Evaluating the representation of Australian East Coast Lows in a regional climate model ensemble

Alejandro Di Luca,[1] Jason P. Evans,[1] Acacia S. Pepler,[1] Lisa V. Alexander[1] and Daniel Argüeso[1,2]

[1] Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, NSW, Australia.
[2] Department of Atmospheric Sciences, SOEST, University of Hawaii at Mānoa, Honolulu, Hawai'i, USA.

(Manuscript received February 2016; accepted April 2016)

Due to their large influence on both severe weather and water security along the east coast of Australia, it is increasingly important to understand how East Coast Lows (ECLs) may change over coming decades. Changes in ECLs may occur for a number of reasons including changes in the general atmospheric circulation (e.g. poleward shift of storm tracks) and/or changes in local conditions (e.g. changes in sea surface temperatures). Numerical climate models are the best available tool for studying these changes however, in order to assess future projections, climate model simulations need to be evaluated on how well they represent the historical climatology of ECLs. In this paper, we evaluate the performance of a 15-member ensemble of regional climate model (RCM) simulations to reproduce the climatology of cyclones obtained using three high-resolution reanalysis datasets (ERA-Interim, NASA-MERRA and JRA55). The performance of the RCM ensemble is also compared to results obtained from the global datasets that are used to drive the RCM ensemble (four general circulation model simulations and a low resolution reanalysis), to identify whether they offer additional value beyond the driving data. An existing cyclone detection and tracking algorithm is applied to derive a number of ECL characteristics and assess results at a variety of spatial scales. The RCM ensemble offers substantial improvement on the coarse-resolution driving data for most ECL characteristics, with results typically falling within the range of observational uncertainty, instilling confidence for studies of future projections. The study clearly highlights the need to use an ensemble of simulations to obtain reliable projections and a range of possible future changes.

1. Introduction

The importance of mid-latitude cyclones on the eastern coast of Australia, known locally as East Coast Lows (ECLs), has been discussed by a number of authors (e.g. Hopkins and Holland 1997, Pepler and Rakich 2010, Dowdy et al. 2014). ECLs cause much of the severe weather that influences this highly-populated region, causing large waves, strong winds, heavy rain, and substantial coastal erosion. However, they are also critically important to water security including long-term dam storages.

For these reasons, there is substantial interest in how the frequency, intensity, location or impacts of ECLs may change in the future. Recent studies (Dowdy et al. 2013, 2014, Ji et al. 2015) have attempted to address this by providing projections

Corresponding author: Alejandro Di Luca, Climate Change Research Centre, Level 4, Mathews Building, University of New South Wales, Sydney NSW 2052, Australia.
Email: a.diluca@unsw.edu.au

of some large-scale atmospheric conditions that were shown to favour the development of ECLs. These studies indicate a mean decline in the frequency of ECLs of 25–42% by the end of the 21st century, mostly due to a reduction of winter ECLs. It should be noted that these studies implicitly assumed a static relationship between the upper-level diagnostic and surface low development over time.

Several studies have shown that the total number and other properties of cyclones, such as their size and intensity, are strongly dependent on the horizontal resolution of the data used to identify and track the low pressure systems (Jung et al. 2006, Greeves et al. 2007, Grose et al. 2012, Colle et al. 2013, Shkolnik and Efimov 2013, Zappa et al. 2013, Di Luca et al. 2015). Some of these articles have shown that increases in the horizontal resolution of climate models improve the representation of cyclones and that these improvements may arise even when comparing results at common lower resolutions due to a better representation of the larger scales through the upscaling of fine-scale phenomena (Greeves et al. 2007, Colle et al. 2013, Shkolnik and Efimov 2013, Zappa et al. 2013). For example, Colle et al. (2013) evaluated the representation of the intensity and frequency of cyclones in 15 global climate models (GCMs) of phase 5 of the Coupled Model Intercomparison Project (CMIP5) over the eastern North America. They showed that six of the top seven best performing models had the highest spatial resolution. Moreover, as shown by Jung et al. (2006), it is expected that higher-resolution simulations might be able to represent some small systems that are absent in lower resolution datasets. For this reason, it is proposed that an ensemble of high-resolution dynamical downscaling simulations can be used to obtain future projections of the frequency and characteristics of ECLs.

As a prerequisite step towards providing future projections of ECL activity, we aim to evaluate the suitability of an ensemble of regional climate model simulations performed in the context of the New South Wales/Australian Capital Territory Regional Climate Modelling (NARCliM) project. Specifically, the purpose of this study is to perform a detailed evaluation of the representation of cyclones in the NARCliM ensemble and to assess the potential advantages of the ensemble compared to the lower resolution datasets used to drive the RCM simulations. The analysis includes the evaluation of various characteristics of ECLs including their frequency, intensity, size and duration. We focus on maritime cyclones that appear over the Tasman Sea to the east of Australia, using a single automated cyclone detection and tracking algorithm that employs 6-hourly sea level pressure (SLP) fields as input data (Browning and Goodwin 2013, Di Luca et al. 2015). While ECL climatologies are somewhat sensitive to the choice of tracking method and identification criteria, we expect that these choices are of second order importance when looking at differences between datasets for which ECLs were identified using the same method. Moreover, Pepler et al. (2015) have shown that the most significant ECL events are identified regardless of method.

Previous research found that the frequency of ECLs identified from reanalyses data is highly sensitive to both the choice of the reanalysis and the horizontal resolution of the input SLP (Di Luca et al. 2015). For this reason, we assess ECL characteristics as simulated by the RCM ensemble at two different resolutions by comparing with three different high-resolution reanalysis datasets, available at horizontal resolutions that are comparable to the RCM simulations. The resolution dependence is assessed by comparing results for ECLs identified using high-resolution SLP (about 50 km) and low-resolution SLP (about 300 km) fields (Di Luca et al. 2015). In addition to the RCM results, we also include in the comparison results from the coarse-resolution driving data in order to assess whether the RCMs provide improvements at coarse scales.

The paper is structured as follows: in Section 2 the RCM ensemble and the reference datasets are briefly described; in Section 3, the methodology is discussed including a description of the identification and tracking algorithm and of the metrics used for evaluation. Section 4 shows results of the mean characteristics, distributions and spatial variability of ECLs when using low-resolution SLP fields while Section 5 presents results for ECLs obtained using the high-resolution SLP fields. Finally the summary and conclusions of the work are presented in Section 6.

2. 6-hourly SLP data

There are three types of 6-hourly SLP fields used in this study. The first group is given by three high-resolution reanalyses, which are used as the reference datasets in the evaluation process. A second group consists of an ensemble of high-resolution RCM simulations while a third group contains low-resolution reanalyses and global simulations used to drive the various members of the RCM ensemble. A brief description of each dataset is given below while an overview of the data can be found in Table 1.

Dataset	Δy by Δx *(km)*	*Analysis (times)*	*Atmospheric model type*	*Reference*
MERRA	∼55x60	00-06-12-18	Finite Volume (2/3°x1/2°)	Rienecker et al. (2011)
ERA-I	∼80x70	00-12	Spectral (T255)	Dee et al. (2011)
JRA55	∼55x55	00-06-12-18	Spectral (T319)	Kobayashi et al. (2015)
WRF-R1 RCM	50x50	00-06-12-18	Grid point	
WRF-R2 RCM	50x50	00-06-12-18	Grid point	Skamarock (2004), Evans et al. (2014)
WRF-R3 RCM	50x50	00-06-12-18	Grid point	
NNRP	∼250x230	00-06-12-18	Spectral (T62)	Kalnay et al. (1996)
CGCM3.1	∼350x350	00-06-12-18	Spectral (T47)	Kim et al. (2002)
ECHAM5	∼250x230	00-06-12-18	Spectral (T63)	Roeckner et al. (2003)
MIROC3.2	∼350x350	00-06-12-18	Spectral (T42)	
CSIRO-MK3.0	∼250x230	00-12	Spectral (T63)	(Gordon et al. 2002)

Table 1 Horizontal (Δx) spacing of the available data together with the type of model for the three reanalyses, the RCM simulations and the driving data used in the study.

