An approach to homogenize daily peak wind gusts: An application to the Australian series

Cesar Azorin-Molina | Jose A. Guijarro | Tim R. McVicar | Blair C. Trewin | Andrew J. Frost | Deliang Chen

1Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden
2State Meteorological Agency (AEMET), Balearic Islands Office, Palma de Mallorca, Spain
3CSIRO Land and Water, Canberra, Australian Capital Territory, Australia
4Australian Research Council Centre of Excellence for Climate System Science, Sydney, New South Wales, Australia
5Australian Bureau of Meteorology, Melbourne, Victoria, Australia
6Australian Bureau of Meteorology, Sydney, New South Wales, Australia

Correspondence Cesar Azorin-Molina, Department of Earth Sciences, University of Gothenburg, Box 460, 405 30 Gothenburg, Sweden. Email: cesar.azorin-molina@gu.se

Funding information H2020 Marie Skłodowska-Curie Actions, Grant/ Award Number: 703733; Ministério de Ciência e Innovación, Grant/Award Number: CGL2014-52901-P, Vetenskapsrådet, Grant/ Award Number: 2017-03780

Daily Peak Wind Gust (DPWG) time series are important for the evaluation of wind-related hazard risks to different socioeconomic and environmental sectors. Yet, wind time series analyses can be impacted by several artefacts, both temporally and spatially, which may introduce inhomogeneities that mislead the study of their decadal variability and trends. The aim of this study is to present a strategy in the homogenization of a challenging climate extreme such as the DPWG using 548 time series across Australia for 1941–2016. This automatic homogenization of DPWG is implemented in the recently developed Version 3.1 of the R package Climatol. This approach is an advance in homogenization of climate records as it identifies 353 break points based on monthly data, splits the daily series into homogeneous subperiods, and homogenizes them without needing the monthly corrections. The major advantages of this homogenization strategy are its ability to: (a) automatically homogenize a large number of DPWG series, including short-term ones and without needing site metadata (e.g., the change in observational equipment in 2010/2011 was correctly identified); (b) use the closest reference series even not sharing a common period with candidate series or presenting missing data; and (c) supply homogenized series, correcting anomalous data (quality control by spatial coherence), and filling in all the missing data. The NCEP/NCAR reanalysis wind speed data were also trialled in aiding homogenization given the station density was very low during the early decades of the record; however, reanalysis data did not improve the homogenization. Application of this approach found a reduced range of DPWG trends based on site data, and an increased negative regional trend of this climate extreme, compared to raw data and homogenized data using NCEP/NCAR. The analysis produced the first homogenized DPWG dataset to assess and attribute long-term variability of extreme winds across Australia.

KEYWORDS Australia, Climatol V3.1, daily peak wind gusts, homogenization

1 | INTRODUCTION

The quality control and homogenization of climate time series must be performed prior to assessing climate change and variability (Venema et al., 2012), or when using observations to estimate future climate scenarios (Laapas and Venäläinen, 2017). This is because there are many artefacts that may substantially alter the meteorological measurements.
Approach

Homogenization Methods of Climate Series: An Integrated Review

et al., (Vose et al., 2003). For instance, for wind speed: (a) station relocations; (b) anemometer height changes; (c) instrument malfunctions; (d) instruments changes; (e) different sampling intervals; and (f) observing environment changes (Azorin-Molina et al., 2014). Therefore, advances in the quality control and homogenization procedures seek to identify and remove all artefacts in the climate time series so that variance in the resultant series is only driven by weather and climate changes (Conrad and Pollack, 1950).

Numerous methods have been developed to identify inhomogeneities in climate series in the form of abrupt (break points) or short-term (local trends) changes in the averages of those series, thereby avoiding artificial trends in the assessment of climate. Peterson et al. (1998) and Aguilar et al. (2003) conducted comprehensive reviews of existing homogenization methods and approaches for developing homogenized data sets. The European COST (Cooperation in Science and Technology) Action ES0601 “Advances in Homogenization Methods of Climate Series: An Integrated Approach” (HOME; 2006–2011; http://www.cost.eu/COST_Actions/essem/ES0601; Accessed 1 December 2018) allowed scientists from over 20 countries to compare and discuss their methodologies developed to perform the challenging task of quality controlling and homogenizing monthly climate series. Venema et al. (2012) reviewed the most suitable homogenization algorithms used in the software HOME for monthly air temperature and precipitation data and concluded that: (a) applying automated homogenization algorithms give better results than user-reliant algorithms when users are not experts; and (b) developing easy-to-use software as operator training is crucial to produce accurate climate series. Recently, Hunziker et al. (2018) demonstrated the effects of undetected data quality issues on the most common relative homogenization approach (i.e., that relies on high spatial correlation with the climate signal of nearby stations; Conrad and Pollack, 1950), by reducing the correlation coefficients of station pairs, deteriorating the performance of data homogenization methods, increasing the spread of individual station trends, and significantly biasing regional air temperature and precipitation trends. However, most previous homogenization studies have focused on two climate variables (air temperature and precipitation; e.g., Štěpánek et al., 2009), with very few homogenization efforts on daily wind extremes (e.g., see review in Table 1 below).

Advances in statistical approaches to homogenize extreme wind time series are of great scientific interest. Furthermore, given a changing climate, assessing trends of this natural hazard from reliable series also has substantial socioeconomic and environmental significance for many sectors (Vose et al., 2014), such as: (a) agriculture and hydrology; (b) long-term wind power generation; (c) wind-related hazards such as wind-driven fires and catastrophes (loss of life, property, and habitat); (d) land and aviation transport; (e) marine and coastal impacts due to wind-driven storm surges and waves; (f) tourism and wind sports; (g) air quality and human health; amongst many others. Consequently, it is of utmost importance to develop adequate data quality control and homogenization approaches for long-term wind studies.

Due to the variety of factors affecting the measurement of wind speed, that is, its high natural short-term variance (Balling and Cerveny, 2005; Jakob, 2010), long-term trends (Vautard et al., 2010; McVicar et al., 2012), and the relatively high spatial variability (Azorin-Molina et al., 2014), as well as the high sensitivity of wind to local site conditions (WMO, 2017), implementing quality control and homogenization procedures on wind series has been challenging. To summarize, only a few approaches have been developed for mean wind speed so far in recent years: (a) Wang (2008) and Wan et al. (2010) used the RHtestV2 data homogenization package (Wang and Feng, 2007) for Canadian monthly mean wind speed data, and Si et al. (2018) applied the RHtestV4 for Tianjin (China) monthly mean wind speed data; (b) Petrović et al. (2008), Li et al. (2011), and Péliné-Németh et al. (2014) applied the Multiple Analyses of Series for Homogenization (MASH; Szentimrey, 1999, 2008) to homogenize daily wind speed series for Ireland, for the greater Beijing area (China) and Hungary, respectively; (c) Štěpánek et al. (2013), Azorin-Molina et al. (2014), and Minola et al. (2016) used the AnClim package (Štěpánek, 2004) to detect sudden break points in monthly wind speed series for the Czech Republic, Spain and Portugal, and Sweden, respectively; (d) Guijarro (2015) and Azorin-Molina et al. (2018b) applied the Climatol package to detect artificial change points and adjust inhomogeneities in monthly wind speed time series in Spain and Portugal and Saudi Arabia, respectively; and (e) Laapas and Venäläinen (2017) and Azorin-Molina et al. (2018a) applied the recently developed relative homogeneity test HOMER (HOMogenization software in R; Mestre et al., 2013) for homogenizing monthly mean wind speed in Finland and the Canary Islands, respectively. Zahradníček et al. (2018) homogenized daily data in the Czech Republic by applying the Standard Normal Homogeneity Test (SNHT, Alexandersson, 1986) and the Maronna-Yohai test to monthly, seasonal, and annual aggregates. None of these methods stands out as being best, so different approaches have been applied thereby justifying the need of benchmarking the performance of homogenization of wind speed data, as Venema et al. (2012) conducted in HOME for air temperature and precipitation. Benchmarking approaches used for temperature studies such as that of Venema et al. (2012) require the construction of a synthetic data set. As no such synthetic data set is known to exist for wind speed, an alternative approach is required.

