ClimPACT2

Indices and software

Lisa Alexander and Nicholas Herold

February 2016

A document prepared on behalf of The Commission for Climatology (CCl) Expert Team on Sector-Specific Climate Indices (ET-SCI)

WORLD CLIMATE PROGRAMME

WORLD CLIMATE SERVICES PROGRAMME

Contents

Acknowledgements

This document and the body of work it represents was made possible through the efforts of The World Meteorological Organisation (WMO) Commission for Climatology (CCl) Open Panel of CCl Experts on Climate Information for Adaptation and Risk Management (OPACE 4) under the guidance of OPACE-4 co-chairs (Rodney Martinez and Andrew Tait); the CCl OPACE 4 Expert Team on Sector-specific Climate Indices (ET-SCI) members: Lisa Alexander (Chair, Australia), Toshiyuki Nakaegawa (co-Chair, Japan), Fatima Zohra El Guelai (Morocco), Amelia Diaz Pablo (Peru), Adam Kalkstein (USA) and Gé Verver (The Netherlands) and the WMO World Climate Applications and Services Programme (Rupa Kumar Kolli and Anahit Hovsepyan). It draws heavily on the input of the Expert Team on Climate Risk and Sector-specific Climate Indices (ET-CRSCI), the predecessor of the ET-SCI and including additional ET-CRSCI members Elena Akentyeva, Alexis Nimubona, G. Srinivasan, Philip Thornton, and Peiqun Zhang. Significant contributions to the development of the ET-SCI indices, software and technical manual also came from Enric Aguilar, Andrew King, Brad Rippey, Sarah Perkins, Sergio M. Vicente-Serrano, Juan Jose Nieto, Sandra Schuster and Hongang Yang. We are also grateful to the other experts and sector representatives who have contributed to the development of indices: Manola Brunet, Albert Klein Tank, Christina Koppe, Sari Kovats, Glenn McGregor, Xuebin Zhang, Javier Sigro, Peter Domonkos, Dimitrios Efthymiadis.

Lisa Alexander and Nicholas Herold contributed significantly to development of this document, the indices and the ***ClimPACT2*** software. The majority of indices in ***ClimPACT2*** are calculated using code from the climdex.pcic R package which was developed by the Pacific Climate Impacts Consortium (PCIC). Input was also provided by James Hiebert of PCIC throughout development of ***ClimPACT2***.

The application of climate indices to the Agriculture sector was undertaken in full cooperation with the WMO Commission for Agricultural Meteorology, through which Brad Rippey and Sergio Vicente Serrano supported the work.

Commission for Climatology experts Glenn McGregor, Christina Koppe and Sari Kovats supported the applications of indices for Climate and Health, in particular for heat waves and health.

The ***ClimPACT2*** software updates ***ClimPACT*** which was based on the RClimDEX software developed by the WMO CCl/WCRP/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI). The CCl Co Chair for the CCl OPACE on Climate Monitoring and Assessment (Manola Brunet), ETCCDI members, Albert Klein Tank and Xuebin Zhang, along with Enric Aguilar, Juan Jose Nieto, Javier Sigro, Peter Domonkos, and Dimitrios Efthymiadis, contributed to development of the indices and software in the previous version of the technical manual.

ClimPACT2 is written in R, a language and environment for statistical computing and graphics and makes use of several R subroutines, including SPEI. R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form (see http://www.r-project.org/).

This work is also supported by WMO grant SSA 3876-12/REM/CNS and the Australian Research Council grant CE110001028 specifically through funding from the New South Wales Office of the Environment and Heritage.

Background Material

1. INTRODUCTION

This document was prepared on behalf of the World Meteorological Organization (WMO) Commission for Climatology (CCl) Expert Team on Sector-specific Climate Indices (ET-SCI). It outlines the background and goals of the ET-SCI and describes indices and software that were developed on their behalf.

The ET-SCI, formerly known as the Expert Team on Climate Risk and Sector-specific Indices (ET-CRSCI) was set up by the Fifteenth session of the WMO Technical Commission for Climatology (CCl-XV, Antalya, Turkey, February 2010), with terms of reference established to support eventual implementation of the Global Framework for Climate Services (GFCS) (for background on GFCS see http://www.wmo.int/hlt-gfcs/downloads/HLT_book_full.pdf). Following the sixteenth World Meteorological Congress in May 2011 where a decision was made by WMO members to implement the GFCS, the ET-SCI held their first meeting in Tarragona, Spain (13-15 July, 2011). See http://www.wmo.int/pages/prog/wcp/ccl/opace/opace4/expertteam.php for more details.

1.1 Role of ET-SCI in GFCS

The ET-SCI sits within CCl under the Open Panel of CCl Experts (OPACE) on Climate Information for Adaptation and Risk Management (OPACE-4). The objective of OPACE-4 is to improve decision-making for planning, operations, risk management and for adaptation to both climate change and variability (covering time scales from seasonal to centennial) and will be achieved through a higher level of climate knowledge, as well as by access to and use of actionable information and products, tailored to meet their needs. Activities primarily focus on the development of tailored climate information, products and services for user application in adaptation and risk management, and building interfaces with user groups to facilitate GFCS implementation.

The work of OPACE-4 is multidisciplinary, and requires close collaboration with experts from various socio-economic sectors. In keeping with the priorities agreed for initial implementation of the GFCS, the core priority sectors for consideration by the OPACE in this present intersessional period are agriculture/food security, water and health. This requires close collaboration with relevant experts in these sectors including seeking guidance and aid from the WMO Technical Commissions for Agricultural Meteorology (CAgM) and Hydrology (CHy) and with the World Health Organisation (WHO).

The ET-SCI Terms of Reference (ToR) and expected deliverables are shown in **Appendix A**. The deliverables include the collection and analysis of existing sector-relevant climate indices in addition to developing the tools required to produce them. At a meeting in Tarragona in 2011, members of the former ET-CRSCI invited sector and Commission representatives to help define a suite of indices that would represent a “core set” that would meet the ToR and deliverables. This manual outlines the rationale behind the creation of those indices and the ***ClimPACT2*** software developed for their calculation. In the next section the development of climate indices and their uses are outlined, followed by a description of ***ClimPACT2*** and instructions on how to run it.

ET-SCI INDICES

2.1 The 'value' of climate indices

Monthly averages of climate data smooth over a lot of important information that is relevant for sectoral impacts. For this reason indices derived from daily data are an attempt to objectively extract information from daily weather observations that answers questions concerning aspects of the climate system that affect many human and natural systems with particular emphasis on extremes. Such indices might reflect the duration or amplitude of heat waves, extreme rainfall intensity and frequency or measures of extremely wet or dry/hot or cold periods that have socio-economic impacts. Climate indices provide valuable information contained in daily data, without the need to transmit the daily data itself.

Much progress has been made in recent decades through internationally agreed indices derived from daily temperature and precipitation that represent more extreme aspects of the climate, overseen by the CCl/WCRP/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI). Development and analyses of these indices has made a significant contribution to the Intergovernmental Panel on Climate Change (IPCC) Assessment Reports.

2.2 Background to ETCCDI, Indices and Software

The ETCCDI started in 1999 and is co-sponsored by the World Climate Research Program (WCRP) and JCOMM. They developed an internationally coordinated set of core climate indices consisting of 27 descriptive indices for moderate extremes (*Alexander et al.* 2006; *Zhang et al. 2011*). These indices were developed with the 'detection and attribution' research community in mind. In order to detect changes in climate extremes, it was important to develop a set of indices that were statistically robust, covered a wide range of climates, and had a high signal-to-noise ratio. In addition, internationally agreed indices derived from daily temperature and precipitation allowed results to be compared consistently across different countries and also had the advantage of overcoming most of the restrictions on the dissemination of daily data that are applied in many countries.

ETCCDI recognized that a two-pronged approach was needed to promote further work on the monitoring and analysis of daily climate records to identify trends in extreme climate events (*Peterson and Manton*, 2008). In addition to the formulation of indices described above, a second prong was to promote the analysis of extremes around the world, particularly in less developed countries, by organizing regional climate change workshops that provided training for the local experts and conducted data analysis. The goals of these workshops are to: contribute to worldwide indices database; build capacity to analyse observed changes in extremes; improve information services on extremes in the region; and publish peer-reviewed journal articles. Most of these articles were directly a result of the regional workshops and included all of the workshop participants as authors (e.g. *Peterson et al.* 2002; *Vincent et al.* 2005; *Zhang et al.* 2005; *Haylock et al.* 2006; *Klein Tank et al.* 2006; *New et al.* 2006; *Aguilar et al, 2006*, *Aguilar et al.* 2009; *Caesar et al.* 2011; *Vincent et al.* 2011).

As part of the workshop development, software called RClimDEX was also developed that could be used at the workshops (thus providing consistent definitions from each workshop and region). Environment Canada provides, maintains, and further develops the R-based software used for the workshops (freely available from http://etccdi.pacificclimate.org).

2.3 Background to Development of ET-SCI Indices

Most ETCCDI indices focus on counts of days crossing a threshold; either absolute/fixed thresholds or percentile/variable thresholds relative to local climate. Others focus on absolute extreme values such as the warmest, coldest or wettest day of the year. The indices are used for both observations and models, globally as well as regionally, and can be coupled with simple trend analysis techniques, and standard detection and attribution methods in addition to complementing the analysis of more rare extremes using Extreme Value Theory (EVT).

One current disadvantage of the ETCCDI indices is that few of them are specifically sector-relevant. While some of these indices may be useful for sector applications (e.g. number of days with frost for agricultural applications, heat waves for health applications) it was realised that it was important to get sectors involved in the development of the ET-SCI indices so that more application-relevant indices could be developed to better support adaptation.

The core set of indices agreed by the ET-SCI (as the ET-CRSCI) at their meeting in Tarragona, Spain in July 2011 were developed in part from the core set of indices that are developed and maintained by ETCCDI. The meeting included technical experts in climate and health and climate and agriculture from CCl and CAgM representing the health representatives from the health, water and agriculture sectors and it was agreed that the initial effort should consider requirements for climate indices relevant to heat waves and droughts. A core set of 34 indices was agreed at that meeting (**Table B1**). In some cases these indices are already part of the core set defined by the ETCCDI. All indices calculated by ***ClimPACT2*** are shown in **Appendix B** and are separated into core and non-core ET-SCI indices. In addition, there is limited scope in the ***ClimPACT2*** software for the user to create their own index based on absolute thresholds.

It should be noted that indices development is an ongoing activity as additional sector-needs arise and other sectors are considered (see **Section 2.5**) within the Terms of Reference and deliverables of the ET-SCI. This should therefore be seen only as the initial step in the continuing work of the ET-SCI.

2.4 Background to Development of ET-SCI Workshops

Given the success of the ETCCDI regional workshops (see **Section 2.1** and *Peterson and Manton* 2008), ET-SCI aims to adapt and further develop this regional workshop model with participants from both National Meteorological and Hydrological Services (NMHSs) and sector groups. ET-SCI will work with sector-based agencies and experts, including those of relevant WMO Technical Commissions, particularly the Commission for Climatology for health, the Commission for

Hydrology (CHy) for water and the Commission for Agricultural Meteorology (CAgM) for agriculture and food security, to facilitate the use of climate information in users' decision-support systems for climate risk management and adaptation strategies. As part of this development, ET-SCI has commissioned the development of software, ***ClimPACT2***, with the aim of producing an easy and consistent way of calculating indices for each user. The development and use of this standardized software for generating sector-specific climate indices is described in **Section 3**.

2.5 Requirements for data quality when computing indices

Before indices can be computed, it is important that any daily input data are checked for quality and homogeneity. Homogeneity implies consistency of a series through time and is an obvious requirement for the robust analysis of climate time series. While many of the time series that are used for indices calculations have been adjusted to improve homogeneity, some aspects of these records may remain inhomogeneous, and this should be borne in mind when interpreting changes in indices. For example, most methods for assessing homogeneity do not consider changes in day-to-day variability or changes in how the series has been derived. It is possible for a change of variance to occur without a change in mean temperature. Two examples of ways in which this could occur are where a station moves from an exposed coastal location to a location further inland, increasing maximum temperatures and decreasing minimum temperatures, or where the number of stations contributing to a composite series changes.

Homogeneity adjustment of daily data is difficult because of high variability in the daily data when compared with monthly or annual data, and also because an inhomogeneity due to a change in station location or instrument may alter behaviour differently under different weather conditions. Homogeneity adjustment of daily data is a very active field of research and there are many methods which could be used. Although many different methods exists, the ETCCDI promote the use of the RHTest software* because it is free and easy to use, making it ideal for demonstration in regional workshops. The software method is based on the penalized maximal t (PMT) or F test (PMF) and can identify, and adjust for, multiple change points in a time series (see *Wang*, 2008 and ETCCDI website for more details). PMT requires the use of reference stations for the homogeneity analysis but PMF can be used as an absolute method (i.e. in isolation or when there are no neighbouring stations to use for comparison). In ***ClimPACT2***, apart from basic quality control, there is currently no means to homogenise data. We therefore assume that the required level of homogeneity testing and/or adjustment has already been applied.

