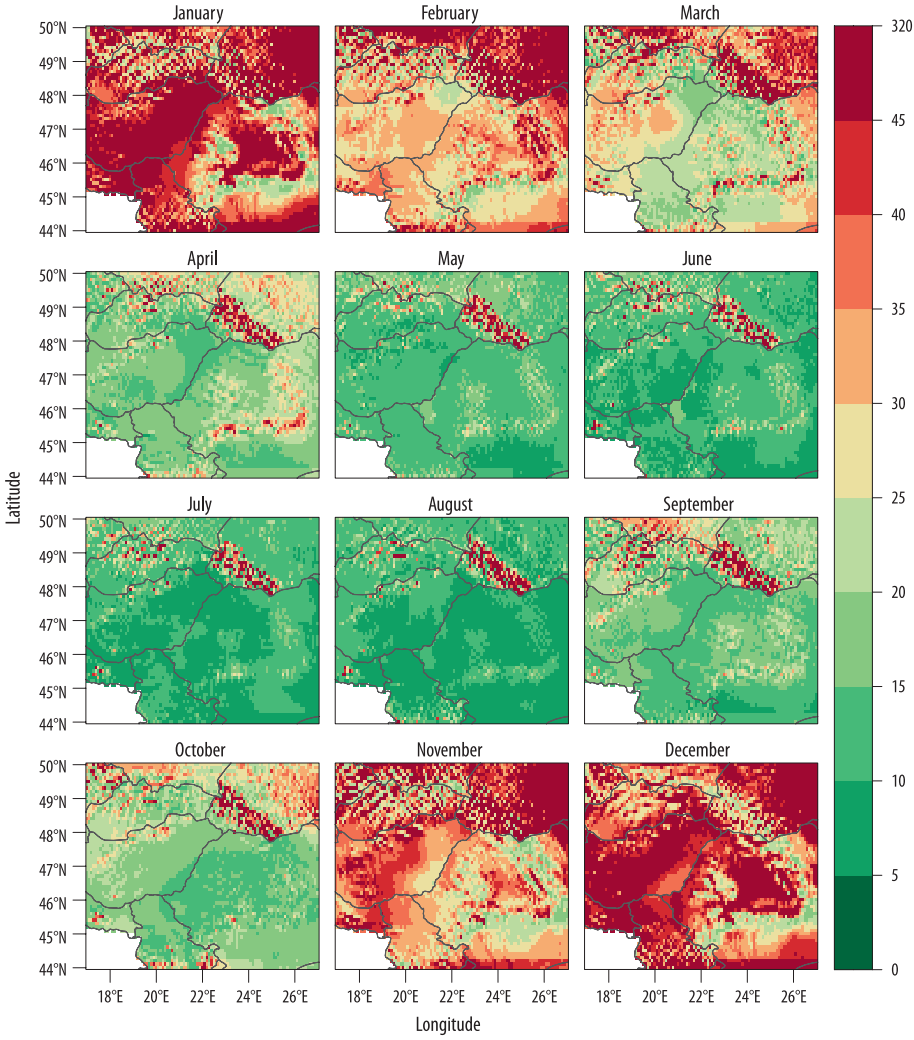


Fig. 2. Performance of the regression model developed by Yin, X. (1999) to estimate monthly time series of the relative sunshine duration (*RSD*, dimensionless), based on the root mean square error normalized by the mean value of observed data (*RRMSE*, in percentages). *RRMSE* values are calculated between the observed and estimated values for the period 1961–2010 using the CarpatClim dataset, separately for each month.

The presence of PFTs is fundamentally dependent on the plant-available moisture, which in the BIOME model is characterized by the α ranging from 0 to 1.26. Its value for the period 1981–2010 is calculated at an annual time scale using the SPLASH v.1.0 model, with a total of six settings *(Figure 4)*. The simulation performed using the measured values of *RSD* and assuming constant monthly means of each meteorological variable over each day of the month is considered as a reference *(Figure 4*, a*)*. Based on this, it can be concluded that there is sufficient moisture in the study area for all woody PFTs related to mid-latitudes (cf. *Figure 4*, a, and *Table 2*), with the spatial resolution and

