



A comparison of the MATCHES and NCEP1 databases for use in Australian east coast low studies

J.L. Gray^{a,*}, J.B.D. Jaffrés^b, D.C. Verdon-Kidd^c, M.G. Hewson^d, J.M. Clarke^e, A. Pepler^f, N. B. English^a

^a School of Health, Medical and Applied Sciences, Central Queensland University, Townsville, Queensland, 4810, Australia

^b C&R Consulting, Townsville, Queensland, 4814, Australia

^c School of Environmental and Life Sciences, University of Newcastle, Callaghan, New South Wales, 2308, Australia

^d School of Education and the Arts Central Queensland University, Rockhampton, Queensland, 4700, Australia

^e CSIRO Climate Science Centre, Aspendale, Melbourne, Victoria, 3195, Australia

^f Australian Bureau of Meteorology, Sydney, New South Wales, 2000, Australia

ARTICLE INFO

Keywords:

East coast lows (ECLs)

Laplacian method

MATCHES

NCEP1

Storms

Eastern seaboard

ABSTRACT

For Australia, there are only a few east coast low (ECL) databases that have been generated to explore aspects of ECL development, movement and subsequent impacts. Improved databases that include ECL track data will enhance future forecasting and damage mitigation on the east coast of Australia. This paper compares ECL track characteristics of a new low-pressure dataset, NCEP1 (1950–2019), to the recently updated MATCHES (Maps and Tables of Climate Hazards of the Eastern Seaboard) database (1950–2019) in order to identify similarities and differences of track characteristics that may be important for future ECL research. To achieve this, defining parameters such as intensity – used to make the MATCHES database – were applied to NCEP1 to ensure a direct comparison of historical ECL events. Although both databases display similar patterns in ECL seasonality and track characteristics, we show that the NCEP1 database identifies additional events not captured in MATCHES and provides improved track morphology of certain well-known historical events (such as the 2007 Pasha Bulker storm and the 1998 Sydney to Hobart Yacht Race). Importantly, this research builds upon Australian ECL research and notes an improvement on the MATCHES database, with NCEP1 offering an almost two-fold out-performance in storm tracking (track length and duration) and greater spatial coverage outside the traditional ECL box.

1. Introduction

An east coast low (ECL) is defined as a cyclone (closed pressure system) that occurs along the southeastern coastline of Australia and is generated from tropical or mid-latitude influences at various levels within the atmosphere (Callaghan and Helman, 2008; Di Luca et al., 2016; Dowdy et al., 2011, 2013a, 2013b, 2013c, 2019; Holland et al., 1987; Kiem et al., 2016; Pepler et al., 2014; Pepler and Coutts-Smith, 2013; Speer et al., 2009; Verdon-Kidd et al., 2010, 2016). Strong ECLs

can have a range of negative impacts, including strong winds, heavy rainfall, strong waves, flash flooding and coastal erosion (Callaghan and Helman, 2008; Dowdy et al., 2011, 2019; Kiem et al., 2016; Speer et al., 2009). However, they are also a key water contributor to freshwater catchments in eastern Australia (Pepler and Rakich, 2010). For example, in June 2007, an ECL event recharged the Greater Sydney dam storage from 33.5% to 55.0% in two weeks (WaterNSW, 2020; Watson et al., 2007).

ECLs form between 25°S and 40°S throughout the year, increasing in

Abbreviations: ECL, East coast low; BOM, Australian Bureau of Meteorology; MATCHES, Maps and Tables of Climate Hazards of the Eastern Seaboard; NCEP, National Centres for Environmental Prediction; NNR1, National Centre for Atmospheric Research Reanalysis 1; 20CRv2c, Twentieth Century Reanalysis; ECMWF, European Centre for Medium Range Weather Forecasts; JRA55, Japanese 55-year reanalysis; MLAP, Maximum Laplacian; CSI, Critical success index; UM, University of Melbourne; CL, Continental lows; IT, Inland trough; MSLP, Mean sea level pressure; MSLPmin, Minimum mean sea level pressure.

* Corresponding author.

E-mail address: j.gray2@cqu.edu.au (J.L. Gray).

<https://doi.org/10.1016/j.wace.2021.100400>

Received 30 April 2021; Received in revised form 12 October 2021; Accepted 29 November 2021

Available online 30 November 2021

2212-0947/© 2021 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

frequency during austral autumn and winter. Heightened periods of activity are influenced by frontal and trough systems. ECL sub-types are specified in [Speer et al. \(2009\)](#), [Browning and Goodwin \(2013\)](#) and [Kiem et al. \(2016\)](#) and include: 1.) inland troughs; 2.) easterly troughs; 3.) frontal lows; 4.) westerlies; 5.) ex-tropical cyclones; and 6.) rapidly intensified lows, also known as bombs. Initial ECL development occurs at the surface from strong baroclinicity (temperature gradient) in the mid to upper tropospheric levels of the atmosphere. The downward extension of the system's development is influenced by lower pressure at or near the surface ([Holland, 1997](#); [Holland and McBride, 1997](#)). During ECL formation, the low-pressure system's western side transports southern-derived cooler air north towards the equator, whilst warmer air on the eastern side is directed south. These systems intensify due to a strong temperature gradient/boundary between the two sides of the low and cloud forcing from coastal orography (e.g. Great Dividing Range; [Holland, 1997](#)). However, the cyclonic development of ECLs is not restricted to the generic cold-core system. Research by [Cavicchia et al. \(2019, 2020\)](#), [Hart \(2003\)](#) and [Quinting et al. \(2019a, 2019b\)](#) has demonstrated the baroclinic characteristics and development of sub-tropical and hybrid ECLs. Hybrid systems have a symmetrical warm core in the lower troposphere, coupled with a cold-core in the upper-troposphere ([Cavicchia et al., 2019, 2020](#); [Hart, 2003](#); [Quinting et al., 2019a, 2019b](#)). Specific types of hybrid ECLs consist of subtropical cyclones ([Evans and Guishard, 2009](#); [Hart, 2003](#); [Quinting et al., 2019a, 2019b](#)), a TC that has undergone extra-tropical transition ([Evans and Guishard, 2009](#); [Hart, 2003](#); [Quinting et al., 2019a, 2019b](#)), and warm-seclusion cyclones ([Hart, 2003](#); [Quinting et al., 2019a, 2019b](#); [Shapiro and Keyser, 1990](#)).

1.1. Australian ECL databases

The interest in ECLs in Australia has led to the development of multiple methods and databases that identify and record events for research and operational purposes ([Table 1](#)). Early ECL databases such as [Callaghan and Helman \(2008\)](#) and [Callaghan and Power \(2014\)](#) used hazard- and impact-based procedures to track storm occurrence with respect to high rainfall events that lead to flooding along the east coast. Databases such as the New South Wales Maritime Low Database (1970–2006) were constructed by identifying events that formed within a specified vicinity of the eastern coastline (the ECL box; [Dowdy et al., 2019](#); [Pepler and Coutts-Smith, 2013](#); [Speer et al., 2009](#)). Although these databases rely on manual and subjective identification of synoptic features, they are still useful as a validation tool for ECL events in more recent databases ([Dowdy et al., 2019](#)).

More recent ECL databases have been developed from reanalysis-model datasets that are regularly updated. Within Australia, reanalysis-based approaches include the Laplacian method ([Pepler and Coutts-Smith, 2013](#)), pressure gradient method ([Browning and Goodwin, 2013, 2016](#)) and upper-level geostrophic vorticity method ([Dowdy et al., 2011, 2013](#)). Briefly, the Laplacian and pressure gradient methods combine gridded reanalysis dataset fields (i.e. mean sea-level pressure [MSLP] or geopotential height), and then use various tracking algorithms to identify and create a track for ECL events ([Dowdy et al., 2019](#); [Speer et al., 2009](#)). This track is then added to the database. The upper-level geostrophic vorticity method uses various isentropic (directional) and isobaric (pressure) levels within the middle and upper troposphere to identify lows with strong cyclonic vorticity ([Dowdy et al., 2019](#)). Each method's focus explores a developmental property of lows that could be further used to define the track and characteristics of an ECL (e.g. the pressure gradient method tracks a closed low in the surface pressure of the atmosphere). More extensive reviews of Australian ECL database construction methods are outlined in [Pepler et al. \(2015\)](#), [Dowdy et al. \(2019\)](#) and summarised in [Table 1](#). As evident by the temporal coverage (cf. [Table 1](#)), the more extensive databases are the Maps and Tables of Climate Hazards of the Eastern Seaboard (MATCHES; 1950–2019) and a new low-pressure database named

NCEP1 ([Pepler, 2020a](#); 1950–2019) derived from the National Centre for Atmospheric Research (NCEP–NCAR) reanalysis-1 (NNR1)¹. The more recent databases, MATCHES and NCEP1, use the Laplacian method exclusively to identify ECL development and tracking.

1.2. The Laplacian method and datasets

The Laplacian method applies the procedure from the University of Melbourne (UM) cyclone tracking scheme ([Murray and Simmonds, 1991](#)). This method detects lows from local maxima of the Laplacian of MSLP ($\nabla^2 p$) and indicates cyclonic vorticity (curvature) – before identifying an associated closed or open low in the MSLP and connecting lows into tracks. The average Laplacian indicates the intensity of the low within a certain radius of the cyclone centre. ECLs are classified using this method by identifying the minimum intensity of 0.25 hPa (deg. lat.)⁻² (0.3 hPa [deg. Lat.]⁻² for extreme events), calculated over a 5° radius.

The most commonly known ECL database constructed from the Laplacian method is MATCHES ([Coutts-Smith et al., 2012](#); [Pepler and Coutts-Smith, 2013](#); [Pepler et al., 2016](#)) that employs the [Jones and Simmonds \(1993\)](#) implementation of the UM tracking scheme. Spanning from 1950 to 2019, the construction of MATCHES combines pressure-based reanalysis data in an interactive user interface. To create the MATCHES database, the LAPv1 tracking algorithm (highlighted in [Pepler et al., 2015](#)) was applied to the NNR1 MSLP dataset ([Coutts-Smith et al., 2012](#); [Pepler et al., 2015](#)), with low-pressure systems required to be closed and occur within an ECL boundary comparable to [Speer et al. \(2009\)](#). Similarly, [Pepler \(2020a\)](#) has constructed NCEP1, which conversely employs the LAPv2 procedure highlighted in [Pepler et al. \(2015\)](#). Importantly, NCEP1 includes ECLs and low-pressure systems throughout the southern hemisphere, with a focus on the Australian continent and surrounding ocean. We refer to this NCEP1 subset from [Pepler \(2020a\)](#) as an ECL database in this study despite the more extensive content.

Other than MATCHES and NCEP1, [Pepler \(2020a\)](#) also applied the Laplacian method on a range of other reanalysis datasets (JRA55, ERAI and ERA5) that are inclusive of low-pressure events up to 2019 ([Table 1](#)) and could also be used for ECL studies. However, due to their shorter temporal coverage, they have not been investigated in this study.

1.3. Study premise and aims

A disadvantage for most ECL databases is their dependence on updates to the reanalysis datasets and manual application of the cyclone tracking algorithms to remain current. Regularly updating ECL databases so that they are inclusive of the most recent events, both geographically and temporally, will ensure risk and resource management protocols are adequate for the future. However, ECL events that have occurred after December 2019 were not included in any of the available ECL databases listed in [Table 1](#) at the time of this study.

Until now, the NCEP1 database has not been applied to explore the possible advantages of ECL track characteristics compared to MATCHES. The objectives of this paper are to 1.) identify ECLs in the NCEP1 database, 2.) compare ECL tracks within the MATCHES and the NCEP1 databases through a range of temporal and spatial analysis and 3.) evaluate the difference in the skill of the MATCHES and NCEP1 databases in reproducing track characteristics of notable ECL events. Thus, this study aims to assess if the track characteristics of storms in the NCEP1 database 1.) are comparable to MATCHES and 2.) offer an

¹ NNR1: NCEP-NCAR reanalysis-1 is the original reanalysis dataset used by [Pepler \(2020a\)](#) to construct NCEP1. NCEP1 database: This is the database of low-pressure systems in the Southern hemisphere created by [Pepler \(2020a\)](#). NCEP1_{ECL} subset: This is a filtered subset of low-pressure systems identified as ECLs in the NCEP1 database ([Pepler, 2020a](#)) used for this study.

