Supplementary Material for
Trouet, V., and Van Oldenborgh, G.J., 2013. KNMI Climate Explorer: A web-based research tool for high-resolution paleoclimatology. *Tree-Ring Research* v. 69 no. 1, p. 3-13.

KNMI CLIMATE EXPLORER (CE) INSTRUCTION MANUAL

VALERIE TROUET[1] and **GEERT JAN VAN OLDENBORGH**[2]
[1] Laboratory of Tree-Ring Research, University of Arizona, 105 W Stadium, Tucson AZ 85721, USA, trouet@ltrr.arizona.edu
[2] Royal Netherlands Meteorological Institute KNMI, PO Box 201, NL-3730 AE De Bilt, Netherlands

1. Introduction

This text is intended as a step-by-step instruction manual for the prospective KNMI CE user. It provides the specific commands required to run analyses and produce figures that are described in the paper. It is worth mentioning that CE supports a comprehensive help overview (http://climexp.knmi.nl/help.cgi) that can be consulted for additional information. We will here show how CE can be used to explore and download available instrumental climate data, to examine the climate signal in uploaded paleoclimate time series, and to investigate the temporal and spatial characteristics of climate reconstructions. This instruction manual will therefore be organized in the same way, including three sub-sections, as the main paper.

2. Exploration of Instrumental Data Sets

After free registration, the researcher can access and investigate climatic data sets by selecting the appropriate time series or fields format on the right-hand side of the web page. Monthly time series and fields are most useful for dendroclimatic analysis, but daily data sets can be transformed to monthly or lower resolution in CE by using the [*Create a lower resolution time series*] function. It is recommended to read the information links available for all data sets ([*i*] tab to the right of each data set) before making a decision on which data set to use. These links also provide the necessary citations for all data sets.

When selecting a time series of monthly **station data** ([*get data*] next to the station meta data), the researcher is given the opportunity to make a selection based on station name (if the station of interest is known) or on vicinity to a location. The second option includes searching for a defined number of stations nearest to a given site or all stations within a defined region. Note that west or south coordinates are indicated by negative numbers. Furthermore, a minimum number of years of data availability can be defined as well as a range of years for which data should be available and an elevation range. Upon station selection, the researcher is offered three data plots and can download the data behind these plots as postscript files, plain text files (raw data), or NetCDF files. Furthermore, the selected time series can be high- and low-pass filtered (using a running mean or LOESS filter with a user-defined window length [cut-off value]).

When selecting a **field**, the researcher is given the choice to retrieve data for a single gridpoint or for an area. For single gridpoint data, only one latitude and one longitude should be entered. Selecting [*halfway grid points*] then provides the gridpoint nearest to the given coordinates, whereas [*interpolated*] provides a simple bilinear interpolation. To obtain data for an area, the corner coordinates for the area should be entered and the researcher can then select either the average or the set of gridpoint data within these boundaries.

All time series or derived time series of station or gridpoint data can be downloaded or named and added to the list of user-defined time series [manipulate this time series/make index] for further analysis.

3. Examination of the Climate Signal in Dendroclimatic Time Series

CE allows uploading of a user-defined time series [*view, upload your time series*] and running a variety of statistical analyses. To upload a time series, a spreadsheet can be created with the years covered by the tree-ring series in the first column and the corresponding tree-ring values in the second column. Headings are recommended but not necessary, undefined data should be represented by the value -999.9. It then suffices to copy both columns in the [*upload text data*] field, name the data set, and upload in CE. It is worth noting that CE also accepts CF-compliant NetCDF files for uploading data time series and grids. Uploaded time series will remain accessible in CE for three days after last use.

CE only correlates data sets with the same temporal resolution (*i.e.* monthly with monthly, annual with annual) and uploaded and existing data sets should therefore be converted to the same resolution. To correlate an annually-resolved tree-ring series with annual instrumental climate data, an annually-resolved climate time series can be derived from monthly (or daily) data. When a comparison of tree-ring data with monthly climate data is desired, the tree-ring series should be converted to monthly resolution. An annual time series (one column after the year) or seasonal one (four columns) can be uploaded and reduced to monthly resolution in CE [*create a monthly time series*], this also allows for a choice for which months the signal is most representative. An alternative (and likely more straightforward) method for calculating seasonal and annual comparisons is described below.

CE allows the user to correlate an uploaded time series with other user-defined time series as well as with system-defined time series of the same temporal resolution [*correlate with other time series*]. Analysis options include (rank) correlation or regression analysis for individual or all months of the calendar year (January-December). Selecting [*together*] will calculate correlation coefficients for the climate data over those twelve months together, including the seasonal cycle. This cycle can be removed by checking [*subtract seasonal cycle*]. The user is also given the option to calculate correlation coefficients using the average of multiple months of climate data. For instance, by [*averaging*] [*all*] [*index selected above*] over [*3*] months of the index, seasonal climate data can be easily compared to the tree-ring series. Given that the tree-ring data set consists of 12 identical values per year, selection of the number of months of the time series is irrelevant except when the season encompasses two calendar years. If the series should be the same across this boundary, another starting month must be chosen in [*Create a monthly time series*]. A single growing season can be selected by selecting a starting month, (*e.g.* April) and the length of the season (*e.g.* [*6*] months). Subtracting the seasonal cycle from the time series will result in identical correlation results in this case.

Furthermore, there is the option of calculating lagged correlations, with positive lags representing the tree-ring series lagging the climate data by a defined number of months. For instance, a lag of [*6*] months applied to [*all*] months of the [*index*] (climate data) will generate coefficients for the correlation between current year growth and climate in the months July through December of the previous year and of the months January through June of the current year. Other selection options include a range of years over which the correlation is run and upper and lower thresholds for both tree-ring and climate data. Selecting one or several threshold values restricts the range of values considered for correlation. Thresholds can be entered as absolute values or as percentages, by adding [*%*] to the provided number. Furthermore, running correlation and regression analyses can be performed with a user-defined time-window. To test the significance of these running correlations, the researcher can choose to replace either the tree-ring series [*timeseries*] or the climatic series [*index*] with either white (high-frequency) or red (low-frequency) noise.

Time series uploaded in CE can not only be compared to other time series, but also to gridded fields of observational and reanalysis climate data. These field analyses typically result in correlation, regression, or composite maps (select [*correlation*], [*regression*], or [*composite*]). For composite map analysis, there are two ways to define the years over which composites are calculated. The first is to define values in the uploaded time series that constitute the upper and/or lower thresholds for the composites ([]< time series <[]). A useful alternative is to define thresholds as time-series percentiles (*e.g.* upper 10%), these can be entered in the same fields by adding a percentage sign (%). The second method is to generate an event-based time series as a bimodal time series (event = 1; no event = 0) and upload this series into CE. Event years can then be composited by selecting 0.5 as the lower threshold [*0.5 < event time series*].

4. Dendroclimatic Time-Series Analysis

Uploaded tree-ring series can be compared in CE to a number of reconstructed climate fields by selecting the appropriate field [*select a field/monthly and seasonal historical reconstructions*]) and comparing it with the uploaded tree-ring series of the same (seasonal or annual) temporal resolution [*investigate this field/correlate with a time series*]. This set-up allows for the development of correlation, regression, and composite maps with similar options as described in section 3.

To complete the spatiotemporal analysis of climate reconstructions, CE provides the opportunity to perform simple spectral, autocorrelation, and wavelet analysis [*view, upload your time series/investigate this time series/ spectrum, autocorrelation function, and /wavelet*].