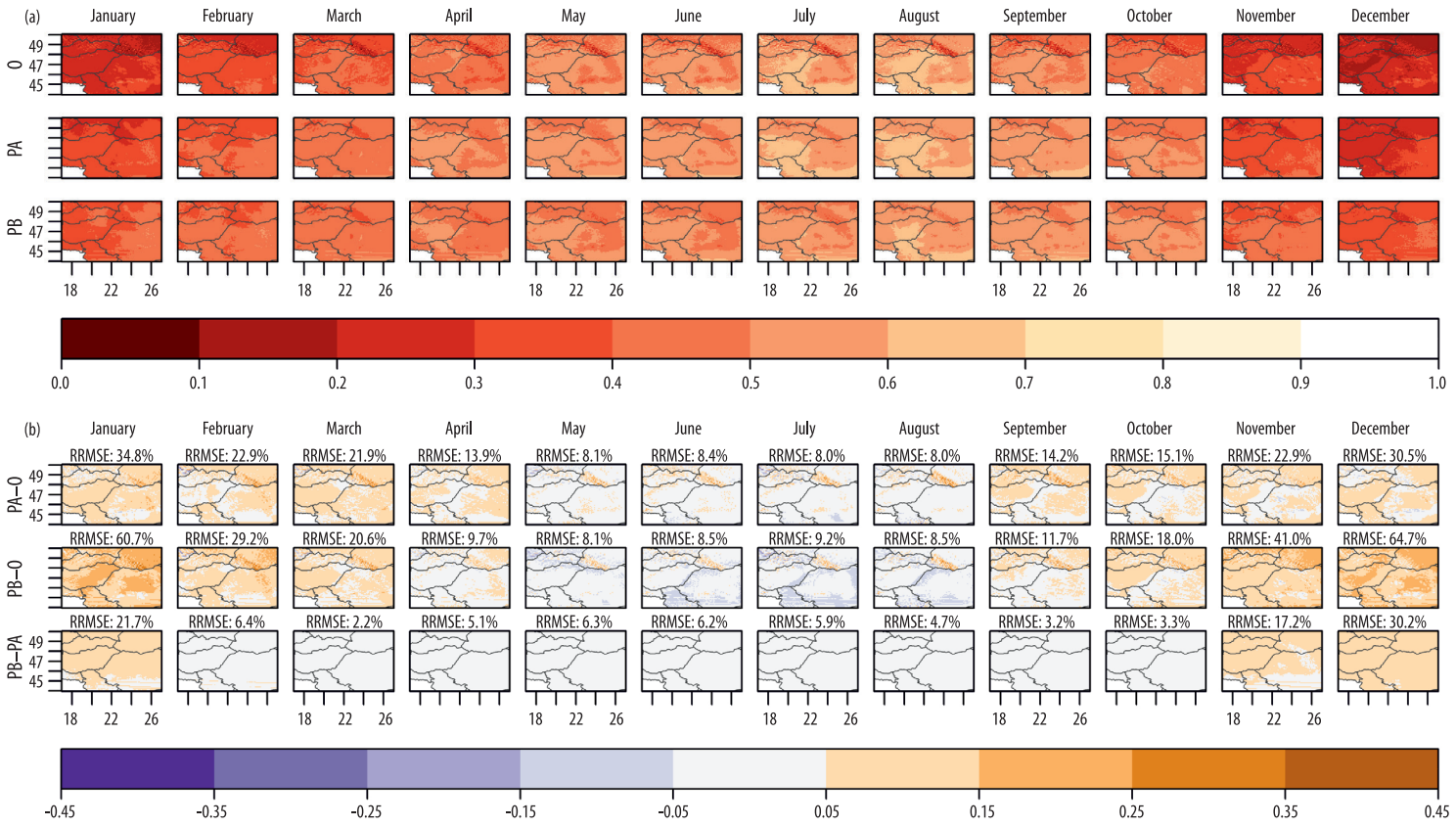
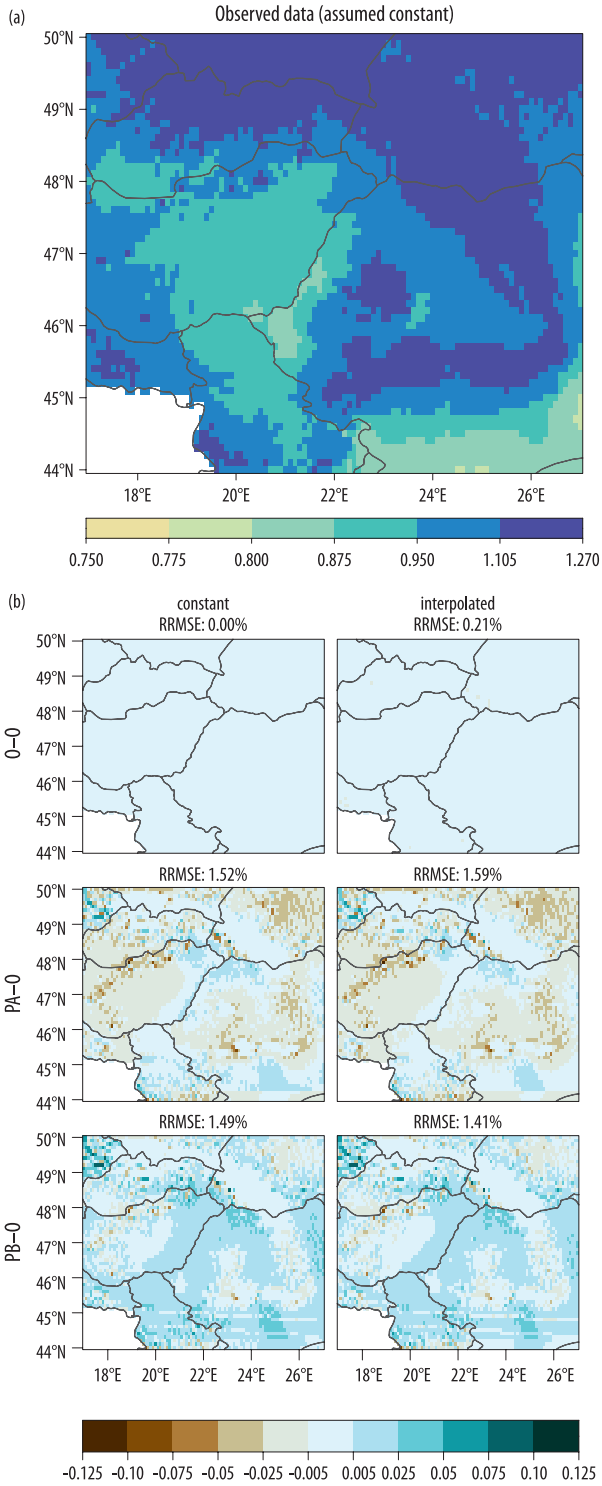