Table 1
Summary of ECL databases for Australia.

Database/ Method	Method of identifying ECLs	Author	Year	Coverage within the study	Reanalysis database variations	Examples of studies	Data accessibility
Callaghan and Helman	A collection of severe storm occurrences from Cape York to Tasmania, composed of a range of analogue sources.	Callaghan and Helman (2008)	2008	1770–2008	The database is constructed from various sources such as ship logs, newspapers, governmental and bureau reports, weather forecasts and reports of storm occurrences within colonial journals.	Callaghan and Helman (2008); Pepler et al. (2016)	Full database accessed from Callaghan and Helman (2008).
Callaghan and Power	A collection of ECLs that have influenced significant flood events from Brisbane to Eden, New South Wales (NSW). Four categories of ECLs are determined based on formation characteristics. Information was sourced from the Australian Bureau of Meteorology (BOM) website and analogue resources.	Callaghan and Power (2014)	2014	1770–2014	The database is constructed from various sources such as ship logs, newspapers, governmental and bureau reports, weather forecasts and reports of storm occurrences within colonial journals.	Callaghan and Helman (2008); Callaghan and Power (2014, 2011); Pepler et al. (2016); Power and Callaghan (2016)	Full database accessed from the supplementary appendix in Callaghan and Power (2015).
East Coast Low Database Southeast Australia	Manually interpreted MSLP charts, reinforced with satellite imagery.	Speer et al. (2021, 2009)	2021	1970–2019	This is an updated version of the NSW Maritime Low Database used in Speer et al. (2009). Data are grouped on the seven synoptic categories outlined in Speer et al. (2009) and their west or east wind orientation from April to September. Other information includes initial date, latitude and longitude, initial central pressure and lowest pressure (maximum intensity; Speer et al., 2021).	Speer et al. (2021)	The updated April–September ECL database 1970–2019 is available at: https://zenodo.org/record/4137957#.YacTBNBBxPY
NSW Maritime Low Database	Manually interpreted MSLP charts, reinforced with satellite imagery.	NSW Climate Services of BOM.	2009	1970–2006	The spatial domain consists of the Australian coast to 160°E and 25°S–40°S. This database contains information on related impacts, dates and positions of low-pressure systems – and is considered subjective based on the identification method for ECLs (Pepler and Coutts-Smith, 2013).	Pepler et al. (2014); Pepler and Coutts-Smith (2013); Speer et al., (2009, 2021)	Accessed by contacting the BOM.
Pressure Gradient Method	The three-stage algorithm is applied to a database, objectively identifying low-pressure systems with a closed cyclonic centre based on surface pressure. Backtracking of the lows to their initial genesis point or the nearest trough is then applied. Lows are classified from their pre-storm directory into ECL subtypes.	Browning and Goodwin (2016, 2013)	2013	1979–2019 1851–2014	The European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis dataset (ERA-Interim) is a global atmospheric model that has a 1.5° latitude/longitude grid of sub-daily, daily and monthly data in, air temperature, geopotential height and precipitation (Berrisford et al., 2011). The Twentieth Century Reanalysis (20CRv2c) has a spatial resolution of 2° latitude/longitude made from 56 ensemble members. 20CRv2c includes MSLP, sea surface temperature and sea ice boundary (Compo et al., 2011).	Browning and Goodwin (2016, 2013); Di Luca et al. (2016); Kiem et al. (2016) Browning and Goodwin (2016); Ji et al. (2018)	ERA-Interim: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim 20CRv2c: https://psl.noaa.gov/data/gridded/data.20thC_ReaV2c.html

(continued on next page)

Table 1 (continued)

Database/ Method	Method of identifying ECLs	Author	Year	Coverage within the study	Reanalysis database variations	Examples of studies	Data accessibility
The Laplacian Method	Lows are detected from maxima in the Laplacian of MSLP (an indication of curvature) before an associated closed low centre is identified, which is then joined into tracks using a probability matching function. The method employs the tracking method from the University of Melbourne cyclone tracking scheme by Murray and Simmonds (1991) and Simmonds and Murray (1999).	Pepler and Coutts-Smith (2013)	2013	1989–2011	ERA-Interim	Ji et al. (2018); Pepler et al. (2021), 2015	ERA-Interim: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim
				1851–2014	20CRv2c	Ji et al. (2018); Pepler et al. (2015)	20CRv2c: https://psl.noaa.gov/data/gridded/data.20thC_ReanV2c.html
				1950–2019	Maps and Tables of Climate Hazards of the Eastern Seaboard (MATCHES) is derived from the National Centres for Environmental Prediction (NCEP) reanalysis. The grid is 2.5° at 6-hourly intervals. This database is used to analyse intensity, decadal variability, frequency and distribution along the east coast. The spatial domain covers the Australian coast to 160°E and 25°S–40°S (Kalnay et al., 1996). MATCHES is considered an objective ECL database.	Ji et al. (2018); Pepler and Coutts-Smith (2013); Pepler et al. (2015)	MATCHES: Accessed by contacting the BOM. NCEP: https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html
		Pepler (2020a, b)	2020	1950–2019	Similar to MATCHES, NCEP1 is derived from the NCEP (at 2.5°). The NCEP reanalysis dataset uses a polar projection, with a resolution of approximately 1.5° at 30°S. The spatial domain of the database covers the area from 10°S to 80°S and 80°E to 190°E (i.e. the Australian continent and surrounding ocean).	Pepler (2020a,b)	NCEP1, JRA55, ERA5 subsets used in Pepler (2020a) can be accessed from Pepler (2020b): https://figshare.com/collection/s/Australian_region_cyclones_1950-2019/4944135/1 JRA55: https://climatedataguide.ucar.edu/climate-data/jra-55
				1958–2019	The Japanese 55-year reanalysis (JRA55) is a global atmospheric dataset, based on four-dimensional variation analysis (Kobayashi et al., 2014, 2015). Data are organised in 6-hourly time-steps in 60 levels up to 0.1 hPa.	Pepler (2020a,b); Pepler and Dowdy (2021)	ERA5: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5
				1979–2019	ERA5 is a global reanalysis dataset projected on a 30 km grid, with a vertical resolution of 137 levels at 0.1 hPa generated every hour (Hersbach et al., 2020). This dataset was designed to replace ERA-Interim.		
Upper-Level Geostrophic Method	ECLs are identified based on the 48-h running maxima of 500 hPa geostrophic vorticity in a region centred to the northwest of the area used for surface lows.	Dowdy et al. (2011)	2011	1989–2011	ERA-Interim	Dowdy et al. (2013a, 2013b, 2013c, 2011); Ji et al. (2018)	ERA-Interim: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim

improvement on currently available ECL track databases. This paper explores subtle but important differences and advantages of these recent datasets, mainly in relation to the quantity and quality of ECL storm tracks.

2. Data and methods

2.1. ECL databases

Both MATCHES and NCEP1 databases use the global NNR1 dataset. The NNR1 dataset is restricted to only include low-pressure systems or

ECLs that meet the desired specifications of the study (i.e. MATCHES: east coast of Australia/ECL box; NCEP1: Southern Hemisphere, focused on the entire Australian region).

2.1.1. MATCHES database

The ECLs in the MATCHES (1950–2019; Coutts-Smith et al., 2012; Pepler and Coutts-Smith, 2013; Pepler et al., 2016) database have six-hourly central pressure and locational data and are identified using the Jones and Simmonds (1993) version of the UM cyclone tracking scheme. The database covers the period of 1950–2019.

2.1.2. NCEP1 subset

The second database (NCEP1_{ECL}) used in this study is derived from the NCEP1 low-pressure database (1950–2019; Pepler 2020a; 2020b; available at <https://doi.org/10.6084/m9.figshare.c.4944135.v1>). NCEP1 employed the Simmonds et al. (1999) version of the UM tracking scheme, which first redistributes the unfiltered NCEP1 pressure fields to a polar projection, with a resolution of approximately 1.5° at 30°S (Pepler, 2020a). The spatial domain covers approximately from 10°S to 80°S and 80°E to 190°E (i.e. the Australian continent and surrounding ocean; Fig. S1) and spans the period 1950 to 2019.

2.2. ECL identification

To fairly compare the performance of each ECL database, the same criteria were applied to both MATCHES and NCEP1_{ECL}. ECL events in both databases are distinguished from regular lows using the following criteria outlined in previous ECL literature:

- The event must be a closed low (Browning and Goodwin, 2013; Dowdy et al., 2013a; Pepler et al., 2014; Speer et al., 2009).
- The event must have a maximum Laplacian (MLAP) value of 0.25 or greater (Pepler, 2020a; Pepler et al., 2014, 2015; Pepler and Coutts-Smith, 2013; Speer et al., 2009)
- The event has at least one track point within the ECL box (25°S to 40°S and eastern coastline to 160°E) outlined by Speer et al. (2009).
- Only tracks that had points with Laplacian values of 0.25 or greater in the ECL box were included.

2.3. ECL track verification

Using the method and equation (Eqn. (2)) in Dowdy et al. (2011) and Pepler et al. (2015), a critical success index (CSI) was applied to MATCHES and NCEP1_{ECL} to verify whether similar events were present in both datasets.

$$CSI = \frac{Hits}{Hits + Misses + False\ Alarms} \quad \text{Eqn.2}$$

Events in the NCEP1_{ECL} database were paired with those present in MATCHES for the 1950–2019 period. Corresponding events of the MATCHES and NCEP1_{ECL} were paired with respect to maximum intensification time (timing of maximum Laplacian value), latitude and longitude. This process was completed twice, with corresponding NCEP1_{ECL} events identified within MATCHES and vice versa. Results from each test were investigated to see if the identified pairing was the same in both cases. A numerical ID system for each probable pairing scenario was derived. These included the following: a.) events present only in NCEP1_{ECL}, not MATCHES (classified as false alarms in NCEP1_{ECL} and misses in MATCHES); b.) events present only in MATCHES, not NCEP1_{ECL} (defined as false alarms in MATCHES and misses in NCEP1_{ECL}); and c.) paired events within both datasets (hits in both).

Cases that required manual intervention involved track investigation through a range of variables. These variables include the time of occurrence, MLAP value, time at MLAP, the distance between the paired tracks, spatial extent, longitude, latitude and, when available, MSLP charts from the Australian Bureau of Meteorology (BOM) archive (BOM, 2021). In addition, some mismatches in pairings revealed that certain tracks in MATCHES were affected by splitting into several segments, an example of which is given in Fig. 1. When manually merging these tracks, several factors were considered, including the distance between two temporally adjacent data points (i.e. last and first track points of two consecutive tracks, respectively) and the tracking behaviour immediately before and after the split. For the CSI calculation, the tracks were sorted into hits, misses and false alarms.

2.4. ECL track length and duration

The differences in track length (km) and duration (days) in the MATCHES and NCEP1_{ECL} databases were examined. The length travelled by the ECL was calculated using the haversine equation (outlined in Laurencin and Misra, 2019). Briefly, the consecutive distance between coordinates was derived and summed together to produce an overall track length. Differences in the duration of ECLs in the MATCHES and NCEP1_{ECL} databases were examined with a paired Wilcoxon signed-rank test (a non-parametric version of a *t*-test).