2.1 Evaluation datasets: high-resolution reanalyses

In order to evaluate the performance of RCMs and driving data at simulating midlatitude cyclones obtained from 6-hourly SLP fields, three different high-resolution reanalyses are used: MERRA, ERA-I and JRA55. The three reanalyses are used to take into account uncertainties in the reanalyses that have been shown to be substantial, particularly when considering ECLs obtained using high-resolution SLP fields (Di Luca et al. 2015).

ERA-Interim (Dee et al. 2011; ERA-I) is a global atmospheric reanalysis provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA-I uses a T255 horizontal resolution, which corresponds to approximately 79 km spacing on a reduced Gaussian grid and a total of 60 sigma-pressure levels. ERA-I also uses a more sophisticated data assimilation scheme (4-D var) and assimilates a greater number of observations than its predecessor ERA-40, leading to a general improvement in the representation of the hydrological cycle and a better temporal consistency on a range of time-scales (Dee et al. 2011). ERA-I SLP fields are available globally every 6 hours from the 1 January 1979 onwards. The data assimilation scheme is applied every 12 hours, at 0000 and 1200 UTC, and so SLP fields valid at 0600 and 1800 correspond to model forecasts.

The NASA Modern Era Retrospective-Analysis for Research and Applications (MERRA) reanalysis (Rienecker et al. 2011) uses version 5.2.0 of the Goddard Earth Observing System (GEOS-5) and the GSI data assimilation system (Rienecker et al. 2008) to construct their analysis of observations. The atmospheric system uses a finite volume model, integrated over a grid with horizontal spacings of 2/3° and 1/2° in longitude and latitude respectively (i.e. equivalent to about 55 km at 45° of latitude) and with 72 sigma vertical levels. SLP analysis is available four times per day at 0000, 0600, 1200 and 1800 UTC; and SLP forecasts are available every hour. Both analyses and forecasts span the period from 1979 to the present.

The Japan Meteorological Agency conducted the second Japanese global atmospheric reanalysis covering the 55 years from 1958 (Kobayashi et al. 2015; JRA55). JRA55 incorporates many improvements compared to the previous version (JRA25), including a revised longwave radiation scheme, a 4-D var assimilation scheme and a variational bias correction scheme for satellite radiances. The JRA55 system is based on the T319 spectral resolution version (about 55-km grid spacing) of the Japan Meteorological Agency global spectral model (JMA 2013). SLP analyses are available four times per day at 0000, 0600, 1200 and 1800 UTC on a 0.5625 degrees grid spacing mesh.

2.2 NARCliM RCM ensemble

The RCM ensemble produced as part of the NARCliM project (Evans et al. 2014) is used here. It is a 15-member RCM ensemble constructed using three versions of the weather research and forecasting (WRF) RCM driven by five different low-resolution global datasets for a period from 1990 to 2009. The three different RCM versions were constructed using version 3.3 of the WRF model by combining different surface/planetary boundary layer, cumulus and atmospheric radiation schemes. Specific schemes used in each of the three versions were selected after an evaluation of model performance and independence over southeast Australia (Evans et al. 2012, Ji et al. 2014). The global driving data (GDD) comprises of the NCEP-NCAR

reanalysis 1 (Kalnay et al. 1996) and four GCMs from the phase 3 of the Coupled Model Intercomparison Project (CMIP3), which were also selected based on model performance and independence.

All simulations were performed using a double nesting approach where the GDD are used to drive the RCM at 50-km grid spacing over a domain that covers the CORDEX-Australasia region (RCM50, see Figure 1). RCM50 is then used to drive the same RCM with a horizontal grid spacing of 10 km over a domain that includes southeast Australia. For further details on the ensemble and the NARCliM experiment design the reader is referred to Evans et al. (2014). Here we focus on the evaluation of present climate results (1990–2009) for the 50-km grid spacing RCM over the outer domain, as it is the resolution most comparable to the reanalyses and allows ECL tracking over a broader region.

2.3 NARCliM global driving data (GDD)

As already stated, five global datasets are used to drive the RCM ensemble. The NCEP-NCAR reanalysis 1 project (Kalnay et al. (1996), hereafter NNRP) was produced through cooperation between the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR). The four CMIP3 GCMs simulations come from the fifth version of ECHAM (ECHAM5; Roeckner et al. (2003)), version 3.1 of the Canadian Centre for Climate Modelling and Analysis Model (CGCM3.1; Kim et al. (2002)), the Model for Interdisciplinary Research on Climate (MIROC3.2) and CSIRO-Mk3.0 (Gordon et al. 2002). Some basic information such as the approximate horizontal grid spacing and the type of model used by each GDD is shown in Table 1. In the case of the CSIRO-Mk3.0 simulation, SLP data were not directly available from the CSIRO modelling group so we have calculated it using the corresponding surface pressure, temperature and humidity. While three dimensional variables are needed to better estimate SLP fields, some of these variables (geopotential height) were not available preventing us from including them in the calculation. Moreover, CSIRO-Mk3.0 fields were available only every 12 hours so we have interpolated the calculated SLP every 6 hours using a simple linear interpolation. Consequently, CSIRO-Mk3.0 SLP fields may include some non-negligable errors. Despite this, we have decided to include them in the analysis.

As already stated, five global datasets are used to drive the RCM ensemble. The NCEP-NCAR reanalysis 1 project (Kalnay et al. (1996), hereafter NNRP) was produced through cooperation between the National Centers for Environmental Prediction and the National Center for Atmospheric Research (NCAR). The four CMIP3 GCMs simulations come from the fifth version of ECHAM (ECHAM5; Roeckner et al. (2003)), version 3.1 of the Canadian Centre for Climate Modelling and Analysis Model (CGCM3.1; Kim et al. (2002)), the Model for Interdisciplinary Research on Climate (MIROC3.2) and CSIRO-Mk3.0 (Gordon et al. 2002). Some basic information such as the approximate horizontal grid spacing and the type of model used by each GDD is shown in Table 1. In the case of the CSIRO-Mk3.0 simulation, SLP data were not directly available from the CSIRO modelling group so we have calculated it using the corresponding surface pressure, temperature and humidity. While three dimensional variables are needed to better estimate SLP fields, some of these variables (geopotential height) were not available preventing us from including them in the calculation. Moreover, CSIRO-Mk3.0 fields were available only every 12 hours so we have interpolated the calculated SLP every 6 hours using a simple linear interpolation. Consequently, CSIRO-Mk3.0 SLP fields may include some non-negligable errors. Despite this, we have decided to include them in the analysis.

3. Methodology

The methodology used in this study follows from Di Luca et al. (2015) and was specifically designed to compare results from datasets with different horizontal resolutions. As mentioned in the introduction, the evaluation of RCM simulations is performed independently for ECLs identified using high-resolution (about 50-km grid spacing) and low-resolution (about 300-km grid spacing) SLP fields.

3.1 Data pre-processing: spatial re-gridding

When considering ECLs identified using high-resolution SLP fields, analyses are performed directly on the native grid of each dataset, noting that this varies between the high-resolution reanalyses and the RCM simulations (see Table 1).

When comparing results at low horizontal resolutions, SLP fields from the high-resolution reanalyses and the RCM simulations are first upscaled onto a common low resolution grid mesh of 300-km grid spacing. That is, 6-hourly (i.e. 0000, 0600, 1200 and 1800 UTC) instantaneous SLP fields from each high-resolution reanalyses and each RCM simulation at their native

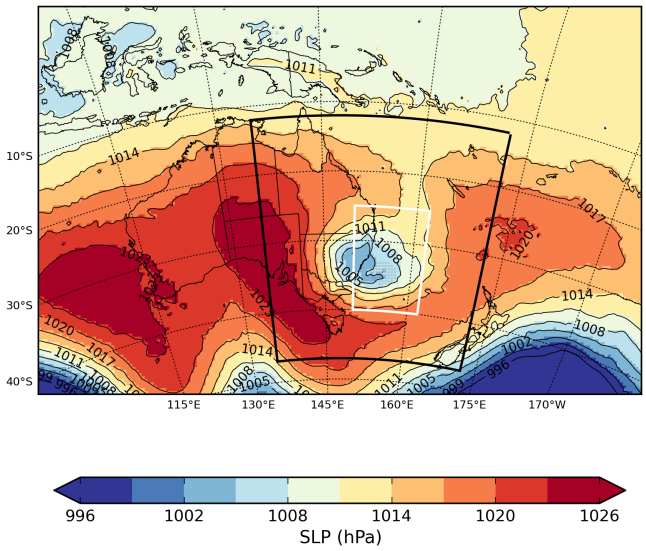

Figure 1 WRF-R2 RCM simulated instantaneous SLP field for a severe ECL at 1800 UTC 8 June 2007 over the outer 50-km NARCliM domain. The black rectangle shows the region where the tracking is applied and the white rectangle the region where the analysis is performed. Only maritime cyclones are considered in the analysis.

resolution are aggregated in space by calculating the average of the high-resolution SLP values inside 300-km grid boxes.