Fig. 3. Spatial distribution patterns of monthly mean of relative sunshine duration (*RSD*, dimensionless) in the Carpathian Region for the period 1981–2010: by averaging the observed time series for each year (O); by averaging the time series estimated using the scheme proposed by Yin, X. (1999) for each year (PA); and by applying the scheme to 30-year averages, with the provision that the term related to the solar irradiance in the model (see R_E in Eq. 3) is calculated for year 1995 using the algorithm used in the SPLASH v.1.0 model (for details see evaluation methodology) (PB). In addition to the values (a), the differences from the various sources are also mapped (b: PA–O, PB–O, and PB–PA), in the latter case showing the root mean square error normalized by the mean value of reference data (*RRMSE*, in percentages) over the whole target domain.



time window used here. The change in the methodology for generating daily weather data has little effect on the spatial distributions of this bioclimatic index over the study period: for only 13 out of the 5,895 grid cells, the value of α changes by more than 0.005 when the daily data required for the simulation are generated using more sophisticated techniques (see the first row of the second column in *Figure 4*, b). Regardless of the settings, the value of α for the period 1981–2010 does not change over at least one-third of the target domain, however, the spatial distribution of these unchanged areas varies depending on the choice of source for

Fig. 4. Spatial distributions of the Priestley–Taylor coefficient (α, dimensionless) in the Carpathian Region for the period 1981–2010: (a) the values of α are simulated by the SPLASH v.1.0 model using the observed values of monthly means of relative sunshine duration (*RSD*, dimensionless), assuming monthly means for each meteorological variable to be constant over each day of the month; and (b) the differences of α values modelled by the initial algorithm and estimated under various model configurations. In each cell of the panel (b), the references are derived from the panel (a). In each row of the panel (b), the estimates are calculated using the *RSD* values derived from various sources: (O) by averaging the single year time series of the observations; (PA) by averaging the time series estimated using the initial scheme for each year; and (PB) by applying the scheme to 30-year averages. In each column of the panel (b), the estimates are calculated using quasi-daily values of each meteorological variable generated by different approaches: (constant) monthly means are assumed constant over each day of the month; and (interpolated) different mean-preserving interpolation techniques are applied (for details, see evaluation methodology). In the panel (b), the root mean square error normalized by the mean value of reference data (*RRMSE*, in percentages) over the whole target domain is shown above each map.

the *RSD* data. When using the model driven by sunshine data estimated using multi-year averages, the unchanged areas are limited to the Carpathians that are the wettest regions of the target domain (see the third row in *Figure 4*, b). At this setting, wetter conditions compared to the reference are simulated over more than two-thirds of the study area (66.9% and 63.2%, respectively, for constant and interpolated daily data). While simulations using estimated single year time series of *RSD* show an underestimation in an area of a similar extent (see the second row in *Figure 4*, b). Overall, changing the configuration settings does not have a significant effect on this bioclimatic index, with the *RRMSE* for this index hovering around 1.5 percent, considering all simulation experiments.