2.5. ECL prevalence across the spatial domain

The frequency and location of ECLs (defined by the criteria in Section 2.2) in the two databases were compared using raster density plots. This

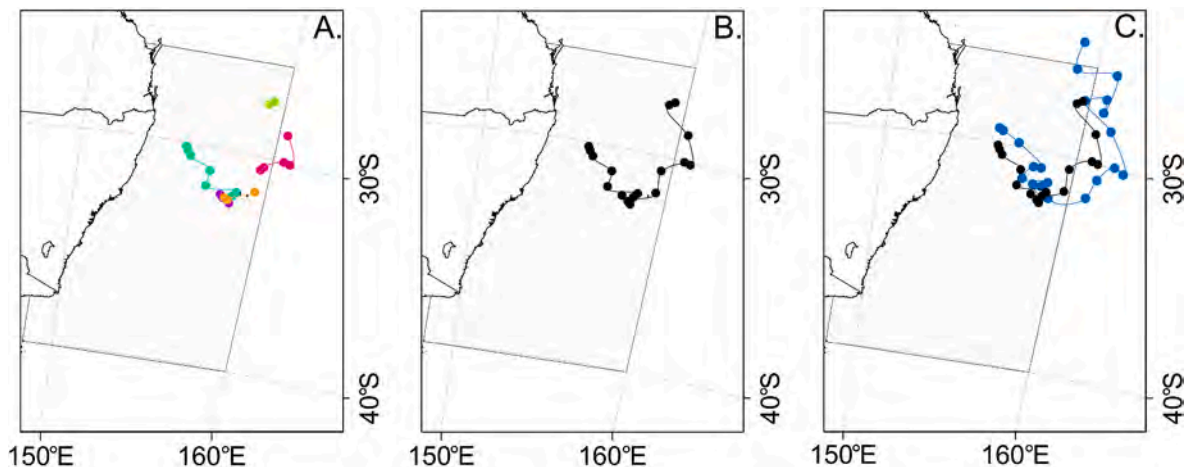


Fig. 1. A.) An example from MATCHES showing the "split" nature of some tracks in the database. B.) Multiple tracks from A.) have been merged into one MATCHES track. C.) The merged MATCHES track compared with its paired NCEP1_{ECL} track.

analysis used the raster count and mapping function in ArcGIS at a cell size of 52 km by 52 km using the Geoscience Australia Lambert 1994 projection (GDA 1994). This procedure was applied to the NCEP1_{ECL} and MATCHES databases, generating a spatial representation of the frequency of all ECLs between 1950 and 2019 at each grid cell location. Each data point within the track was counted once by classifying the raster feature from the track ID (after split tracks had been merged).

2.6. Long-term and seasonal trends in the ECL databases

Using the method outlined in Gray et al. (2020), monthly spaghetti plots of ECL tracks were created to compare MATCHES and NCEP1_{ECL} from 1950 to 2019. Monthly seasonality was compared through a box-plot distribution and a timeseries of the total number of ECLs (per month and per season). The statistical differences between the two databases were tested using a Wilcoxon signed-rank test (for linear trends) and a two-sample Kolmogorov-Smirnov test (difference in distribution; Corder and Foreman, 2014; NCSS Statistical Software, 2020).

A CUSUM distribution-free test was applied to the seasonal timeseries (three-month period; Fig. S7) using the TREND eWater Toolkit (Chiew and Siriwardena, 2005). The CUSUM method split each timeseries into the pre- and post-satellite period, combined with 1000 bootstrapping iterations for each sample. A negative (positive) test statistic indicates that the mean of the latter section is higher (lower), suggesting an increasing (decreasing) trend (Chiew and Siriwardena, 2005).

2.7. Case studies

Two examples of well-known ECLs were chosen to compare and contrast the individual tracks represented in MATCHES and NCEP1_{ECL}. The Pasha Bulker storm has been extensively examined (Callaghan and Helman, 2008; Verdon-Kidd et al., 2010, 2016) and highlighted by the media because of the extensive damage it caused. The 1998 Boxing Day ECL during the Sydney to Hobart Yacht Race is included as a comparison due to its widespread media coverage, the associated fatalities (Australian Institute for Disaster Resilience, 2020; Greenslade, 2001; Mills, 2001) and its more unusual, relatively southern location. The ECLs that corresponded to the same date and location were retrieved and compared from both databases. An unequal variance *t*-test was applied to determine if the mean track distribution of each storm was statistically significantly different (Sharma et al., 2019). This test took the mean of each paired set of latitude and longitude points and compared them at a significance threshold of 0.05. The minimum MSLP (MSLP_{min}) of each of these four tracks was compared with synoptic charts from the BOM (1998, 2007, 2021) and Buckley and Leslie (2000).

3. Results

3.1. ECL track comparisons

The NCEP1_{ECL} has an overall total of 1996 tracks over the

Table 2
Number of tracks, average track length and track duration for MATCHES and NCEP1_{ECL} databases. Data denoted by the asterisk (*) are for NCEP1_{ECL}.

	MATCHES	NCEP1
Number of unfiltered tracks	3889	56361
Number of tracks in each subset (filtered for ECL, unmerged)	2158	1996*
Number of tracks (merged)	1928	1996*
Mean track duration (days)	2.49	4.27*
Mean track length (km)	1782.31	3690.76*
Mean max Laplacian	0.48	0.70*
Mean track duration in ECL box (days)	1.013	1.088*
Mean track length in ECL box (km)	566.84	679.37*
Mean max Laplacian of tracks in the ECL box	0.44	0.54*

1950–2019 period (Table 2). Before the track merging, MATCHES had 2158 individual tracks (inclusive of smaller events with two data points). However, this decreased to 1928 tracks as 230 were merged through our validation and cross-reference process. NCEP1_{ECL} tracks are more extensive in duration, length and power than those in MATCHES (overall – and more subtly within the ECL box; Table 2).

The results of the events pairing from 1950 to 2019 in the MATCHES and NCEP1_{ECL} were investigated using a CSI (Table 3). The CSI from 1950 to 2019 for MATCHES was higher (80.29%) than for NCEP1_{ECL} (77.92%). This can be attributed to the larger number of tracks in NCEP1_{ECL}. NCEP1_{ECL} had an equal number of hits with an increased number of misses and false alarms than MATCHES (Table 3). A CSI score was calculated for the pre-satellite (1950–1978) and the post-satellite (1979–2019) period to ensure there was no bias (Table S1). Both periods showed a negligible difference when compared. Importantly, tracks categorised as false alarms may be correctly identified tracks that were missed by the other dataset. However, they were not included as a hit for the CSI. For our CSI classification, a false alarm in NCEP1_{ECL} is a miss in MATCHES, and a false alarm in MATCHES is a miss in NCEP1_{ECL}.

Basic characteristics of the false alarms and misses (an event present in one database but not in the other) were investigated with respect to the CSI in order to see if there were prominent differences between each database (Fig. 2 and Table 4). False alarms and misses were more (less) prominent in NCEP1_{ECL} (MATCHES), with 449 (381) tracks being classified during the CSI process. Track length, duration and MLAP values are shorter/lower for missed and false alarm tracks when compared to the hits (events that occur within both databases; Table 4). For both categories (hits vs false alarms/misses), NCEP1_{ECL} tracks were more extensive (longer and more powerful) than those in MATCHES, in agreement with the results for the full datasets (Table 2). False alarms in the NCEP1_{ECL} increase between July to November. Conversely, in MATCHES, more false alarms were evident in June and December (Fig. 2; Table S6).

The distributions for track length and track duration for both datasets are right-skewed (Fig. 3). MATCHES (NCEP1_{ECL}) had a higher track count with a shorter (longer) duration (Fig. 3A) and length (Fig. 3B). In addition, the top panels of Fig. 3A and B displays the difference in length and duration of paired ECL tracks, with the attributes of individual, paired MATCHES ECL events subtracted from the corresponding NCEP1_{ECL} ECLs. The negative values of the top-panel curves in Fig. 3A and B illustrate the instances where the length (16.68%) or duration (15.19%) of the MATCHES track exceeded the corresponding NCEP1_{ECL} track. In contrast, NCEP1_{ECL} exceeded the track duration (75.89%) and length (83.32%) of the paired track within MATCHES in most instances. A two-sided Wilcoxon signed-rank test was used to examine the statistical significance of whether NCEP1_{ECL} tracks were different in length and duration compared to their paired MATCHES counterparts. NCEP1_{ECL} tracks are longer with respect to length and time than those within the MATCHES database (p-value of <0.001, two-sided Wilcoxon signed-rank test).

3.2. Density distribution of ECLs in MATCHES and NCEP1_{ECL}

The spatial extent of ECLs represented in the MATCHES and NCEP1_{ECL} databases are similar (Fig. 4). The area with the highest count of ECL tracks is located in the southeastern section of the ECL box for both MATCHES and NCEP1_{ECL}. However, the higher counts are located more centrally in the ECL box of MATCHES than in NCEP1_{ECL}. In contrast, the highest density region is located in the eastern part of the ECL box for

Table 3
The CSI index for each dataset.

	Hits	Misses	False Alarms	CSI (%)
MATCHES vs NCEP1 _{ECL}	1547	209	172	80.24
NCEP1 _{ECL} vs MATCHES	1547	328	121	77.51

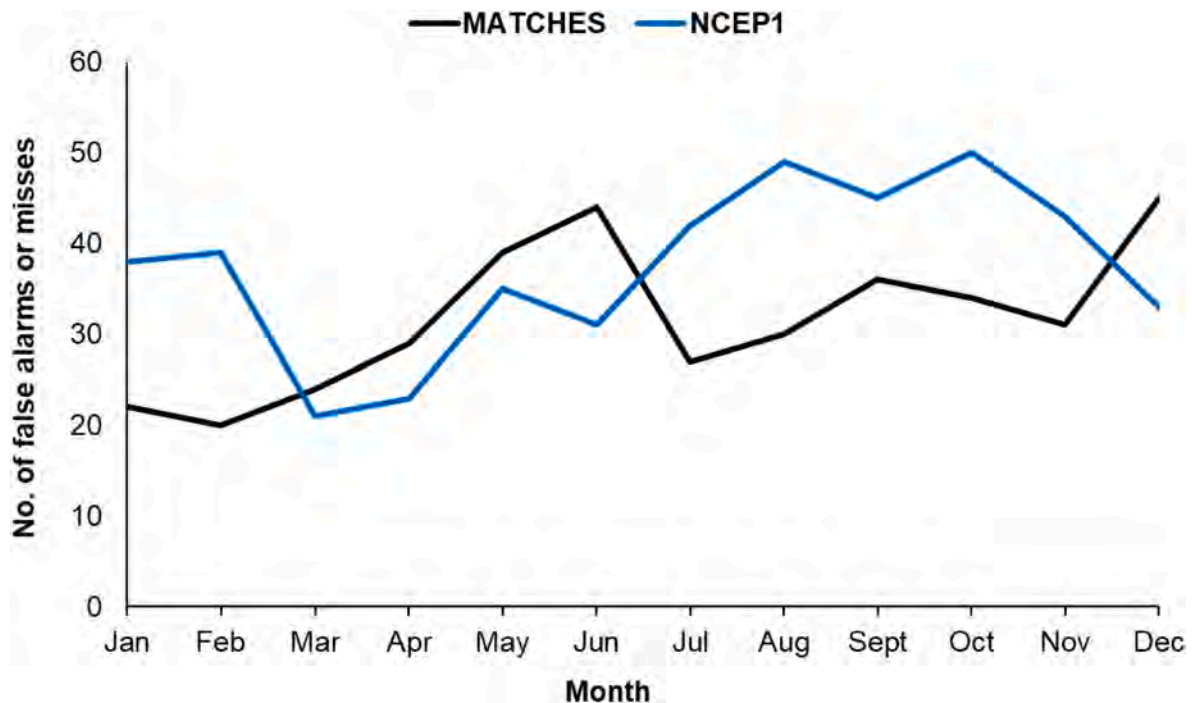


Fig. 2. Monthly distribution of both false alarms and misses in MATCHES (black) and NCEP1_{ECL} (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4

The comparison of track duration and track length for false alarms and misses against hits for both datasets.

	Hits		False Alarms/Misses	
	MATCHES	NCEP1	MATCHES	NCEP1
Mean track length (km)	1985.01	3859.34	959.29	3109.91
Mean track duration (days)	2.78	4.59	1.31	3.16
Mean max Laplacian	0.51	0.73	0.37	0.61
Mean track length in ECL box (km)	566.84	679.37	257.82	324.92
Mean track duration in ECL box (days)	1.01	1.09	0.45	0.38
Mean max Laplacian of tracks in the ECL box	0.44	0.54	0.32	0.36

NCEP1_{ECL}. In addition, there are a higher number of events in NCEP1_{ECL} that have a tropical (tracks that transgress north of 25°) and western (events that start and travel from west of 140°E) influence. The NCEP1 low-pressure database (rather than NCEP1_{ECL}) includes many additional low-pressure systems over a larger areal extent (Fig. S1) and may be useful if the definition of ECLs evolves in the future. Higher density areas within NCEP1 (Fig. S1B) correlate with the Pilbara (Western Australia) and Cloncurry (Queensland) heat lows (area of decreased surface pressure that is caused by intense daytime heating; Lavender, 2017; Sturman and Tapper, 2005; Tapper and Hurry, 1933).