Finally, the last step of the pre-processing consists of interpolating all datasets, no matter their spatial resolution, onto a high-resolution (50 km) regular grid mesh that encompasses the region of interest. This interpolation is performed using a cubic spline algorithm, which allows for the likely variation on the SLP fields between grid points, as argued by Murray and Simmonds (1991).

3.2 Detection and tracking algorithm

ECLs are identified and tracked within the region denoted with a black line in Figure 1. The detection and tracking algorithm is based on the method developed by Browning and Goodwin (2013) with modifications that were detailed in Di Luca et al. (2015). The main characteristics of the detection algorithm used in this study are described below.

Lows are identified by searching for both a local minima in the SLP field and a SLP gradient ($\overline{\nabla p}$) around the local minima that exceeds a given threshold. The use of the local minima criterion implies that only closed lows are identified by the algorithm. The $\overline{\nabla p}$ is computed by averaging differences between the central SLP and the SLP in grid points located within a radius of 200 km around the central pressure ($\overline{\nabla p}_{200}$). The value of the 200-km SLP gradient threshold was chosen to be 0.6 hPa/100 km and is thus very similar to the threshold used by Browning and Goodwin (2013) of 1 hPa/1.5 degree.

Once lows have been detected for individual time steps, cyclone events are generated by grouping lows that are close in both time and space. Tracks are constructed by a nearest neighbour search in the following 6-hourly SLP field around a cyclone position. The search extends to a maximum distance that depends on the temporal resolution of the data assuming that a cyclone will not move faster than 60 km/h. The resulting search radius thus extends to about 360 km for 6-hourly data.

A number of lows appear to be quasi-stationary features that might be associated with either heat lows or with uncertainties in extrapolating the atmospheric surface pressure to mean sea level. In this analysis, we filter out some of these quasi-stationary systems by discarding cyclones that move at an average speed less than 5 km/h over the total duration of the event inside the region of interest (i.e. 120 km for 24 h events). The number of events that are filtered when imposing this condition is generally around 5% when using high-resolution SLP fields and around 10% when using low resolution SLP fields. While this may falsely remove some ‘true’ ECL events, it is necessary to allow proper comparison of the different datasets. Also, as in Di Luca et al. (2015), we only retain events that last at least two consecutive 6-hourly time steps over the region of interest.

Using the parameters specified above, we identify an exceptionally large number of events in the CSIRO Mk3.0 simula-

tion, with nearly five times more events than in the ERA-I dataset. A visual inspection of individual events suggests that a large number of them are still associated with stationary lows that appear near coastal regions, presumably due to some artificial lows created in the interpolation process. As a consequence, for this particular simulation, we have increased the cyclone's averaged speed from 5 km/h to 20 km/h. With this modified threshold, the total number of ECL events appears to be comparable to other GCM or reanalysis results.

3.3 Evaluation metrics

A comprehensive evaluation of ECLs in RCM simulations and GDDs is performed by assessing their representation of the duration, intensity and size of ECL events together with a measure of their frequency.

All statistics are calculated for the common 20-year period between 1990 and 2009 for which data from RCM simulations and reanalyses are available. Based on the annual cycle of the number of events at different resolutions presented in Di Luca et al. (2015), we have decided to show results for two six-month blocks. The cold season is defined from April to September, corresponding to the peak season of ECL impacts in Hopkins and Holland (1997), and the warm season from October to March. An advantage of using results over the whole year is that it maximises the number of cyclones, thus compensating for the relatively short 20-year periods used for the analysis. Also, ECL statistics are calculated within the region encompassed by 25–40° S and 162° E and the Australian coast (see white rectangle in Figure 1) thus only considering maritime ECLs.

The intensity of cyclones will be evaluated by the same quantity adopted for the tracking (i.e. mean value of the 200-km SLP gradient along the track of the cyclone). Alternative measures of the intensity such as the event-maximum 200-km SLP gradient or the event-mean SLP Laplacian calculated at the centre of the cyclone lead to similar results. The duration of the event is obtained by counting the number of time steps the cyclone spends in the region of analysis and multiplying by the time step of 6 hours. That is, the duration metric used here does not measure the duration of the whole event but the time that a given low spends over the region of interest. Clearly, event durations calculated in this way will be lower than those obtained if considering the whole life time of a given low.

Following Rudeva and Gulev (2007), the radius of a given cyclone is estimated by calculating the area inside the so-called 'last closed isobar'. First, we determine the locations for which the first radial derivative of SLP falls to zero across eight radial lines (N, NE, E,...) that pass through the centre of the cyclone. The first derivative is calculated using the 50-km grid mesh for a distance of 850 km from the centre. The locations where the derivative is zero are used to determine a 'critical SLP' in each direction and the last closed isobar corresponds to the minimum of all the 'critical' values. In those cases where the derivative is never zero, we set the radius to a maximum value of 850 km and the critical SLP corresponds to the SLP value at 850 km from the centre. This occurs in only 2–3% of cases. Once the SLP of the last closed isobar is known, distances of this isobar to the centre are obtained by interpolation thus obtaining a radius in each radial direction. We then calculate the area encompassed by the last closed isobar by summing over the triangles obtained using individual radii in different directions. Finally, the radius of the cyclone is obtained by calculating the radius of a circumference that has the same area as the one obtained from the triangle summation.

Statistical confidence of correlation coefficients is tested by assuming that correlations are distributed following a Student's t-distribution. At the 5% significance level, the correlation coefficient can be calculated as $r_{sig} = \sqrt{\frac{1}{1+(N-2)/t^2_{95,N-2}}}$ with N the size of the sample.

4. Comparison using ECLs derived from low-resolution SLP fields

Figure 2 shows scatter plots of the mean number of events versus their mean duration (top panels) and the event-mean intensity versus their mean size (bottom) in warm (left) and cold (right) seasons. Each individual plot shows event-mean results obtained using the smooth version (i.e., about 300-km grid mesh) of the SLP fields from all datasets: high-resolution reanalyses, RCM simulations (i.e. 15 simulations) and driving data. In order to summarise results obtained by the various groups of datasets, three mean values are also included in each plot: the mean value across the three high-resolution reanalyses (red square), the mean value across RCM simulations driven by GCM data (black square) and the mean value across GCM results (grey square). That is, the RCM and GDD ensemble mean values exclude results that use the NCEP/NCAR reanalysis data in order to evaluate ensemble-mean results that can be used for future ECL projections.