***NB** Daily adjustments, especially with absolute methods, must be applied with extreme care as – if incorrectly applied – they can damage the statistical distribution of the series. Therefore, data require careful post-workshop analysis in concert with metadata (where available) and as such ET-SCI recommend that any homogeneity software used at regional workshops is for demonstration purposes only.

2.6 Future prospects for the Indices

At present the core set of indices are defined using only daily maximum temperature (TX), daily minimum temperature (TN) and daily precipitation (PR). It is acknowledged that for sector applications, these variables (and the related indices) are all required, but users have also indicated a need for additional variables including: humidity (important for both agricultural and health indices); wind speed and direction (important for health indices, building design, energy, transportation, etc.); Sea Surface Temperatures (SSTs; useful for marine applications and in relation to the onset and variability of the El Niño-Southern Oscillation (ENSO)); onset and cessation dates for monsoon; rain periods, snow fall, snow depth, snow-water equivalent, days with snowfall and hydrological parameters (particularly important for mid-and high latitude applications). Some of these (e.g. onset dates) may require considerable study and available systematic long-term data. Furthermore, in a subsequent phase of the work of the Team, addition of 'event statistics' such as days with thunderstorms, hail, tornados, number of consecutive days with snowfall, etc., for expanded studies of hazards could be considered. The ET-SCI will consider (at a later date) whether to add these new variables to the dataset as a second level priority.

ET-SCI also feels that it is important to add several complex indices to this initial effort (for example heat waves), but recognized that more could be demanded by (or may be in current use by) sectors, once they are consulted on the process and through training. The development of indices to assess multi-day temperature extremes (e.g., prolonged heat waves) has been particularly challenging, as the occurrence of such events depends not just on the frequency distribution of daily temperatures, but also on their persistence from day to day. The existing ETCCDI indices measure the maximum number of consecutive days during events with six or more consecutive days above a specified percentile value or anomaly, vary widely in frequency across climates, describe events that occur rarely or not at all in many climates, and are poor discriminators of very extreme events. The ET-SCI are therefore recommending some new heat wave indices (see **Appendix B**; *Perkins and Alexander*, 2013 and *Perkins et al*. 2012) that have been added as a supplement to the core set in this initial phase of the software. This range of indices is defined for most climates and has a number of other desirable statistical properties, such as being approximately normally distributed in many climates.

Also drought indices have been included following ET-SCI recommendations. Since drought severity is difficult to quantify and is identified by its effects or impacts on different types of systems (e.g. agriculture, water resources, ecology, forestry, economy), different proxies have been developed based on climatic information. These are assumed to adequately quantify the degree of drought hazard exerted on sensitive systems. Recent studies have reviewed the development of drought indices and compared their advantages and disadvantages (*Heim*, 2002; *Mishra and Singh*, 2010; *Sivakumar et al*., 2010). Currently ***ClimPACT2*** includes the Standardized Precipitation Index (SPI), proposed by *McKee et al.* (1993), and accepted by the WMO as the reference drought index for more effective drought monitoring and climate risk management (*World Meteorological Organization*, 2012), and the Standardized Precipitation Evapotranspiration Index (SPEI), proposed by *Vicente-Serrano et al.* (2010), which combines the sensitivity to changes in evaporative demand, caused by temperature fluctuations and trends, with the simplicity of calculation and the multi-temporal nature of the SPI.
In a subsequent phase, ET-SCI will investigate additional complex indices combining meteorological variables (e.g. temperature and humidity for physiological comfort), and could

consider indices that combine meteorological/hydrological parameters with sector-based information including measures of vulnerability.

It is also acknowledged that updating indices is problematic for many regions and some regions would need specific indices to cope with their particular needs to provide climate services. As GFCS stresses the importance of the global, regional and local scales, ET-SCI acknowledges that support for this could come from Regional Climate Centers (RCCs) or Regional Climate Outlook Forums (RCOFs) etc. In addition, there are constraints on access to daily data. It is a considerable challenge to assemble worldwide datasets which are integrated, quality controlled, and openly and easily accessible. There is tension between traceability (access to the primary sources) and data completeness (use whatever available). Also a problem arises through the use of specified climatological periods which vary from group to group and which are used for base period calculations for percentile-based indices. In the first iteration of the software we use the base period of 1971-2000 but recognise that this will need to be amended for countries that do not have records covering this period. The software has been written in such a way that the user can specify the climatological base period which is most suitable for their data.

Users are invited to view ***ClimPACT2*** as 'living software' in that it can and will be amended as additional user needs arise.

2.7 Differences between *ClimPACT* and *ClimPACT2*

The main difference between ***ClimPACT* and *ClimPACT2*** is that in the latter the R package climdex.pcic is used as the base for the software rather than RClimDex. The advantages of this are that the software is faster and cleaner and also it is designed to work with model data in addition to point-based station measurements. Several bug fixes have been included and the software is now being updated within a version control environment. Apart from changes to some indices due to the bug fixes, the indices remain essentially unchanged.

The next section describes the development of the software and instructions on how to use it.

RUNNING ClimPACT2

3. DEVELOPMENT OF THE *ClimPACT2* SOFTWARE

ClimPACT2 is an R software package for calculating the 34 core ET-SCI indices, and additional non-core ET-SCI indices (**Appendix B**). It provides two methods for calculating these indices. The first is through a Graphical User Interface (GUI) designed to calculate the indices for time-series data stored in ASCII format for a specific location in space (e.g. a weather station). Through this GUI a limited number of parameters relating to the indices can be adjusted and various plots and quality control diagnostics are produced. The second method for calculating the ***ClimPACT2*** indices is through the climpact.loader function and is meant for advanced users familiar with the R programming language. The climpact.loader function has no GUI and does not produce plots or diagnostics. However, it offers significant flexibility in adjusting index parameters and is designed to calculate the ***ClimPACT2*** indices on gridded netCDF datasets. Thus, users who have three dimensional datasets (time x latitude x longitude) in netCDF format are able to have all of the indices included in this software package calculated with minimal effort.

ClimPACT2 should be run with version 3.0.2 of R or above. Users familiar with the RClimDEX[1] software will notice some similarities in the output of ***ClimPACT2*** GUI. This is because ***ClimPACT2*** was developed using the basic code from climdex.pcic[2] which was modelled after RClimDEX. This means that the same simple data quality control procedure is implemented (see **Section 2.4** for details on the importance of QC) along with a bootstrapping procedure to determine climatological percentile thresholds (see **Appendix E**). At present there is no facility to homogenise data, so it is assumed that the daily weather information provided by users will be of a high standard and free from artificial inconsistencies. The external software *RHtest* provides a free option for checking inhomogeneities that also works within the R programming language[1].

This section provides step-by-step instructions on running the ***ClimPACT2*** GUI. Section 4 presents a brief introduction to the climpact.loader function.

Note that Windows users are only able to use the GUI functionality of ClimPACT2. To calculate the ClimPACT2 indices on gridded netCDF files a Unix-based environment is required where access to additional R packages is available.

3.1 Requirements for running the *ClimPACT2* GUI

The only requirement to run the ***ClimPACT2*** GUI is that the R software package is installed (see **APPENDIX D** for details). Once R is installed, download the latest version of ***ClimPACT2*** from github (https://github.com/ARCCSS-extremes/climpact2/archive/master.zip) and extract the files to a new directory. Input files intended for ***ClimPACT2*** must be in the format described in **Appendix F** and a sample file is provided in the above .zip file.

[1] http://etccdi.pacificclimate.org/software.shtml

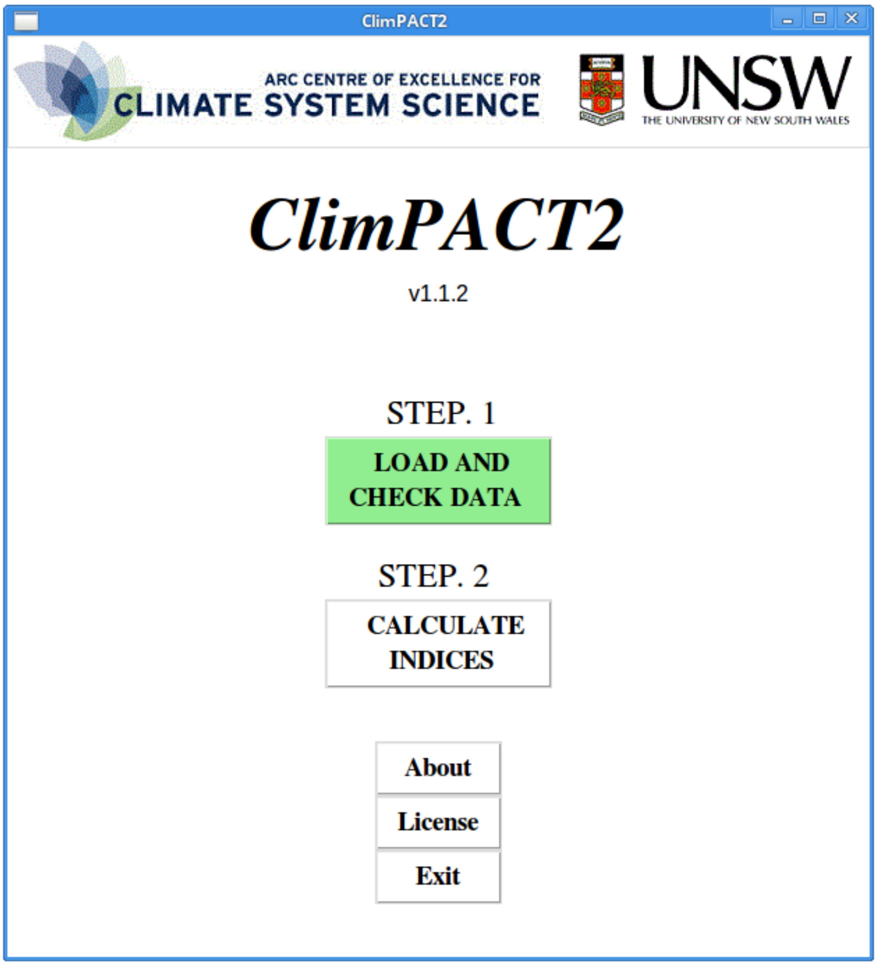
[2] https://pacificclimate.github.io/climdex.pcic/

3.2 Running the *ClimPACT2* GUI

In a **Microsoft Windows operating system**, select the 'R' icon that was created during the installation process. Once in R, from the drop down menu click on "File>Change dir…" and choose the folder where you have downloaded and extracted the ***ClimPACT2*** software (this is where the climpact2.GUI.r file resides). Then, within the R console prompt ">", type *source("climpact2.GUI.r")*.

In a **Unix-based operating system (e.g. MacOS, Debian, Ubuntu, Red Hat etc.)**, open a terminal window, navigate to the directory where you have downloaded the ***ClimPACT2*** software (this is where the climpact2.GUI.r file resides). Enter R (by typing *R* at the terminal prompt) and then type *source("climpact2.GUI.r")*.

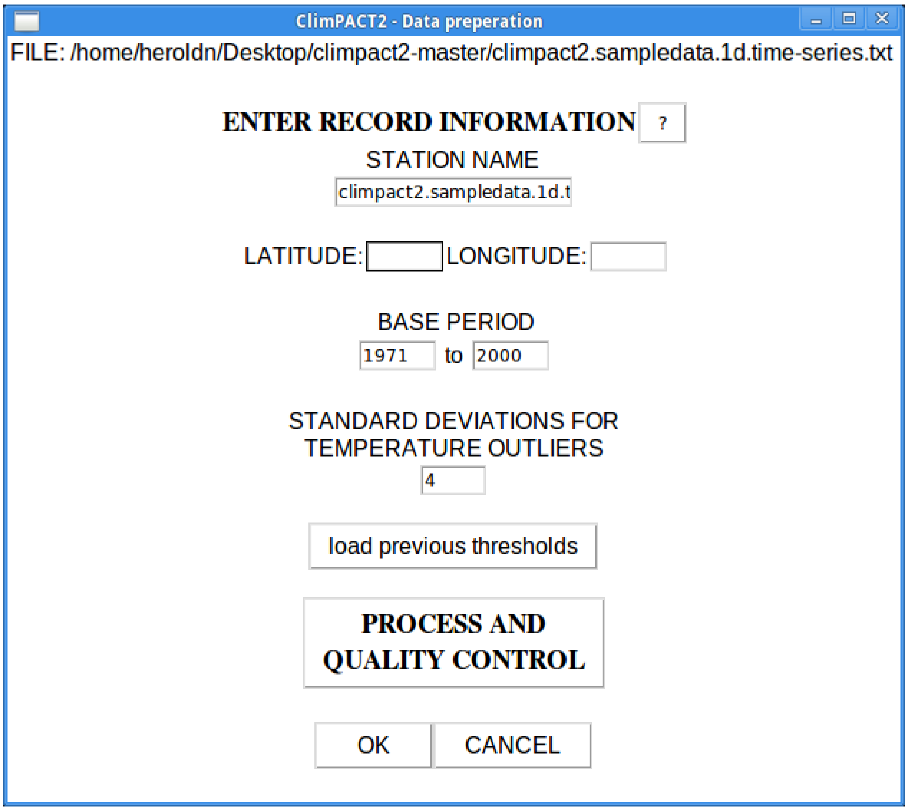
The first time *climpact2.GUI.r* is called, required R packages will be downloaded and installed. This may take some time but will only occur once. During this process you may be asked to select the geographical location of the closest "mirror" to download these packages from (see figure below). You may select any location, though the closest location will offer the fastest download speed.