3.3. Long-term and seasonal trends in ECLs

Monthly track plots illustrate the distribution of ECL events from 1950 to 2019 (Fig. 5). Tracks within the warmer months (February; Fig. 5A) have a more tropical origin than the cooler seasons (August; Fig. 5B) when tracks are more westerly oriented. Tracks with an inland origin are commonly observed in transitional months (Fig. 5C). For these transition months, many westerly tracks are only found in NCEP1_{ECL} and not in MATCHES (all months shown in Fig. S2).

The monthly track results reveal differences in the number of ECLs

forming either inside or outside the ECL box for the two databases (Fig. 6, Table 5). Events within the MATCHES database have a higher monthly percentage of ECL tracks with a genesis in the ECL box. This is expected, as the NCEP1_{ECL} includes more, longer ECL tracks that travel through the ECL box but do not originate within that area. Within the ECL box, there are no apparent patterns or differences in the locations of ECL genesis between the two datasets. However, there is a slight shift of ECL genesis to the south in both databases in the colder months (Fig. 6C and D). In both databases, the percentage of tracks that start within the box increases in the austral autumn and winter months (Table 5).

ECL seasonality is evident in both datasets, with ECLs increasing in frequency in the colder months (April–September; Fig. 7). This seasonality is more prominent in MATCHES than NCEP1_{ECL}. Outliers are more conspicuous within NCEP1_{ECL} boxplots than MATCHES, especially in June and November. During these months, ECL numbers in NCEP1_{ECL} vary relatively consistently between one and four (compared to the larger range registered by MATCHES). The median number of ECLs during January (0.254), July (0.0257), August (0.0452) and December (0.461) were significantly different between the two datasets at $\alpha = 0.05$ (Table 6).

Both NCEP1_{ECL} and MATCHES show comparable periods of enhanced and reduced ECL activity from 1950 to 2019 (Fig. 8). For example, both NCEP1_{ECL} and MATCHES exhibit reduced ECL activity since approximately 2005. No significant long-term linear trends for either timeseries was identified by the Wilcoxon signed-rank test. Further, no discernible difference in the distribution of the two datasets was identified, as confirmed by the non-significant result of the two-sample Kolmogorov-Smirnov test. Additional tests for step-changes (CUSUM) applied to the seasonal (three-month; Fig. S7) ECL timeseries also conveyed no significant trend (Table S2 and Table S3).

3.4. Case studies of significant notable ECL events

Comparing the representations of tracks for two significant ECL storms (the 1998 Sydney to Hobart Yacht Race and the Pasha Bulker storm ECLs) in the MATCHES and NCEP1_{ECL} databases reveals functional differences in track length and characterisation (Table 7, Fig. 9,

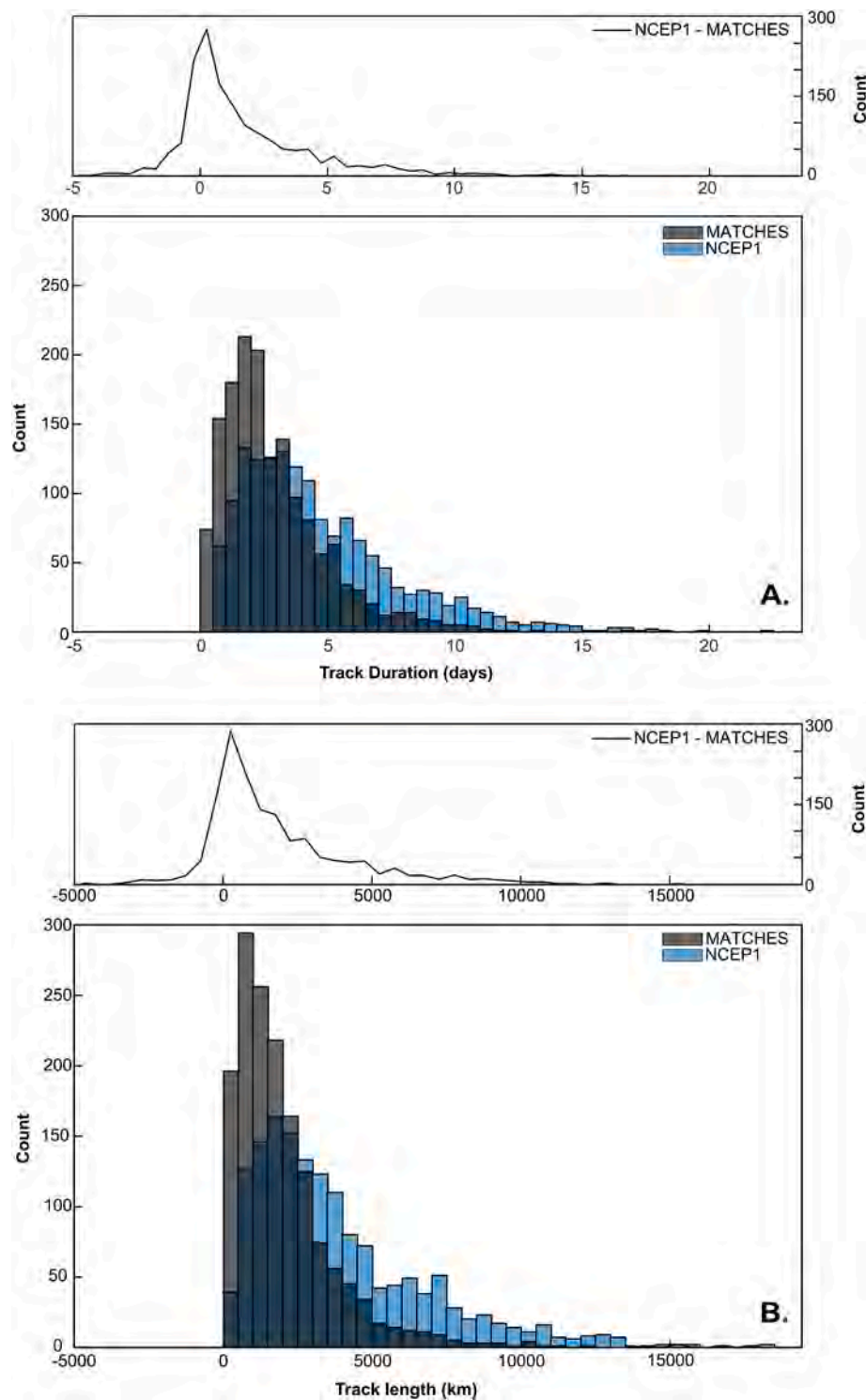


Fig. 3. Histogram of A.) the track duration (days) and B.) the track length (km) of NCEP1_{ECL} (blue) and MATCHES (black). The distribution of the difference between NCEP1_{ECL} and MATCHES is highlighted by the black line (top panels). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table S7). For both storms, the MLAP value of the MATCHES tracks is lower than in NCEP1_{ECL}, thus suggesting less intense ECLs in MATCHES. However, for the Pasha Bulker storm, MSLP_{min} is higher (less intense) in the NCEP1_{ECL} than the paired event in MATCHES.

For these two storms, there is a slight difference (~3 hPa) in the MSLP_{min} derived from MATCHES and NCEP1_{ECL}. For the 2007 Pasha Bulker storm, minimum MSLPs on synoptic charts (998 hPa; BOM (2007, 2021)) correspond more closely to the MATCHES MSLP_{min} value

(998.82 hPa), whereas NCEP1_{ECL} (1001.1 hPa) underestimates the intensity at the time minimum pressure occurred (8 June 2007 at 6 p.m.).

The ECL event during the 1998 Sydney to Hobart Yacht Race, as derived from NCEP1_{ECL}, peaks at the last point of the track (29 December 1998 at 6 p.m.). By that time, the ECL moved outside the area shown on Australian MSLP charts, and the associated MSLP_{min} of NCEP1_{ECL} could thus not be compared to the synoptic maps. Therefore, the track comparison for the 1998 Sydney to Hobart Yacht Race analysis

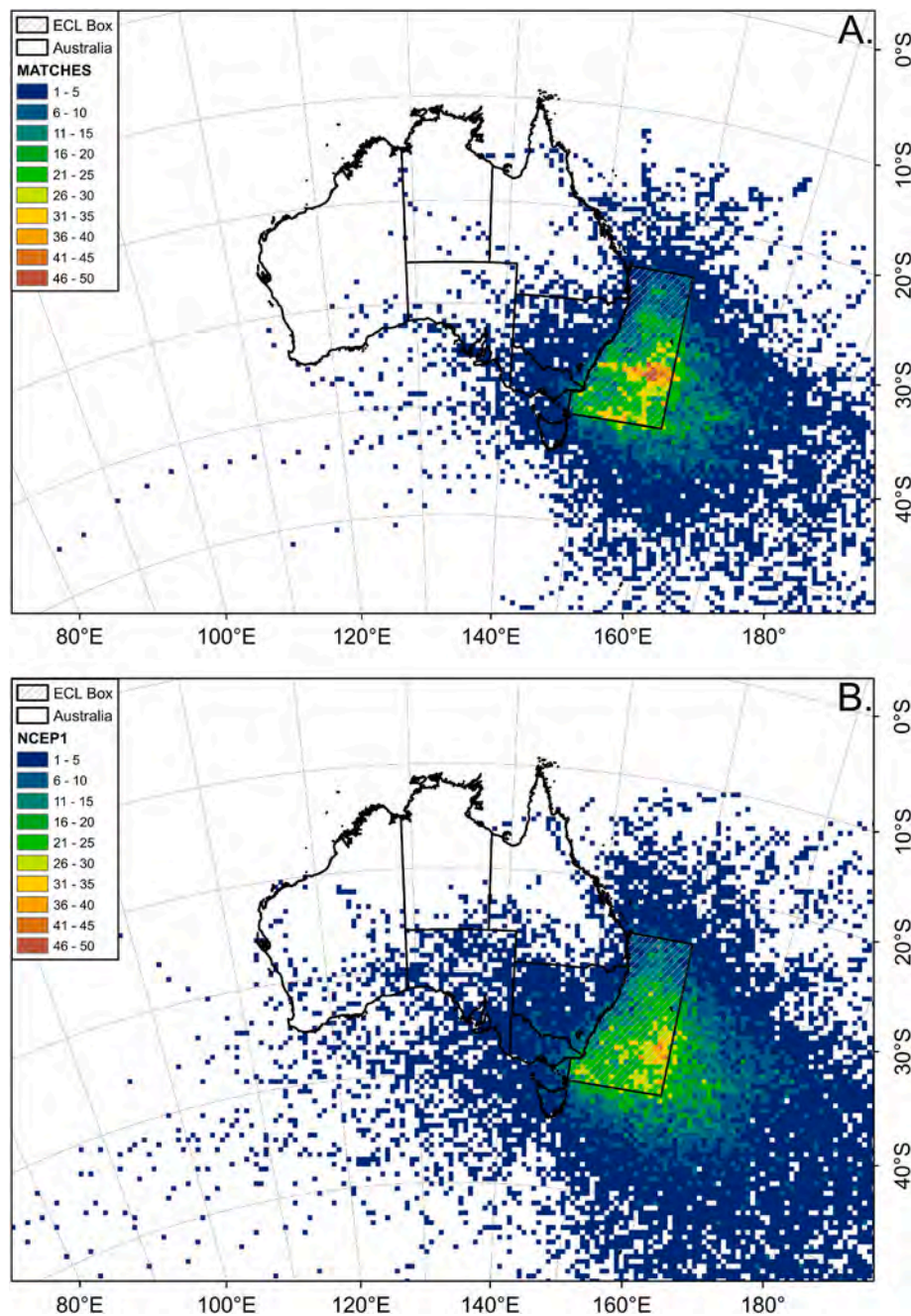


Fig. 4. Density (count per cell) of the ECL-filtered datasets of MATCHES A.) and NCEP1_{ECL} B.) from 1950 to 2019. Cells are coloured based on the number of points that occur within the area. Areas with a higher count are highlighted by warmer colours. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

uses the time-step when the $MSLP_{min}$ occurred in MATCHES (27 December at 12 a.m.), that timing of peak intensity also supported by Buckley and Leslie (2000). Synoptic charts from the BOM (1998) approximated the pressure of the low to slightly below 988 hPa. Conversely, Buckley and Leslie (2000) stated that the peak intensity of the storm had a central pressure of 978 hPa. Pressure within the MATCHES (984.66 hPa) database is closer to the Buckley and Leslie (2000) central pressure, whereas NCEP1_{ECL} (987.92 hPa) is nearer to the BOM. With respect to this storm, the pressure difference between MATCHES and NCEP1_{ECL} was minor, especially as the likelihood of either database capturing the moment of maximum strength is relatively low given the spatial and temporal resolution of NNR1 (Di-Luca et al., 2015; Eichler and Gottschalck, 2013).