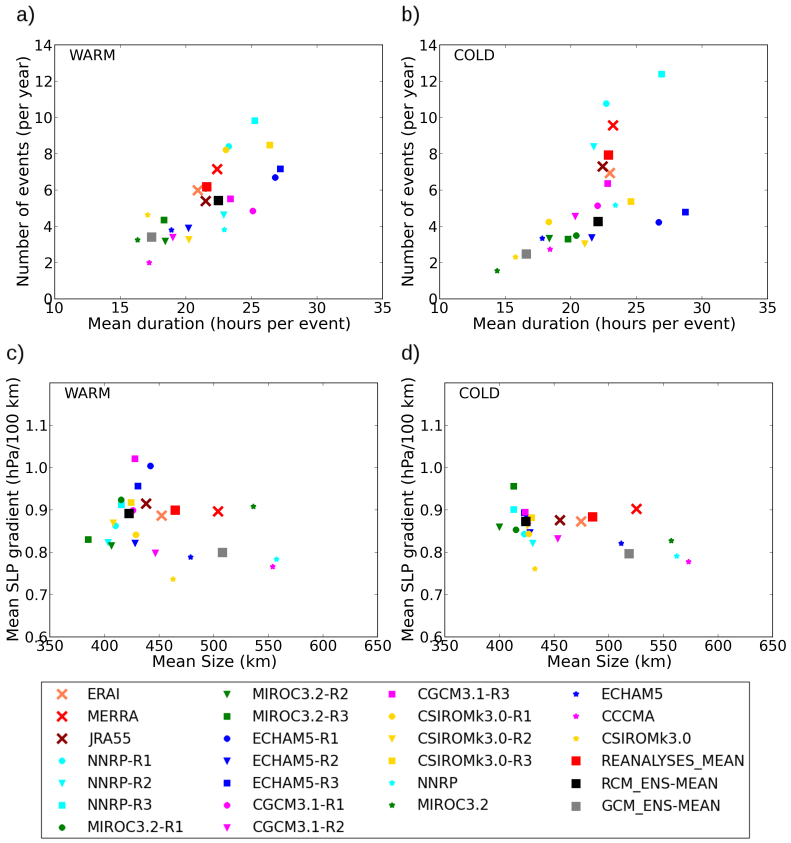

Figure 2 Scatter plots of the number of events Vs. their mean duration (a and b) and the event-mean 200-km SLP gradient Vs. event-mean size (c and d). Top panels show results for the warm season events and bottom panels for the cold season events. Ensemble mean values are shown in red, black and grey for the reanalyses, the RCM and the GDD ensemble respectively. The RCM and GDD ensemble mean values exclude results using the NCEP/NCAR reanalysis data. Only results obtained using the low resolution SLP fields (i.e. at the common 300-km grid mesh) are shown.

It is first worth noting that regardless of the metric considered, results from the three high-resolution reanalyses are in good agreement with differences across reanalyses being much smaller than the range of responses from the RCM and GDD ensembles. The RCM-ensemble mean value generally performs well compared to the reanalyses although it is clear that individual simulations can show quite poor performances with some showing nearly twice (e.g. NNRP-R3 in the warm season) and others less than half (e.g. CSIROMk3.0-R2 in the cold season) the number of events in the reanalyses. The large spread across RCM simulations is explained by both the choice of the RCM version and the choice of the GDD although in general the choice of the RCM version has a larger impact on the results particularly during the warm season.

There are, however, some systematic differences between the RCM ensemble mean and the high-resolution reanalysis results. In the cold season, RCM simulations largely underestimate the number of events compared with the reanalyses ensemble mean. An exception to this underestimation is given by simulations driven by the NCEP/NCAR reanalysis that produce a much larger number of events compared to GCM-driven simulations and even too many events compared to the reanalyses. Also, in both seasons, the mean size of ECLs tends to be underestimated compared to the three reanalyses by about 40 to 60 km.

The GDD ensemble also shows reasonable results for the number of events and their mean duration although it consistently shows poorer performance compared to the RCM ensemble mean. As for the RCM ensemble, the performance of the GDD-ensemble mean results from a large compensation between the individual GDDs. For example, when considering the number of events in warm season, the CSIRO-MK3.0 GCM shows about 6 events per year while CGCM3.1 shows less than 2 events per year. Compared to the reanalyses, the GDD systematically shows too few cyclones that are too weak, too short and too

large. A general underestimation of the intensity of cyclones by coarse resolution datasets was also found by Zappa et al. (2013) and Grieger et al. (2014), especially when considering the coarsest horizontal resolutions GCMs.

Figure 3 shows the relative frequency of events as a function of the event-mean intensity for high-resolution reanalyses, individual RCM simulations and GDDs. In both seasons but particularly in the cold season (right panels in Figure 3), the RCM ensemble improves the representation of the frequency distributions compared to the GDD ensemble for the intensity and the durations of events. More importantly, Figure 3 shows that the performance of the RCM ensemble to represent the various ECL metrics does not deteriorate as we consider more extreme ECLs, suggesting that the advantages of the RCM ensemble compared to the GCM ensemble appear more evident as we consider more extreme ECLs. For example, while the GCM ensemble does not produce any events with an intensity of 1.2 hPa/100 km and a duration longer than 12 hours, both the reanalyses and the RCM ensemble have about 6 events in 20 years for such kinds of cyclone.

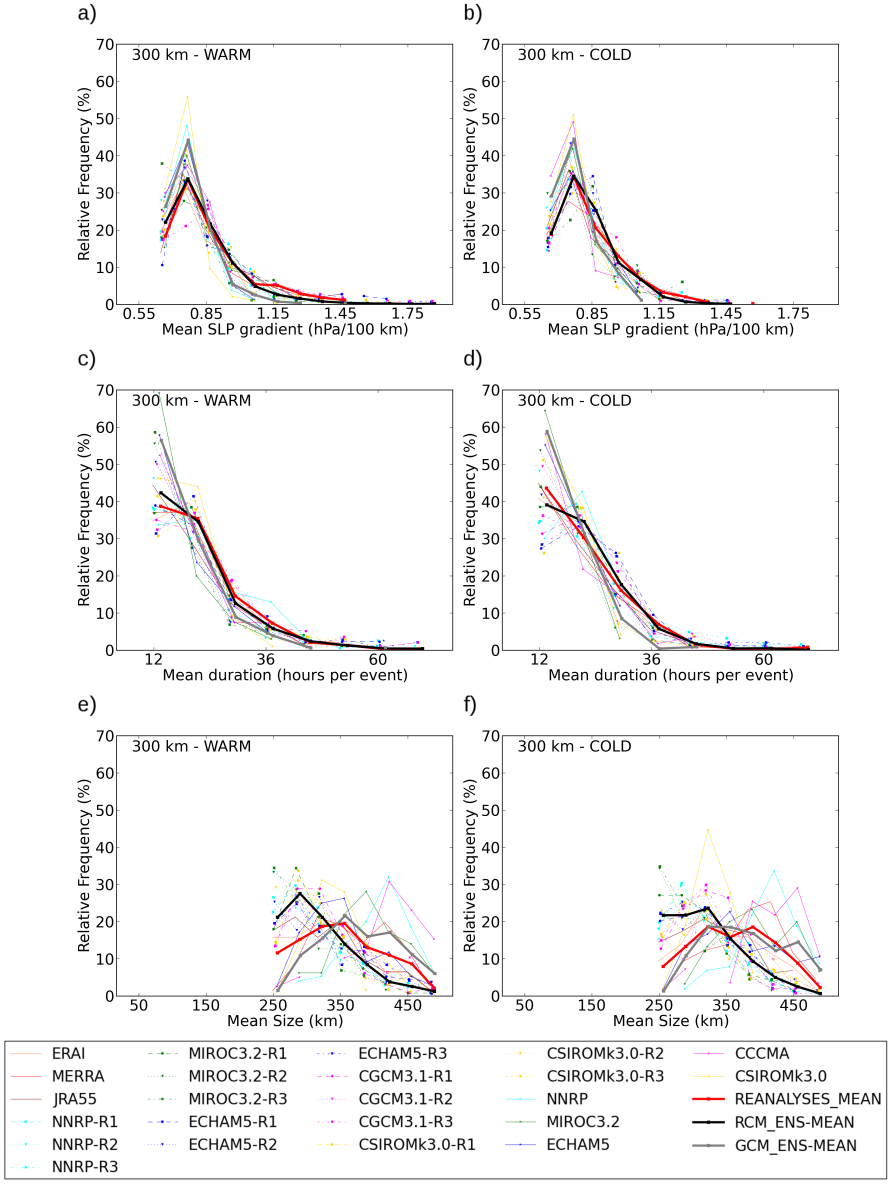

Figure 3 Frequency distribution of the intensity, duration and size for ECL events identified at the common 300-km grid mesh for the warm season (left panels) and the cold season (right panels). Different colours show the different datasets available in each case. Red, black and grey thick lines show results for the mean values across the reanalyses, the RCM ensemble and the GCM ensemble respectively.