The detection of break points in daily series such as daily peak wind gusts (hereafter DPWG) is more difficult than when using monthly mean series, as the higher variability of
| Study # | Source                          | Data details | Location           | Period       | Software | Methodological approach                                                                 | Reference series | Observed data                                      | Time-scale |
|--------|--------------------------------|--------------|--------------------|--------------|----------|----------------------------------------------------------------------------------------|------------------|--------------------------------------------------|------------|
| 1      | Klink (1999)                    | 187/10 m     | United States      | 1961–1990    | No       | Quality of the wind data was presumed to be good according to the standard meteorological quality control procedures applied by the National Renewable Energy Laboratory | No               | Mean monthly maximum wind speed                  | Monthly    |
| 2      | Sweeney (2000)                  | 1/10 m       | Dublin             | 1903–1999    | No       | No                                                                                      | No               | Daily maximum wind gusts > 30 m/s                | Daily      |
| 3      | Pirazzoli and Tomasin (2003)    | 17/10 m      | Italy              | 1955–1996    | No       | Quality controls to remove outliers                                                     | No               | Daily maximum wind gust                        | Daily      |
| 4      | Alexander et al. (2005)         | 28           | United Kingdom and Iceland | 1950–2003    | No       | Quality control to remove outliers                                                     | No               | 3-h mean sea level pressure data to detect storms | Hourly–daily |
| 5      | Cabalar-Fuentes (2005)          | 3/10 m       | NW Spain           | 1961–2001    | No       | Helmert criterion and run tests                                                        | No               | Daily peak wind gust                            | Daily      |
| 6      | Smits et al. (2005)             | 13/10 m      | The Netherlands    | 1962–2002    | No       | Correction for observational changes                                                   | No               | Mean wind speed                                 | Hourly–daily |
| 7      | Xu et al. (2006)                | 305/10 m     | China              | 1969–2000    | No       | No                                                                                      | No               | Windy days (daily mean wind speed > 5 m/s)      | Daily      |
| 8      | Fujii (2007)                    | 150/variables m | Japan              | 1966–2005    | No       | Basic quality controls based on metadata                                               | No               | Mean wind speed and peak gust speed            | Hourly–daily |
| 9      | Pryor et al. (2009)             | 2 data sets/10 m | U.S.              | ~1948–2006   | No       | NCDC data quality control procedures                                                   | No               | Mean wind speed                                 | Hourly–daily |
| 10     | Jiang et al. (2010)             | 535/10 m     | China              | 1956–2004    | No       | Quality tests by the China Meteorological Administration (CMA)                        | No               | Days of strong winds; maximum wind              | Daily      |
| 11     | Kruger et al. (2010)            | 94/10 m      | South Africa       | 1993–2008    | No       | Visual inspections of plots                                                            | No               | Maximum wind gusts                             | Daily      |
| 12     | Usbeck et al. (2010)            | 1/10 m       | Canton Zurich, Switzerland | 1891–2007    | No       | Basic quality controls                                                                | No               | Gust wind speed                                 | Daily      |
| 13     | Vautard et al. (2010)           | 822/10 m     | Global, with a northern hemisphere focus | 1979–2008    | Yes      | Validation and quality control                                                        | No               | Strong winds > certain thresholds              | Hourly–daily |
| 14     | Hewston and Dorling (2011)      | 43/10 m      | United Kingdom     | 1980–2005    | No       | Quality control by the UK met Office and the authors                                  | No               | Maximum wind gusts                             | Hourly–daily |
| 15     | Péliné-Németh et al. (2011)     | 26/10 m      | Hungary            | 1997–2010    | No       | Basic quality controls                                                                | No               | Wind gust                                       | Hourly–daily |
| 16     | Cusack (2013)                   | 5/10 m       | The Netherlands    | 1910–2010    | No       | Relative homogenization—based on Wan et al. (2010)                                      | Average of winds at four other neighbouring stations | Maximum wind speed                             | Daily      |
| 17     | Vose et al. (2014)              | 2 data sets/10 m | U.S.              | 1973–2006    | Yes (see Groisman, 2002)                                                               | A homogeneity-adjusted dataset by Groisman (2002) | No               | 90th percentile 10-m wind speed                 | Hourly–daily |
| 18     | Klink (2015)                    | 39/variables m | U.S.              | 1993–2009    | RHtestsV4 | Relative homogenization—PMTred algorithm                                                | Yes               | Fastest 2-min wind speed                        | Hourly–daily |
daily data series decreases the signal-to-noise ratio (Szentimrey, 2013). Table 1 reviews the studies of long-term trends of wind speed extremes from anemometer observations and the methodological approaches applied to quality control and homogenize series. This shows that almost all studies have only applied limited quality control and homogenization based on few metadata. The generally accepted strategy to overcome this difficulty has been to homogenize the monthly aggregates first, and then adjust the daily series by applying the monthly corrections interpolated to the daily resolution (Vincent et al., 2002). This approach, formerly applied to homogenize daily air temperature and precipitation data series (Lakatos et al., 2011) and wind speed (Péliné-Németh et al., 2014), was trialled with DPWG time series from Spain and Portugal (Azorin-Molina et al., 2016). However, these daily adjusted results from the interpolation of monthly corrections were discarded as the resultant series had remaining inhomogeneities, and it was decided to apply a direct homogenization of the daily series. This direct daily homogenization approach adopted by Azorin-Molina et al. (2016) detected a smaller number of break points, and achieved a better correction as errors from the interpolation of monthly corrections were avoided whilst the most substantial inhomogeneities were detected.

In view of this state of the art, the overall aim of this study is (a) to present a strategy and advances in the homogenization of a challenging climate extreme such as the DPWG using the Version 3.1 of the R package Climatol (Guijarro, 2016). The scientific innovation implements the use of break points detected at monthly basis to split the daily series, which are then homogenized without needing the monthly corrections. This homogenization approach is described in detail and demonstrated through application to a large DPWG dataset (548 stations) covering Australia for 1941–2016. Our secondary objectives are: (b) investigating the ability of reanalysed mean wind-speed data (from reanalysis) to aid the detection of break points when used as reference series in the absence of neighbouring stations; and finally (c) creating a long-term quality-controlled and homogenized DPWG dataset for the assessment and attribution of the variability of extremes across Australia for future studies.

## 2 | DATA AND METHODS

### 2.1 | Observed DPWGs and metadata

The original dataset contains 706 DPWG series across Australia spanning from 1 April 1939 until 31 December 2016. The long-term temporal coverage, the different spatial density of stations across the country and throughout the study period, and the fact that the DPWG series have not been comprehensively quality controlled, homogenized and analysed before, encouraged us to choose this Australian
dataset as a suitable one to test the homogenization approach presented herein. Observed DPWG data were recorded and supplied by the Australian Bureau of Meteorology (BoM; http://www.bom.gov.au/; last accessed 1 December 2018) at heights of ~10 m above the land surface. DPWG data were measured by different types of anemometers, with Dines pressure tube anemometers used until the ~1980s and replaced with Synchrotac 706 rotating cup anemometers in the Automatic Weather Stations (AWS) installed in the last two to three decades (for details of instruments, see Jakob, 2010; Cechet and Sanabria, 2012). This general replacement is the major cause of inhomogeneities in the observed series (Jakob, 2010), with Synchrotac anemometers showing a tendency towards to “overspeed” compared to the Dines ones; i.e., the sensor’s inertia causes the anemometer to continue spinning after the wind speed has decreased (Gorman, 2004).

The BoM defines the observed DPWG data sets (in m/s) as the highest fluctuation of wind speed recorded in 24 hr (0000–0000 local time; local time in Australia ranges from UTC + 8 in Western Australia to UTC + 10 in the eastern states, with UTC + 11 in summer in some states; UTC: Coordinated Universal Time), following the requirements adopted by the World Meteorological Organization (WMO, 2017). The gust duration (i.e., measure of the duration of the observed peak gust; Azorin-Molina et al., 2016) is established as ~3 s as recommended by the WMO (1987). Prior to the introduction of automatic weather stations, daily maximum wind gust was determined by visual inspection of autographic charts.