3.3 Using the *ClimPACT2* GUI

Once *climpact2.GUI.r* has installed the required packages, the ***ClimPACT2*** GUI will open. The user will be presented with the ***ClimPACT2*** home screen shown above. Here, two main options are presented, with the labels "STEP. 1" and "STEP. 2", indicating the order in which the user should proceed to calculate the climate indices, using their data. The green highlighting of the "STEP. 1" button indicates which step the user currently needs to complete and thus which option they should select.

Selecting "STEP. 1" presents a prompt where the user can choose an ASCII file containing their climate data (refer to **Appendix F** for the required format of this file) and which is subsequently checked for formatting. The filename should be of the form "stationname.txt". For this manual the file climpact2.sampledata.1d.time-series.txt will be used as an example and the user is encouraged to use this sample file as a template for their own data. Once this file is selected a progress bar may briefly appear indicating progress in scanning for comma delimiters and replacing any with white space, checking that years are in the correct order, and substituting missing values of -99.9 with NA (the R nomenclature for a missing value). If any errors occur in reading the chosen file ***ClimPACT2*** will display the error message and the user must check their file for the correct formatting.



3.4 Load and check data

Once the chosen climate data file has been successfully read by ***ClimPACT2*** the above screen will appear, displaying the chosen file across the top (in this case climpact2.sampledata.1d.time-series.txt) and a series of input text boxes and buttons below. In this window metadata for the chosen ASCII file is input for the calculation of the indices. Selecting the '?' icon at the top of the screen will provide a summary of each input on this screen.

The first input text box allows the user to customise the **station name** of the record (the default being the filename). This should be informative and will be used to name files produced by ***ClimPACT2*** (these include output index .csv files, plots and diagnostic files).

Below this the user must specify the **latitude** and **longitude** of the station. This is required for some indices to approximate radiation balance for the site (latitude only). The valid latitude range is -90 to +90 and the valid longitude range is -180 to + 180.

The **base period** input text boxes refer to the years that the user wishes percentile indices to be measured against. For example, in a record from 1950 to 2010, the user may wish percentiles to be calculated for the years 1961 to 1990.

The last input text box allows the user to specify the number of '**standard deviations for temperature outliers'**, a value that ***ClimPACT2*** will use to identify outliers in the user's temperature data. Specifically, outliers are defined as the mean plus or minus *n* times the standard deviation of the value for the day, that is, [mean – *n**std, mean+*n**std]. Here std represents the standard deviation for the day, n is an input from the user and mean is the computed climatology for the day in question (input at next step).

Below the input text boxes is the button **'load previous thresholds'** which allows the user to **optionally** load climate thresholds that have been previously calculated by ***ClimPACT2***. Selecting this button will present a prompt allowing the user to select a *_thres.csv file that has been previously created. This option may be desired when a reference period is outside of the time period of the chosen data. For example, the climpact2.sampledata.1d.time-series.txt sample data contains values for 1931 to 2010, however, if thresholds were previously calculated for this station for the period 1911 to 1930, for example, then these thresholds can be loaded here. ***ClimPACT2*** automatically calculates thresholds for each dataset every time the 'PROCESS AND QUALITY CONTROL' button is selected. If a separate threshold file is desired it must be selected *before* selecting 'PROCESS AND QUALITY CONTROL' (see below). Lastly, when a threshold file is loaded, the specified base period is ignored, since no new thresholds will be calculated.

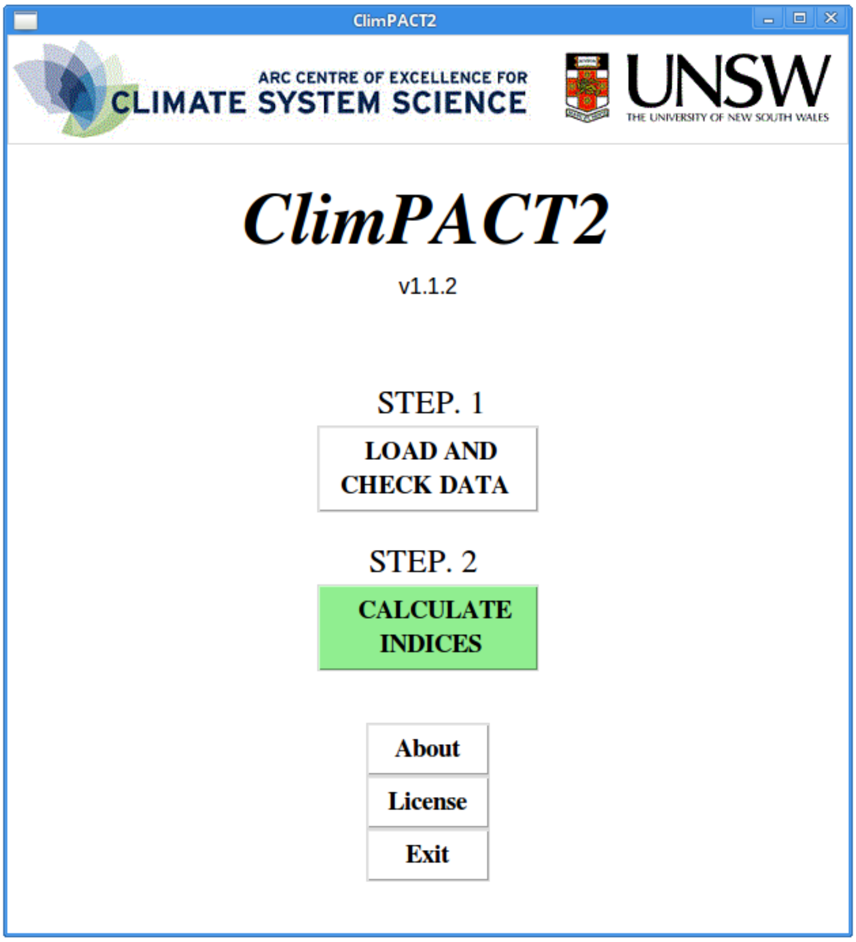
After all input text boxes have been filled out and previous thresholds have been loaded (if desired), select '**PROCESS AND QUALITY CONTROL'**. This step takes one to two minutes and a progress bar will appear. This step is mandatory and the user will not be allowed to select 'OK' at the bottom of the screen (and thus be prevented from proceeding to 'STEP. 2' of the ***ClimPACT2*** process) until this process has been completed. During this step ***ClimPACT2*** may stop if it detects errors in the data or the user's preferences. Specifically, ***ClimPACT2*** will stop if the latitude and longitude values are not valid, if the base period years are not valid or compatible with the data or if negative precipitation values are found. Upon completion, a message stating "QUALITY CONTROL COMPLETE" will be shown, along with a message asking the user to evaluate some quality control diagnostic files produced in the /qc subdirectory (this is located in the same directory as the station data file that the user selects, in this case climpact2.sampledata.1d.time-series.txt).

Quality control diagnostics: The files in the /qc subdirectory are very important to evaluate before calculating the climate indices. **Refer to Appendix G for details of the contents of this directory.**

Once the user has evaluated the contents of the /qc directory and the user's data has been processed successfully, the '**OK**' button at the bottom of the screen may be selected. At this point the user will be returned to the ***ClimPACT2*** home screen.

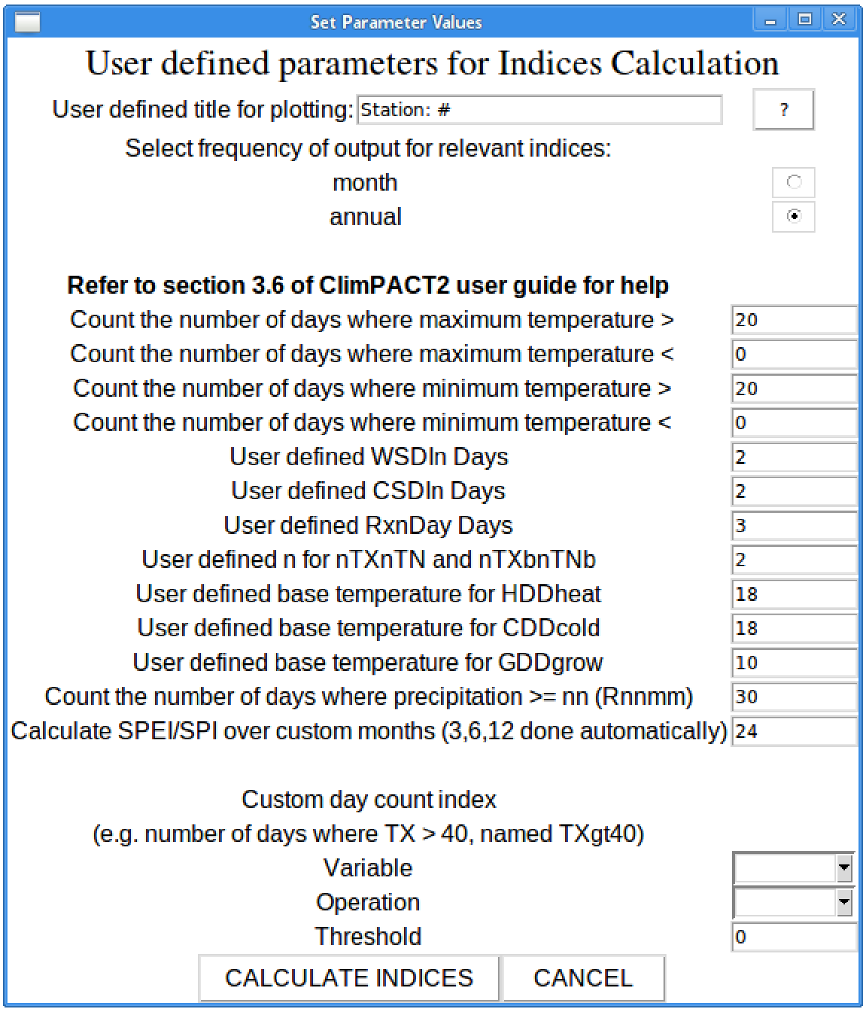
3.5 Calculating the indices

After STEP. 1 has been completed successfully, the “STEP. 2” button on the ***ClimPACT2*** home screen will be highlighted green to indicate that the user is now able to calculate the indices, as shown below. Select the “CALCULATE INDICES” button.



3.6 Parameter values for indices

This screen allows the user to set parameters relevant to some of the indices. Some indices can be calculated over months or years (e.g. the number of frost days 'FD' can be summed monthly or annually) and this can be specified with the radio buttons at the top of the screen.



The "Count the number of days where maximum temperature > n (sun)" allows the user calculate an index where the number of days with maximum temperatures greater than a set amount (n) is counted. This index will be called 'sun', where 'n' is the temperature set by the user.

The "Count the number of days where maximum temperature < n (idn)" allows the user calculate an index where the number of days with maximum temperatures less than a set amount (n) is counted. This index will be called 'idn', where 'n' is the temperature set by the user.

The "Count the number of days where minimum temperature > n (trn)" allows the user calculate an index where the number of days with minimum temperatures greater than a set amount (n) is counted. This index will be called 'trn', where 'n' is the temperature set by the user.

The "Count the number of days where minimum temperature < n (fdn)" allows the user calculate an index where the number of days with minimum temperatures less than a set amount (n) is counted. This index will be called 'fdn', where 'n' is the temperature set by the user.

The “User defined WSDIn Days” sets the number of days which need to occur consecutively with a TX > 90^{th} percentile to be counted in the WSDIn index.

The “User defined CSDIn Days” sets the number of days which need to occur consecutively with a TN < 10^{th} percentile to be counted in the CSDIn index.