Considering all five bioclimatic indices used in the BIOME model, in addition to α, the growing degree-days can also be significantly influenced by the approach used to generate daily temperature values. Thus, a sensitivity analysis is also performed for these two indices. Values of GDD_5 and GDD_0 for the period 1981–2010 are calculated using the two approaches described above, and their spatial distribution *(Figure 5*) is plotted (mostly) using the thresholds used in the BIOME model (see *Table 1*). Except for the highest peaks of the Carpathians, in the target domain, the value of GDD_5 exceeds the threshold of

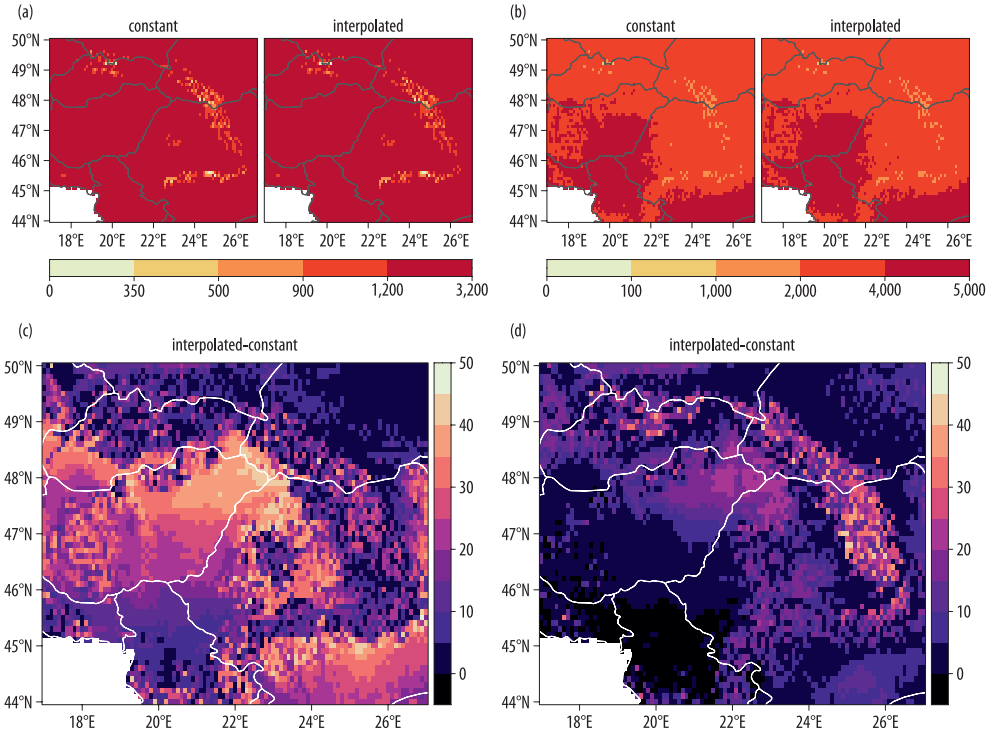


Fig. 5. Spatial distributions of growing degree-days above base temperatures of 5 °C and 0 °C (GDD_5 and GDD_0, in °C day) in the Carpathian Region for the period 1981–2010. Panels (a) and (b) show the spatial distribution of respectively GDD_5 and GDD_0 depending on how the quasi-daily temperature values required to compute these bioclimatic indices are constructed: (constant) monthly means are assumed constant over each day of the month; and (interpolated) the 'harmonic' interpolation technique described by Epstein, E.S. (1991) is applied. Panels (c) and (d) show the spatial distributions of the differences between growing degree-days calculated from interpolated and constant daily values, respectively, for GDD_5 and GDD_0.

1,200 °C day, regardless of the settings (see *Figure 5*), and thus, in these regions, the PFT "temperate summer-green tree" (te.s.t) can occur (see *Table 2*), due to the availability of sufficient moisture (see *Figure 4*). Although, in relation to the spatial distribution of these indices (*Figure 5*, a and b), a large difference between the two approaches cannot be observed, inter alia due the thresholds used to plotting, but the distribution maps for the differences (*Figure 5*, c and d) provide additional valuable information. In the case of GDD_5 (GDD_0), the areas with a difference of greater than 30 °C day are mostly located at altitudes lower (higher) than 500 m a.s.l. (cf. *Figure 5*, c and d, and *Figure 1*). For growing degree-days, the difference between the two algorithms is greater in grid cells where the monthly mean temperatures in the first and last months of the growing season are spread around the given base temperature and the annual diurnal range is relatively large. It is easy to understand that under a typical annual temperature course, in cases where the monthly mean temperature is equal to or slightly greater than 5 °C in both April and November, for the GDD_5, a larger amount of heat can be generated from interpolated values than from constant daily data.