The first quality control applied to the raw DPWG dataset consisted in removing all data with quality control flags as “wrong,” “suspect” or “inconsistent” from the BoM (i.e., 6,427 data; 0.17%), noting that DPWG data are not currently systematically quality controlled by the BoM (David Sinclair, pers. comm.), with the flags that exist being part of an ad hoc process, mostly based on identifying values that exceed gross thresholds. Due to incomplete years at the beginning of the DPWG series, our homogenization approach focused on 1 January 1941 until 31 December 2016 (i.e., a 76-year period). After removing series having less than 1,826 observations (i.e., 5 years) and 12 stations located too far away to be useful as references (e.g., located on islands), 548 DPWG series remained in the dataset. Figure 1a displays the temporal evolution of DPWG data availability, with three noteworthy steps: (a) a steady increase up to around 100 series from 1941 till the 1980s; (b) a stabilization or weak decline of series till 2000s; and (c) an abrupt increase in 2003 up to >400 series due to the installation of automatic weather stations capable of measuring daily wind gusts. Automatic weather stations progressively entered the Australian observing network from the late 1980s onwards, yet the software used by most such stations prior to 2003 was not capable of transmitting DPWG data.

2.2 Reanalysed wind speed data
A challenging feature for applying a relative homogenization strategy was that relatively few wind observation sites were available until the 2000s across Australia. Given this, we evaluate the ability of gridded-reanalysis data to aid the detection of break points in the observed DPWG dataset. The National Centres for Environmental Prediction (NCEP) and the National Centre for Atmospheric Research (NCAR) Reanalysis (hereafter NCEP/NCAR reanalysis; Kalnay et al., 1996) was chosen as McVicar et al. (2008) documented that it most optimally captured Australian wind speed observed climatology and trends, compared to other reanalysis products. The 6-hourly $u$ (zonal velocity in m/s, that is, the component of the horizontal wind towards east) and $v$ (meridional velocity in m/s, that is, the component of the horizontal wind towards north) components from the global NCEP/NCAR reanalysis were downloaded from the NOAA server: ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis/surface_gauss/ (last accessed 1 December 2018). We created a land mask to select 245 grid-cells across Australia, for which mean wind speeds (in m/s) were computed from the $u$ and $v$ components at each grid cell, and then the highest wind speeds of the four synoptic times at 00:00, 06:00, 12:00, 18:00 hours UTC plus the 00:00 of the following day were added to the dataset; the 00:00 value may be the maximum on either one or both of the two consecutive days when not exceeded by any other values. The assessment of the potential of spatial gridded data in detecting break points in the observed dataset is restricted to the NCEP/NCAR reanalysis time-period, that is, 1948–2016 (69-years), with data from 791 daily series in total: 245 reanalysis grid-cells plus 546 observed stations. The geographical distribution of the stations and centre points of the NCEP/NCAR reanalysis grid-cells across Australia are shown in Figure 1b.

3 AN APPROACH TO HOMOGENIZE DPWG SERIES

3.1 The Climatol package and overview of the homogenization approach
Climatol (Guijarro, 2018) is an R (R Core Team, 2015) package for homogenizing climate series and producing automated products (e.g., homogenized time series) derived from them. This package can be freely downloaded from https://CRAN.R-project.org/package=climatol (last accessed 1 December 2018) and further information about the documentation and improvements of the most recent version can be found in http://www.climatol.eu/ (last accessed...
1 December 2018. Climatol strongly benefited from participating in the COST Action HOME activity and improved its algorithms in its Version 2. In the last few years Climatol has been widely applied to homogenize a variety of monthly climate databases, particularly air temperature and precipitation series (Guijarro, 2008, 2013a; Hernández et al., 2012 -at daily basis-; Luna et al., 2012; and Mamara et al., 2013), but also wind speed (Azorin-Molina et al., 2018b) and DPWG (Azorin-Molina et al., 2016). Due to the limitations encountered by Azorin-Molina et al. (2016) related to the interpolation of monthly corrections to daily series with a positively skewed probability distribution such as wind, in the following subsections we present a strategy that advances DPWG homogenization. This further improvement used the break points detected on the monthly scale to split the daily series into their homogeneous subperiods, which are then reconstructed by infilling all missing data with their estimated values, as explained in Section 3.2.1 below. This strategy was developed in Version 3.1 of the Climatol package (Guijarro, 2016).

Figure 2 is a flowchart summarizing the homogenization procedure implemented in Climatol in which, after reading the DPWG series and performing some initial checks and calculation of means and standard deviations, three stages follow: (a) detection of inhomogeneities by applying the SNHT on stepped overlapping temporal windows, splitting the most inhomogeneous series by the detected break points, and repeating the process until all series appear homogeneous; (b) as in (a), but applying the SNHT to the entire series (i.e., 1941–2016); and (c) generation of the final homogeneous series by filling all missing data by weighted ratios of nearby stations in the homogeneous subseries.

Finally, homogenization outputs from Climatol are delivered in a R binary file containing the original and homogenized series, a document with a wealth of diagnostic graphics, and three text files with a log of the process plus lists of the corrected outliers and break points. A detailed description of this homogenization approach is given in the following subsections.

3.2 | Description of the homogenization approach

3.2.1 | Series estimation from near neighbour data

Prior to the application of Stages 1–3, Climatol’s homogenization approach relies heavily on the estimation of complete series from their nearest available data. This method was adapted from Paulhus and Kohler’s (1952) strategy to interpolate missing precipitation records to complete reports in U. S. Weather Bureau bulletins. They simply divided the precipitation values by the station precipitation mean and estimated the missing values by averaging the closest three values around the site to be interpolated. In this way, undoing the normal-ratio estimations by multiplying by the mean precipitation at each location yielded the interpolated precipitation values. This approach was chosen because it can be applied even when neighbour series do not have any common period of observation, hence enabling the use of short series that otherwise would have to be disregarded, and Climatol extends the method by offering three data normalization options being: (a) subtracting the mean; (b) dividing by the mean (as in Paulhus and Kohler, 1952); and (c) subtracting the mean and dividing by the standard deviation. Therefore, the user can choose the most appropriate normalization depending on the climatic variable studied,
whether it has a limiting zero value and an L-shaped probability distribution (e.g., precipitation and wind; where case (b) above is more appropriate) or bears a near-normal distribution (e.g., air temperature and air pressure; where case (c) or even (a) above are more suitable). Moreover, the number of reference series can be set by users, defaulting to 10 when building reference series and to 4 during the final series reconstruction.

The main drawback of this method is the need for accurate estimates of means and standard deviations for the full time series (including days missing). An iterative approach is used here to estimate these statistics. As missing data are common in observational datasets, Climatol first computes these statistical parameters from all available data in each series and, after filling in the missing data with first-guess estimations, means and standard deviations are calculated again. Then, using these new means and standard deviations, a second normalization of the series and missing data estimation is performed, and the procedure is repeated (see Figure 2) until the maximum difference between means from the previous iteration becomes lower than a prescribed threshold (half the precision of the data by default). The default number of iterations is set to 999. All normalized series can be estimated in this way from an (optionally weighted) average of their available (at each time step) normalized nearby data (i.e., up to 10 stations in Stages 1 and 2, and up to 4 in Stage 3), and the estimated series are used to: (a) fill in the missing data of the observed series; (b) compute series of spatial anomalies (observed data minus Climatol estimations); and (c) calculate root mean square errors (RMSE) of the Climatol estimations, which can serve to compute confidence intervals for the homogenized series. By default, no weighting is applied to estimate the series in Stages 1 and 2, because nearby series may contain inhomogeneities, but in the final reconstruction of all series from the homogeneous subperiods (Stage 3) neighbour data are weighted by the inverse distance function:

\[ 1/(1 + \frac{d^2}{h^2}) \]

where \( d \) is the distance between the stations and \( h \) the distance at which the weight becomes 0.5. The parameter \( h \), expressed in km, can also be set by the user, defaulting to 100.