The “User defined RxnDay Days” sets the monthly maximum consecutive *n*-day precipitation to be recorded by the Rxnday index.

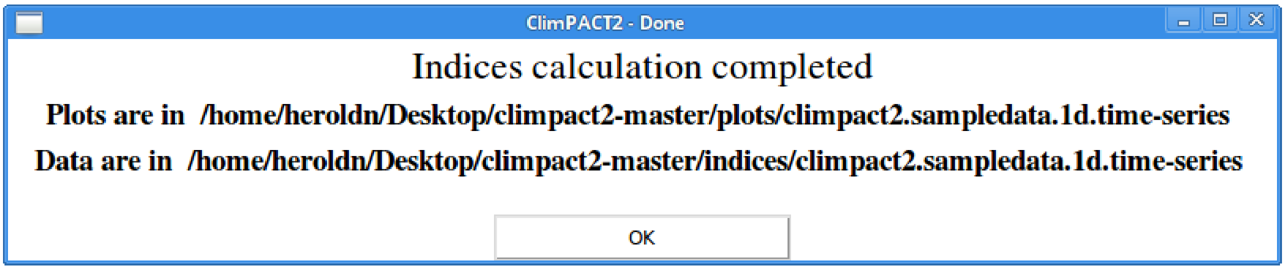
The “User defined n for nTXnTN and nTXbnTNb” sets the number of consecutive days required for the nTXnTN and nTXbnTNb indices.

The “User defined base temperature” for HDDheat, CDDcold and GDDgrow set the temperature to be used in the subtraction in these indices.

The “Count the number of days where precipitation >= nn (Rnnmm)” allows the user to calculate an index where the number of days with precipitation greater than or equal to a set amount is counted. This index will be called ‘rnnmm’, where ‘nn’ is the precipitation set by the user.

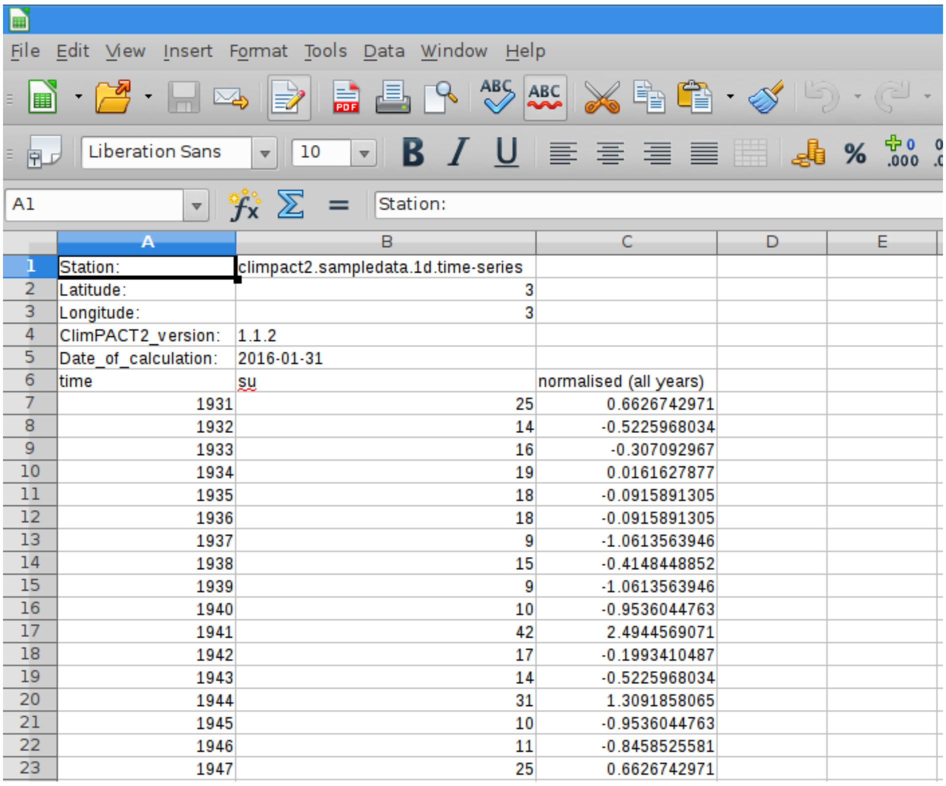
Lastly, the user has the option to create their own index based on the number of days crossing a specified threshold for daily maximum temperature (TX), minimum temperature (TN), diurnal temperature range (DTR) or precipitation (PR). To calculate a custom index, the user must select one of these variables, an operator (<,<=,>,>=) and a constant. For example, selecting TX, the ‘>=’ operator and specifying ‘40’ as a constant will produce an index that counts the number of days where TX is greater than or equal to 40°C. ***ClimPACT2*** output will refer to the index as TXge40. Operators are abbreviated in text with lt, le, gt and ge for <, <=, > and >=, respectively.

Once this step is completed, click “OK”. A progress bar will appear to indicate the time remaining. This should take less than a minute. A pop-up window will appear once the indices are computed.



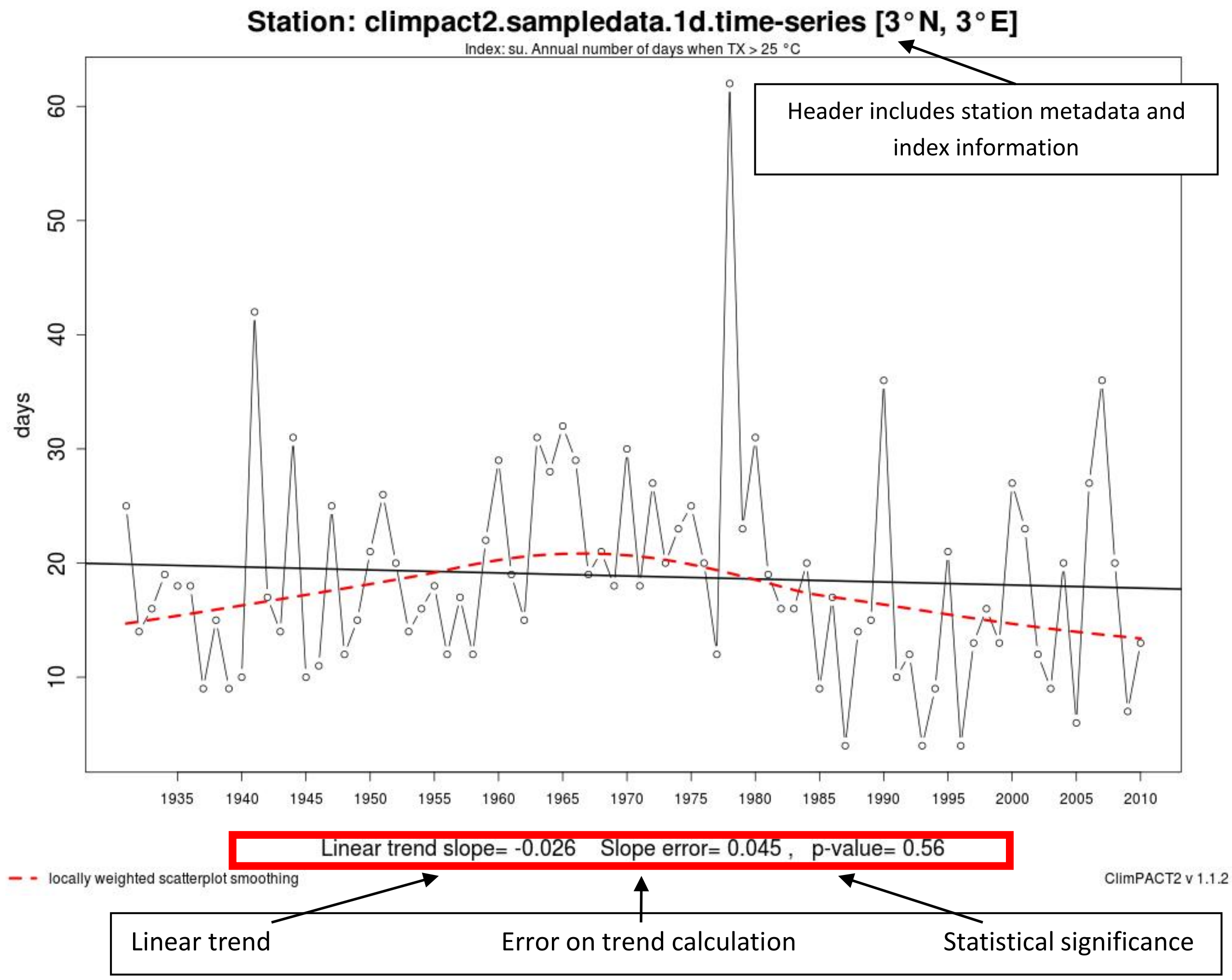
3.7 Examining *ClimPACT2* output

ClimPACT2 produces two files for each index calculated. One JPEG file (.jpg) containing a plot of the index, and one .csv file containing the index values. These are stored in the /plots and /indices subdirectories, respectively. The .csv files can be opened in Microsoft Excel, Open Office Calc or with a text editor. The index files have names "climpact2.sampledata.1d.time-series_XXX_YYY.csv" where XXX represents the name of the index and YYY is either ANN, MON or DAY depending on whether the index has been calculated annually or monthly. A sample .csv file for su is shown below. Prior to time-series information being printed, some basic information relating to the station, as well as the version of ***ClimPACT2*** being used, is written. There is one value for each year the index is calculated. Note that for indices calculated monthly there will be one value per month. A column containing normalised values is also written for most indices (these values are normalised using ALL available data).

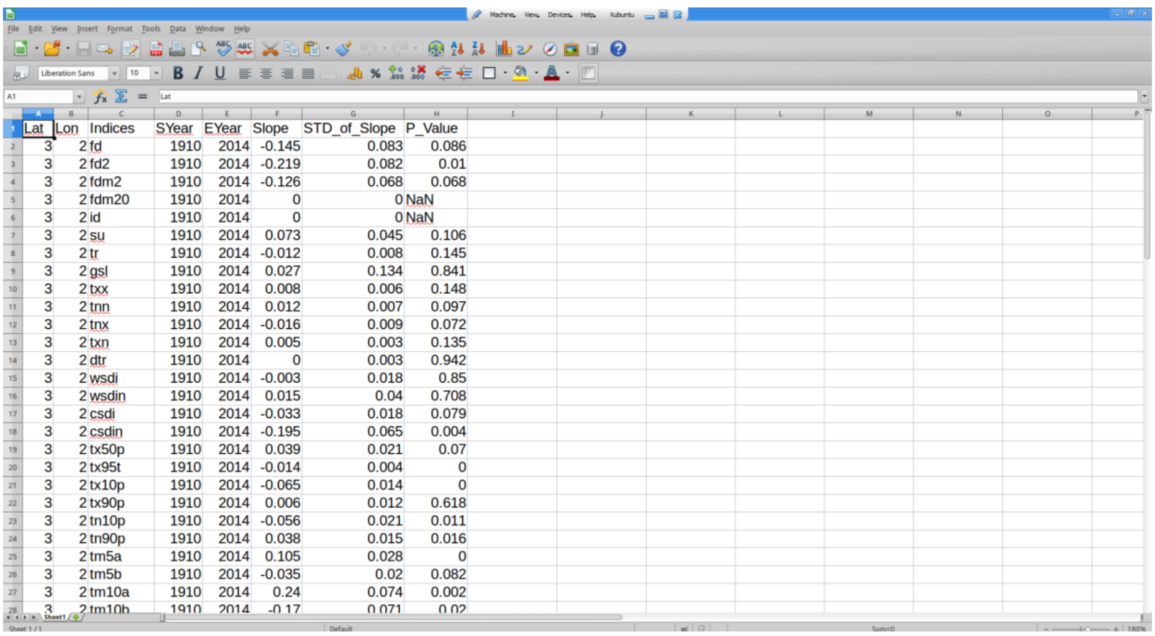


	A	B	C	D	E
1	Station:	climpact2.sampledata.1d.time-series			
2	Latitude:	3			
3	Longitude:	3			
4	ClimPACT2_version:	1.1.2			
5	Date_of_calculation:	2016-01-31			
6	time	su	normalised (all years)		
7	1931	25	0.6626742971		
8	1932	14	-0.5225968034		
9	1933	16	-0.307092967		
10	1934	19	0.0161627877		
11	1935	18	-0.0915891305		
12	1936	18	-0.0915891305		
13	1937	9	-1.0613563946		
14	1938	15	-0.4148448852		
15	1939	9	-1.0613563946		
16	1940	10	-0.9536044763		
17	1941	42	2.4944569071		
18	1942	17	-0.1993410487		
19	1943	14	-0.5225968034		
20	1944	31	1.3091858065		
21	1945	10	-0.9536044763		
22	1946	11	-0.8458525581		
23	1947	25	0.6626742971		

An example of a plot for the index su is shown below. The plot of each index is shown with a locally weighted linear regression (dashed line) to give an indication of longer-term variations. Statistics of the linear trend fitting are displayed on the plots (see below for an example of tm10b, the annual number of days when TM (daily mean temperature) is below 10°C). In addition, one .pdf file, *YYY_all_plots.pdf (climpact2.sampledata.1d.time-series_ANN_all_plots.pdf in our case), containing all plots is produced in the sub-directory /plots, where YYY is either ANN or MON depending on whether the 'month' or 'annual' radio button was selected in section 3.6.