As a final step in this study, it is checked how changing the configuration settings affects the biome designation. Thus, the main results of this study include the distribution maps of biomes under different configuration settings *(Figure 6)*. Here again, the reference simulation is prepared using the measured values of *RSD* and assuming constant monthly means (the first row of the first column in *Figure 6*). For the period 1981–2010, 5 out of the 14 extratropical biome types used by the BIOME model can be observed in the Carpathian Region, at a horizontal resolution of 0.1°. More than half (55%) of the target area is covered by the biome type "temperate deciduous forest" (TEDE), mostly limited to the lowlands (elevation < 250 m a.s.l.). With an areal proportion of 40.7 percent, the second most dominant biome type in the target domain is the "cool mixed forest" (COMX), covering significant areas in Slovakia, Ukraine and the mountains of Romania. The types "taiga" (TAIG) and "cool conifer forest" (COCO) together cover a total of 4.2 percent of the study area. As shown in *Figure 6*, these types appear most markedly in the Eastern Carpathians. The type "tundra" (TUND) covers slightly more than 0.1 percent of the target area (7 grid cells). Comparing the simulation experiments, it can be found that the choice of source for the time series of *RSD* has no effect on the biome distribution under given space and time conditions: values of the Kappa statistic between the maps derived from the reference simulation and from the remaining two experiments are equal to one, i.e., the spatial distribution patterns of biomes are completely identical (see the first column in *Figure 6*). Biome maps generated using interpolated daily values are consistent with each other (see the second column in *Figure 6*). Comparing them to the reference map, only a slight mismatch can be found (Kappa statistic = 0.9923). There is a disagreement between biome maps for only 24 of the 5,895 grid cells derived using different daily weather data. In all cases, this mismatch is explained by the discrepancy in the spatial distribution of GDD_5 (see *Figure 5*, c). The difference is ultimately due to the fact that in some grid cells of the Carpathians, certain heat-demanding woody PFTs (e.g., te.s.t) can occur in the case of the reference simulation, in contrast to the experiments using interpolated daily values. In summary, the BIOME model is not sensitive to modify configuration settings considered here.

Conclusions

In this paper, as a case study for the Carpathian Region, we inspect the efficiency of biome distribution simulation using only monthly temperature and precipitation climatologies. The biome maps were constructed by using a simple process-based vegetation model, the BIOME model, with one minor amendment: the water balance module of the model was