### 3.2.2 Splitting series into homogeneous subperiods (Stages 1–2) and fill in of missing data (Stage 3)

The estimated series are then used as references to obtain series of anomalies (observed minus estimated values) on which to detect outliers and changes in the average (break points) by means of the well-established SNHT. The SNHT test demonstrated good performance to detect break points in the series (Guijarro, 2013b) defined by the highest difference of means before and after any time step, yet it could give misleading results in the presence of more than one inhomogeneity in the series (Begert et al., 2008; Rienzner and Gandolfi, 2011). To minimize this possibility, the

---

**FIGURE 2** Overview of the homogenization procedure in the Climatol package with the operations performed within each stage (Guijarro, 2018). Red arrows show the progress from the main process (in blue) to the secondary one (in yellow). Shapes refer as follows: Rectangles (processes), parallelograms (input/output data), and diamonds (conditional processes).
detection of shifts in the mean is performed in Stage 1 on overlapping stepped temporal windows, followed by a Stage 2 in which the SNHT is applied to the entire series, as intended by Alexandersson (1986). The statistic $T$ measures the inhomogeneity of a series in Alexandersson’s SNHT (the lower the $T$, the higher the homogeneity). After computing $T$ for all series, those with higher inhomogeneities are split into two subseries in decreasing $T$ order, as long as no reference series has been split in the same iteration with a similar $T$ value, and the procedure is repeated until the maximum $T$ value in all series is lower than a threshold, set to 25 by default. This figure is conservative when compared to published values of $T$, which lie around 12 and 15 for significant levels of 0.05 and 0.01, respectively, and sample sizes over 10,000 (Khaliq and Ouarda, 2007), hence trying to avoid false detections at the cost of letting pass the smaller inhomogeneities. However, these values were computed on Monte Carlo ideal simulations, but in real series experience shows that this test can give very different values depending on the climatic variable studied, the degree of correlation between the series and their temporal frequency (Guijarro, 2018). Therefore, the Climatol package lets the user adaptively set $T$ thresholds after inspecting anomalies graphs and histograms produced by a first run of the homogenization function; in the package, and in the rest of this text, SNHT’s $T$ statistic is referred simply as SNHT. The two break-point detection stages (i.e., Stage 1 and 2) are followed by a Stage 3 devoted to fill in all missing data by replacing them with their estimated values. Therefore, no correction is applied in the two first stages, devoted only to detect the break-points and split the series accordingly. Then, the missing data infilling in Stage 3 builds complete series from every homogeneous subperiod without the need of computing correction terms. In this case, the convergence criterion to halt the iterative process of calculation of means is applied to the individual data rather than to the means. This procedure has proved to give good results in some benchmarking comparisons (Guijarro, 2011; and in the MULTITEST project http://www.climatol.eu/MULTITEST/, last accessed 1 December 2018).

### 3.3 Application of Climatol to the homogenization of the DPWG series

As previously explained, the detection of break points in daily series is compromised by the high variability of wind gusts encountered at this fine timescale. To overcome this issue, we first aggregated the observed DPWG series into monthly time series for 1941–2016, on which Climatol was applied using the normal ratio normalization (i.e., dividing all values by their means, as previously explained in Section 3.2.1). The same procedure was adopted after adding the reanalysis series to the dataset, but as they begin in 1948, homogenizations were performed for 1948–2016, also with and without the NCEP/NCAR reference series. The results of the homogenization are shown in Section 4.1.

The potential benefits of using reference series from the NCEP/NCAR reanalysis were evaluated by calculating mean correlations and distances between every single station and their 4 and 10 closest references with a minimum of 10 years of common data, with and without the support of reanalysis data as reference. As these benefits are expected to be higher when the density of observed series is lower, mean correlations and distances were computed for the whole common record of 1 January 1948 to 31 December 2016 (69 years) and also for three shorter periods: (a) 1 January 1948 to 31 December 2002 (55 years); (b) 1 January 1948 to 31 December 1982 (35 years); and (c) 1 January 1948 to 31 December 1965 (18 years). The assessment of the added value of reanalysis series in the homogenization of DPWG is described in Section 4.2. Whilst the use of reanalysis series as a reference implicitly assumes their homogeneity, major inhomogeneities in reanalysis series typically coincide with major changes in observation systems (Rienecker et al., 2011), such as the widespread introduction of satellite data in the late 1970s, and there are no such major changes known to have a substantial effect in wind data around the 2003 break points referred to in Section 4.2.

The detected break points were carefully checked against the metadata from 35 stations selected to maximize the temporal and spatial coverage (Figure 3). All metadata about wind speed equipment history in these stations were extracted from http://www.bom.gov.au/climate/datasets/about-data-observations.shtml (last accessed 1 December 2018), where BoM offers abundant and outstanding information about the location, data inventories, site and skyline sketches at fixed times, and history of changes in the instruments of any observatory, amongst other observation details. These metadata are extensive from 1997 onwards, although they still only capture changes proximal to the near vicinity (typically within 50–100 m) of the observation site, whereas changes further away can be important for wind speed, to a limited extent. They are much more limited prior to 1997. The comparison of the break points against real metadata is shown in Section 4.3.

Trends of the DPWG series before and after the homogenization are difficult to compare using the complete observed dataset, as most of the stations operated over a short period. Therefore, we also limit the comparison of the impact of the homogenization on the DPWG trends to the 35 selected series. This was evaluated by computing Ordinary Least Squares (OLS; Wilks, 2011) linear adjustments with time (independent variable or X axis) for: (a) the raw series; (b) the homogenized series with and without the NCEP/NCAR reanalysis references; and (c) the 35 NCEP/NCAR reanalysis series of the closest grid-cell centres to the locations of the 35 selected observatories. DPWG trends are
expressed in metres per second per decade (m s\(^{-1}\) dec\(^{-1}\)) and the nonparametric correlation coefficient of Mann-Kendall’s tau-b (Kendall and Gibbons, 1990) was used to assess the statistical significance of linear trends at \( p < 0.05 \). The assessment of trends between the raw and the homogenized datasets was conducted for two periods: (a) 1948–2016; and (b) 1971–2016, to quantify the impact of the density of DPWG observations during the first decades on trends. The comparison is described in Section 4.4.

Finally, comparison of all these results conducted to choosing the most suitable lists of break points to use as input to a last application of Climatol, this time to the original DPWG series, to split them at the shift dates and proceed to the reconstruction of all homogeneous subperiods. Homogenized and complete DPWG were obtained by applying the infilling method implemented in Climatol. The final series are presented in Section 4.5.

4 | RESULTS

4.1 | Homogenization of DPWG

Figure 4 shows an example of the homogenization of DPWG series for station # 9789-Esperance (southwest of Australia; location provided on Figure 3), with Figure 4a displaying the detection of a break point with SNHT value = 40 in October 1989 (another split was found in a later iteration in December 2009) and Figure 4b illustrating the reconstruction of whole series from the three homogeneous subperiods by the infilling procedure described above.

Along with this example, Figure 5 summarizes the number of splits applied to the DPWG series at the detected break points, showing the split frequencies of the 1948–2016 homogenization without (Figure 5a) and with (Figure 5b) the support of the NCEP/NCAR reanalysis references. Both homogenizations coincide in detecting a maximum of break points (>40 splits) in 2011, also showing a high frequency of splits between 2009 and 2012. However, we found a secondary prominent maximum of break points in 2003 (>40 splits) when NCEP/NCAR reanalysis references are used, particularly during 2002–2005, which is absent without using them. We carefully compared the 69 splits detected in 2010–2011 with the metadata of the stations, and 56 of them (i.e., 81.2% of break points) can be explained by a change of automatic weather station type in part of the network, from the Almos AWS to the Telmet 320. For instance, dates of the detected break points by Climatol occurred in most cases within a month of the annotated metadata. Yet no clear reason was found for the 2003 maximum detection when NCEP/NCAR references are used,
other than a sharp increment in the number of available data (see Figure 1b) due to a change in the software that allowed AWS to begin reporting DPWG data.

Table 2 displays the statistical summary of applying the SNHT and RMSE metrics to the entire 548 DPWG dataset for 1941–2016, and to the 546 DPWG series without and with the use of the NCEP/NCAR references for 1948–2016. The lower number of detected break points in the first (1941–2016) homogenization compared to the second one (1948–2016), whilst having two more series and seven more years, is associated with the higher SNHT threshold (i.e., 30 instead of the default 25) applied to the overlapping stepped windows in the Stage 1 of the homogenization approach to avoid spurious break point detection at the extremes of the windows. However, statistical summaries of the SNHT and RMSE metrics do not much impact of these lower number of break points when compared to the 1948–2016 homogenization without the use of the NCEP/NCAR references. In contrast, some differences are found when the NCEP/NCAR references are used. For instance, the mean SNHT is lower (i.e., 9.47 with NCEP/NCAR versus 8.59 without), which is expected as DPWG series have been split more (i.e., 401 with NCEP/NCAR versus 384 without). Furthermore, the RMSE of the estimations of the DPWG series when compared to the observed values are also lower, which suggests weaker correlations between the DPWG series and the reanalysis series than with the nearest DPWG series, even in case these series may be more far away.

