E-Clic – easy climate data converter

Pedro Tarroso and Hugo Rebelo

P. Tarroso (ptarroso@mail.icav.up.pt), CIBIO/UP-Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Rua Padre Armando Quintas, PT-4485-661 Vairão, Portugal, and Equipe Paléoenvironnements et Paléoclimats, Inst. des Sciences de l’Évolution de Montpellier, UMR 5554, Univ. Montpellier 2, Bat.22, CC061, Pl. Eugène Bataillon, FR-34095 Montpellier cedex 5, France.
– H. Rebelo, School of Biological Sciences, Univ. of Bristol, Woodland Road, Bristol, BS8 1UG, UK, and CIBIO/UP-Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Rua Padre Armando Quintas, PT-4485-661 Vairão, Portugal.

There are an increasing number of studies that are now focusing on the influence of climate change on species’ distributions. However, access to predictive climatic datasets for future scenarios is difficult due to their specific formats and/or the need to be geographically downscaled. The TYN dataset is freely available to users and provides a synthetic format with several climatic models and IPCC future climate scenarios. Moreover, the CRU historical dataset (1901–2000) is also available which allows users to create baseline models for current climatic variables. E-clic is a free, user-friendly software package that offers three different ways to convert these two datasets into a spatially explicit raster format which is compatible with the most common geographic information systems and usable on different platforms.

Climatic variables provide key input parameters when modelling species’ distributions (Guisan and Zimmermann 2000). Despite the range of different variables that are frequently used in ecological modelling, climate data is commonly used due to its continuous influence in shaping and limiting species’ distributions, and also because of its correlation with other biologically meaningful variables (e.g. land cover) (Huntley and Webb 1989, Thomas et al. 2004, Thuiller et al. 2004). The changing dynamics of climatic forces are highly correlated with the shift of species’ distributions, as shown in the past by the repeated moving patterns of populations during the climate oscillations of the glacial and interglacial periods (Huntley and Webb 1989). The current forecast of rapid climate change will force a shift in species’ distributions and, in some cases, will increase their risk of extinction (Parmesan and Yohe 2003, Thomas et al. 2004). Predicting changes in species’ distributions and the likelihood of their extinction is an important measure in understanding the possible impacts of climate change on biodiversity (Hannah et al. 2002). Recently, several studies have been carried out on a number of species, including plants, birds, marine animals, amphibians and reptiles (Thuiller et al. 2005, Araújo et al. 2006, Hawkes et al. 2007, Lemoine et al. 2007). Consequently, achieving accurate distribution models using present day climatic variables is the first step in projecting these models for future predicted climate scenarios. This will allow for more informed decisions regarding the possible loss of biodiversity.

As one of the greatest political, social and scientific concerns of our time, research on climate change has generated a multidisciplinary interest which has produced a wide range of models that combine different social-economic scenarios to predict several climatic variables in the future (Nakicenovic and Swart 2000). Climatic models are available at different geographic scales but usually have a coarse resolution and can be stored in a difficult format that is not easily integrated into the most commonly used geographic information systems (GIS) (e.g. ArcMap) and ecological modelling software (e.g. MaxEnt). The TYN dataset (Mitchell et al. 2004) provides several downscaled models that are frequently used to study the impact of climate change. Specifically, it consists of monthly data broken down into five variables (precipitation, daily mean temperature, diurnal temperature range, vapour pressure, cloud cover) for five global circulation models (CGCM2, CSIRO mk 2, DOE PCM, HadCM3, ECHam4) over 100 yr intervals (from 2001 to 2100). It also includes four special reports on emission scenarios (A1FI, A2, B1 and B2) resulting from the Intergovernmental Panel on Climate Change (IPCC) meeting (Nakicenovic and Swart 2000). These high-resolution datasets are freely available from the Climate Research Unit web page (<www.cru.uea.ac.uk/cru/data/hrg.htm>) at two different scales: the TYN SC 1.0 with European data downscaled to 10’ resolution and the TYN SC 2.0 with global coverage at a 0.5° resolution. Although new future climate models and datasets occasionally emerge, the TYN dataset remains as a stable and complete source of future climate predictions for ecological
studies. In addition to the TYN dataset, the CRU TS 1.0 dataset (Mitchell et al. 2004) is also freely available to researchers and comprises of historical climate data ranging from 1901 to 2000 for the same variables, coverage and resolution as TYN SC 1.0. This data is highly useful because it provides researchers the opportunity to widen the timeframe of their studies and to test the accuracy of their models if a species range has been well documented over the last century. The equivalent historical dataset for the TYN SC 2.0, the CRU TS 2.10 (Mitchell and Jones 2005) is also available, including four new variables (monthly average daily maximum temperature, monthly average daily minimum temperature, wet day frequency, frost day frequency).