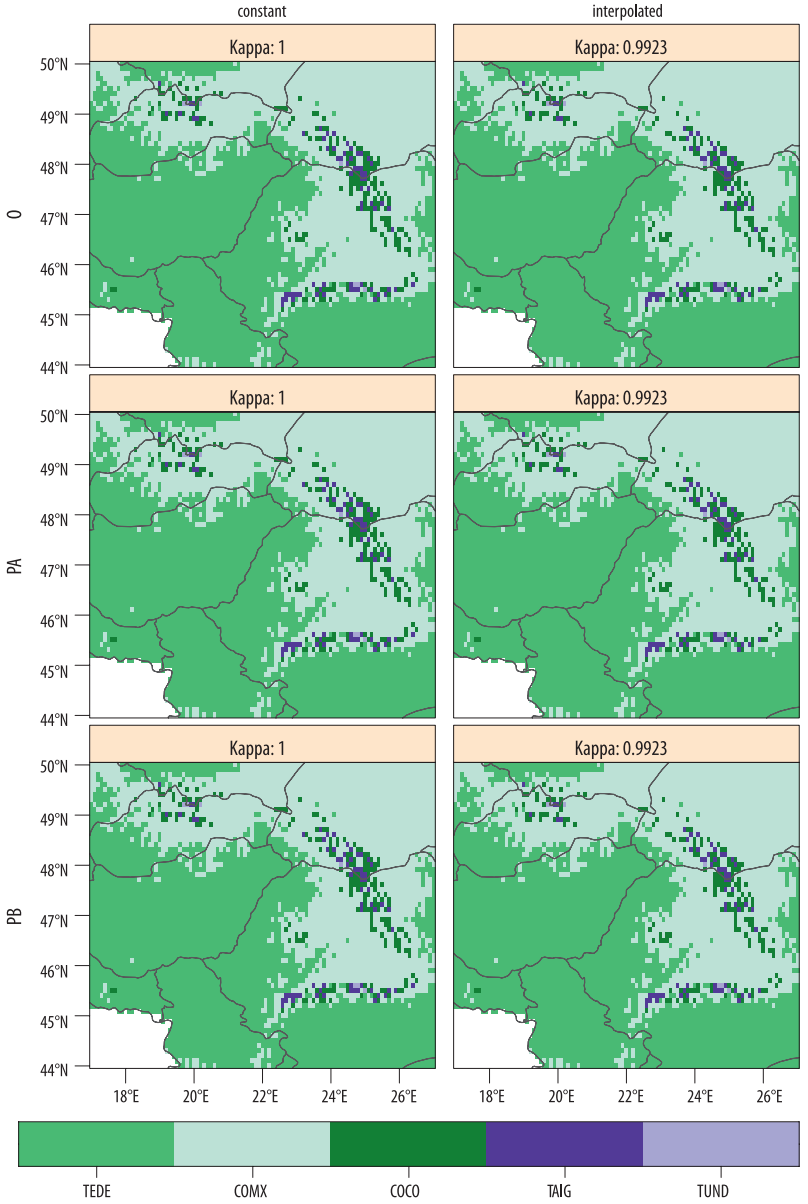


Fig. 6. Spatial distribution patterns of biomes simulated by the BIOME model in the Carpathian Region for the period 1981–2010. In each row, the biomes are derived using the sunshine duration data derived from different sources: (O) by averaging the single year time series of the observations; (PA) by averaging the time series estimated using the initial scheme for each year; and (PB) by applying the scheme to 30-year averages. In each column, the biomes are derived using quasi-daily values of each meteorological variable generated by different approaches: (constant) monthly means are assumed constant over each day of the month; and (interpolated) different mean-preserving interpolation techniques are applied (for details, see evaluation methodology). Above each map, the Kappa statistic reflecting the degree of similarity of the distribution patterns is shown, using the map in the first column of the first row as a reference in each case. The abbreviations of biome types can be found in *Table 2*.

replaced by the SPLASH v.1.0 model, thus switching to analytical expressions for calculating daily radiation, evapotranspiration and soil moisture. Monthly temperature and precipitation data, which were required to create the biome maps, were taken from the CarpatClim dataset for the period 1981–2010. The relative BSD data required to run the SPLASH v.1.0 model were taken from the CarpatClim dataset and also estimated by the scheme proposed by Yin, X. (1999) using the above-mentioned temperature and precipitation data. Comparisons between the observed and estimated relative BSD time series for the period 1961–2010 showed that the estimation procedure performed relatively well from late spring to early autumn, i.e., in the most important period in terms of the evapotranspiration processes. It was also examined the effects of two modifications justified by the applicability of the estimation method to paleoclimate datasets: (a) to apply the scheme of Yin, X. (1999) to multi-year averages instead of single year time series, and (b) to calculate the term related to the solar irradiance in the scheme by applying the algorithm used in the SPLASH v.1.0 model under changing orbital parameters of the Earth. It was found that although the magnitude of overestimation for the modified algorithm is significant in the winter period, the proposed procedure performs similarly well as the initial procedure in the period from March to October. When modelling the distribution of biomes, simulation experiments were performed to assess the effects of modifying some configuration settings of the model: (a) the generation of relative BSD data, and (b) the algorithm used to create quasi-daily weather data from the monthly climatologies. We found that under both the recent humidity conditions of the study region and the spatial resolution of the climate dataset used, the results can be considered sufficiently robust, regardless of the configuration settings tested. The choice of source for BSD data had no effect on the results, while the choice of the method for temporal downscaling of monthly temperature data had little effect on the distribution patterns of biomes. Thus, the main message of this paper is that using climate data available for Quaternary studies (i.e., monthly temperature and precipitation climatologies), the spatial distribution of biomes can be properly simulated via more sophisticated biome models than BCMs. We believe that by applying the modelling framework outlined here to the data provided by the CHELSA-TraCE21k v1.0, the evolution of biomes over the past millennia can be properly mapped.

***Acknowledgements*:** The work of the first author was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'. The work of the second author was partly financed by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. The authors would like to thank the creators of the CarpatClim dataset (http://www.carpatclim-eu.org/) used in this study for sharing their datasets with the scientific community. The authors are also grateful to the software developers of the SPLASH model (https://bitbucket.org/labprentice/splash) for making available to everyone their source codes.

REFERENCES

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A. and Hegewisch, K.C. 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data* 5. 170191. Available at https://doi.org/10.1038/sdata.2017.191

Allen, R.G. 1996. Assessing integrity of weather data for reference evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering* 122. (2): 97–106. Available at https://doi.org/10.1061/(ASCE)0733-9437(1996)122:2(97)

Bellocchi, G., Acutis, M., Fila, G. and Donatelli, M. 2002. An indicator of solar radiation model performance based on a fuzzy expert system. *Agronomy Journal* 94. (6): 1222–1233. Available at https://doi.org/10.2134/agronj2002.1222

Berger, A. and Loutre, M.F. 1991. Insolation values for the climate of the last 10 million years. *Quaternary Science Reviews* 10. (4): 297–317. Available at https://doi.org/10.1016/0277-3791(91)90033-Q

Beyer, R.M., Krapp, M. and Manica, A. 2020. High-resolution terrestrial climate, bioclimate and vegetation for the last 120,000 years. *Scientific Data*

7. 236. Available at https://doi.org/10.1038/s41597-020-0552-1

Brock, T.D. 1981. Calculating solar radiation for ecological studies. *Ecological Modelling* 14. (1–2): 1–19. Available at https://doi.org/10.1016/0304-3800(81)90011-9

Castellvi, F. 2001. A new simple method for estimating monthly and daily solar radiation. Performance and comparison with other methods at Lleida (NE Spain), a semiarid climate. *Theoretical and Applied Climatology* 69. (3–4): 231–238. Available at https://doi.org/10.1007/s007040170028