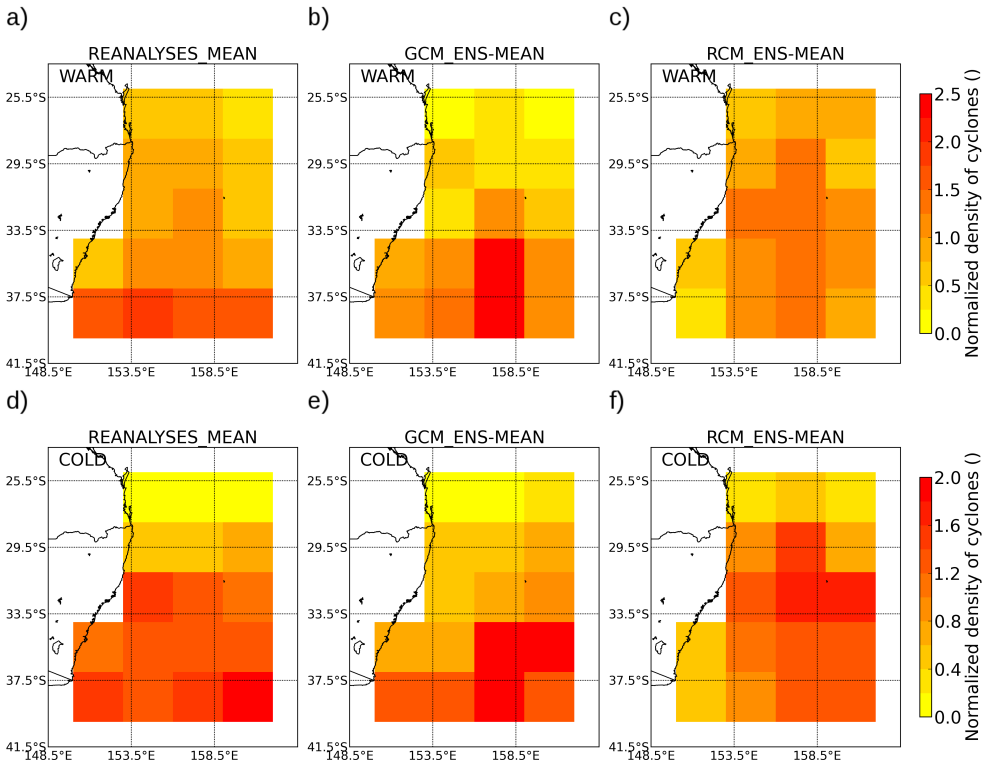

Figure 4 Normalised number of lows for the mean across reanalyses (a and d), the mean across the driving GCMs (b and e) and the mean across RCM members (c and f). Top panels show results for the warm season events and bottom panels for the cold season events. All values are normalised by the seasonal mean number of lows over the whole region.

Figure 3 also shows that GDD members generally overestimate the size of events (see bottom panels), the largest of which come from CCCMA3.1 and NNRP suggesting that the horizontal resolution is a clear limiting factor to get the correct size of lows. On the contrary, in both seasons, most RCM members produce too many small lows and too few large systems compared with the high-resolution reanalyses.

To analyse the spatial variability of ECLs over the region of interest, Figure 4 shows the spatial distribution of the number of lows for the reanalysis ensemble mean (a and d), the GDD ensemble mean (b and e) and the RCM ensemble mean (c and f) for the warm (top panels) and the cold (bottom panels) seasons. In both seasons, the high-resolution reanalyses (left panels in Figure 4) show a clear north–south gradient with about seven times more cyclones in the south than in the north. The GDD and RCM ensembles represent the north/south gradient in cold season reasonably well. However, in the warm season, the GDD ensemble shows a stronger north–south gradient while the RCM ensemble shows a maxima over the centre of the region with no clear north/south gradient.

The ability of individual RCM members and GDDs to represent the spatial distribution of ECLs is quantified in Figure 5. It shows mean square errors and spatial correlations between the reanalysis ensemble mean and all other datasets. Clearly, in both seasons, there is a large agreement across the high-resolution reanalyses with correlations between individual reanalyses and the mean reanalysis always higher than 0.9. Spatial correlations between the reanalyses mean and individual members of the RCM and GDD ensembles span a very large range of values. In the warm season (Figure 5a), with the exception of a few simulations (e.g. R2 RCM simulations), correlations are generally small and not statistically significant suggesting the poor ability of the models to represent the spatial variability of cyclones in this season. It is worth noting that RCMs tend to improve GDDs errors, but worsen their spatial correlation.

In the cold season (Figure 5b), spatial correlations are now statistically significant for most members of the RCM and GCM ensembles. The RCM ensemble improves mean square errors compared with GCMs although without improvements in the spatial distribution of the number of lows.

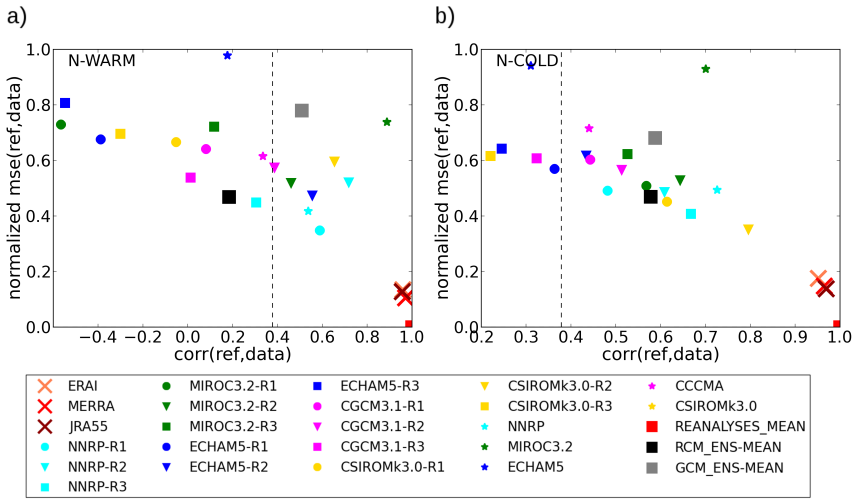

Figure 5 Mean square errors as a function of spatial correlations for the number of events in individual datasets compared with the mean across the reanalyses. Warm season results are shown in a) and cold season results in b).

5. Comparison using ECLs derived from high-resolution SLP fields

Figure 6 shows scatter plots of the number of events versus their mean duration (top panels) and the event-mean intensity versus their mean size (bottom panels) for results obtained using the high-resolution SLP fields. The mean value across the three reanalyses and across GCM-driven RCM simulations (i.e. excluding simulations driven by the NNRP) are also shown by the red and black squares respectively.

In the warm season (left panels in Figure 6), the three high-resolution reanalyses show large agreement with an average number of 27 events per year. In this season, the RCM ensemble shows a good representation of the event-average duration although it shows a large overestimation of the number of events with an average of near 60 events per year, thus about twice the number of events than reanalyses. ECLs in the RCM ensemble are also generally slightly stronger (by about 0.1 hPa/100 km) and substantially smaller (by about 70 km) than in the reanalyses.