4.2 | Assessment of the added value of NCEP/NCAR reanalysis series

The use of the NCEP/NCAR reanalysis series as reference in one of the tested homogenization approaches was to compensate for the low correlation coefficients found between sites in the DPWG dataset. Figure 6 shows a rapid degradation of correlation coefficients as distance increases, with \( r \) values <0.5 at distances greater than 1,000 km, being even nearly 0 or negatives for a high proportion of cases. However, Table 3 summarizes that correlation coefficients between DPWG series and the NCEP/NCAR reanalysis ones are consistently lower than with other DPWG series for all combinations and periods. The only exception is the mean correlation in the first 18 years (i.e., 1948–1965), but computed with only four reference series (relatively far apart compared to the much closer grid-cell outputs).

4.3 | Comparison with station metadata

In addition to having checked the metadata associated with the 69 splits detected in 2010–2011 (see Section 4.1), Figure 7 compares the dates of metadata related to changes in wind instrumentation with break points detected in the homogenization of the 35 selected series (see Figure 3 for locations) with and without the support of the NCEP/NCAR reanalysis references. The joint plot of metadata and break point dates shows a general disagreement in both homogenization approaches, which can be attributed to the fact that many changes in wind sensors do not necessarily introduce...
noticeable changes in the measurements (e.g., replacements of masts, vanes or even cups anemometers of the same brand). Furthermore, as noted in Section 3.3, metadata are sparse prior to 1997, and not necessarily complete even after that date. Therefore, comparison of break points with the metadata of these 35 stations is of limited use for the homogenizations.

4.4 | Comparison of DPWG trends

DPWG trends are assessed using the set of 35 long-term stations identified in Section 3.3 (and located in Figure 3). Figure 8 displays the box plots of their OLS trends before (raw series) and after the homogenization without and with the support of NCEP/NCAR reanalysis series, also including trends of the 35 NCEP/NCAR reanalysis grid-cell series closest to the 35 selected stations. For the longest period (i.e., 1948–2016; statistics in Table 4), the major result when applying Climatol to homogenize the observed DPWG series is the reduction in the range of site trends, ranging from \(-0.648\) to \(0.405\) m s\(^{-1}\) dec\(^{-1}\) (i.e., \(1.053\) m s\(^{-1}\) dec\(^{-1}\)) before the homogenization to \(-0.168\) to \(0.004\) m s\(^{-1}\) dec\(^{-1}\) (i.e., \(0.172\) m s\(^{-1}\) dec\(^{-1}\)) after applying it, hence increasing the spatial consistency of those trends. In terms of the statistical significance of the reported DPWG trends, we also found that after applying the homogenization there is an increase in the number of stations showing negative trends, and the proportion with significant negative trends, with 48.6% before (37.1% significant at \(p < 0.05\)) and 97.1% after (77.1% significant at \(p < 0.05\)) the homogenization; the opposite is true for the site positive trends, with a lower percentage of stations and significance after applying the homogenization. In fact, the declining trend of DPWG for the regional series is larger in magnitude after \((-0.073\) m s\(^{-1}\) dec\(^{-1}\)) than before \((-0.005\) m s\(^{-1}\) dec\(^{-1}\)) the homogenization. Furthermore, our assessment of using the NCEP/NCAR reanalysis revealed a strong impact on the magnitude of the reported DPWG trends with both the homogenized series with reanalysis (\(+0.040\) m s\(^{-1}\) dec\(^{-1}\)) and the 35 NCEP/NCAR reference
series (+0.020 m s\(^{-1}\) dec\(^{-1}\)) having an opposite positive tendency of DPWG, with larger percentage of stations showing positive and significant trends than negative ones. For the shorter period (i.e., 1971–2016; statistics in Table 5), the major finding is that differences in the mean trends between the four datasets are lower compared to the longest 1948–2016 period; also showing a reduction in the range of site trends for the homogenized DPWG data compared to the raw one.

### 4.5 Homogenized DPWG dataset across Australia

Results presented in the former sections suggest that the best strategy to homogenize DPWG series is applying a homogenization approach without the NCEP/NCAR reanalysis as references. This is due to the strong influence of reanalysis on the resultant homogenized series, positively biasing the spatiotemporal variability of the resultant DPWG trends. Therefore, the list of the 353 break points detected on the monthly scale without the NCEP/NCAR references was provided to a last application of the Climatol package: (a) split original series at the shift dates; (b) reconstruct all homogeneous subperiods; and (c) infill the DPWG series. Our approach applying the Climatol V3.1 software resulted in the first homogenized and complete DPWG dataset (at daily basis) with 548 series across Australia spanning the 1941–2016 (76-years) period. Running annual means of DPWG monthly averages for the homogenized dataset are shown in Figure 9, together with the raw, homogenized with and without the NCEP/NCAR reanalysis, and the NCEP/NCAR reanalysis itself. It is noteworthy that prior to the ~1970s, large differences between the four datasets exist (i.e., observations vs. reanalysis), which forces to different long-term trends as shown in Table 4.

### 5 DISCUSSION

Previously, homogenization of daily/monthly wind speed series was assessed in five ways, being: (a) checking for a good spatial consistency of trends of the resultant homogenized series, which were also in agreement with trends estimated from independent datasets (Wan et al., 2010); (b) comparison of the detected break points with a list of station relocations (Li et al., 2011); (c) increase of correlation coefficients between candidate and reference series after applying adjustments (Štěpánek et al., 2013); (d) confidence in the reliability of reference series derived from reanalysis or mesoscale climate model outputs (Azorin-Molina et al., 2014); and (e) comparison of homogeneity test statistics before and after the homogenization (Péliné-Németh et al., 2014; Guijarro, 2015; Azorin-Molina et al., 2016). When compared to the previous methodologies to homogenize DPWG (see review in Table 1), the major advantages of our approach using Version 3.1 of the R package Climatol are its ability to: (a) automatically homogenize a large number of series, including short-term ones; (b) use the closest reference series even though they do not have a common observation period with the candidate series or present missing data; and (c) supply homogenized series, correcting

### TABLE 3

| Period     | Years | No. of stations | NCEP/NCAR | 4 references | 10 references |
|------------|-------|-----------------|----------|--------------|---------------|
|            |       |                 |          | Mean \(r\)    | Mean dist. (in km) | Mean \(r\) | Mean dist. (in km) |
| 1948–2016  | 69    | 139             | With     | 0.575         | 108            | 0.546       | 180        |
|            |       |                 | Without  | 0.635         | 227            | 0.577       | 355        |
| 1948–2002  | 55    | 86              | With     | 0.554         | 116            | 0.512       | 204        |
|            |       |                 | Without  | 0.580         | 291            | 0.517       | 463        |
| 1948–1982  | 35    | 58              | With     | 0.547         | 122            | 0.513       | 215        |
|            |       |                 | Without  | 0.592         | 356            | 0.519       | 542        |
| 1948–1965  | 15    | 7               | With     | 0.558         | 131            | 0.512       | 220        |
|            |       |                 | Without  | 0.501         | 626            | 0.532       | 1,244      |
anomalous data (quality control by spatial coherence) and filling in all the missing data. In this study, our novel strategy has been tested using 548 stations across Australia covering the 76-year 1941–2016 study period; with raw DPWG series showing a diversity of spatiotemporal coverage and data characteristics.

The evaluation of the most suitable approaches to homogenize raw observational series (in particular wind speed and peak wind gusts) is challenging in climate research, as true solutions are unknown by the homogenization community, and only benchmarking exercises may serve as a guidance; it thereby justifies, in a way, the different approaches proposed in the last few years (e.g., Wan et al., 2010; Azorin-Molina et al., 2014; Minola et al., 2016; amongst others). In addition, metadata provide information about when changes in instrumentation and observational practices occurred that may produce measurement bias, yet such metadata are often incomplete and can explain only part of the detected inhomogeneities (Dunn et al., 2014). For instance, in this study even in well documented and easily accessible metadata, such as those provided by the Bureau of Meteorology for Australia, quantifying the impact of changes in the surroundings like new buildings or growing trees or the cup anemometer working health (Azorin-Molina et al., 2018c) is challenging. This is because not all metadata produce break points in climate series, and/or not all metadata are reported, especially for older data. In particular, wind observations can be affected by obstructions up to several hundred metres from the observation site, and such distant influences are often not captured in metadata. Laapas and Venäläinen (2017) found the same for wind speed data in Finland, as less than half of the detected inhomogeneities were verified by metadata. This means that homogenization assessment methods solely based on metadata are not able to be implemented herein. Metadata can be crucial for adjusting shifts in series, particularly in cases where only a small number of well-documented stations are involved; this is the case in a series of studies currently being conducted in New Zealand for DPWG in Auckland and Wellington airports (Safaei Pirooz et al., 2018). The main factor that has affected the measurements in these cases at both stations is the instrument and observing practice changeover in the 1990s. By using random process and linear system theory, Safaei Pirooz and Flay (2018) proposed a set of gust-factor ratios that converts the DPWG measurements of the former recording system (Mark II Munro cup anemometer; gust duration ~1 s) to an equivalent 3-s gust recorded by the new anemometers (Vector A101 and Vaisala WAA151). This is a clear example of the importance of reporting metadata for climate studies.