The 1998 Sydney to Hobart ECL event had the most significant difference in track morphology represented in the two databases (Fig. 9A). The MATCHES track for this event was very short in length compared to NCEP1_{ECL} and contained no data points within the ECL box. As a result, this iconic event would not have been included in studies that define ECLs based on the presence of a single point in the arbitrary ECL box. Conversely, the NCEP1_{ECL} track length and duration for this event was more extensive and contained one point in the ECL box.

In contrast, the tracks of the Pasha Bulker ECL are similar in the MATCHES and NCEP1_{ECL} databases. Despite their similarities, there are still intriguing differences in the track representations. For example, the NCEP1_{ECL} track for the Pasha Bulker storm (Fig. 9B) remains offshore and is positioned a little further northward than the MATCHES track that

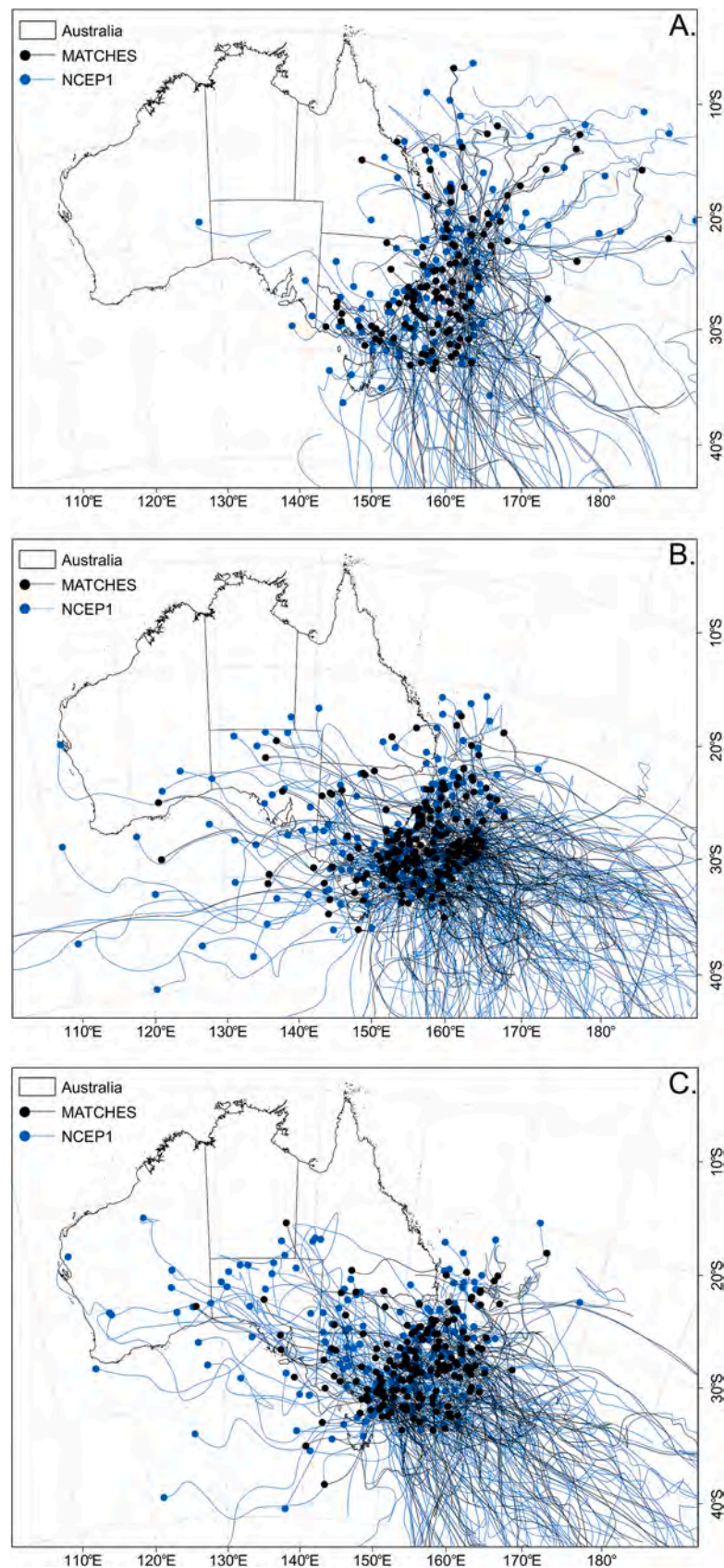


Fig. 5. Examples of seasonal tracks from NCEP1_{ECL} (blue) and MATCHES (black) from 1950 to 2019. A.) demonstrates track characteristics for warmer months (February). B.) conveys track characteristics for cooler months (August), and C.) illustrates the nature of tracks in transitional months (November). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

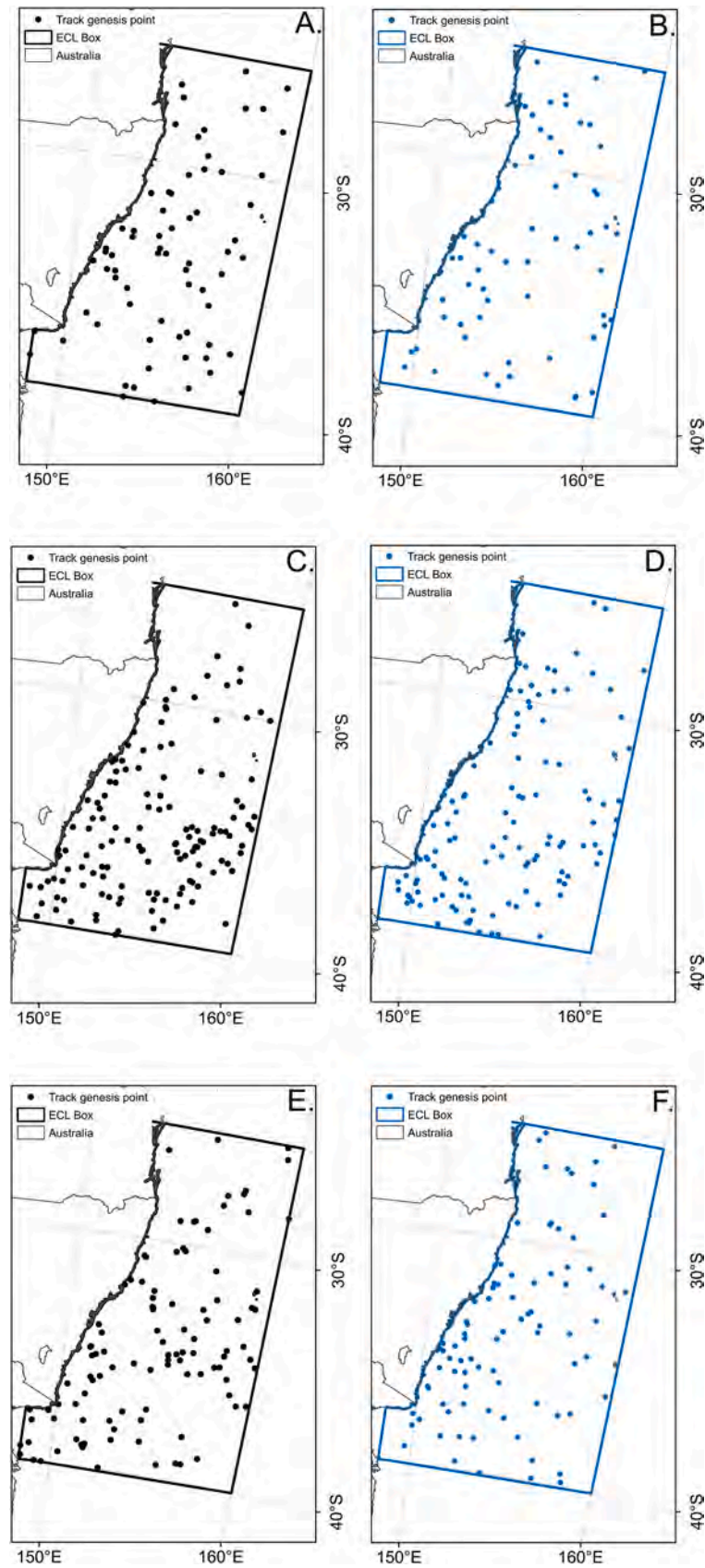


Fig. 6. February (A. and B.), August (C. and D.) and November (E. and F.) genesis points of monthly ECL events in the MATCHES (A., C. and E.) and NCEP1_{ECL} (B., D. and F.) database from 1950 to 2019. Genesis points are only shown for tracks that originated within the ECL box. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 5
Percentage of tracks that start in the ECL box for each database.

Month	MATCHES (%)	NCEP1 (%)	% Difference
January	56	48	8
February	61	47	14
March	66	53	13
April	69	52	17
May	73	57	16
June	73	56	17
July	65	52	13
August	69	54	15
September	57	42	15
October	51	46	5
November	61	44	17
December	58	44	14

crosses onto land.

The mean latitudinal positioning of the 1998 Sydney to Hobart ECL tracks was statistically different between MATCHES and NCEP1_{ECL} (Table S7; $p < 0.001$, $t_{\text{stat}} = 4.029$, two-sample t -test of means). No statistically significant difference between the mean track longitude or latitude coordinates between the two datasets was established for the Pasha Bulker storm. The northerly bias of the 1998 Sydney to Hobart NCEP1_{ECL} track is consistent with the observation that MATCHES ECLs are located more centrally in the ECL box compared to NCEP1_{ECL}.

4. Discussion

The aim of this study was to compare and contrast the representation of ECL tracks in the NCEP1 and MATCHES databases. The results highlighted that, although the MSLP_{min} and MLAP values in the MATCHES and NCEP1_{ECL} databases are similar, the length and duration of ECL tracks are greater in NCEP1_{ECL}. These differences are primarily due to the improved cyclone tracking scheme used by NCEP1_{ECL} (LAPv2; Pepler et al., 2015). MATCHES and NCEP1_{ECL} are both relatively easy to update and contain accurate representations of storm strength. However, the NCEP1_{ECL} database affords ECL researchers the following advantages: 1) the ability to easily expand the scope of their studies beyond the ECL box; 2) more extensive track details in NCEP1_{ECL} in terms of ECL

track length and 3) the tracks in NCEP1_{ECL} are complete, without the need for manual merging of multiple tracks.

The approach taken in the comparison of the MATCHES and NCEP1_{ECL} databases was conservative. Cross-comparison of the datasets was based on the pairing of events with respect to the time of MLAP values (intensity), latitude and longitude. The initial cross-validation process yielded multiple events paired from the MATCHES dataset to one event in NCEP1_{ECL}, a feature rarely observed when the test was reversed. This issue resulted from multiple small tracks ("split tracks") in the MATCHES database belonging to one, longer event, the tracks thus needing to be manually merged. The merging process included consideration of time (consecutive tracks with a limited temporal gap), location (latitude and longitude), distance from track A (last point) to track B (first point) and velocity (longer distance permitted for faster systems). In addition, the potentially split tracks were also compared with MSLP charts when available. Once merged, these previously split MATCHES tracks (Fig. 1) were then paired with one NCEP1_{ECL} track.

The pairing of multiple MATCHES tracks to the same NCEP1_{ECL} track is attributable to the formation of a secondary low during the NCEP1_{ECL} track. Ultimately, the merging of split tracks decreased the number of

Table 6
Wilcoxon signed-rank test for paired monthly counts of ECL events in NCEP1_{ECL} and MATCHES. Statistical significance ($\alpha = 0.05$) is outlined by an asterisk.