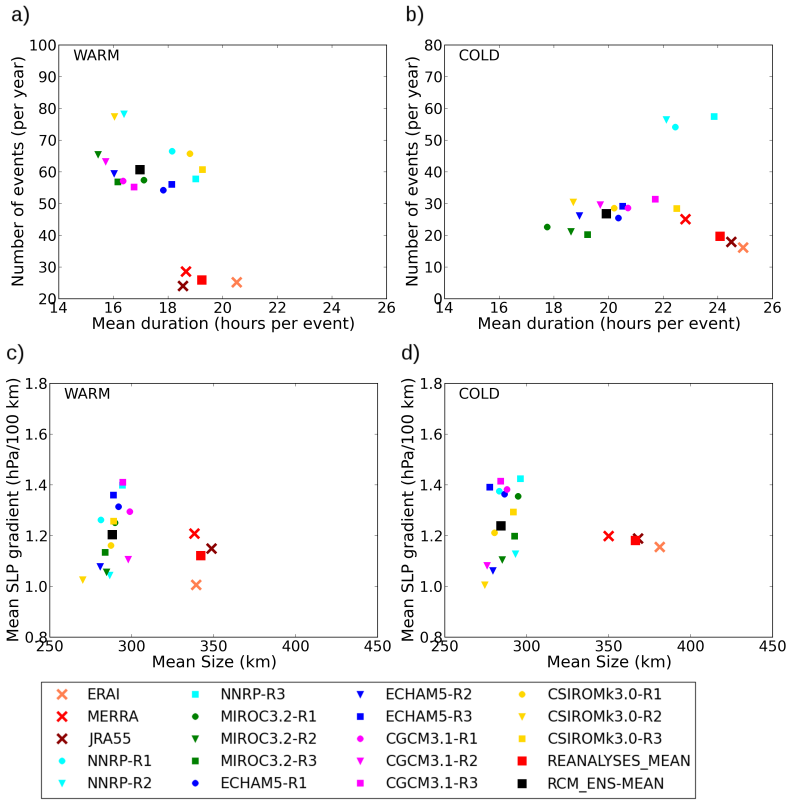

Figure 6 Same as in Figure 2 but for results obtained using SLP fields at their native resolution (see Table 1 for details on the horizontal resolution of individual datasets).

In the cold season (right panels in Figure 6), the RCM ensemble represents well the total number of events compared with high-resolution reanalyses. Surprisingly, simulations driven by the NNRP reanalysis tend to largely overestimate the number of events and, on average, they perform worse than the GCM-driven simulations. That is, simulations driven by the more realistic large-scale conditions provided by the NNRP reanalysis do not necessarily produce results that are closer to the high resolution reanalyses than those simulations driven by GCMs.

In agreement with the high-resolution reanalyses, all RCM members produce less events with longer durations in cold rather than warm seasons. RCM members do not show an increase in the size of these lows in the cold season, a characteristic that the three reanalyses show. As in the warm season, the RCM ensemble substantially underestimates the mean size and duration of ECLs when compared to the reanalyses.

In both seasons, as found for low-resolution SLP results, the variability across the RCM ensemble depends on the choice of both the RCM version and the driving data. In the warm season, R2 RCM simulations produce the largest number of events together with the weakest mean intensities and smallest sizes suggesting that this version tend to develop relatively weak ECL events more often than any other version.

Top panels in Figure 7 show the relative frequency of events as a function of the event-mean intensity for individual reanalyses and RCM simulations and their mean values. In both seasons, the RCM ensemble mean distribution (thick black line) is very close to the mean reanalyses distribution (thick red line) with only a small underestimation of the number of very weak ECLs in the warm season (Figure 7a). Importantly, the RCM ensemble seems to reproduce the ECL climatology of the most intense events well.

In both seasons, but particularly in the warm season, the uncertainty across reanalyses is quite large with the largest discrepancies arising between the ERA-I and the MERRA reanalyses, with the JRA55 reanalysis usually showing an intermediate

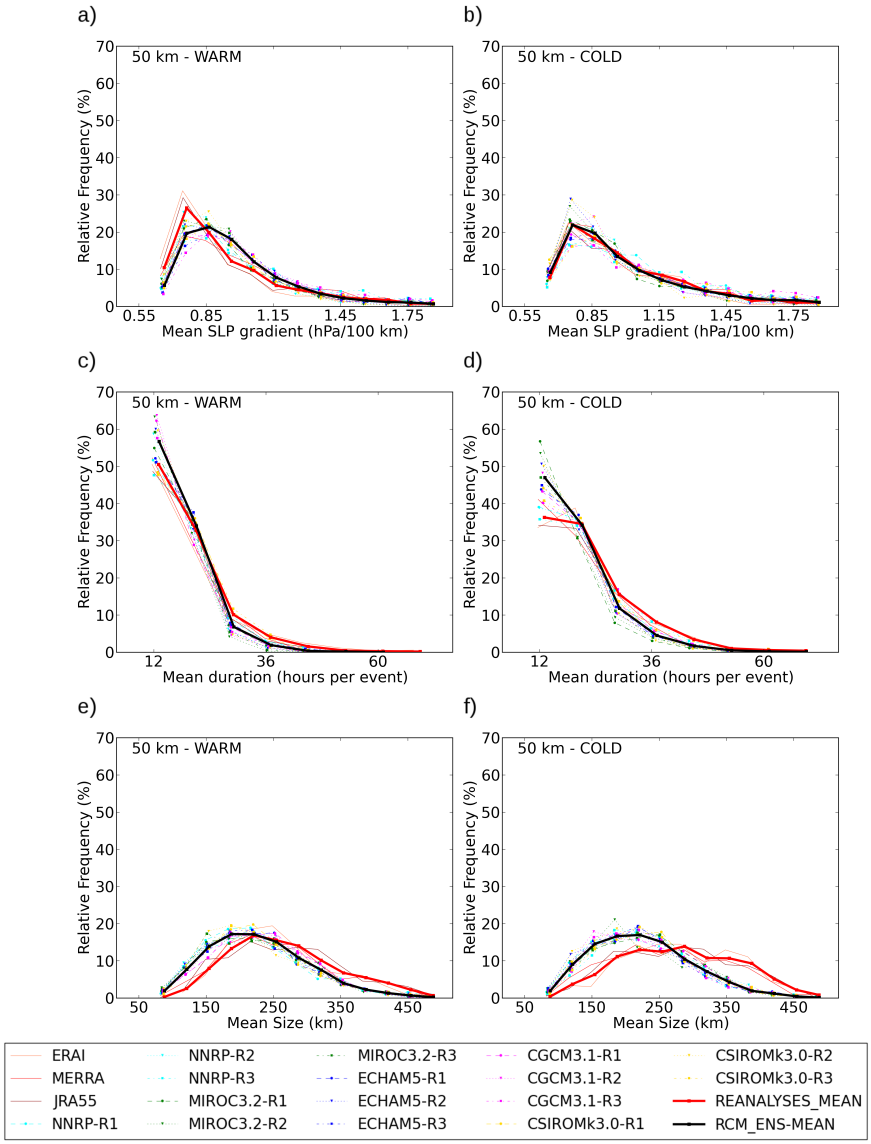

Figure 7 Same as in Figure 3 but for results obtained using SLP fields at their native resolution (see Table 1 for details on the horizontal resolution of individual datasets).

behaviour. As a consequence, although the RCM simulations show a substantial range of responses, their frequencies are generally within the reanalyses uncertainty. Again, as we found when looking at the average across all events, frequencies for the various intensities depend on both the choice of the RCM and the driving data.

The relative frequency of events as a function of mean duration in cold and warm seasons is presented in Figures 7c and 7d respectively. In both seasons (Figures 7c,d), the RCM ensemble tends to produce too many short duration events and too few long lasting (i.e., longer than 24 hours) ECL events compared with the three reanalyses. NNRP-driven simulations tend to show a better distribution of event-mean durations than most GCM-driven simulations.

Finally, the distribution of mean sizes across the reanalyses and simulations are shown in Figures 7e and 7f. In both seasons the RCM ensemble shows a small range of responses with all RCM members producing too many small cyclones compared with the three high-resolution reanalyses. The last result appears to be consistent with the systematic underestimation of long-lasting ECL events suggesting that RCMs produce too few, large and slow-moving ECLs particularly in the cold season.

Figure 8 shows the spatial distribution of the number of events for the high-resolution reanalyses mean (left panels in Figure 8) and the RCM ensemble mean (right panels) in warm (top panels) and cold (bottom panels) seasons. In the warm season, the spatial distribution of cyclones shows a distinctive characteristic given by a maximum near the eastern coast of Australia consistent across the various reanalyses. As discussed in Di Luca et al. (2015), this feature appears to be consistent with a larger number of ECL events of the 'inland trough' type which develop within coastal surface troughs (see the annual cycle and the spatial distribution of inland trough lows in Speer et al. (2009)). The much larger number of events near the coast is also evident in the RCM ensemble mean field. Overall correlations between RCM members and the mean across the reanlyses (see Figure 9a) are always statistically significant ($P < 0.05$) and strongly dependent on the choice of the driving data.