Previous COST Action “HOME” (Venema et al., 2012) and ongoing MULTITEST (Guijarro, 2016) method comparisons have focused their attention on monthly air temperature and precipitation series, and Killick (2016) benchmarked daily air temperature series. Both climatic variables, that is, air temperature and precipitation, are commonly taken as paradigms of variables with quasi-normal
and zero-limited L-shaped Probability Density Functions, respectively (Venema et al., 2012). However, synthetic daily wind databases and inhomogeneities introduced to use them for benchmarking homogenization methods should be carefully prepared to reliably reproduce the inhomogeneities found in real series around the world. Therefore, until this benchmarking of wind data is undertaken, the approaches summarized in Table 1 are the only ones available to assess the homogenization of wind speed observations, with most previous studies only applying basic quality controls.

As wind is highly variable in time and space, quality control and homogenization methods applicable for other climate elements, such as the use of reference sites, may be inappropriate (Jakob, 2010). Azorin-Molina et al. (2014; pp. 3697) stated “in areas of complex topography surrounded by ocean/sea surfaces, wind is not solely driven by surface pressure gradients (i.e., it is also governed by Earth’s surface friction force), and the spatial dependency among observatories can markedly degrade over short distances.” Recently, the influence of network density on homogenization has been

---

**TABLE 4** For all 35 selected stations the mean, maximum, minimum, and range of daily peak wind gust (DPWG) trends (all having units of m s⁻¹ dec⁻¹), and % of positive and negative (and significant) stations with DPWG trends computed from the monthly aggregates of the original (raw) series and the homogenized series with (Homog. + NCEP/NCAR) and without (Homog.) reanalysis references, plus trends of the 35 NCEP/NCAR reference series drawn from the closest grid cell centres to the selected station locations. Statistics are reported for 1948–2016

| Datasets          | Mean (m s⁻¹ dec⁻¹) | Max. (m s⁻¹ dec⁻¹) | Min. (m s⁻¹ dec⁻¹) | Range (m s⁻¹ dec⁻¹) | + trends (% of 35) | + and sig. trends (% of 35) | − trends (% of 35) | − and sig. trends (% of 35) |
|-------------------|-------------------|-------------------|-------------------|---------------------|-------------------|--------------------------|-------------------|--------------------------|
| Raw               | −0.005            | 0.405             | −0.648            | 1.053               | 51.4              | 34.3                     | 48.6              | 37.1                     |
| Homog.            | −0.073            | 0.004             | −0.168            | 0.172               | 2.9               | 0.0                      | 97.1              | 77.1                     |
| Homog. + NCEP/NCAR| 0.040             | 0.219             | −0.126            | 0.345               | 71.4              | 45.7                     | 28.6              | 8.6                      |
| NCEP/NCAR         | 0.020             | 0.147             | −0.058            | 0.205               | 62.9              | 40.0                     | 37.1              | 8.6                      |

**TABLE 5** As in Table 4, but for 1971–2016

| Datasets          | Mean (m s⁻¹ dec⁻¹) | Max. (m s⁻¹ dec⁻¹) | Min. (m s⁻¹ dec⁻¹) | Range (m s⁻¹ dec⁻¹) | + trends (% of 35) | + and sig. trends (% of 35) | − trends (% of 35) | − and sig. trends (% of 35) |
|-------------------|-------------------|-------------------|-------------------|---------------------|-------------------|--------------------------|-------------------|--------------------------|
| Raw               | 0.039             | 0.442             | −0.681            | 1.123               | 54.3              | 40.0                     | 45.7              | 17.1                     |
| Homog.            | 0.006             | 0.148             | −0.083            | 0.231               | 45.7              | 8.6                      | 54.3              | 0.0                      |
| Homog. + NCEP/NCAR| −0.011            | 0.272             | −0.205            | 0.477               | 45.7              | 22.9                     | 54.3              | 28.6                     |
| NCEP/NCAR         | −0.017            | 0.171             | −0.102            | 0.273               | 22.9              | 11.4                     | 77.1              | 25.7                     |

---

**FIGURE 9** Running annual means of daily peak wind gust (DPWG) monthly averages for raw, homogenized with and without the NCEP/NCAR reanalysis across Australia for 1948–2016. The same for the NCEP/NCAR reanalysis but using the mean wind speed (i.e., with lower values compared to DPWG)
quantified (Gubler et al., 2017). For instance, for maximum wind gusts, Brázdil et al. (2017) stated that spatial correlation coefficients between stations decrease more strongly in relation to station distance rather than elevation and are higher (i.e., better) in winter than in summer, although the seasonal results are specific to their study area (the Czech Republic) and could not be automatically assumed to apply in for example, tropical or subtropical climate. For Australia, these factors come into play. Low correlations (i.e., typically $r < 0.60$) effectively compromise the reliability of reference series. Yet previous applications of Climatol (albeit using earlier versions) to the homogenization of wind series, it has been shown that inhomogeneities can be detected despite such low correlations (Azorin-Molina et al., 2016, 2018b). This is because, in general, current homogenization procedures detect changes in the mean, and errors in this parameter are much lower than those of its individual components. In the case of low correlations, missing data in the original series will be infilled with estimated data that may suffer from substantial errors, but their probability distribution is expected to be adjusted to the population characteristics, in the same way as climate model projections try to predict climate rather than day-to-day weather of the forthcoming decades (IPCC, 2014).

The approach demonstrated here identifies trends in monthly data first (using a three-stage iterative process to infill missing records and identify break points in the time series according to a moving window then whole of record approach), then using the identified break points apply the same approach to the daily time series to produce an homogeneous infilled daily time series. The results show that the range of site trends have been substantially reduced, indicating that the homogenized records are potentially more reliable (assuming the homogenization procedure is robust). Moreover, the procedure identifies a regional negative bias an order of magnitude greater than otherwise found. This discrepancy found between raw trends before the homogenization (i.e., $-0.005 \text{ m s}^{-1} \text{ dec}^{-1}$) and after applying our approach (i.e., $-0.073 \text{ m s}^{-1} \text{ dec}^{-1}$) for 1948–2016, can be due to the low density of observations at the beginning of the series (1940s–1960s); e.g., trend differences between raw (i.e., $+0.039 \text{ m s}^{-1} \text{ dec}^{-1}$) and the homogenized (i.e., $+0.006 \text{ m s}^{-1} \text{ dec}^{-1}$) DPWG dataset are of lesser magnitude for the recent 1971–2016 period (see Figure 8 and Table 5).

In addition to the straight application to site data alone, we also assessed the homogenization of DPWG series with and without the support of reanalysis data (i.e., the NCEP/NCAR reanalysis as reference). The results of the two homogenization approaches have been compared by assessing their residual SNHT and RMSE of the estimations of all data from their closest reference series (i.e., neighbouring stations or NCEP/NCAR reanalysis output). Comparison of the break points with wind instrumentation metadata at 35 near-complete stations did not help determine which of the two procedures was optimal. Their DPWG trends before and after homogenization (Figure 8 / Table 4) showed the expected reduction in the dispersion of their values if no reanalysis had been used, with the NCEP/NCAR series showing opposite positive trends in DPWG. These results, together with the low correlations between DPWG series and their closest NCEP/NCAR series, lead us to advocate homogenization without use of the NCEP/NCAR reanalysis output as supporting reference series. The poor performance of the NCEP/NCAR series as reference is because reanalyzed data represent space-averaged and time-averaged wind speeds (i.e., instead of DPWGs), and the shortcomings in the simulation of near-surface layer processes have pointed out by for example, McVicar et al. (2008) and Pryor et al. (2009). This allowed us to perform homogenization to the entire 1941–2016 period (if the NCEP/NCAR reanalysis output was deemed suitable, then homogenization could only commence in 1948 from when these reanalysis outputs are available). Therefore, the dates of the break points detected with this procedure at monthly basis were used to split the daily gust series, which are then homogenized without needing the monthly corrections. The development of this method produced a homogenized dataset of 548 DPWG series with all their missing data infilled for the 76-year 1941–2016 period. The new DPWG database will undoubtedly help to better understand the spatiotemporal variability in the frequency and magnitude of DPWG across Australia. This enhanced understanding is important for several applications as mentioned in Section 1. Future work entails developing relationships between DPWG and average wind speed with the aim of developing DPWG grids at the same daily time step and 0.01° resolution as the already available all-Australian daily average wind speed grids (McVicar et al., 2008).