Month	p-value
January	0.0254*
February	0.0551
March	0.3574
April	0.6077
May	1.0000
June	0.0918
July	0.0257*
August	0.0452*
September	0.2762
October	0.1220
November	0.0646
December	0.0461*

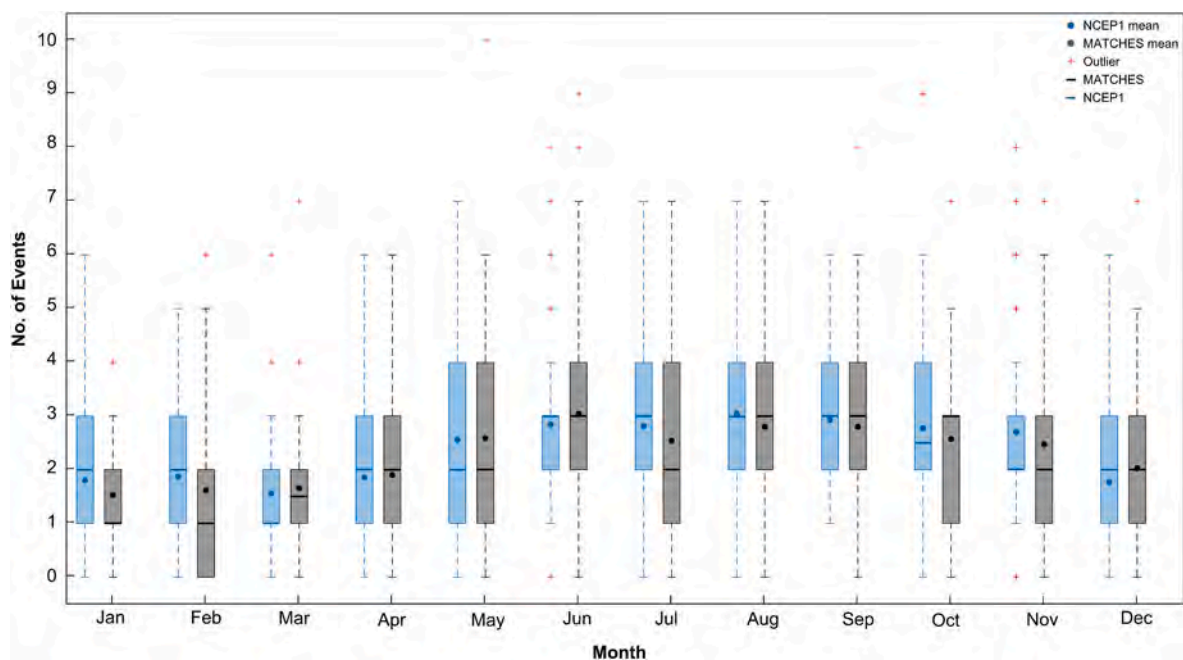


Fig. 7. The distribution of ECLs per month for NCEP1_{ECL} (blue) and MATCHES (black). Dots represent the mean of the spread, and red crosses are the outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

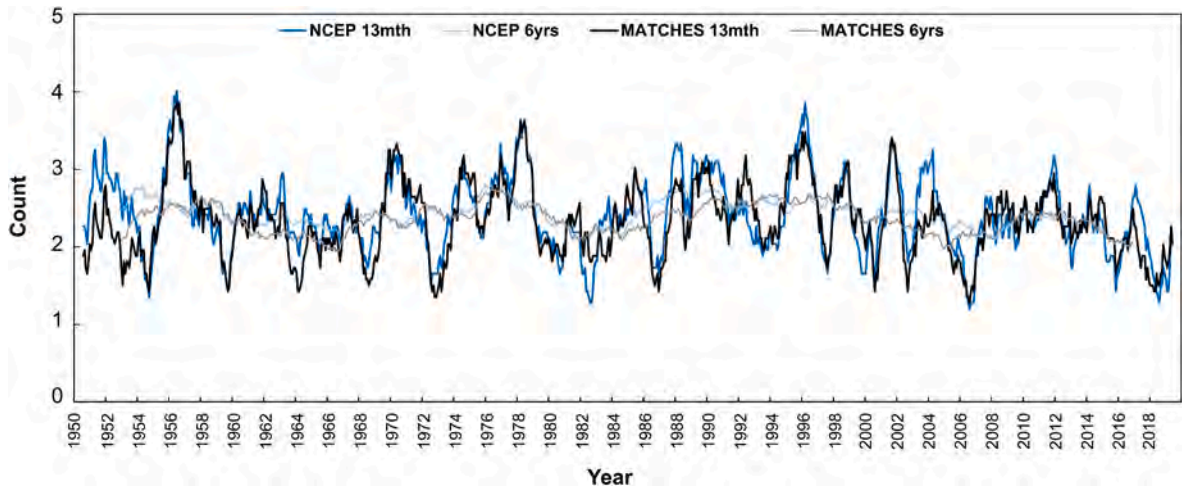


Fig. 8. The time series of the NCEP1_{ECL} (blue) and MATCHES (black) datasets until 2019, shown as a centred running mean (13 months), accompanied by a 6-year running mean. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 7

Track properties corresponding with the timing of major ECL events in both databases. The identification (ID) number, maximum Laplacian value (MLAP) and minimum mean sea level pressure (MSLP_{min}) are listed.

ECL event	Duration	MATCHES			NCEP_1		
		ID	MLAP	MSLP _{min}	ID	MLAP	MSLP _{min}
1998 Sydney to Hobart Yacht Race	25 – 28 December 1998	2829	0.82	984.66	40401	1.12	971.17
Pasha Bulker	7 – 9 June 2007	3273	0.54	998.82	46729	0.63	1001.1

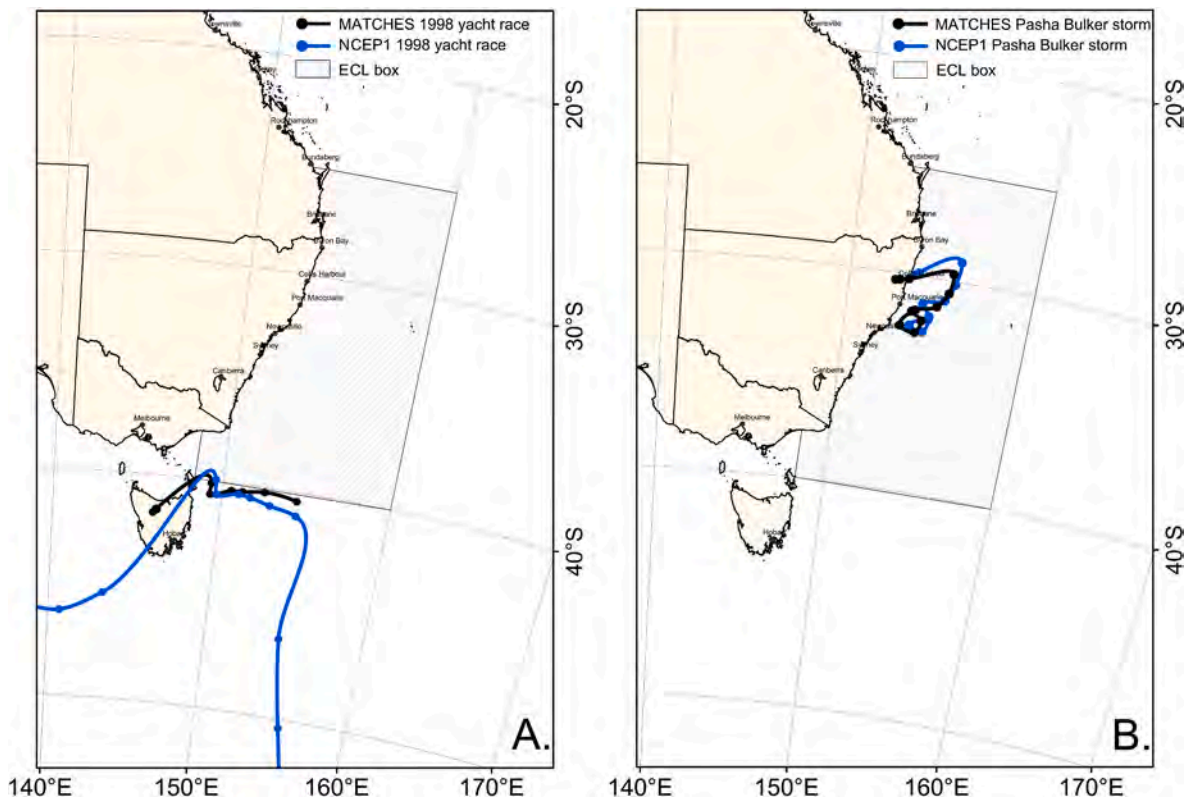


Fig. 9. Storm tracks from the NCEP1_{ECL} and MATCHES database for two significant ECL events. The two events include the 1998 Sydney to Hobart Yacht Race (A.) and the 2007 Pasha Bulker storm (B.). Events from the MATCHES (NCEP1_{ECL}) database are in black (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

ECL events present within MATCHES and influenced the CSI and comparison statistics by increasing the number of hits while decreasing misses and false alarms. Tracks in MATCHES have a greater likelihood to be paired with a corresponding NCEP1_{ECL} track because NCEP1_{ECL} has more events than MATCHES (i.e. potential false alarms) and is thus more likely to contain the events that MATCHES has. Conversely, NCEP1_{ECL} is more challenging to pair with MATCHES because MATCHES is missing many of the events contained in NCEP1_{ECL} (misses). Counterintuitively, the lower CSI of NCEP1_{ECL} is the result of NCEP1_{ECL} having more and longer ECL tracks.

There are caveats associated with some tracks that were classified as false alarms or misses in the CSI test. The strict adherence to three categories (hit, miss and false alarm) and the right area (i.e. the ECL box) leaves no flexibility or consideration for varied scenarios. An example of this is evident in the 1998 Sydney to Hobart Yacht Race storm event. The associated NCEP1_{ECL} track is originally classified as a false alarm (and a miss in MATCHES). This issue is problematic because the Sydney to Hobart Yacht Race caused a great deal of damage and therefore is an important event from an impact perspective. However, if this storm were to be classified in terms of the CSI test for NCEP1_{ECL} (and the ECL box for MATCHES), it would not be included as an ECL, unlike the Pasha Bulker storm.

Both NCEP1_{ECL} and MATCHES demonstrated similarities in monthly seasonality from 1950 to 2019. The comparison of monthly occurrences highlighted and reinforced some patterns in ECL seasonality mentioned throughout literature. The production of similar track variability to previous studies demonstrates consistency across ECL databases that incorporate differing cyclone tracking methods. These findings induce confidence in the ability of the NCEP1_{ECL} to reproduce many of the sub-annual characteristics found in the MATCHES database and may even include additional information not contained in MATCHES.

Within this study, both datasets conveyed a prominence of ECL events in the cooler months (April–September; Fig. 7). This pattern can be influenced by an increase in frontal and westerly systems being active across southern Australia (Fig. 5B. and C.; Browning and Goodwin, 2013). Continental lows (CLs) within this study occur within the cooler and transition months (mainly spring). This result is more evident in the NCEP1_{ECL} dataset than in MATCHES. A similar outcome is reflected in Browning and Goodwin (2013), where CLs and inland troughs (ITs) were more prominent in the austral spring and early summer months. However, CLs and ITs were less dominant during austral winter in their study. Summer and early autumn ECL tracks were composed of more northern-derived lows (Fig. 5A and Fig. S2), another finding mirrored in ECL variability in Browning and Goodwin (2013). Specifically, ECL occurrence during summer was commonly linked to tropical cyclones.

Although not statistically significant, there were a number of ECL events present in the NCEP1_{ECL} in December and January (Fig. 7) that were not contained in MATCHES. Many of these summer storms in NCEP1_{ECL} are characterised by inland genesis (Fig. S4; Table S4 and Table S5), a crucial trait that could aid in storm and disaster management because of the significant impact of these systems (e.g. the 1998 Sydney to Hobart storm event). The increased ability of NCEP1_{ECL} to resolve summer storms compared to the MATCHES database allows for these types of ECLs to be included in future research and enhance our knowledge and understanding of their potential drivers. Although previous studies such as Browning and Goodwin (2016) have outlined a decrease of “southern” lows and more “easterly trough” lows, there was no significant overall trend for the annual (Fig. 8) and seasonal (Fig. S7) timeseries for either ECL dataset. This finding was consistent with results from Dowdy et al. (2019) and Pepler and Dowdy (2021). Statistical tests (Kolmogorov-Smirnov and the CUSUM in the supplementary section; Table S2 and Table S3) for annual and seasonal timeseries (Fig. S7) also showed no significant shift from 1950 to 2019.