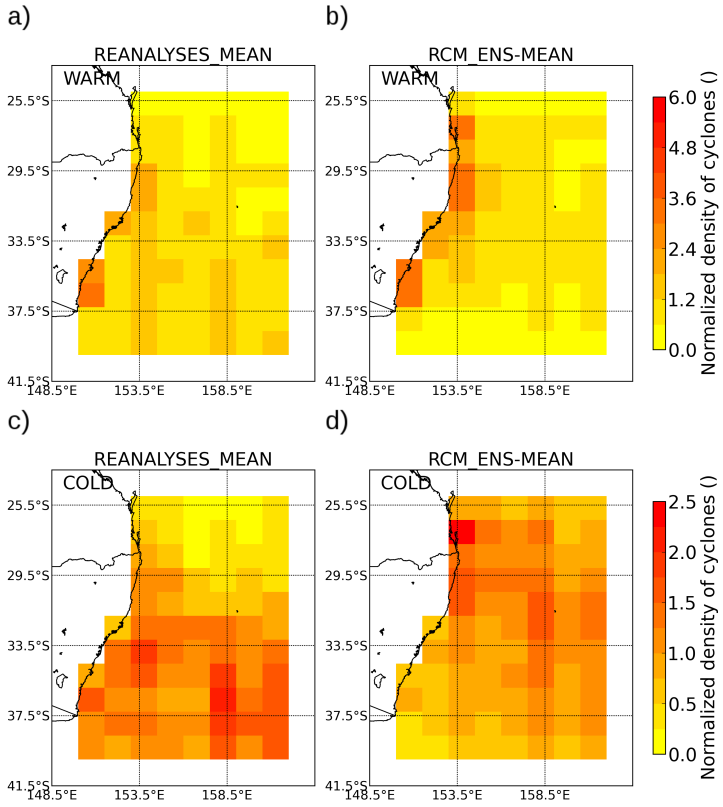

Figure 8 Same as in Figure 4 but for results obtained using SLP fields at their native resolution (see Table 1 for details on the horizontal resolution of individual datasets).

In the cold season (bottom panels in Figure 8), the mean-reanalyses spatial distribution show relatively low spatial variability although all reanalyses show a clear north-south gradient and two secondary maxima, one near the coast and another off

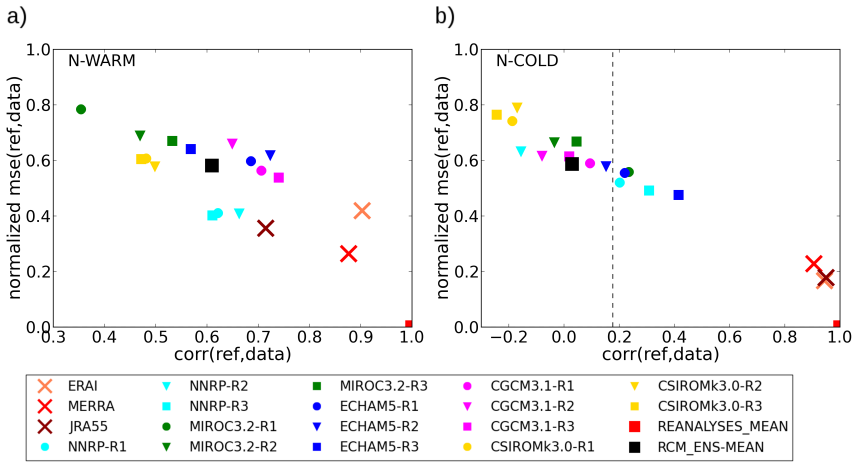

Figure 9 Same as in Figure 8 but for results obtained using SLP fields at their native resolution (see Table 1 for details on the horizontal resolution of individual datasets).

the coast at about 37° S. Most RCM members do not show such a latitudinal gradient in the number of events and produce too many events in the northern and central part of the region of interest, with too few events along the south coast where these are a major contributor to cool season heavy rainfall. As a consequence, spatial correlations between individual RCM simulations and the reanalyses mean are quite low (see Figure 9b).

6. Summary

In this study we have evaluated the representation of ECLs in a high-resolution 15-member RCM ensemble and in the low resolution global datasets used to drive the RCM ensemble. The evaluation was performed separately for ECLs derived using SLP fields at low (about 300 km) and high (about 50 km) resolutions. The analysis made using low-resolution SLP fields allows direct comparison of results obtained from RCM simulations with results obtained from the global data used to drive the RCMs (i.e. low resolution reanalysis or GCMs).

Regardless of the resolution of the SLP fields and the specific metric considered for evaluation (i.e. intensity, duration, size or number of events), we have shown that the RCM ensemble spread is affected by both the choice of the RCM version and the choice of the driving data although it is generally dominated by the former. This result makes clear that the simulation of ECLs is strongly dependent on the physical representation of various subgrid scale processes in climate models. Specifically, considering some large differences across different RCM versions, it seems that the cumulus parameterisation scheme plays a dominant role in agreement with previous work (Ji et al. 2014, Gilmore et al. 2015).

In general, the results found in this study highlight the importance of using a multi-RCM, multi-driving data approach to simulate the climatology of ECLs. Although individual RCM simulations sometimes show important discrepancies compared to the reference datasets, the ensemble mean consistently performs well and outperforms most individual simulations. Although this result is not particularly surprising given the results of previous studies, it is worth emphasising the usefulness of the RCM ensemble.

The evaluation of the ECL climatology obtained using low-resolution SLP fields shows that the RCM ensemble performs substantially better than the driving data ensemble mean, regardless of the evaluation metric considered. In particular, the coarse resolution GCMs tend to produce too few ECLs that are too weak and too short compared with the reanalyses ECLs. The intensity of ECL events and their duration are then systematically improved when considering the spatially upscaled RCMs instead of the lower resolution driving data. Moreover, the added value of the RCM ensemble is even more evident when considering subsets of more extreme ECLs (i.e. more intense or longer duration) for which the coarse models sometimes do not produce any events.

Evaluating the ECL climatology using high-resolution SLP fields shows that RCMs perform reasonably well at representing

some aspects of the climatology such as the number of events in cold season and, more generally, frequency distributions in terms of intensities, sizes and durations. Some systematic differences were apparent, however, such as an overestimation of the number of events in the warm season, an underestimation of the size of events in the cold season and substantial discrepancies in the spatial distribution of the cyclones.

Importantly, the evaluation shows that the RCM ensemble performs well even when considering more extreme subsets of ECLs. This constitutes a key result suggesting that the present RCM ensemble can be used to study a range of ECLs, specifically those most extreme ECLs that are generally associated with substantial impacts. For example, as shown by Callaghan and Power (2014), ECLs and their associated fronts are responsible for about 57% of all major floods along coastal catchments in eastern Australia. In this regard, future research plans include the consideration of metrics that quantify the cyclone's strength from a social and economical perspective including the maximum precipitation rate and the amount of rainfall accumulated along the cyclone's lifecycle.

It is worth noting that the climatology of ECLs appears substantially different depending on the resolution of the SLP fields used to identify and track ECLs. Apart from a much larger number of ECL events, the use of high-resolution SLP fields leads to a larger proportion of stronger (e.g. larger pressure gradient values) and longer lasting ECLs suggesting that the high-resolution fields are more suitable to analyse more extreme ECLs.

This assessment also highlights the quite complex relationship between RCMs and the global boundary conditions where simulations driven by the NCEP/NCAR reanalysis do not necessarily produce results that are closer to the reference datasets compared to those driven by GCMs. Clearly, further work is needed in order to understand the sources of these differences.

7. Acknowledgments

This work is supported by funding from the NSW Environmental Trust and the NSW Office of Environment and Heritage for the ESCCI-ECL project through the Australian Research Council grant LP120200777. Jason Evans was supported by the Australian Research Council Future Fellowship FT110100576. Research was also supported by the Australian Research Council Centre of Excellence for Climate System Science, grant CE110001028. The authors would like to acknowledge Stuart Browning for his provision of the identification and tracking scheme code and for advice on running it.