Resulting trends for all indices are stored in the subdirectory /trend in a single .csv file. There is one file for all indices per station with the name "climpact2.sampledata.1d.time-series_trend.csv". Columns represent latitude, longitude, start year for trend calculation, end year for trend calculation, trend per year, standard error on trend calculation and the significance of the trend (< 0.05 indicates significance at the 5% level). See below for an example.



Lat	Lon	Indices	SYear	EYear	Slope	STD_of_Slope	P_Value
3	2	fd	1910	2014	-0.145	0.083	0.086
3	2	fd2	1910	2014	-0.219	0.082	0.01
3	2	fdm2	1910	2014	-0.126	0.068	0.068
3	2	fdm20	1910	2014	0	0	NaN
3	2	id	1910	2014	0	0	NaN
3	2	su	1910	2014	0.073	0.045	0.106
3	2	tr	1910	2014	-0.012	0.008	0.145
3	2	gsl	1910	2014	0.027	0.134	0.841
3	2	txx	1910	2014	0.008	0.006	0.148
3	2	tnn	1910	2014	0.012	0.007	0.097
3	2	tnx	1910	2014	-0.016	0.009	0.072
3	2	txn	1910	2014	0.005	0.003	0.135
3	2	dtr	1910	2014	0	0.003	0.942
3	2	wsdi	1910	2014	-0.003	0.018	0.85
3	2	wsdin	1910	2014	0.015	0.04	0.708
3	2	csdi	1910	2014	-0.033	0.018	0.079
3	2	csdin	1910	2014	-0.195	0.065	0.004
3	2	tx50p	1910	2014	0.039	0.021	0.07
3	2	tx95t	1910	2014	-0.014	0.004	0
3	2	tx10p	1910	2014	-0.065	0.014	0
3	2	tx90p	1910	2014	0.006	0.012	0.618
3	2	tn10p	1910	2014	-0.056	0.021	0.011
3	2	tn90p	1910	2014	0.038	0.015	0.016
3	2	tm5a	1910	2014	0.105	0.028	0
3	2	tm5b	1910	2014	-0.035	0.02	0.082
3	2	tm10a	1910	2014	0.24	0.074	0.002
3	2	tm10b	1910	2014	-0.17	0.071	0.02

4. CALCULATING THE *ClimPACT2* INDICES ON THREE DIMENSIONAL DATASETS

Users who have three-dimensional datasets (time x latitude x longitude) of daily temperature and precipitation may utilise the climpact.loader function contained in climpact2.r to calculate the ***ClimPACT2*** indices. This functionality is intended for users familiar with R. This section provides a brief overview of this functionality.

If you do not have the ***ClimPACT2*** code, you can download the latest version from github (https://github.com/ARCCSS-extremes/climpact2/archive/master.zip). The file climpact2.wrapper.r provides an example R script that calls the climpact.loader function in climpact2.r. Users are encouraged to modify this script to their needs (running this simply requires "Rscript climpact2.wrapper.r" from the linux command line). Prior to running ***ClimPACT2*** for the first time, be sure to run the climpact2.checker.r script, this script will check and install all required packages, including a modified version of the climdex.pcic R package.

climpact.loader takes three dimensional netCDF files in CF (Climate and Forecast) format as its input. The list of parameters that can be specified when calling climpact.loader are shown below and can be found in the climpact2.r source code.

- tsminfile: min temperature netCDF file. Must be gridded daily data.
- tsmaxfile: max temperature netCDF file. Must be gridded daily data.
- precfile: daily precipitation netCDF file. Must be gridded daily data.
- tsminname: name of min temperature variable in tsminfile
- tsmaxname: name of max temperature variable in tsmaxfile
- precname: name of daily precipitation variable in precfile
- timename: name of time variable in input files
- indices: a list specifying which indices to calculate. See index.master.list for a list of supported indices. Specify "all" to calculate all indices.
- identifier: an optional string to aid identifying output files (e.g. this may be the particular model/dataset the indices are being calculated on).
- lonname: name of longitude dimension.
- latname: name of latitude dimension.
- baserange: reference period for calculating percentile indices.
- freq: frequency at which to calculate relevant indices.
- write_quantiles: boolean specifying whether to write percentiles to a file for later use.
- quantile_file: netCDF file created from a previous execution of climpact.loader with write_quantiles set to TRUE.
- cores: specify the number of cores to use for processing. Default is to use one core.
- tempqtiles: temperature percentiles to calculate.
- precqtiles: precipitation percentiles to calculate.
- time_format: if the time variable in tsminfile, tsmaxfile and precfile are not in "units since YYYY-MM-DD" then this parameter must be passed specifying the date format. Uses standard notations as specified in http://stat.ethz.ch/R-manual/R-devel/library/base/html/strptime.html

- max.missing.days: maximum missing days under which indices will still be calculated.
- min.base.data.fraction.present: minimum fraction of data required for a quantile to be calculated for a particular day.
- output_dir: directory where index files will be created.
- hw_ehf: either “PA13” (default) or “BOM”. Specifies the method for calculating the excess heat factor (EHF). The default method (PA13) calculates this using the method set out in **Appendix C**, while specifying “BOM” uses the method endorsed by the Australian Bureau of Meteorology (Nairn and Fawcett, 2013).
- ... additional parameters: any parameters that are defined for specific indices in climdex.pcic can be specified by prefixing the index name followed by an underscore. For example, the spells.can.span.years parameter for the index wsdi can be specified by passing wsdi_spells.can.span.years.

A strength of ***ClimPACT2*** is that given it is built upon climdex.pcic functionality, any parameters that can be passed to index functions in climdex.pcic can be also passed to climpact.loader. This function can also be executed on multiple cores to increase performance. This is achieved by setting *cores* to some value greater than 1.

Important things to note:

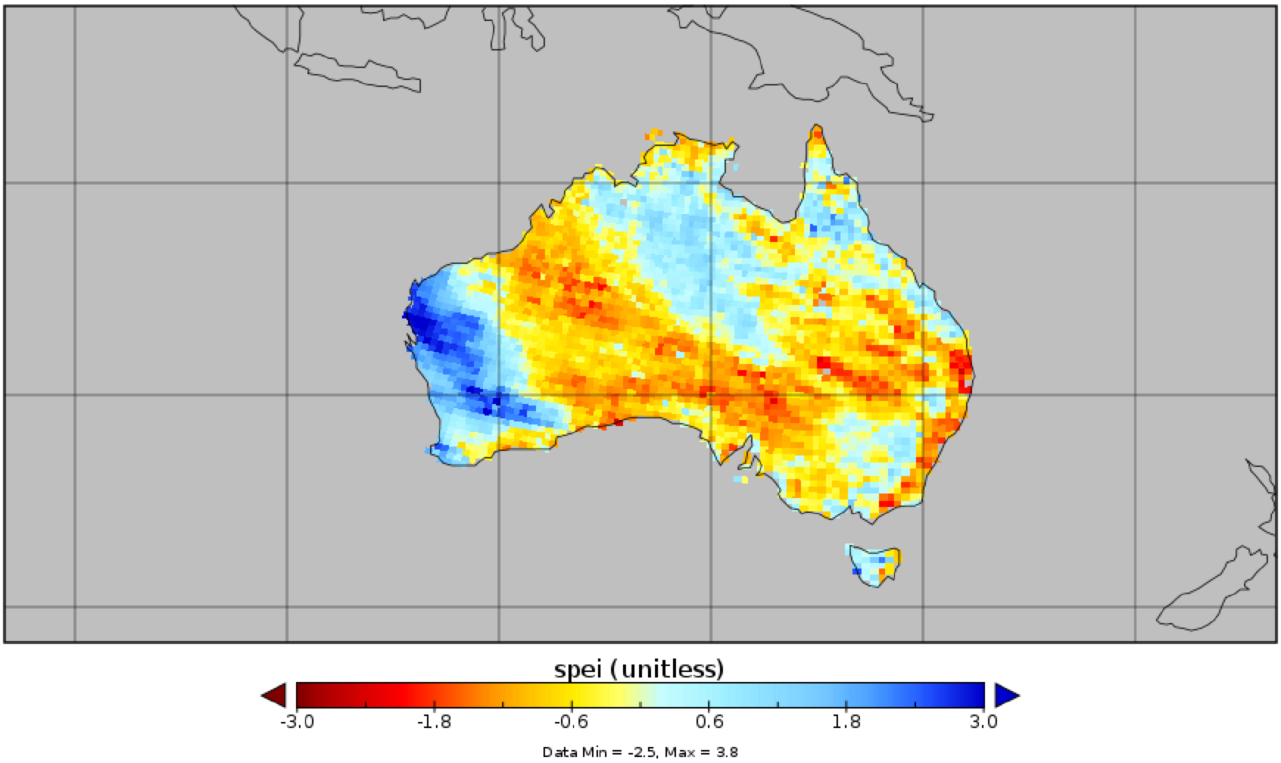
- Users should know the limits of their hardware. Currently, operating on two cores, calculating all indices on a 20 year, 144 x 215 grid takes approximately 12 hours.

- If the user wishes to record quantiles for later use they must set *write_quantiles* to “TRUE”. When doing this the user must ensure that *tempqtiles* and *precqtiles* have all of the percentiles that will be required for later calculations. In other words, the percentiles specified in *tempqtiles* and *precqtiles* are the percentiles that will be recorded when *write_quantiles* is set to “TRUE”. If not specified these parameters will be set to their default values (which can be found in climpact2.r). The percentiles specified in climpact2.wrapper.r are all that are required for calculating all indices in ***ClimPACT2***.

- ***ClimPACT2*** assumes the time dimension of the input data is of units "units since YYYY-MM-DD". If it is not, then the time_format parameter needs to be specified (see above list of climpact.loader parameters).

An example output of the Standardised Precipitation-Evapotranspiration Index (SPEI) calculated for a gridded dataset is shown on the next page. We recommend using Ncview (http://meteora.ucsd.edu/~pierce/ncview_home_page.html) or Panoply (http://www.giss.nasa.gov/tools/panoply/) for viewing netCDF output.

spei



APPENDICES

APPENDIX A: Goals and terms of reference of the ET-SCI

At the first meeting of the ET-SCI in Tarragona, Spain in July 2011, the following Terms of Reference (ToR) and deliverables were agreed as follows are:

- Develop methods and tools including standardized software for, and to generate, sector-specific climate indices, including their time series based on historical data, and methodologies to define simple and complex climate risks;

- Promote the use of sector-specific climate indices to bring out variability and trends in climate of particular interest to socio-economic sectors (e.g., droughts), with global consistency and to help characterize the susceptibility of various sectors to climate;

- Develop the training materials needed to raise capacity and promote uniform approaches around the world in applying these techniques;

- Work with sector-based agencies and experts, including those of relevant WMO Technical Commissions, particularly the Commission for Hydrology (CHy) and the Commission for Agricultural Meteorology (CAgM), to facilitate the use of climate information in users' decision-support systems for climate risk management and adaptation strategies;

- Submit reports in accordance with timetables established by the OPACE 4 co-chairs.

In addition various deliverables were proposed for consideration by the Team. These included:

- A collection and analysis of existing climate indices with particular specific sectoral (agriculture, water, health and Disaster Risk Reduction (DRR)) applications at national and regional scales;

- Technical publication on climate indices for sectoral application in risk assessment and adaptation;

- Methods and tools, standardized software and associated training materials required to produce sector-specific climate indices for systematic assessment of the impact of climate variability and change and to facilitate climate risk management and adaptation (to be done in collaboration with WMO Technical Commissions, particularly CCl OPACE-2 and with relevant agencies and organizations if required;

- Pilot training workshop (at least one region) on development of the indices;

- Workshop Report/Publication.

APPENDIX B: Tables of *ClimPACT2* indices

To calculate all indices time-series of daily minimum temperature (TN), maximum temperature (TX) and precipitation (PR) are required. Daily mean temperature (TM) is calculated from TM = (TX + TN)/2. Diurnal temperature range (DTR) is calculated from DTR = TX – TN.

Note four additional indices are available in the ***ClimPACT2*** GUI that are not available when using the climpact.loader function (see section 4). These indices are variants of pre-existing absolute threshold indices and are detailed in section 3.6. They allow the user to count the number of days where TX/TN is greater than or less than a specified value. The ***ClimPACT2*** GUI also allows users to create their own absolute threshold index as detailed in section 3.6.