The tracks in the NCEP1_{ECL} database are on average over twice as long and almost twice the duration as the paired event in MATCHES. Track similarity was investigated in detail using two significant ECL

events (Fig. 9, Table 7). There was no statistical difference between the NCEP1_{ECL} and MATCHES tracks of the Pasha Bulker storm. However, the 1998 Sydney to Hobart Yacht Race tracks showed statistically significant differences in latitude, length and duration. The intensity of storm tracks is another dissimilarity between the MATCHES and NCEP1_{ECL} databases. The MLAP value for NCEP1_{ECL} (MATCHES) tracks was recorded with a higher (lower) intensity. However, the MSLP_{min} for the tracks was lower (higher) in MATCHES (NCEP1_{ECL}). These features could potentially result from using a different version of the NCEP1 reanalysis dataset and tracking algorithm (Pepler et al., 2015). Comparing MSLP_{min} for both storms (of the case studies) to synoptic charts conveyed that the MATCHES pressure levels are closer to true values than those within NCEP1_{ECL}. However, the differences are minor (~3 hPa).

As commonly seen in TC studies, there is always a degree of error in the visual representation of an event (Chand et al., 2019). However, due to the lack of ECL track studies, there is ambiguity around the most representative track pathway. This was evident in the track maps of the two case studies, particularly for the 1998 Sydney to Hobart storm event. The visual representation of a track smoothing function in the mapping software (ArcGIS) does imply that the MATCHES-derived Sydney to Hobart yacht track does indeed pass through the box, although all positional data points are outside the area. The split-track nature of the MATCHES database could also influence how this track is represented. With the additional length seen in the NCEP1_{ECL} track, it is plausible that additional relevant, clipped tracks are present in MATCHES that were not connected to the main track segment. This is also an excellent example of the arbitrary nature of the ECL box – and how it may hinder ECL studies by omitting important ECL events.

The extensive analysis that went into pairing events contained within MATCHES and NCEP1_{ECL} identified some problematic characteristic of both databases that ultimately may have influenced the detailed results of this study – but not the conclusions. In some cases, manual merging, pairing and unpairing had to be undertaken. A range of track characteristics (MLAP, latitude, longitude, velocity), combined with MSLP charts, were utilised to ensure that the track merging was appropriate for every individual case. Although a set of consistent filters and tests were applied to identify ECLs in this study, there may be a residual bias of ECL events in MATCHES due to a higher level of cross-validation with other studies. Another consideration is that several analyses (particularly the August and December boxplots) produced a borderline statistically significant result (p-values close to 0.05; Fig. 7).

To reduce the complexity of this study, two ECL datasets were compared that used the same reanalysis (NNR1) database but different cyclone tracking schemes. However, it should be noted that there are a vast number of different reanalysis datasets that can and have been used for cyclone tracking (e.g. 20CR, JRA55, ERAI, ERA5, BARRA), and NNR1 may not be the ideal choice for ECL track studies (Fu et al., 2016). Low-pressure or cyclone comparison studies (e.g. Di-Luca et al., 2015; Eichler and Gottschalck, 2013) have demonstrated that, in comparison to other reanalysis datasets, NNR1 has a coarser resolution. Low-pressure systems or cyclones identified within the NNR1 dataset showed a lower mean intensity and a larger average size compared to the higher resolution datasets. The number of storm observations was also higher (lower) in austral summer (winter) for NNR1 (Di-Luca et al., 2015). With respect to the present study, the monthly counts for summer ECLs in NCEP1_{ECL} could be artificially inflated because of the reanalysis dataset used and require further study. Response to data assimilation schemes and how information is extracted from observations in both hemispheres was superior again in newer, higher resolution datasets than in NNR1 (Di-Luca et al., 2015; Eichler and Gottschalck, 2013). However, the advantages of using higher-resolution datasets for ECL studies are offset by the relatively short time span of these databases.

For this study, the coarse resolution and other assimilation issues of the NNR1 reanalysis are accepted because NCEP1 and MATCHES are not restricted to the satellite period – encompassing a longer timeseries

(1950–2019). The ability to use data prior to the satellite period is important for extending future cyclone studies further back in time (Pepler, 2020a). Additionally, the CSI scores reinforced a negligible difference from the pre- and post-satellite period for the two datasets (Table S1), demonstrating the usability of the NNR1 and its derivatives before 1979. Other studies also support that all reanalysis datasets (including NNR1) for southern Australia are similar for cyclones throughout the satellite period (Pepler, 2020a). Ultimately, NCEP1 allows the user to stipulate their own study-defined parameters for identifying ECL/high-intensity low-pressure systems (e.g. NCEP1_{ECL}), including cyclonic systems outside the ECL box. However, to compare results from various ECL databases and studies, a consistent application of defining parameters is required. Thus, future ECL comparison studies could adopt already existing defining parameters (as seen in 2.2) to provide consistency (Kiem et al., 2016) and include extensions of these parameters to include larger geographic areas. In addition, a replication of this study using a higher resolution dataset such as ERA5 (when the preliminary version of 1950–1978 has been finalised) could prove advantageous for updating ECL databases in the future.

5. Conclusion

Based on the analysis presented in this paper, we suggest that the NCEP1 database (Pepler, 2020a) is suitable and, in some cases, preferable as an updated ECL database for Australia. Importantly, the ability of NCEP1 to recreate ECL tracks was comparable if not better than those produced in the MATCHES database. NCEP1_{ECL} of this study contained more tracks than MATCHES, and paired events were approximately twice as long in length and duration. The nature of the NCEP1 database allows greater flexibility in defining ECL parameters – including areas outside the ECL box – and examining the origin and characteristics of the ECLs. Events in the NCEP1 database also require less manual processing to produce continuous, unbroken ECL tracks. The NCEP1 database and its derivatives provide more opportunities – and more accessible methods – for studying the nature of these destructive and often drought-breaking weather systems.

Credit author statement

Jessie L. Gray: Conceptualisation, Formal analysis, Investigation, Methodology, Software, Writing – Original Draft, Visualisation, Jasmine B.D. Jaffrés: Formal analysis, Investigation, Methodology, Software, Supervision, Writing – review and editing, Danielle C. Verdon-Kidd: Conceptualisation, Funding acquisition, Supervision, Writing – review and editing, Validation, Michael G. Hewson: Writing – review and editing, John M. Clarke: Supervision, Funding acquisition, Writing – review and editing, Acacia Pepler: Resources, Validation, Writing – review and editing, Nathan B. English: Funding acquisition, Supervision, Writing – review and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

J. Gray acknowledges the financial support of Central Queensland University (CQU), financial support from a postgraduate scholarship from the Commonwealth Scientific and Industrial Research Organisation (CSIRO, John Clarke supervising) and C&R Consulting, Townsville. The authors would like to acknowledge the traditional owners of the land, the Bindal and Wulgurukaba, on which this research was conducted. We also pay our respects to the indigenous elders of the past, present and emerging.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.wace.2021.100400>.

References

- Australian Bureau of Meteorology, 2021. Analysis Chart Archive. <http://www.bom.gov.au/australia/charts/archive/index.shtml>. (Accessed 17 March 2021).
- Australian Bureau of Meteorology, 2007. Monthly Weather Review: New South Wales. June 2007, Melbourne, Australia.
- Australian Bureau of Meteorology, 1998. Monthly Weather Review: New South Wales December 1998 (Melbourne, Australia).
- Australian Institute for Disaster Resilience (AIDR), 2020. Sydney to Hobart Yacht Race. <https://knowledge.aidr.org.au/resources/maritimecoastal-sydney-to-hobart-yacht-race-1998>. (Accessed 10 September 2020), 1998.
- Berrisford, P., Dee, D., Pole, P., Brugge, R., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., Uppala, S., Simmons, A., 2011. The ERA-Interim Archive: Version 2.0, ERA Report Series.
- Browning, S., Goodwin, I., 2016. Large-scale drivers of Australian east coast cyclones since 1851. *J. South. Hemisph. Earth Syst. Sci.* 66, 125–151. <https://doi.org/10.22499/3.6602.004>.
- Browning, S.A., Goodwin, I.D., 2013. Large-scale influences on the evolution of winter subtropical maritime cyclones affecting Australia's east coast. *Mon. Weather Rev.* 141, 2416–2431. <https://doi.org/10.1175/MWR-D-12-00312.1>.
- Buckley, B.W., Leslie, L.M., 2000. The Australian Boxing Day storm of 1998 - synoptic description and numerical simulations. *Weather Forecast.* 15, 543–558. [https://doi.org/10.1175/1520-0434\(2000\)015<0543:TABDSO>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0543:TABDSO>2.0.CO;2).
- Callaghan, J., Helman, P., 2008. Severe Storms on the East Coast of Australia 1770–2008. Bureau of Meteorology, Griffith Centre for Coastal Management.
- Callaghan, J., Power, S., 2015. Supplementary Appendix. *Aust. Meteorol. Oceanogr. J.* 3, 925–927.
- Callaghan, J., Power, S.B., 2014. Major coastal flooding in southeastern Australia 1860–2012, associated deaths and weather systems. *Aust. Meteorol. Oceanogr. J.* 64, 183–213. <https://doi.org/10.22499/2.6403.002>.
- Callaghan, J., Power, S.B., 2011. Variability and decline in the number of severe tropical cyclones making land-fall over eastern Australia since the late nineteenth century. *Clim. Dynam.* 37, 647–662. <https://doi.org/10.1007/s00382-010-0883-2>.
- Cavicchia, L., Pepler, A., Dowdy, A., Evans, J., Di Luca, A., Walsh, K., 2020. Future changes in the occurrence of hybrid cyclones: the added value of cyclone classification for the east Australian low-pressure systems. *Geophys. Res. Lett.* 47 <https://doi.org/10.1029/2019GL085751>.
- Cavicchia, L., Pepler, A., Dowdy, A., Walsh, K., 2019. A physically based climatology of the occurrence and intensification of Australian east coast lows. *J. Clim.* 32, 2823–2841. <https://doi.org/10.1175/JCLI-D-18-0549.1>.
- Chand, S.S., Dowdy, A.J., Ramsay, H.A., Walsh, K.J.E., Tory, K.J., Power, S.B., Bell, S.S., Lavender, S.L., Ye, H., Kuleshov, Y., 2019. Review of tropical cyclones in the Australian region: climatology, variability, predictability, and trends. *Wiley Interdiscip. Rev. Clim. Chang.* 10, 1–17. <https://doi.org/10.1002/wcc.602>.
- Chiew, F., Siriwardena, L., 2005. Trend/change Detection Software: User Manual, eWater Toolkit: Documentation.
- Compo, G., Whitaker, J., Sardeshmukh, P., Matsui, N., Allan, R., Yin, X., Gleason, E., Vose, R., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R., Grant, A., Groisman, P., Jones, P., Kruk, M., Kruger, A., Marshall, G., Maugeri, M., Mok, H., Nordli, O., Ross, T., Trigo, R., Wang, X., Woodruff, S., Worley, S., 2011. Review article the twentieth century reanalysis project. *R. Meteorol. Soc.* 137, 1–28.
- Corder, G.W., Foreman, D.I., 2014. Comparing two unrelated samples: the Mann-Whitney U-test. In: *Nonparametric Statistics for Non-statisticians: A Step-by-step Approach*. John Wiley & Sons, Incorporated, Somerset, pp. 57–78. <https://doi.org/10.1002/9781118165881.ch4>.
- Coutts-Smith, A., Gamble, F., Rakich, C., Schweitzer, M., 2012. Eastern seaboard climate hazard tool - MATCHES. In: *AMOS 18th Annual Conference : Connections in the Climate System: General Information, Programme and Abstracts Handbook*. University of New South Wales, p. 143, 31 Jan to 3 Feb 2012.
- Di-Luca, A., Evans, J.P., Pepler, A., Alexander, L., Argüeso, D., 2015. Resolution sensitivity of cyclone climatology over Eastern Australia using six reanalysis products. *J. Clim.* 28, 9530–9549. <https://doi.org/10.1175/JCLI-D-14-00645.1>.
- Di Luca, A., Evans, J., Pepler, A., Alexander, L., Argüeso, D., 2016. Evaluating the representation of Australian East Coast Lows in a regional climate model ensemble. *J. South. Hemisph. Earth Syst. Sci.* 66, 108–124. <https://doi.org/10.22499/3.6602.003>.
- Dowdy, A.J., Mills, G.A., Timbal, B., 2013a. Large-scale diagnostics of extratropical cyclogenesis in eastern Australia. *Int. J. Climatol.* 33, 2318–2327. <https://doi.org/10.1002/joc.3599>.
- Dowdy, A.J., Mills, G.A., Timbal, B., 2011. Large-scale Indicators of Australian East Coast Lows and Associated Extreme Weather Events (Melbourne).
- Dowdy, A.J., Mills, G.A., Timbal, B., Griffiths, M., Wang, Y., 2013b. Understanding rainfall projections in relation to extratropical cyclones in eastern Australia. *Aust. Meteorol. Oceanogr. J.* 63, 355–364. <https://doi.org/10.22499/2.6303.001>.
- Dowdy, A.J., Mills, G.A., Timbal, B., Wang, Y., 2013c. Changes in the risk of extratropical cyclones in eastern Australia. *J. Clim.* 26, 1403–1417. <https://doi.org/10.1175/JCLI-D-12-00192.1>.
- Dowdy, A.J., Pepler, A., Di Luca, A., Cavicchia, L., Mills, G., Evans, J.P., Louis, S., McInnes, K.L., Walsh, K., 2019. Review of Australian east coast low pressure systems