Here we present a new software program to convert public climatic datasets (TYN SC and CRU TS) into formats that are more commonly used and therefore can be directly utilized in GIS. “Easy climate data converter” (E-clic) (<http://purl.oclc.org/eclic/>) is a free, open-source application which is written in Python and can be used in different operating systems such as Windows, Mac OS or Linux (Fig. 1). It can read both TYN SC and CRU TS datasets and is able to convert them to ASCII raster maps—a format easily readable in most GIS and ecological modelling software (e.g. Arcgis, Idrisi, Maxent).

Prior to data conversion, users should download and decompress the TYN and/or CRU datasets to a chosen folder. The CRU dataset is available in two different forms; one single file that includes all the CRU TS range data, and as separate files containing data for each decade. The single file that includes all data must be used. Unpacking of the data from the files done by E-clic follows the procedure described online with the dataset (see dataset internet site for more details). In short, it consists of searching the data for the defined period of time, converting it into real units given by the model/scenario, and then testing it with the allowed minimum and maximum limits for the chosen variable. The user can choose one model, scenario and variable from all of the TYN supporting data. The output is the average model for the chosen time period and, optionally, a raster for each month within the same period.

The output created by E-clic is easily integrated into most GIS packages and modelling software. The ability to query a spatial and temporal dataset and create raster data from them avoids, in most cases, the need to post-process the raster datasets. Although there are other software packages available to convert TYN SC and CRU TS data
into different formats (see the dataset web page for more software and Solymosi et al. 2008), they lack some of the features presented in E-Clic such as a user-friendly interface, ability to be used on different platforms and/or the option to export data into a raster format. As E-clic is written in Python, it benefits from the multi-platform availability of this programming language (Fig. 1). Moreover, it is not dependent upon external python modules and can be run directly after python installation (<www.python.org>), which is already available as default on some platforms (e.g. most of Linux distributions and Mac OS). In addition, this software also offers three different interfaces to convert data, depending on the user’s preferences. E-clic may be used at the command line with data inputted in a strict sequence format (see the online instructions in E-clic website for more details on how to use E-clic at the command line). The command line feature may be used to build quick batches to produce large quantities of data. The built-in graphical user interface (Fig. 1) is displayed when the user runs E-clic without any additional command, the interface is launched displaying all the possible options making the process of data conversion between formats relatively simple. Finally, it may also be integrated into ArcGIS (ESRI, Redlands, CA, USA) as a toolbox (Fig. 2) where the graphical interface provided by the GIS package can be used to access all of E-clic’s functions. When it is used within the GIS or when it detects a valid license, it will also output the raster in a geoTIF format. All result files are

Figure 3. Average daily temperature forecasts for the years 2025, 2050, 2075 and 2100 for four available scenarios (A1FI, A2, B1 and B2). The plots represent the frequency of pixels with the same value of temperature. The year 2025 is plotted with a solid line (---) 2050 with a dashed line (--), 2075 with a dash and dot line (.-) and 2100 with a dotted line (..).
saved to the chosen output directory and the names are explicit to the data they contain.

We predict that the main application of the output raster maps created by E-Clic will be in species distribution modelling for future climate scenarios. We present here a simple example of the output maps for an area covering the Iberian Peninsula and part of North Africa (Fig. 3). With TYN_SC_1.06 data, we have built annual mean values for four daily mean temperature scenarios (ranging from the more extreme A1FI, A2, B2 to the less severe B1) for the years 2025, 2050, 2075 and 2100. We synthesized the data using histograms where the frequency of map cells with the same value is plotted against the temperature. The results show that all the scenarios predict an increasing number of cells with higher temperatures with time and, as expected, the A1FI scenario has the greatest increase in temperature. The increase of the maximum values of the predicted mean temperatures for this area between 2025 and 2100 reaches the highest value of 6.7°C with the A1FI scenario, whereas the lowest is 2.7°C for the B1 scenario.

Even though there are a limited number of climatic data conversion software packages available, the majority of them require some knowledge of computer programming. Since the majority of users will not be familiar with this type of programming, especially those who come from different scientific fields such as ecology, accessing climatic data can be difficult. Here we present a user-friendly application that allows anyone with the most basic GIS knowledge to easily access climatic data for future scenarios. Given that climate change studies are becoming more important, we hope that this software will help more researchers to have access to the core datasets for their studies.

To cite E-Clic or acknowledge its use, cite this software note as follows, substituting the version of the application that you used for “Version 0”:


**Acknowledgements** – The authors are grateful to Jon Flanders for his extensive review of the manuscript that greatly improved its quality. The authors are also grateful to James Harris, Anna Perera and two anonymous referees for their comments and suggestions.

PT and HR were funded by the “Fundação para a Ciência e Tecnologia” doctoral grants SFRH/BD/42480/2007 and SFRH/BD/17755/2004, respectively.

**References**