TABLE B1: Core ET-SCI indices (AS AGREED JULY 2011).

Short name	**Long name**	**Definition**	**Units**
FD0	Frost days 0	Annual number of days when TN < 0 °C	days
FD2	Frost days 2	Annual number of days when TN < 2 °C	days
FDm2	Hard freeze	Annual number of days when TN < -2 °C	days
FDm20	Very Hard freeze	Annual number of days when TN < -20 °C	days
ID	Ice days	Annual number of days when TX < 0 °C	days
SU25	Summer days	Annual number of days when TX > 25 °C	days
TR	Tropical nights	Annual number of days when TN > 20 °C	days
GSL	Growing season Length	Annual number of days between the first occurrence of 6 consecutive days with TM > 5 °C and the first occurrence of 6 consecutive days with TM < 5 °C	days
TXx	Max TX	Warmest daily TX	°C
TNn	Min TN	Coldest daily TN	°C
WSDI	Warm spell duration indicator	Annual number of days with at least 6 consecutive days when TX > 90th percentile	days
WSDIn	User-defined WSDI	Annual number of days with at least n consecutive days when TX > 90th percentile	days
CSDI	Cold spell duration indicator	Annual number of days with at least 6 consecutive days when TN < 10th percentile	days
CSDIn	User-defined CSDI	Annual number of days with at least n consecutive days when TN < 10th percentile	days
TX50p	Above average Days	Percentage of days of days where TX > 50th percentile	%
TX95t	Very warm day threshold	Value of 95th percentile of TX	°C
TM5a	TM above 5°C	Annual number of days when TM >= 5 °C	days
TM5b	TM below 5°C	Annual number of days when TM < 5 °C	days
TM10a	TM above 10°C	Annual number of days when TM >= 10 °C	days

TM10b	TM below 10°C	Annual number of days when TM < 10 °C	days
SU30	Hot days	Annual number of days when TX >= 30 °C	days
SU35	Very hot days	Annual number of days when TX >= 35 °C	days
nTXnTN	User-defined consecutive number of hot days and nights	Annual count of n consecutive days where both TX > 95th percentile and TN > 95th percentile, where n >= 2 (and max of 10)	Number of events
HDDheat	Heating degree Days	Annual sum of Tb - TM (where Tb is a user-defined location-specific base temperature and TM < Tb)	°C
CDDcold	Cooling degree Days	Annual sum of TM - Tb (where Tb is a user-defined location-specific base temperature and TM > Tb)	°C
GDDgrow	Growing degree Days	Annual sum of TM - Tb (where Tb is a user-defined location-specific base temperature and TM > Tb)	°C
CDD	Consecutive dry days	Maximum annual number of consecutive dry days (when PR < 1.0 mm)	days
R20mm	Number of very heavy rain days	Annual number of days when PR >= 20 mm	days
PRCPTOT	Annual total wet-day PR	Annual sum of daily PR >= 1.0 mm	mm
R95pTOT	Contribution from very wet days	100*r95p / PRCPTOT	%
R99pTOT	Contribution from extremely wet days	100*r99p / PRCPTOT	%
RXnday	User-defined consecutive days PR amount	Maximum n-day PR total	mm
SPI	Standardised Precipitation Index	Measure of "drought" using the Standardised Precipitation Index on time scales of 3, 6 and 12 months. See McKee et al. (1993) and the WMO SPI User guide (World Meteorological Organization, 2012) for more details.	unitless
SPEI	Standardised Precipitation Evapotranspiration Index	Measure of "drought" using the Standardised Precipitation Evapotranspiration Index on time scales of 3, 6 and 12 months. See Vicente-Serrano et al. (2010) for more details.	unitless

TABLE B2: Non-core ET-SCI indices also calculated by *ClimPACT2*.

Short name	Long name	Definition	Units
nTXbnTNb	User-defined consecutive number of cold days and nights	Annual number of n consecutive days where both TX < 5th percentile and TN < 5th percentile where 10 >= n >=2	Number of events
TNx	Max TN	Warmest daily TN	°C
TXn	Min TX	Coldest daily TX	°C
DTR	Daily temperature range	Mean difference between daily TX and daily TN	°C
TMm	Mean TM	Mean daily mean temperature	°C
TXm	Mean TX	Mean daily maximum temperature	°C
TNm	Mean TN	Mean daily minimum temperature	°C
TX10p	Amount of cool days	Percentage of days when TX < 10th percentile	%
TX90p	Amount of hot days	Percentage of days when TX > 90th percentile	%
TN10p	Amount of cold nights	Percentage of days when TN < 10th percentile	%
TN90p	Amount of hot nights	Percentage of days when TN > 90th percentile	%
CWD	Consecutive wet days	Maximum annual number of consecutive wet days (when PR >= 1.0 mm)	days
R10mm	Number of heavy rain days	Annual number of days when PR >= 10 mm	days
Rnnmm	Number of customised rain days	Annual number of days when PR >= n	days
SDII	Daily PR intensity	Annual total PR divided by the number of wet days (when total PR >= 1.0 mm)	mm/day
R95p	Total annual PR from heavy rain days	Annual sum of daily PR > 95th percentile	mm
R99p	Total annual PR from very heavy rain days	Annual sum of daily PR > 99th percentile	mm
Rx1day	Max 1-day PR	Maximum 1-day PR total	mm
Rx5day	Max 5-day PR	Maximum 5-day PR total	mm
HWN (EHF/Tx90/Tn90)	Heatwave number (HWN) as defined by either the Excess Heat Factor (EHF), 90th percentile of TX or the 90th percentile of TN	The number of individual heatwaves that occur each summer (Nov – Mar in southern hemisphere and May – Sep in northern hemisphere). A heatwave is defined as 3 or more days where either the EHF is positive, TX > 90^{th} percentile of TX or where TN > 90^{th} percentile of TN. Where percentiles are calculated from base period specified by user in section 3.4. See Appendix C and Perkins and Alexander (2013) for more details.	events
HWF (EHF/Tx90/Tn90)	Heatwave frequency (HWF) as defined by either the	The number of days that contribute to heatwaves as identified by HWN.	days

	Excess Heat Factor (EHF), 90th percentile of TX or the 90th percentile of TN	See Appendix C and Perkins and Alexander (2013) for more details.	
HWD (EHF/Tx90/Tn90)	Heatwave duration (HWD) as defined by either the Excess Heat Factor (EHF), 90th percentile of TX or the 90th percentile of TN	The length of the longest heatwave identified by HWN. See Appendix C and Perkins and Alexander (2013) for more details.	days
HWM (EHF/Tx90/Tn90)	Heatwave magnitude (HWM) as defined by either the Excess Heat Factor (EHF), 90th percentile of TX or the 90th percentile of TN	The mean temperature of all heatwaves identified by HWN. See Appendix C and Perkins and Alexander (2013) for more details.	°C (°C^2 for ECF/EHF)
HWA (EHF/Tx90/Tn90)	Heatwave amplitude (HWA) as defined by either the Excess Heat Factor (EHF), 90th percentile of TX or the 90th percentile of TN	The peak daily value in the hottest heatwave (defined as the heatwave with highest HWM). See Appendix C and Perkins and Alexander (2013) for more details.	°C (°C^2 for ECF/EHF)
ECF_HWN	Heatwave number (HWN) as defined by the Excess Cold Factor (ECF).	The number of individual 'coldwaves' that occur each year. See Nairn and Fawcett (2013) for more information.	events
ECF_HWF	Heatwave frequency (HWF) as defined by the Excess Cold Factor (ECF).	The number of days that contribute to 'coldwaves' as identified by ECF_HWN. See Nairn and Fawcett (2013) for more information.	days
ECF_HWD	Heatwave duration (HWD) as defined by the Excess Cold Factor (ECF).	The length of the longest 'coldwave' identified by ECF_HWN. See Nairn and Fawcett (2013) for more information.	days
ECF_HWM	Heatwave magnitude (HWM) as defined by the Excess Cold Factor (ECF).	The mean temperature of all 'coldwaves' identified by ECF_HWN. See Nairn and Fawcett (2013) for more information.	°C^2
ECF_HWA	Heatwave amplitude (HWA) as defined by the Excess Cold Factor (ECF).	The minimum daily value in the coldest 'coldwave' (defined as the 'coldwave' with lowest ECF_HWM). See Nairn and Fawcett (2013) for more information.	°C^2

APPENDIX C: Heatwave calculation

The heatwave calculations used in ***ClimPACT2*** are based off Perkins and Alexander (2013), hereafter PA13, with some slight modifications to the EHF (Perkins personal comms 2015). See PA13 for background information.

According to the framework of PA13, three heatwave definitions are used in ***ClimPACT2***. Neither is more 'correct' than the other, and all are provided for the user to apply with the appropriate discretion. These definitions are based on the 90th percentile of TN (minimum daily temperature) designated Tn90, the 90th percentile of TX (maximum daily temperature) designated Tx90, and the EHF (Excess Heat Factor).

Under the above **three heatwave definitions (Tn90, Tx90 and EHF)** a heatwave is defined as any length of three or more days where;

- Tn90 definition: TN > 90th percentile of TN.
- Tx90 definition: TX > 90th percentile of TX.
- EHF definition: the EHF is positive.

The percentiles for Tn90 and Tx90 are calculated within a user-specified base period, over the calendar year and using a 15 day running window. Thus there is a unique percentile for each calendar day.

The EHF is a combination of two *excess heat indices (EHI)*;

$EHI(accl.) = [(TM_i + TM_{i-1} + TM_{i-2})/3] – [(TM_{i-3} + … + TM_{i-32})/30]$

$EHI(sig.) = [(TM_i + TM_{i-1} + TM_{i-2})/3] – TM_{90}$

Where TM_i represents the average daily temperature for day i and TM_{90} is the 90th percentile of TM which is also calculated within a user-specified base period, over the calendar year and using a 15 day running window. TM is calculated via TM = (TX + TN)/2.

The difference between the above definitions and those in PA13 lie in the EHF. In PA13 the EHF was defined as in Nairn and Fawcett (2013), using the climatological 95th percentile of TM over the base period (i.e. one percentile for the calendar year, not a unique percentile for each day).

As stated in section 4, when running ***ClimPACT2*** on a netCDF dataset, the alternate version of the EHF (Nairn and Dawcett, 2013) can be specified.

APPENDIX D: Installation and running of R

R is a language and environment for statistical computing and graphics. It is a GNU implementation of the S language developed by John Chambers and colleagues at Bell Laboratories (formerly AT&T, now Lucent Technologies). S-plus provides a commercial implementation of the S language.

F.1 How to install R

ClimPACT2 requires the base package of R. The installation of R involves a very simple procedure. 1) Connect to the R project website at http://www.r-project.org, 2) Follow the links to download the most recent version of R for your computer operating system from any mirror site of CRAN.

For Microsoft Windows, download the Windows setup program. Run that program and R will be automatically installed in your computer, with a short cut to R on your desktop.

For Linux, download proper precompiled binaries and follow the instruction to install R. For other unix systems, you many need to download source code and compile it yourself.

F.2 How to run R

Under the Windows environment, double click the R icon on your desktop, or launch it through Windows "start" menu. This usually gets you into the R user interface. For some computers, you may need to first setup an environment variable called "HOME". See R for Windows FAQ for details if you have any problems.

Under a unix and mac environment, simply type *R* at the command line and you will be brought to the R console.

Exit from R by entering *q()* in the R console under both Windows and unix. Under Windows, you may also click "File" menu and then "Exit".

APPENDIX E: Threshold estimation and base period temperature indices calculation

Empirical quantile estimation:

The quantile of a distribution is defined as

$Q(p) = F^{-1}(p) = \inf\{x : F(x) \geq p\}$, 1<*p*<1,

where *F(x)* is the distribution function. Let $\{X_{(a)},..., X_{(n)}\}$ denote the order statistics of $\{X_1,..., X_n\}$ (i.e. sorted values of {*X*}), and let $\hat{Q}_i(p)$ denote the *i*th sample quantile definition. The sample quantiles can be generally written as:

$\hat{Q}_i(p) = (1-\gamma)X_{(j)} + \gamma X_{(j+1)}$.

Hyndman and Fan (1996) suggest a formula to obtain medium un-biased estimate of the quantile by letting $j = \mathrm{int}(p * n + (1+p)/3))$ and letting $\gamma = p * n + (1+p)/3 - j$, where int(*u)* is the largest integer not greater than *u*. The empirical quantile is set to the smallest or largest value in the sample when *j<1* or *j> n* respectively. That is, quantile estimates corresponding to *p<1/(n+1)* are set to the smallest value in the sample, and those corresponding to *p>n/(n+1)* are set to the largest value in the sample.

