Weather and Climate Extremes 9 (2015) 6–16

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

Systematic investigation of gridding-related scaling effects on annual statistics of daily temperature and precipitation maxima: A case study for south-east Australia

Francia B. Avila [a,*], Siyan Dong [b], Kaah P. Menang [c], Jan Rajczak [d], Madeleine Renom [e], Markus G. Donat [f,g], Lisa V. Alexander [f,g]

[a] *Manila Observatory, Philippines*
[b] *National Climate Center, China Meteorological Administration, China*
[c] *National Department of Meteorology, Ministry of Transport, Yaoundé, Cameroon*
[d] *Institute for Atmospheric and Climate Science, ETH Zurich, Zürich, Switzerland*
[e] *Atmospheric Science Department, University of the Republic, Montevideo, Uruguay*
[f] *Climate Change Research Centre, University of New South Wales, Australia*
[g] *ARC Centre of Excellence for Climate System Science, University of New South Wales, Australia*

ARTICLE INFO

Article history:
Received 29 December 2014
Received in revised form
13 May 2015
Accepted 15 June 2015
Available online 19 June 2015

Keywords:
Temperature and precipitation extremes
Gridded datasets
Observations
Interpolation methods
Scaling issues

ABSTRACT

Using daily station observations over the period 1951–2013 in a region of south-east Australia, we systematically compare how the horizontal resolution, interpolation method and order of operation in generating gridded data sets affect estimates of annual extreme indices of temperature and precipitation maxima (hottest and wettest days). Three interpolation methods (natural neighbors, cubic spline and angular distance weighting) are used to calculate grids at five different horizontal gridded resolutions ranging from 0.25° to 2.5°. In each case the order of operation in which the grid values of the hottest and wettest day are calculated is varied: either they are estimated from daily grids or from station points and then gridded. We find that the grid resolution-despite showing more regional detail at high resolution – has relatively limited effect when considering regional averages. However, the interpolation method and the order of operation can substantially influence the actual gridded values. And while the difference due to the order of operation is not substantial when using natural neighbor and cubic spline interpolation, it is particularly apparent for indices calculated from daily gridded estimates using the angular distance weighting method. As expected given the high spatial variability of precipitation fields, precipitation extremes are most sensitive to method, but temperature extremes also exhibit substantial differences. For the annual maximum values averaged over the study area, the differences may be up to 2.8 °C for temperature and 60 mm (about a factor 2) for precipitation. Differences are seen most prominently in return period estimates where a 1 in 100 year return value calculated using the angular distance weighting daily gridded method is equivalent to about a 1 in 5 year return value in most of the other methods. Despite substantial differences in the actual values of gridded extremes, analyses suggest that the impact on long-term trends and inter-annual variability is small.

1. Introduction

Changes in the mean climate state are often used as indicators of a changing climate, but the impacts of climate change are particularly experienced through climate extremes (such as floods, droughts and heat waves). Indices of temperature and precipitation extremes, such as the hottest or wettest days of the year, can be used to monitor when changes in climate extremes occur (e.g., Kiktev et al., 2003; Alexander et al., 2006; Donat et al., 2013a) as well as to examine changes in the present and future climate and its impacts (e.g., Sillmann and Roeckner, 2008; Alexander and Arblaster, 2009; Orlowsky and Seneviratne, 2011; Sillmann et al., 2013).

The need to distinguish extreme behavior on the basis of the space and time scales involved becomes paramount when comparison between observed and modeled extremes is performed, often in order to assess the processes driving extremes or to investigate future changes. To adequately compare observations and models, observations generally have to be ‘gridded’ (i.e., converted

---

[*] Correspondence to: Manila Observatory, Ateneo de Manila University Campus, Loyola Heights, Quezon City 1108, Philippines.
E-mail address: avila.fb@gmail.com (F.B. Avila).

http://dx.doi.org/10.1016/j.wace.2015.06.003

from point observations to values on a latitude–longitude grid) or climate models have to be ‘downscaled’ (i.e., data values relevant for observation sites have to be inferred from gridded values). Thus, care must be taken to distinguish between gridded products whose values may either represent regularly spaced, point locations or area averages.

At the core of the problem is the fundamental mismatch between the spatial representativeness of *in situ* observations on the one hand, which are necessarily collected at observation sites (points), and that of gridded climate model output which represents area mean values, on the other hand. In general, global observational datasets of climate extremes (e.g. Alexander et al., 2006; Donat et al., 2013a) are constructed by calculating indices at observation sites and then interpolating them to produce grids of similar dimension to global climate models. This makes them structurally different from the climate model simulated fields of the same indices which are usually calculated from daily gridded fields, creating an ‘issue of scale’ when comparing the observed and modeled datasets (Alexander and Tebaldi, 2012). Scale mismatch, often referred to as the ‘problem of a change in support’ by statisticians, more importantly affects phenomena such as precipitation whose spatial features are discontinuous, or extremes calculated from daily or sub-daily data. In addition, the current spatial resolution of global climate models is generally insufficient to easily provide detail on extremes at local and regional levels. Being aware of these scaling issues is important to avoid the misinterpretation of the results when comparing observed and modeled extremes.