6 | SUMMARY AND CONCLUSION

The main conclusions of this study can be summarized as follows:

1. The approach for DPWG homogenization is based on the detection of break points on the monthly scale, the splitting of the daily series into homogeneous sub-periods, and their homogenization without needing the monthly corrections.

2. In spite of the low intersite correlations, the homogenization procedure implemented using Climatol Version 3.1 detected 353 break points in the 548 Australian series with DPWG data for 1941–2016. This resulted in an increase in the number of stations showing significant negative trends, with the regional DPWG trend being larger in magnitude after ($-0.073 \text{ m s}^{-1} \text{ dec}^{-1}$) than before ($-0.005 \text{ m s}^{-1} \text{ dec}^{-1}$) the homogenization; and
the range of site trends was significantly reduced indicating increased reliability of estimated trends.

3. Correlations between the observed DPWG series decrease strongly as the distance increases, but the use of the NCEP/NCAR reanalysis series as references does not address this issue adequately. This is because reanalysis series represent mean wind speed (both spatially and temporally) instead of gust winds observed at a station, and it is also related to the quality of the reanalysis.

4. The quality of the detection of break points in DPWG time series cannot be only based on the availability of metadata. Only a small proportion of the detected break points are supported by metadata because not all metadata necessarily produce a shift in the mean, and not all relevant metadata are reported (especially in the early decades). However, metadata served to support the reliability of the Climatol in detecting 69 shifts in 2010–2011, supporting the use of this approach in identifying these inhomogeneities.

5. This study has produced the first quality-controlled, homogenized, and large (548 series) DPWG dataset that will serve for assessing long-term variability and trends of this extreme weather-related hazard across Australia for 1941–2016, with direct implications for many socio-economic and environmental applications.

ACKNOWLEDGEMENTS

C.A.-M. received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 703733 (STILLING project). This work was also supported by the project “Detection and attribution of changes in extreme wind gusts over land” (2017-03780) funded by the Vetenskapsrådet, and the MULTITEST (Multiple verification of automatic software homogenizing monthly temperature and precipitation series; CGL2014-52901-P) project, funded by the Spanish Ministry of Economy and Competitiveness. We acknowledge the Australian Bureau of Meteorology for the provision of the data, in particular to David Sinclair, and the two anonymous reviewers for their detailed and helpful comments to the original manuscript.

ORCID

Cesar Azorin-Molina https://orcid.org/0000-0001-5913-7026
Jose A. Guijarro https://orcid.org/0000-0002-9527-9758
Tim R. McVicar https://orcid.org/0000-0002-0877-8285
Deliang Chen https://orcid.org/0000-0003-0288-5618

REFERENCES

Aguilar, E., Auer, I., Brunet, M., Peterson, T.C. and Wieringa, J. (2003) Guidelines on climate metadata and homogenization. World Meteorological Organisation, 52. Available at: http://www.wmo.int/databook/documents/WCDMP-53_1.pdf [Accessed 1 December 2018].

Alexander, L.V., Tett, S.F.B. and Jonsson, T. (2005) Recent observed changes in severe storms over the United Kingdom and Iceland. Geophysical Research Letters, 32(13), L13704. https://doi.org/10.1029/2005GL022371.

Alexandersson, H. (1986) A homogeneity test to precipitation data. International Journal of Climatology, 6(6), 661–675. https://doi.org/10.1002/joc.3370060607.

Azorin-Molina, C., Vicente-Serrano, S.M., McVicar, T.R., Jerez, S., Sanchez-Lorenzo, A., Lopez-Moreno, J.J., Revuelto, J., Trigo, R.M., Lopez-Bustins, J.A. and Espirito-Santo, F. (2014) Homogenization and assessment of observed near-surface wind speed trends over Spain and Portugal, 1961–2011. Journal of Climatology, 27(10), 3692–3712. https://doi.org/10.1175/JCLI-D-13-00652.1.

Azorin-Molina, C., Guijarro, J.A., McVicar, T.R., Vicente-Serrano, S.M., Chen, D., Jerez, S. and Espirito-Santo, F. (2016) Trends of daily peak wind gusts in Spain and Portugal, 1961-2014. Journal of Geophysical Research—Atmospheres, 121(3), 1059–1078. https://doi.org/10.1002/2015JD024485.

Azorin-Molina, C., Menendez, M., McVicar, T.R., Acvedo, A., Vicente-Serrano, S.M., Cuevas, E., Minola, L. and Chen, D.L. (2018a) Wind speed variability over the Canary Islands, 1948–2014: focusing on trend differences at the land-ocean interface and below-above the trade-wind inversion layer. Climate Dynamics, 50(11–12), 4061–4081. https://doi.org/10.1007/s00382-017-3861-0.

Azorin-Molina, C., Rehman, S., Guijarro, J.A., McVicar, T.R., Minola, L., Chen, D. and Vicente-Serrano, S.M. (2018b) Recent trends in wind speed across Saudi Arabia, 1978–2013. International Journal of Climatology, 38 (Supplement 1), e966–e984. https://doi.org/10.1002/joc.5423.

Azorin-Molina, C., Asin, J., McVicar, T.R., Minola, L., Lopez-Moreno, J.I., Vicente-Serrano, S.M. and Chen, D. (2018c) Evaluating anemometer drift: a statistical approach to correct biases in wind speed measurement. Atmospheric Research, 203(1), 175–188. https://doi.org/10.1016/j.atmosres.2017.12.010.

Ball ing, R.C. and Cerveny, R.S. (2005) Winds and wind systems. In: Oliver, J.E. (Ed.) Encyclopedia of World Climatology, Neth erlands: Springer, pp. 813–819.

Begert, M., Zenklusen, E., Häberli, C., Appenzeller, C. and Klok, L. (2008) An automated procedure to detect discontinuities; performance assessment and application to a large European climate data set. Meteorologische Zeitschrift, 17(5), 663–672. https://doi.org/10.1127/0941-2948/2008/0314.

Brázdil, R., Hostýnek, J., Řezníčková, L., Zahradníček, P., Tolasz, R., Dobrovolný, P. and Štěpánek, P. (2017) The variability of maximum wind gusts in the Czech Republic between 1961 and 2014. International Journal of Climatology, 37(4), 1961–1978. https://doi.org/10.1002/joc.4827.

Caballar-Fuentes, M. (2005) Los temporales de lluvia y viento en Galicia; Propuesta de clasificacion y analisis de tendencias (1961–2001). Investigaciones Geograficas, 36, 103–118 (in Spanish with an English abstract).

Cechet, R.P. and Sanabria, L.A. (2012) Statistical comparison of coincident wind gust measurements from Dines and cup anemometers. 15th Australasian Wind Engineering Society Workshop, Vol. 1, Sydney, Australia, 23–24 February 2012.

Conrad, V. and Pollack, L.W. (1950) Methods in Climatology. Cambridge, MA: Harvard University Press, p. xiv+459.

Cusack, S. (2013) A 101 year record of windstorms in the Netherlands. Climatic Change, 116(3), 693–704. https://doi.org/10.1007/s10584-012-0527-0.

Dunn, R.J.H., Willett, K.M., Morice, C.P. and Parker, D.E. (2014) Pairwise homogeneity assessment of HadISD. Climate of the Past, 10(4), 1501–1522. https://doi.org/10.5194/cp-10-1501-2014.

Fujii, T. (2007) On geographical distributions and decadal changes of the annual maximum wind speeds caused by typhoons in Japan. Journal of Natural Disaster Science, 26(3), 267–277 in Japanese with an English abstract.

Gorman, J. (2004) The conversion equation of the Synoptic 706 anemometer. Instrument test report 677, Bureau of Meteorology. Available at: https://www.wmo.int/pages/prog/www/IMOP/WebPortal-AWS/Tests/TR677.pdf [Accessed 1 December 2018].