Bootstrap procedure for the estimation of exceedance rate for the base period:

It is not possible to make an exact estimate of the thresholds due to sampling uncertainty. To provide temporally consistent estimate of exceedance rate throughout the base period and out-of-base period, we adapt the following procedure (Zhang et al. 2005) to estimate exceedance rate for the base period.

a) The 30-year base period is divided into one “out-of-base” year, the year for which exceedance is to be estimated, and a “base-period” consisting the remaining of 29 years from which the thresholds would be estimated.
b) A 30-year block of data is constructed by using the 29 year “base-period” data set and adding an additional year of data from the “base-period" (i.e., one of the years in the “base-period” is repeated). This constructed 30-year block is used to estimate thresholds.
c) The “out-of-base” year is then compared with these thresholds and the exceedance rate for the “out-of-base” year is obtained.
d) Steps (b) and (c) are repeated for an additional 28 times, by repeating each of the remaining 28 in-base years in turn to construct the 30-year block.
e) The final index for the “out-of-base” year is obtained by averaging the 29 estimates obtained from steps (b), (c) and (d).

APPENDIX F: Input Data Format for *ClimPACT2*

The input data file has several requirements which are listed below. We recommend that users use the sample input file provided with ***ClimPACT2*** as a template for their own data.

1. ASCII text file
2. Columns as following sequences: Year, Month, Day, P, TX, TN (NOTE: P units = millimeters and Temperature units= degrees Celsius)
3. The format as described above must be space delimited (e.g. each element separated by one or more spaces).
4. For data records, missing data **must** be coded as -99.9; data records **must** be in calendar date order. Missing dates allowed.

Example data Format for the initial data file (e.g. used in the 'Quality Control' step):

```
1901    1    1    -99.9   -3.1   -6.8

1901    1    2    -99.9   -1.3   -3.6

1901    1    3    -99.9   -0.5   -7.9

1901    1    4    -99.9   -1     -9.1

1901    1    7    -99.9   -1.8   -8.4
```

APPENDIX G: Quality control diagnostics

The text in this appendix is adapted from text written by Enric Aguilar and Marc Prohom for the R functions they created to perform quality control, which have been integrated into the ***ClimPACT2*** *software.*

Once the user selects '**PROCESS AND QUALITY CONTROL**' ***ClimPACT2*** will take a minute or two to calculate thresholds and perform quality control checks. At the end of this process a dialogue box will appear telling the user to check the /qc subdirectory created in the directory where their climate information is stored.

The /qc folder contains the following files (where "mystation" refers to the name of the user's station file):

7 .pdf files, with graphical information on data quality:

- mystation_tminPLOT.pdf
- mystation_tmaxPLOT.pdf
- mystation_dtrPLOT.pdf
- mystation_prcpPLOT.pdf
- mystation_boxes.pdf
- mystation_boxseries.pdf
- mystation_rounding.pdf

10 .csv files with numerical information on data quality

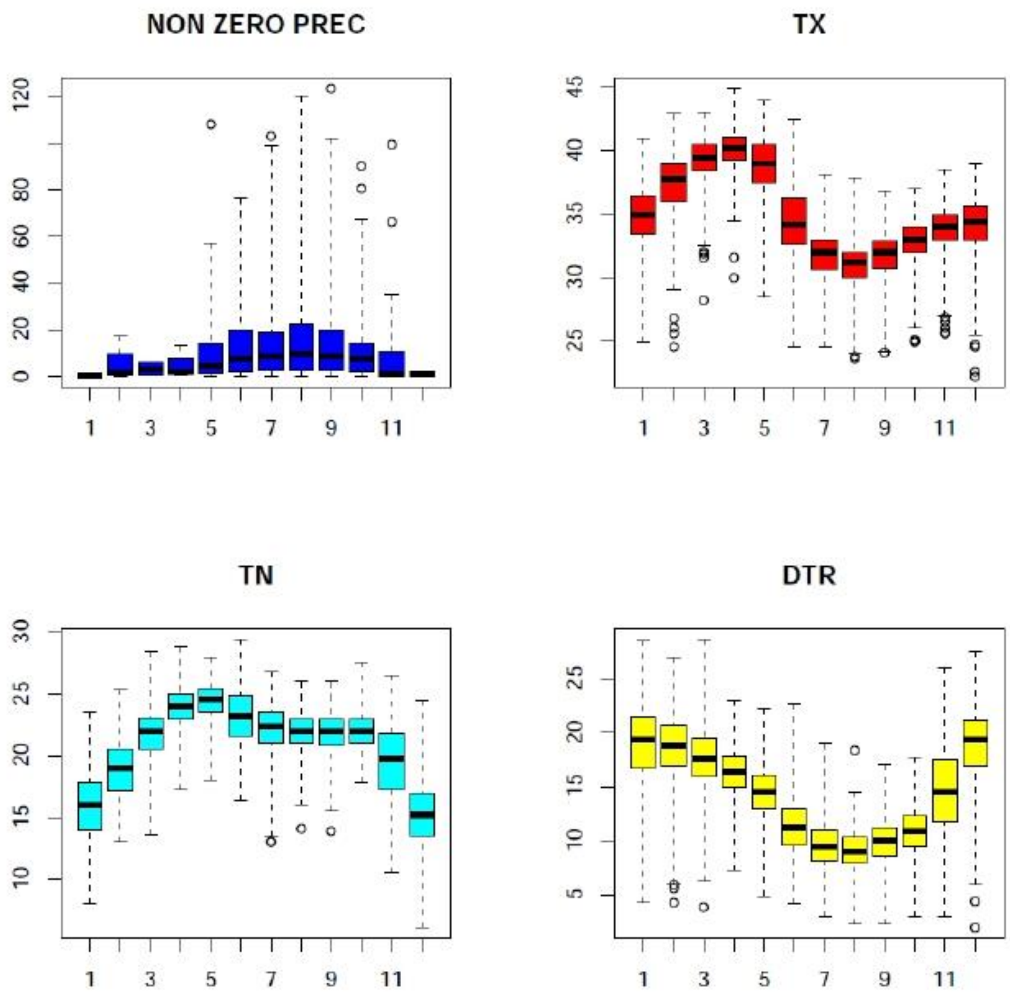
- mystation_duplicates.csv
- mystation_outliers.csv
- mystation_tmaxmin.csv
- mystation_tx_flatline.csv
- mystation_tn_flatline.csv
- mystation_toolarge.csv
- mystation_tx_jumps.csv
- mystation_tn_jumps.csv
- mystation_temp_stddev_QC.csv
- mystation_temp_nastatistics.csv

mystation_tminPLOT.pdf
mystation_tmaxPLOT.pdf
mystation_dtrPLOT.pdf
mystation_prcpPLOT.pdf

These files contain simple plots of the daily time-series of minimum temperature, maximum temperature, diurnal temperature range and precipitation, respectively. This allows the user to view the data and identify obvious problems by eye such as missing data (indicated by red circles) or unrealistic values.

mystation_boxes.pdf

This file identifies potential outliers based on the interquartilic (IQR). The IQR is defined as the difference between the 75th (p75) and the 25th (p25) percentiles. As can be seen in the example below, the mystation_boxes.pdf file contains boxplots of temperature and precipitation data flagging as outliers (round circles) all those temperature values falling outside a range defined by p25 – 3 interquartilic ranges (lower bound) and p75 + 3 interquartilic ranges (upper bound). For precipitation, 5 IQR are used.



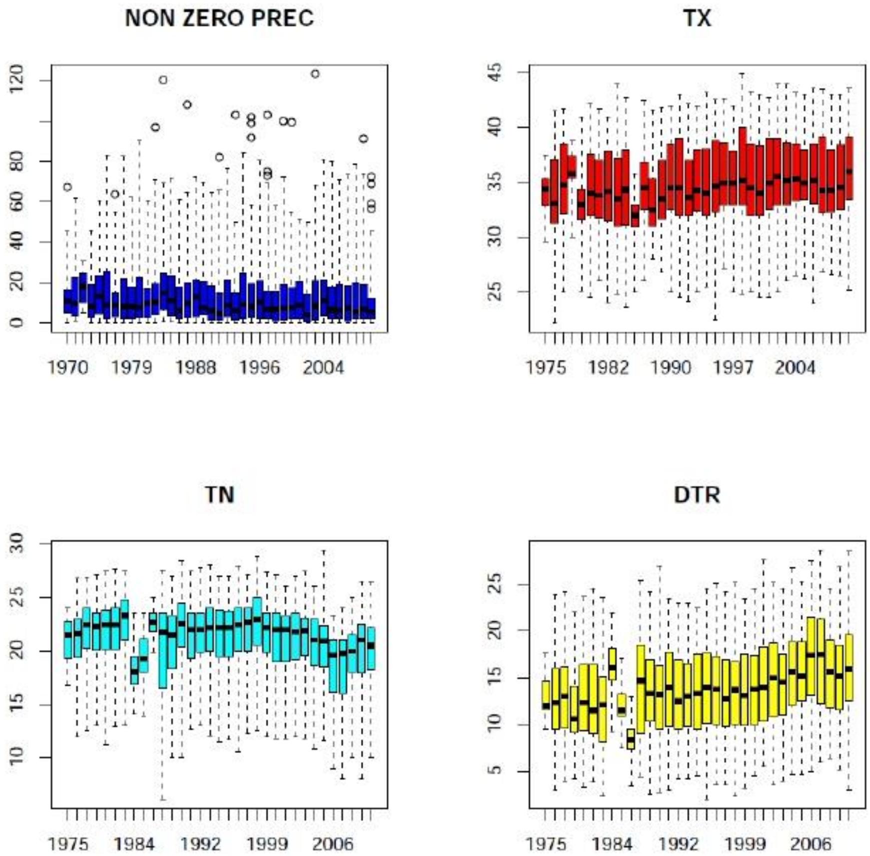
The values identified by this graphical quality control, are sent to a .csv file, **mystation_outliers.csv**. This file lists the outliers grouped under the variable that produced the inclusion of the record in the file and specifying the margin (upper bound or lower bound) that is surpassed. So, under ‘Prec up’ appear those values that represent a precipitation outlier; under ‘TX up’ are those that represent a maximum temperature higher than p75+3*IQR; under ‘TX low’ are outliers that represent an observation lower than p25-3*IQR. The explanation given for TX, also applies to TN and DTR. The advantage of this approach is that the detection of this percentile based outliers is not affected by the presence of larger outliers, so ONE RUN OF THE PROCESS IS ENOUGH!

Date	Prec	TX	TN	DTR
Prec up				
2/01/1951	31.8	14.3	10.2	4.1
12/01/1961	47.5	23.4	11.4	12
5/04/1963	42.8	19.2	13.6	5.6
18/04/1967	29.1	20.2	11.8	8.4

	19/04/1969	28.2	27.7	17.9	9.8
	19/04/1973	53.6	14.8	11.1	3.7
	21/11/1991	55.9	11.4	7.8	3.6
	11/11/1995	32.1	18.4	13.5	4.9
	1/12/2000	31.6	18.6	12.6	6
	31/12/2001	32.1	16	9.4	6.6
	15/12/2005	30.2	22.1	13.3	8.8
TX up					
TX low					
TN up					
TN low					
	30/10/1972	2.5	-11.2	-23.4	12.2
	31/10/1972	4.3	-4.8	-24.8	20
DTR up					
DTR low					