While it may be difficult at present to produce observational datasets that are fully comparable with model output, previous studies which have tried to address at least some aspects of these issues of scale (e.g. Chen and Knutson, 2008, Yin et al., 2014, Donat et al., 2013b) suggest that they are important when it comes to understanding and assessing changes in extremes.

To date, however, relatively little work has been done in systematically assessing these scaling issues with respect to their impacts on the estimation of extremes of climate variables. Studies which have discussed its effects, such as those on extreme temperatures globally (e.g. Donat et al., 2014), have been limited by the issue that the station networks used for calculating the daily and annual extremes grids have not been identical. And although previous studies that assess aspects of the sensitivity of daily extremes to interpolation method and station network density do exist (e.g. Hofstra et al., 2010), we are not aware of previous studies investigating scaling issues related to grids of precipitation and temperature extremes based on identical station networks. We therefore assess how these scaling issues impact various statistics of extremes, and in particular how much the order of operation in which extremes are calculated matters (i.e., extremes calculated from daily grids, as would be the case when calculating extremes from climate model data, versus gridded point-based extremes, as is the case for many gridded observational datasets of extremes). Focusing on a small region with relatively good observational coverage, we explore the sensitivity of typical applications when analyzing extremes as a function of grid resolution, interpolation method and order of operation when calculating grids of annual extremes. These applications include representation of long-term changes, inter-annual variability, spatial patterns, and extreme value analysis. It should be noted, however that the purpose of this study is not to find the ‘best’ method for gridding extremes data but rather to show how large interpolation and scaling errors could be when using a range of techniques that are commonly used in climate science. Our hope is that the magnitude of the errors displayed here can be used to inform detection and attribution and model evaluation studies.

2. Data and methods

For this study we focus on extremes with potentially significant impacts i.e. the hottest and wettest days of the year: termed TXx (unit: °C) and Rx1day (unit: mm) respectively (Zhang et al., 2011), two of the 27 core indices recommended by the Expert Team on Climate Change Detection and Indices (ETCCDI). To investigate systematic changes related to scaling issues we calculate grids for a 15°x15° region in south-east Australia spanning 139 °E to 154 °E longitude and 24°S to 39°S latitude, a region with a reasonably high density of station data for the period 1951 to 2013 (Fig. 1a and 1b). Data were obtained from the Global Historical Climatology Network (GHCN)-Daily dataset (Menne et al., 2012) and the stations chosen were those used in the GHCNDEX dataset (Donat et al., 2013b) for TXx and Rx1day. Stations within 2° around the chosen domain were also used to ensure that there is sufficient support to calculate grid values in the cells at the edges of the domain. There is maximum of 129 stations available with temperature data while the number of stations with precipitation data was as high as 2173. However, the actual number of available station data could be substantially lower than these for certain time periods for both temperature and precipitation indices, particularly prior to 1960 and after 2000 (Fig. 1c and 1d), due to stations opening or closing or not having made enough measurements throughout the year to calculate values for TXx or Rx1day. Despite the changes in the number of available stations, we make use of the full dataset in our analysis as we are less concerned with the uncertainty related to network density but rather with the uncertainty related to issues of scale and the structural uncertainty related to methodological framework. Note that, as we are using an identical set of input stations for all grids, changes in the station network density would affect all of the gridded fields and therefore do not affect the comparisons discussed in this study.

Note also that the station data have not been homogenized for this study. We aimed to investigate the methodological uncertainties related to the order of operation using the densest possible station network and using only homogeneous stations would have substantially reduced the number of stations we could work with. Furthermore, as we use exactly the same input stations for all methods, any potential inhomogeneities will affect all of the constructed grids and therefore would not affect our conclusions.

Statistical methods used to assess our results throughout the paper are as follows. When calculating differences we calculate significance at the 10% level using the Student *t*-test. Linear trends are calculated using Sen's trend estimator (Sen, 1968) and trend significance is estimated at the 5% level using the Mann–Kendall test (Mann, 1945; Kendall, 1975). Correlations with the El Niño-Southern Oscillation (ENSO) index are calculated using the Spearman rank-order method. Calculations of the pattern correlation were centered (i.e. the mean over the study area is removed prior to comparison of two patterns). Applied methods of extreme value analysis and considered interpolation methods, grid resolutions and aggregation methods used are outlined in the following sections.

2.1. Grid size and order of operation

To determine the sensitivity to horizontal resolution, the station data were interpolated at five different resolutions (latitude by longitude): 0.25°x0.25°, 0.5°x0.5°, 1.0°x1.0°, 1.5°x1.5° and 2.5° × 2.5° (Fig. 1a and b).

To determine how the order of calculation affected the estimates of TXx and Rx1day, two approaches for calculating the gridded annual extremes were implemented. The first approach involved calculating the annual extremes of the indices from the