Groisman, P. Ya. (2002) Data documentation for dataset TD-6421: enhanced hourly wind station data for the contiguous United States. Version 1.1, National Climatic Data Center, 18 pp.

Gubler, S., Hunziker, S., Begert, M., Croci-Maspoli, M., Konzelmann, T., Brönnimann, S., Schwierz, C., Ortí, C. and Rosas, G. (2017) The influence of observed near-surface wind speed trends over Spain and Portugal, 1961-2011. Journal of Geophysical Research—Atmospheres, 121(3), 1059–1078. https://doi.org/10.1002/2015JD024485.

Guijarro, J.A.. (2008) Homogenization of a dense thermo-pluviometric monthly database in the Balearic Islands using the free contributed R package
Guijarro, J.A. (2016) Automatización de la homogeneización de series climáticas: nuevas funciones del paquete Climatol 3.0. In: Ocina, J., Rico, A.M., y Moltó, E. (Eds.) Clima, sociedad, riesgos y ordenación del territorio. Universidad de Alicante (Spain), Asociación Española de Climatología. ISBN 978-84-16724-19-2, 153–160. (in Spanish with an English abstract).

Guijarro, J.A. (2018) Homogenization of climate series with Climatol. Available at: http://www.climatol.eu/homog_climatol-en.pdf [Accessed 1 December 2018].

Hernández, E.M., García-Valero, J.A., Palenzuela, J.E. and Belda, F. (2012) Homogenization of annual mean wind speed monthly series. In: Kakani, V.G., Frew, R. and Kong, L. (Eds.) Interpolation of missing precipitation data. Wiley, Chichester, pp. 277–292.

Hewston, R. and Dorling, S.R. (2011) An analysis of observed daily maximum wind gusts in the UK. Journal of Wind Engineering and Industrial Aerodynamics, 98(8), 845–856. https://doi.org/10.1016/j.jweia.2011.06.004.

Hunziker, S., Brönnimann, S., Calle, J., Moreno, I., Andrade, M., Ticona, L., Jiang, Y., Luo, Y., Zhao, Z. and Tao, S. (2010) Changes in wind speed over China during 1956–2008. Geophysical Research Letters, 37(1), 205. https://doi.org/10.1029/2009GL041067.

IPCC. (2014) In: Core Writing Team, Pachauri, R.K. and Meyer, L.A. (Eds.) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva: IPCC, p. 151.

Jakob, D. (2010) Challenges in developing a high-quality surface wind-speed data set for Australia. Australian Meteorological and Oceanographic Journal, 60, 227–236.

Jiang, Y., Luo, Y., Zhao, Z. and Tao, S. (2010) Changes in wind speed over China during 1956–2004. Theoretical and Applied Climatology, 99(3–4), 421–430. https://doi.org/10.1007/s00704-009-0152-7.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, W., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, J., Jenne, R. and Joseph, D. (1996) The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society, 77(3), 377–407. https://doi.org/10.1175/1520-0477(1996)077<0377:TNRNP>2.0.CO;2.

Kendall, M.G. and Gibbons, J.D. (1990) Rank Correlation Methods. New York, NY: Oxford University Press, p. 272.

Khaliq, M.N. and Ouarda, T.B.M.J. (2007) On the critical values of the Standard Normal Homogeneity Test (SNHT). International Journal of Climatology, 27(5), 681–687. https://doi.org/10.1002/joc.1438.

Killick, R.E. (2016) Benchmarking the performance of homogenisation algorithms on daily temperature data. PhD Thesis, University of Exeter, 249 pp. Available at: http://hdl.handle.net/10871/23095 [Accessed 1 December 2018].

Klink, K. (2015) Seasonal patterns and trends of fastest 2-min winds at coastal stations in the conterminous USA. International Journal of Climatology, 35(14), 4167–4175. https://doi.org/10.1002/joc.4275.

Kruger, P.A., Goliger, A.M., Retief, J.V. and Sekele, S. (2010) Strong wind climatic zones in South Africa. Wind and Structures, 13(1), 37–55. https://doi.org/10.1299/was.2010.13.1037.

Laapas, M. and Venäläinen, A. (2017) Homogenization and trend analysis of monthly mean and maximum wind speed time series in Finland, 1959–2015. International Journal of Climatology, 37(14), 4803–4813. https://doi.org/10.1002/joc.5124.

Lakatos, M., Szentimrey, T., Bihari, Z. and Szalai, S. (2011) Homogenization of daily data series for extreme climate indices calculation. In: Proceedings of COST-ES0601 (HOME) Action Management Committee and Working Groups and Sixth Seminar for Homogenization and Quality Control in Climatological Databases, Budapest, 26–30 May 2008. WDCMD-No. 76, WMO/TC-No. 1576, pp. 100–109.

Li, Z., Yan, Z., Tu, K., Liu, W. and Wang, Y. (2011) Changes in wind speed and extremes in Beijing during 1960–2008 based on homogenized observations. Advances in Atmospheric Sciences, 28(2), 408–420. https://doi.org/10.1007/s00703-010-0018-z.

Luna, M.L., Guijarro, J.A. and López, J.A. (2012) A monthly precipitation database for Spain (1851–2008): reconstruction, homogeneity and trends. Advances in Science and Research, 8(1), 1–4. https://doi.org/10.5194/asr-8-1-2012.

Mamara, A., Argiriou, A.A. and Anadramistakis, M. (2013) Homogenization of mean monthly temperature time series of Greece. International Journal of Climatology, 33(12), 2649–2666. https://doi.org/10.1002/joc.3614.

McVicar, T.R., Van Niel, T.G., Li, L.T., Roderick, M.L., Rayner, D.P., Ricciardulli, L. and Donohue, R.J. (2008) Wind speed climatology and trends for Australia, 1975–2006: capturing the shifting phenomenon and comparison with near-surface reanalysis output. Geophysical Research Letters, 35(20), L20403. https://doi.org/10.1029/2008GL035627.

McVicar, T.R., Roderick, M.L., Donohue, R.J., Li, L.T., Van Niel, T.G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N.M., Mescherskaya, A.V., Kruger, A.C., Rehman, S. and Dnupashoh, Y. (2012) Global review and synthesis of trends in observed terrestrial near-surface wind speeds: implications for evaporation. Journal of Hydrology, 416, 182–205. https://doi.org/10.1016/j.jhydrol.2011.10.024.

Mestre, O., Domonkos, P., Picard, F., Auer, I., Robin, S., Lebarbier, E., Boehm, R., Aguilar, E., Guijarro, J., Vertachnik, G., Klanar, M., Dubusson, B. and Štěpánek, P. (2013) HOMER: a homogenization software—methods and applications. Időjárás, 117(1), 47–67.

Minola, L., Azorin-Molina, C. and Chen, D. (2016) Homogenization and assessment of observed near-surface wind speed trends across Sweden, 1956–2013. Journal of Climate, 29(20), 7397–7415. https://doi.org/10.1175/JCLI-D-15-0636.1.

Paulhus, J.L.H. and Kohler, M.A. (1952) Interpolation of missing precipitation records. Monthly Weather Review, 80(8), 129–133. https://doi.org/10.1175/1520-0493(1952)080<0129:IOMPR>2.0.CO;2.

Péliné-Németh, C., Radics, K. and Bartholy, J. (2011) Seasonal variability of wind climate in Hungary. Acta Silvatica & Lignaria Hungarica, 7, 39–48.

Péliné-Németh, C., Bartholy, J. and Pongrácz, R. (2014) Homogenization of Hungarian daily wind speed data series. Időjárás, 118(2), 119–132.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincint, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Firlend, E.J., Hansen-Bauer, I., Alexanderson, H., Jones, P.D. and Parker, D. (1998) Homogeneity adjustments of in situ atmospheric climate data: a review. International Journal of Climatology, 18(13), 1493–1517. https://doi.org/10.1002/(SICI)1050-0448(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T.

Petrović, P., Curley, M., Eíreann, M. and Toši, I. (2008) Detected inhomogeneities in wind direction and speed data from Ireland. In: Proceedings. Meeting of COST-ES0601 (HOME) Action Management Committee and Working Groups and the Sixth Seminar for Homogenization and Quality Control in Climatological Databases, Hungarian Meteorological Service, Budapest, Hungary, 28pp.

Pirazolli, P.A. and Tomasin, A. (2003) Recent near-surface wind changes in the Central Mediterranean and Adriatic areas. International Journal of Climatology, 23(8), 963–973. https://doi.org/10.1002/joc.925.