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20. (1): 37–46. Available at https://doi.org/10.1177/001316446002000104

Davis, T.W., Prentice, I.C., Stocker, B.D., Thomas, R.T., Whitley, R.J., Wang, H., Evans, B.J., Gallego-Sala, A.V., Sykes, M.T. and Cramer, W. 2017. Simple process-led algorithms for simulating habitats (SPLASH v.1.0): robust indices of radiation, evapotranspiration and plant-available moisture. *Geoscientific Model Development* 10. (2): 689–708. Available at https://doi.org/10.5194/gmd-10-689-2017

Doorenbos, J. and Pruitt, W.O. 1977. *Guidelines for predicting crop water requirements. FAO Irrigation and Drainage Paper No. 24, Technical Report*. Rome, Food and Agriculture Organization of the United Nations.

Duffie, J.A. and Beckman, W.A. 1991. *Solar Engineering of Thermal Processes.* Second Edition. New York, Wiley-Interscience.

Elith, J. and Leathwick, J.R. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40. 677–697. Available at https://doi.org/10.1146/annurev.ecolsys.110308.120159

Epstein, E.S. 1991. On obtaining daily climatological values from monthly means. *Journal of Climate* 4. (3): 365–368. Available at https://doi.org/10.1175/1520-0442(1991)004<0365:OODCVF>2.0.CO;2

Harris, I., Osborn, T.J., Jones, P. and Lister, D. 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* 7. 109. Available at https://doi.org/10.1038/s41597-020-0453-3

Harrison, S.P. 2013. Data–Model Comparisons. In *Encyclopedia of Quaternary Science.* Second Edition. Eds.: Elias, S.A. and Mock, C.J., Amsterdam, Elsevier, 135–146. Available at https://doi.org/10.1016/B978-0-444-53643-3.00007-8

He, F. 2011. *Simulating Transient Climate Evolution of the Last Deglaciation with CCSM3*. Dissertation, Madison, University of Wisconsin–Madison.

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

Hoyt, D.V. 1977. Percent of possible sunshine and the total cloud cover. *Monthly Weather Review* 105. (5): 648–652. Available at https://doi.org/10.1175/1520-0493(1977)105<0648:POPSAT>2.0.CO;2

Iqbal, M. 1983. *An Introduction to Solar Radiation*. London, Academic Press.

Kandirmaz, H.M., Kaba, K. and Avci, M. 2014. Estimation of monthly sunshine duration in Turkey using artificial neural networks. *International Journal of Photoenergy* 2014. 680596. Available at https://doi.org/10.1155/2014/680596

Kaplan, J.O. 2001. *Geophysical Applications of Vegetation Modelling*. Dissertation, Lund, Sweden, Lund University.

Karger, D.N., Nobis, M.P., Normand, S., Graham, C.H. and Zimmermann, N.E. (in review) CHELSA-TraCE21k v1.0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past Discussion*. Available at https://doi.org/10.5194/cp-2021-30

Köppen, W. 1936. Das geographische System der Klimate. In *Handbuch der Klimatologie*. Hrsg.: Köppen, W. und Geiger, R., Berlin, Verlag von Gebrüder Borntraeger, 1–44. (in German)

Krapp, M., Beyer, R.M., Edmundson, S.L., Valdes, P.J. and Manica, A. 2021. A statistics-based reconstruction of high-resolution global terrestrial climate for the last 800,000 years. *Scientific Data* 8. 228. Available at https://doi.org/10.1038/s41597-021-01009-3

Lüdeke, M.K.B., Badeck, F.W., Otto, R.D., Häger, C., Dönges, S., Kindermann, J., Würth, G., Lang, T., Jäkel, U., ... Kohlmaier, G.H. 1994. The Frankfurt Biosphere Model: A global process-oriented model of seasonal and long-term CO_2 exchange between terrestrial ecosystems and the atmosphere. I. Model description and illustrative results for cold deciduous and boreal forests. *Climate Research* 4. (2): 143–166. Available at https://doi.org/10.3354/cr004143

Mauri, A., Davis, B.A.S., Collins, P.M. and Kaplan, J.O. 2014. The influence of atmospheric circulation on the mid-Holocene climate of Europe: a data–model comparison. *Climate of the Past* 10. (5): 1925–1938. Available at https://doi.org/10.5194/cp-10-1925-2014

Miller, P.A., Giesecke, T., Hickler, T., Bradshaw, R.H.W., Smith, B., Seppä, H., Valdes, P.J. and Sykes, M.T. 2008. Exploring climatic and biotic controls on Holocene vegetation change in Fennoscandia. *Journal of Ecology* 96. (2): 247–259. Available at https://doi.org/10.1111/j.1365-2745.2007.01342.x