- and associated extremes. *Clim. Dynam.* 53, 4887–4910. <https://doi.org/10.1007/s00382-019-04836-8>.
- Eichler, T.P., Gottschalck, J., 2013. A comparison of Southern Hemisphere cyclone track climatology and interannual variability in coarse-gridded reanalysis datasets. *Adv. Meteorol.* <https://doi.org/10.1155/2013/891260>, 2013.
- Evans, J.L., Guishard, M.P., 2009. Atlantic subtropical storms. Part I: diagnostic criteria and composite analysis. *Mon. Weather Rev.* 137, 2065–2080. <https://doi.org/10.1175/2009MWR2468.1>.
- Fu, G., Charles, S.P., Timbal, B., Jovanovic, B., Ouyang, F., 2016. Comparison of NCEP-NCAR and ERA-interim over Australia. *Int. J. Climatol.* 36, 2345–2367. <https://doi.org/10.1002/joc.4499>.
- Gray, J. L., Verdon-Kidd, D. C., Callaghan, J. C., English, N. B., 2020. Journal of Southern Hemisphere Earth System Science 70, 180–192. <https://doi.org/10.1002/joc.6368>.
- Greenslade, D.J.M., 2001. A wave modelling study of the 1998 Sydney to Hobart yacht Race. *Aust. Meteorol. Mag.* 50, 53–63.
- Hart, R.E., 2003. A cyclone phase space derived from thermal wind and thermal asymmetry. *Mon. Weather Rev.* 131, 585–616. [https://doi.org/10.1175/1520-0493\(2003\)131<0585:ACPSDF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0585:ACPSDF>2.0.CO;2).
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146, 1999–2049. <https://doi.org/10.1002/qj.3803>.
- Holland, G., 1997. The east coast cyclone. In: Webb, E.K. (Ed.), *Windows in Meteorology: Australian Perspective*. CSIRO Publishing, Collingwood, pp. 242–245.
- Holland, G., McBride, J., 1997. Tropical cyclones. In: Webb, E.K. (Ed.), *Windows in Meteorology: Australian Perspective*. CSIRO Publishing, Collingwood, pp. 200–216.
- Holland, G.J., Lynch, A.H., Leslie, L.M., 1987. Australian east-coast cyclones. Part I: synoptic overview and case study. *Mon. Weather Rev.* 2 [https://doi.org/10.1175/1520-0493\(1987\)115<3024:AECCPI>2.0.CO](https://doi.org/10.1175/1520-0493(1987)115<3024:AECCPI>2.0.CO).
- Ji, F., Pepler, A.S., Browning, S., Evans, J.P., Di Luca, A., 2018. Trends and low frequency variability of East Coast Lows in the twentieth century. *J. South. Hemisph. Earth Syst. Sci.* 68, 1–15. <https://doi.org/10.22499/3.6801.001>.
- Jones, D.A., Simmonds, I., 1993. A climatology of Southern Hemisphere extratropical cyclones. *Clim. Dynam.* 9, 131–145. <https://doi.org/10.1007/BF00209750>.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society* 77 (3), 437–472. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- Kiem, A., Twomey, C., Lockart, N., Willgoose, G., Kuczera, G., Chowdhury, A.K., Parana Manage, N., Zhang, L., 2016. Links between East Coast Lows and the spatial and temporal variability of rainfall along the eastern seaboard of Australia. *J. South. Hemisph. Earth Syst. Sci.* 66, 162–176. <https://doi.org/10.22499/3.6602.006>.
- Kobayashi, C., Endo, H., Ota, Y., Kobayashi, S., Onoda, H., Harada, Y., Onogi, K., Kamahori, H., 2014. Preliminary results of the JRA-55c, an atmospheric reanalysis assimilating conventional observations only. *Sci. Online Lett. Atmos.* 10, 78–82. <https://doi.org/10.2151/sola.2014-016>.
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Kiyotoshi, T., 2015. The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteorol. Soc. Japan* 93, 5–48. <https://doi.org/10.2151/jmsj.2015-001>.
- Laurencin, C.N., Misra, V., 2019. Characterising the variations of the motion of the north Atlantic tropical cyclones. *Meteorol. Atmos. Phys.* 131, 225–236. <https://doi.org/10.1007/s00703-017-0566-1>.
- Lavender, S.L., 2017. A climatology of Australian heat low events. *Int. J. Climatol.* 37, 534–539. <https://doi.org/10.1002/joc.4692>.
- Mills, G.A., 2001. Mesoscale cyclogenesis in reversed shear - the 1998 Sydney to Hobart yacht Race storm. *Aust. Meteorol. Mag.* 50, 29–52.
- Murray, R., Simmonds, I., 1991. A numerical scheme for tracking cyclone centres from digital data. Part II: Application to January and July general circulation model simulations. *Australian Meteorological Magazine* 39 (3), 167–180.
2020. NCSS Statistical Software: Chapter 206: Two-sample t-test. (Accessed 5 October 2020).
- Pepler, A., 2020a. Record lack of cyclones in southern Australia during 2019. *Geophys. Res. Lett.* 47, 1–9. <https://doi.org/10.1029/2020GL088488>.
- Pepler, A., 2020b. Australian region cyclones, 1950–2019. Rec. lack cyclones South. Aust. Dur. <https://doi.org/10.6084/m9.figshare.c.4944135.v1>, 2019.
- Pepler, A., Coutts-Smith, A., 2013. A new, objective, database of East Coast Lows. *Aust. Meteorol. Oceanogr.* J. 63, 461–472. <https://doi.org/10.22499/2.6304.001>.
- Pepler, A., Coutts-Smith, A., Timbal, B., 2014. The role of East Coast Lows on rainfall patterns and inter-annual variability across the East Coast of Australia. *Int. J. Climatol.* 34, 1011–1021. <https://doi.org/10.1002/joc.3741>.
- Pepler, A., Dowdy, A., 2021. Intense east coast lows and associated rainfall in eastern Australia. *J. South. Hemisph. Earth Syst. Sci.* 71, 110–122.
- Pepler, A.S., Di Luca, A., Ji, F., Alexander, L.V., Evans, J.P., Sherwood, S.C., 2015. Impact of identification method on the inferred characteristics and variability of Australian East Coast Lows. *Mon. Weather Rev.* 143, 864–877. <https://doi.org/10.1175/mwr-d-14-00188.1>.
- Pepler, A.S., Dowdy, A.J., Hope, P., 2021. The Differing Role of Weather Systems in Southern Australian Rainfall between 1979–1996 and 1997–2015. *Clim. Dyn.* <https://doi.org/10.1007/s00382-020-05588-6>.
- Pepler, A.S., Imielska, A., Coutts-Smith, A., Gamble, F., Schweitzer, M., 2016. Identifying East Coast Lows with climate hazards on the eastern seaboard. *J. South. Hemisph. Earth Syst. Sci.* 66, 97–107.
- Pepler, A.S., Rakich, C.S., 2010. Extreme inflow events and synoptic forcing in Sydney catchments. *IOP Conf. Ser. Earth Environ. Sci.* 11, 012010 <https://doi.org/10.1088/1755-1315/11/1/012010>.
- Power, S., Callaghan, J., 2016. The frequency of major flooding in coastal Southeast Australia has significantly increased since the late 19th century. *J. South. Hemisph. Earth Syst. Sci.* 66, 2–11. <https://doi.org/10.22499/3.6601.002>.
- Quinting, J.F., Catto, J.L., Reeder, M.J., 2019a. Synoptic climatology of hybrid cyclones in the Australian region. *Q. J. R. Meteorol. Soc.* 145, 288–302. <https://doi.org/10.1002/qj.3431>.
- Quinting, J.F., Reeder, M.J., Catto, J.L., 2019b. The intensity and motion of hybrid cyclones in the Australian region in a composite potential vorticity framework. *Q. J. R. Meteorol. Soc.* 145, 273–287. <https://doi.org/10.1002/qj.3430>.
- Shapiro, M.A., Keyser, D., American Meteorological Society, 1990. Fronts, jet streams and the tropopause. In: *Extratropical Cyclones: the Erik Palmen Memorial Volume*, pp. 167–191. https://doi.org/10.1007/978-1-944970-33-8_10. Boston, MA.
- Sharma, K. K., Verdon-Kidd, D. C., Magee, A. D., 2019. Decadal variability of tropical cyclogenesis and decay in the southwest Pacific. *International Journal of Climatology* 39 (15), 1–19. <https://doi.org/10.1002/joc.6368>.
- Simmonds, I., Murray, R., 1999. Southern extratropical cyclone behaviour in ECMWF analyses during the FROST special observing periods. *Weather and Forecasting* 14, 878–891. [https://doi.org/10.1175/1520-0434\(1999\)014<0878:SECBIE>2.0.CO;2](https://doi.org/10.1175/1520-0434(1999)014<0878:SECBIE>2.0.CO;2).
- Simmonds, I., Murray, R., Leighton, R., 1999. A refinement of cyclone tracking methods with data from FROST. *Australian Meteorological Magazine* 48, 35–49.
- Speer, M., Leslie, L., Hartigan, J., Macnamara, S., 2021. Changes in frequency and location of east coast low pressure systems affecting southeast Australia. *Climate* 9, 1–21. <https://doi.org/10.3390/cli9030044>.
- Speer, M.S., Wiles, P., Pepler, A., 2009. Low pressure systems off the New South Wales coast and associated hazardous weather: establishment of a database. *Aust. Meteorol. Oceanogr.* J. 58, 29–39. <https://doi.org/10.22499/2.5801.004>.
- Sturman, A., Tapper, N., 2005. *The Weather and Climate of Australia and New Zealand, second ed.* Oxford University Press, Melbourne Australia.
- Tapper, N., Hurry, L., 1933. *Australia's Weather Patterns: an Introductory Guide.* Dellasta PTY LTD, Mount Waverley, Victoria.
- Verdon-Kidd, D., Kiem, A., Willgoose, G., 2016. East coast lows and the Pasha Bulker storm - lessons learned nine years on. *J. South. Hemisph. Earth Syst. Sci.* 66, 152–161. <https://doi.org/10.22499/3.6602.005>.
- Verdon-Kidd, D., Kiem, A., Willgoose, G., Haines, P., 2010. east coast lows and the newcastle/central coast Pasha bulker storm. *Gold Coast* 61.
- WaterNSW, 2020. Water Supply Historical Timeline: 2007–2009. <https://www.watnsw.com.au/supply/heritage/timeline>. (Accessed 5 October 2021).
- Watson, P., Lord, D., Kulmar, M., McLuckie, D., James, J., 2007. Analysis of a Storm – June 2007.