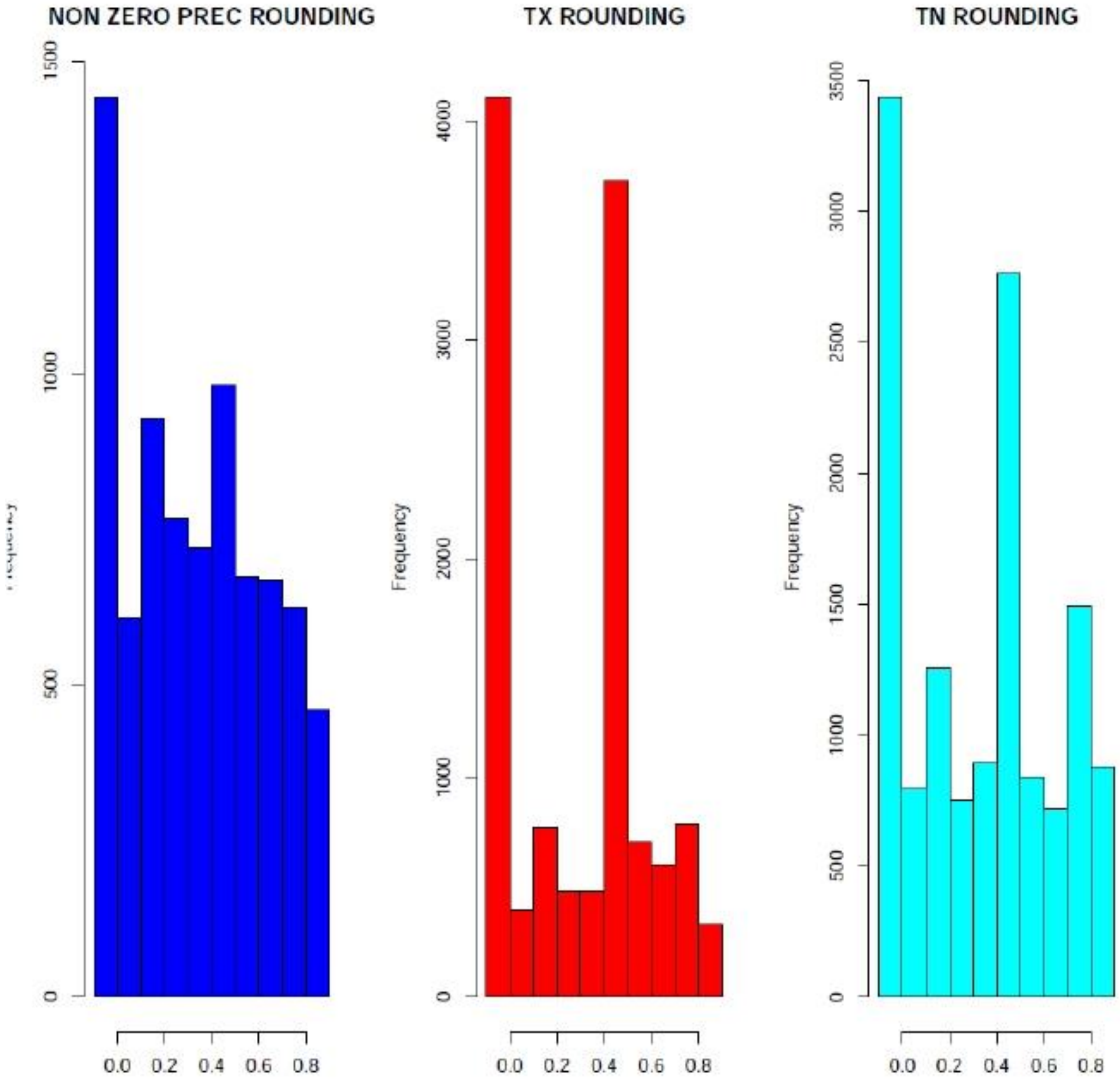
mystation_boxseries.pdf

The graphic file boxseries.pdf (which does not have a numerical counterpart) produces annual boxplots. This file is useful to have a panoramic view of the series and be alerted of parts of the series which can be problematic (see values around 1984 in the example figure below).



mystation_rounding.pdf

This file looks at rounding problems by plotting the frequency of values after each decimal point. It shows how frequently each of the 10 possible values (.0 to .9) appears. It is expected that .0 and .5 will be more frequent (although there is no statistical reason for this!).



mystation_tn_flatline.csv
mystation_tx_flatline.csv

The mystation_tn_flatline.csv and mystation_tx_flatline.csv files report occurrences of 4 or more equal consecutive values in, respectively, TX and TN. A line for each sequence of 4 or more consecutive equal values is generated. In the example below all sequences are 4 values long (i.e. each corresponding value has been repeated 3 extra times). The date specified belongs to the end of the sequence.

Date	TX	Number of duplicates
4/09/1937	18	3
28/11/1937	16.9	3

Looking at the data, the first sequence identified by the QC test is shown below.

1937	9	1	0	16.4	11.6
1937	9	2	0	18	10.2
1937	9	3	0	18	8.6
1937	9	4	0	18	7

mystation_duplicates.csv

The file mystation_duplicates.csv includes all dates which appear more than once in a datafile. In the listing below, one can see that 1958/08/26 occurs twice, and thus will be reported in mystation_duplicates.csv.

1951 8 24
1951 8 25
1951 8 26
1951 8 26
1951 8 28
1951 8 29
1951 8 30
1951 8 31

mystation_toolarge.csv

The file mystation_toolarge.csv reports precipitation values exceeding 200 mm (this and any other threshold can be easily reconfigured before execution) and temperature values exceeding 50 ºC.

mystation_tx_jumps.csv
mystation_tn_jumps.csv

The files mystation_tx_jumps.csv and mystation_tn_jumps.csv will list those records where the temperature difference with the previous day is greater or equal than 20 ºC.

mystation_tmaxmin.csv

The mystation_tmaxmin.csv file, records all those dates where maximum temperature is lower than minimum temperature.

mystation_temp_stddev_QC.csv

The mystation_temp_stddev_QC.csv file contains dates where TX, TN or DTR are more than n standard deviations away from their respective means, where n is a user-specified value entered in the ***ClimPACT2*** GUI (see section 3.4). Given that successive corrections to spurious outlying values will alter a stations standard deviation, this process may need to be repeated several times.

mystation_temp_nastatistics.csv

This file lists the number of missing values that exists for each variable (TX, TN, PR) for each year.

APPENDIX H: FAQs

1. What should be the length of the baseline period?

The answer to this question depends on your application and the length of data you have. At the moment the default period is 1971 – 2000 but this can easily be amended within the software (see Section 3.1) to shorter or longer periods.

2. How are missing data handled by *ClimPACT2*?

Missing data need to be stored as -99.9 in the input data files (see APPENDIX F) but are converted to an internal format that R recognises (NA, not available).

3. How can *ClimPACT2* results be analysed further or used to produce customised graphics using other popular packages?

ClimPACT2 produces its own plots of each index (in the “plots” folder) once the software has completed running (see Section 3.1). However, all of the indices output data are stored in the “indices” directory in .csv format. Many graphics packages are able to handle this file format so you can produce your own customised packages easily with your favourite software package.

4. Can I add additional indices to *ClimPACT2* myself?

At the moment there is no easy method to do this but if you are familiar with the R programming language you can amend the code to add additional indices if you require. This is a good solution if you have very specific sector requirements that are not covered by the current suite of indices.

5. Can I recommend additional indices to be added to the ET-SCI core set?

Yes, but any indices added to the core set have to be agreed by the members of the ET-SCI.

APPENDIX I: Software licence agreement

ClimPACT2 Software Licence

All source code developed by this project is provided under the following licence:

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

* Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

* Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

* The WMO, the Commission for Climatology (CCl), the CCl Expert Team on Climate Risk and Sector Specific Climate Indices (CCl ET-SCI), or the names of any contributors to the ***ClimPACT2*** software and the related technical manual may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDER AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

In addition to the licence, if you redistribute or create derived works from this software, it is requested that you acknowledge the WMO and the ET-SCI and we would be very grateful if you would notify us of any publications resulting from the use of this software.

REFERENCES

Aguilar E, Barry A, Brunet M, Ekang L, Fernandes A, Massoukina M, Mbah J, Mhanda A, Nascimento D, Peterson TC, et al. Changes in temperature and precipitation extremes in western Central Africa, Guinea Conakry and Zimbabwe, 1955–2006. 2009. J Geophys Res, 114:D02115. doi:10.1029/2008JD011010.

Aguilar E, et al. Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. 2006. J Geophy, 110:D23107. doi:10.1029/2005JD006119

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Kumar Kolli R, Revadekar JV, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez Aguirre JL. 2006. Global observed changes in daily climate extremes of temperature and precipitation. Journal of Geophysical Research-Atmospheres 111: D05109, doi:10.1029/2005JD006290

Caesar J, Alexander LV, Trewin B, Tse-ring K, Sorany L, Vuniyayawa V, Keosavang N, Shimana A, Htay MM, Karmacharya J, et al. Changes in temperature and precipitation extremes over the Indo-Pacific region from 1971 to 2005. Int J Climatol 2011, 31:791–801. doi:10.1002/joc.2118.

Haylock MR, et al. Trends in total and extreme South American rainfall 1960–2000 and links with sea surface temperature. J Climate 2006, 19:1490–1512.

Heim, R.R., (2002): A review of twentieth-century drought indices used in the United States. Bulletin of the American Meteorological Society, 83, 1149-1165. Hyndman, R.J., and Y. Fan, 1996: Sample quantiles in statistical packages. The American Statistician, 50, 361-367.

Klein Tank AMG, et al. Changes in daily temperature and precipitation extremes in central and south Asia. J Geophys Res 2006. doi:10.1029/2005JD006316.

McKee, T.B.N., J. Doesken, and J. Kleist, (1993): The relationship of drought frecuency and duration to time scales. Eight Conf. On Applied Climatology. Anaheim, CA, Amer. Meteor. Soc. 179-184.

Mishra, A. K., and V. P. Singh (2010), A review of drought concepts, J. Hydrol., 391, 202–216.

Nairn J R and Fawcett R G (2013), Defining heatwaves: heatwave defined as a heat-impact event servicing all community and business sectors in Australia (Centre for Australian Weather and Climate Research), CAWCR Technical Report No. 060.

New M, et al. Evidence of trends in daily climate extremes over southern and west Africa. J Geophys Res 2006, 111:D14102. doi:10.1029/2005JD006289.

Perkins SE, Alexander LV. 2013. On the measurement of heat waves. Journal of Climate, in press

Perkins SE, Alexander LV. Nairn J. 2012. Increasing frequency, intensity and duration of observed global heat waves and warm spells. Geophysical Research Letters, 39, 20, doi:10.1029/2012GL053361

Peterson TC, et al. Recent changes in climate extremes in the Caribbean region. J Geophys Res 2002, 107:D214601. doi:10.1029/2002JD002251.

Peterson TC, Manton MJ. Monitoring changes in climate extremes: a tale of international collaboration. Bull Am Meteorol Soc 2008, 89:1266–1271.

Sivakumar, M. V. K., R. P. Motha, D. A. Wilhite, and D. A. Wood (Eds.) (2010): Agricultural Drought Indices: Proceedings of an Expert Meeting, 2–4 June 2010, Murcia, Spain, 219 pp., World Meteorol. Org., Geneva, Switzerland.

Vicente-Serrano S.M., Santiago Beguería, Juan I. López-Moreno, (2010) A Multi-scalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index – SPEI. Journal of Climate 23: 1696-1718.Vincent LA, et al. Observed trends in indices of daily temperature extremes in South America 1960–2000. J Climate 2005, 18:5011–5023.

Vincent LA, Anguilar E, Saindou M, Hassane AF, Jumaux G, Roy D, Booneeady P, Virasami R, Randriamarolaza LYA, Faniriantsoa FR, et al. Observed trends in indices of daily and extreme temperature and precipitation for the conutries of the western Indian Ocean, 1961–2008. J Geophys Res 2011, in press.

Wang, X. L. L., 2008: Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. J. App. Met. Climatol., 47, 2423-2444.

World Meteorological Organization (2012): Standardized Precipitation Index User Guide (M. Svoboda, M. Hayes and D. Wood). (WMO-No. 1090), Geneva.

Zhang X, Alexander L, Hegerl GC, Jones P, Klein Tank A, Peterson TC, Trewin, B, Zwiers FW. 2011. Indices for monitoring changes in extremes based on daily temperature and precipitation data. WIREs Climate Change doi:10.1002/wcc.147.

Zhang X, G Hegerl, FW Zwiers and J Kenyon 2005. Avoiding inhomogeneity in percentile-based indices of temperature extremes. J. Climate, 18, 1641-1651

Zhang X, Aguilar E, Sensoy S, Melkonyan H, Tagiyeva U, Ahmed N, Kutaladze N, Rahimzadeh F, Taghipour A, Hantosh TH, et al. Trends in Middle East climate extremes indices during 1930–2003. J Geophys Res 2005, 110:D22104.

Meeting reports

Final report of the Meeting of the Commission for Climatology (CCl) Expert Team on Climate Risk and Sector-Specific Climate Indices (ET-CRSCI) (Tarragona, Spain 13-15 July 2011):

http://www.wmo.int/pages/prog/wcp/ccl/opace/opace4/meetings/documents/ET_CRSCI_FinalReport_Tarragona.pdf

Final report of the High Level Task Force on the Global Framework for Climate Services:
http://www.wmo.int/hlt-gfcs/downloads/HLT_book_full.pdf

The Abridged final report with resolutions of the Sixteenth World Meteorological Congress (WMO-No. 1077):
ftp://ftp.wmo.int/Documents/PublicWeb/mainweb/meetings/cbodies/governance/congress_reports/english/pdf/1077_en.pdf

The Abridged final report with resolutions and recommendations of the Fifteenth session of the WMO Commission for Climatology (WMO-No. 1054):
ftp://ftp.wmo.int/Documents/PublicWeb/mainweb/meetings/cbodies/governance/tc_reports/english/pdf/1054_en.pdf