Phillips, S.J., Anderson, R.P. and Schapire, R.E. 2006. Maximum entropy modelling of species geographic distributions. *Ecological Modelling* 190. (3–4): 231–259. Available at https://doi.org/10.1016/j.ecolmodel.2005.03.026

Prentice, K.C. 1990. Bioclimatic distribution of vegetation for general circulation model studies. *Journal of Geophysical Research: Atmospheres* 95. (D8): 11811–11830. Available at https://doi.org/10.1029/jd095id08p11811

Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A. and Solomon, A.M. 1992. A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19. (2): 117–134. Available at https://doi.org/10.2307/2845499

Prentice, I.C., Sykes, M.T. and Cramer, W. 1993. A simulation model for the transient effects of climate change on forest landscapes. *Ecological Modelling* 65. (1–2): 51–70. Available at https://doi.org/10.1016/0304-3800(93)90126-D

Prentice, I.C., Harrison, S.P., Jolly, D. and Guiot, J. 1998. The climate and biomes of Europe at 6000 year BP: comparison of model simulations and pollen-based reconstructions. *Quaternary Science Reviews* 17. (6–7): 659–668. Available at https://doi.org/10.1016/S0277-3791(98)00016-X

Prentice, I.C., Bondeau, A., Cramer, W., Harrison, S.P., Hickler, T., Lucht, W., Sitch, S., Smith, B. and Sykes, M.T. 2007. Dynamic global vegetation modelling: Quantifying terrestrial ecosystem responses to large-scale environmental change. In *Terrestrial Ecosystems in a Changing World*. Eds.: Canadell, J.G., Pataki, D.E. and Pitelka, L.F., Berlin and Heidelberg, Springer, 175–192. Available at https://doi.org/10.1007/978-3-540-32730-1_15

Spinoni, J., Szalai, S., Szentimrey, T., Lakatos, M., Bihari, Z., Nagy, A., Németh, Á., Kovács, T., Mihic, D., … Vogt, J. 2015. Climate of the Carpathian Region in the period 1961–2010: Climatologies and trends of 10 variables. *International Journal of Climatology* 35. (7): 1322–1341. Available at https://doi.org/10.1002/joc.4059

Stanghellini, C. 1981. A simple method for evaluating sunshine duration by cloudiness. *Journal of Applied Meteorology and Climatology* 20. (3): 320–323. Available at https://doi.org/10.1175/1520-0450(1981)020<0320:ASMFES>2.0.CO;2

Szabóné André, K., Bartholy, J., Pongrácz, R. and Bór, J. 2021. Local identification of persistent cold air pool conditions in the Great Hungarian Plain. *Időjárás* 125. (2): 167–192. Available at https://doi.org/10.28974/idojaras.2021.2.1

Tapiador, F.J., Moreno, R., Navarro, A., Sánchez, J.L. and García-Ortega, E. 2019. Climate classifications from regional and global climate models: Performances for present climate estimates and expected changes in the future at high spatial resolution. *Atmospheric Research* 228. 107–121. Available at https://doi.org/10.1016/j.atmosres.2019.05.022

Valdes, P.J., Armstrong, E., Badger, M.P.S., Bradshaw, C.D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J.J., Farnsworth, A., … Williams, J.H.T. 2017. The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0. *Geoscientific Model Development* 10. (10): 3715–3743. Available at https://doi.org/10.5194/gmd-10-3715-2017

Webb, III T., Anderson, K.H., Bartlein, P.J. and Webb, R.S. 1998. Late Quaternary climate change in eastern North America: a comparison of pollen-derived estimates with climate model results. *Quaternary Science Reviews* 17. (6–7): 587–606. Available at https://doi.org/10.1016/S0277-3791(98)00013-4

Yates, D.N., Kittel, T.G.F. and Cannon, R.F. 2000. Comparing the correlative Holdridge model to mechanistic biogeographical models for assessing vegetation distribution response to climatic change. *Climatic Change* 44. (1–2): 59–87. Available at https://doi.org/10.1023/A:1005495908758

Yin, X. 1997a. Optical air mass: Daily integration and its applications. *Meteorology and Atmospheric Physics* 63. (3–4): 227–233. Available at https://doi.org/10.1007/BF01027387

Yin, X. 1997b. Calculating daytime mean relative air mass. *Agricultural and Forest Meteorology* 87. (2–3): 85–90. Available at https://doi.org/10.1016/S0168-1923(97)00029-4

Yin, X. 1998. The albedo of vegetated land surfaces: Systems analysis and mathematical modelling. *Theoretical and Applied Climatology* 60. (1–4): 121–140. Available at https://doi.org/10.1007/s007040050038

Yin, X. 1999. Bright sunshine duration in relation to precipitation, air temperature and geographic location. *Theoretical and Applied Climatology* 64. (1–2): 61–68. Available at https://doi.org/10.1007/s007040050111