References

Browning, S.A. and Goodwin, I.D. 2013. Large scale influences on the evolution of winter subtropical maritime cyclones affecting australia's east coast. *Monthly Weather Review*, 141, 2416–2431. ISSN 0027-0644, 1520-0493. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-12-00312.1.

Callaghan, J. and Power, S. 2014. Major coastal flooding in southeastern australia 1860–2012, associated deaths and weather systems. *Australian Meteorological and Oceanographic Journal*, 64, 183–213.

Colle, B.A., Zhang, Z., Lombardo, K.A., Chang, E., Liu, P. and Zhang, M. 2013. Historical Evaluation and Future Prediction of Eastern North American and Western Atlantic Extratropical Cyclones in the CMIP5 Models during the Cool Season. *Journal of Climate*, 26(18), 6882–6903. ISSN 0894-8755, 1520-0442. URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00498.1.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N. and Vitart, F. 2011. The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. ISSN 00359009. URL http://doi.wiley.com/10.1002/qj.828.

Di Luca, A., Evans, J., Pepler, A., Alexander, L. and Argüeso, D. 2015. Resolution sensitivity of cyclone climatology over eastern australia using fivereanalysis products. *J Clim*, 28, 9530–9549.

Dowdy, A.J., Mills, G.A., Timbal, B. and Wang, Y. 2013. Changes in the risk of extratropical cyclones in eastern australia. *Journal of Climate*, 26(4), 1403–1417. ISSN 0894-8755, 1520-0442. URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00192.1.

Dowdy, A.J., Mills, G.A., Timbal, B. and Wang, Y. 2014. Fewer large waves projected for eastern australia due to decreasing storminess. *Nature Clim Change*, 4, 283–286. .

Evans, J.P., Ekström, M. and Ji, F. 2012. Evaluating the performance of a WRF physics ensemble over south-east australia. *Climate Dynamics*, 39(6), 1241–1258. ISSN 0930-7575, 1432-0894. . URL http://link.springer.com/10.1007/s00382-011-1244-5.

Evans, J.P., Ji, F., Lee, C., Smith, P., Argüeso, D. and Fita, L. 2014. Design of a regional climate modelling projection ensemble experiment – NARCliM. *Geosci. Model Dev.*, 7(2), 621–629. ISSN 1991-9603. URL http://www.geosci-model-dev.net/7/621/2014/.

Gilmore, J.B., Evans, J.P., Sherwood, S.C., Ekström, M. and Ji, F. 2015. Extreme precipitation in wrf during the newcastle east coast low of 2007. *Theor Appl Climatol*, 1–19.

Gordon, H.B., Rotstayn, L.D., McGregor, J.L., Dix, M.R., Kowalczyk, E.A., O'Farrell, S.P., Waterman, L.J., Hirst, A.C., Wilson, S.G., Collier, M.A., Watterson, I.G. and Elliott, T.I. 2002. The csiro mk3 climate system model. NARCliM Technical Note 60, Aspendale: CSIRO Atmospheric Research, Aspendale, Australia. URL http://www.dar.csiro.au/publications/gordon_2002a.pdf.

Greeves, C.Z., Pope, V.D., Stratton, R.A. and Martin, G.M. 2007. Representation of Northern Hemisphere winter storm tracks in climate models. *Climate Dynamics*, 28(7-8), 683–702. ISSN 0930-7575, 1432-0894. URL http://link.springer.com/10.1007/s00382-006-0205-x.

Grieger, J., Leckebusch, G., Donat, M., Schuster, M. and Ulbrich, U. 2014. Southern Hemisphere winter cyclone activity under recent and future climate conditions in multi-model AOGCM simulations. *International Journal of Climatology*, 34(12), 3400–3416. ISSN 08998418. URL http://doi.wiley.com/10.1002/joc.3917.

Grose, M.R., Pook, M.J., McIntosh, P.C., Risbey, J.S. and Bindoff, N.L. 2012. The simulation of cutoff lows in a regional climate model: reliability and future trends. *Climate Dynamics*, 39(1-2), 445–459. ISSN 0930-7575, 1432-0894. URL http://link.springer.com/10.1007/s00382-012-1368-2.

Hopkins, L.C. and Holland, G.J. 1997. Australian heavy-rain days and associated east coast cyclones: 1958-92. *Journal of Climate*, 10(4), 621–635. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1997)010%3C0621:AHRDAA%3E2.0.CO%3B2.

Ji, F., Ekström, M., Evans, J.P. and Teng, J. 2014. Evaluating rainfall patterns using physics scheme ensembles from a regional atmospheric model. *Theor Appl Climatol*, 115(1-2), 297–304.

Ji, F., Evans, J.P., Argueso, D., Fita, L. and Di Luca, A. 2015. Using large-scale diagnostic quantities to investigate change in East Coast Lows. *Clim Dyn*, 45, 2443–24531.

JMA. 2013. Outline of the operational numerical weather prediction at the japan meteorological agency. appendix to wmo technical progress report on the global data-processing and forecasting system (gdpfs) and numerical weather prediction (nwp) research. Tech. rep., Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research, JMA, Japan. URL http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/nwp-top.htm.

Jung, T., Gulev, S.K., Rudeva, I. and Soloviov, V. 2006. Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Quarterly Journal of the Royal Meteorological Society*, 132(619), 1839–1857. ISSN 00359009, 1477870X. . URL http://doi.wiley.com/10.1256/qj.05.212.

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., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3), 437–471.

Kim, S.J., Flato, G., Boer, G. and McFarlane, N. 2002. A coupled climate model simulation of the last glacial maximum, part 1: transient multi-decadal response. *Climate Dynamics*, 19(5-6), 515–537. ISSN 0930-7575, 1432-0894. URL http://link.springer.com/10.1007/s00382-002-0243-y.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K. 2015. The jra-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, 93.

Murray, R.J. and Simmonds, I. 1991. A numerical scheme for tracking cyclone centres from digital data. part i: Development and operation of the scheme. *Australian Meteorological Magazine*, 39, 155–166.

Pepler, A. and Rakich, C. 2010. Extreme inflow events and synoptic forcing in sydney catchments. *IOP Conf. Ser.: Earth Environ. Sci.*, 11, 012010.

Pepler, A.S., Di Luca, A., Ji, F., Alexander, L.V., Evans, J.P. and Sherwood, S.C. 2015. Impact of identification method on the inferred characteristics and variability of Australian East Coast Lows. *Monthly Weather Review*, 143, 864–877.

Rienecker, M., Suarez, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H.C., Gu, W., Sienkiewicz, M. Koster, R., Gelaro, R., Stajner, I. and Nielsen, J. 2008. The geos-5 data assimilation system—documentation of versions 5.0.1 and 5.1.0, and 5.2.0. Tech. rep., NASA Tech. Rep. Series on Global Modeling and Data Assimilation, NASA/TM-2008-104606.

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M. and Woollen, J. 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24(14), 3624–3648. ISSN 0894-8755, 1520-0442. URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00015.1.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schulzweida, U. and Tompkins, A. 2003. The atmospheric general circulation model echam5—part i: Model description. Tech. Rep. 349, Max-Planck-Institut für Meteorologie, Hamburg, Germany.

Rudeva, I. and Gulev, S.K. 2007. Climatology of cyclone size characteristics and their changes during the cyclone life cycle. *Monthly Weather Review*, 135(7), 2568–2587. ISSN 0027-0644, 1520-0493. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR3420.1.

Shkolnik, I.M. and Efimov, S.V. 2013. Cyclonic activity in high latitudes as simulated by a regional atmospheric climate model: added value and uncertainties. *Environmental Research Letters*, 8(4), 045007. ISSN 1748-9326. URL http://stacks.iop.org/1748-9326/8/i=4/a=045007?key=crossref.b4d20d8c59ae6a9db00bcaf79e187b65.

Skamarock, W.C. 2004. Evaluating mesoscale NWP models using kinetic energy spectra. *Mon Wea Rev*, 132(12), 3019–3032. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR2830.1.

Speer, M.S., Wiles, P. and Pepler, A. 2009. Low pressure systems off the new south wales coast and associated hazardous weather: establishment of a database. *Australian Meteorological and Oceanographic Journal*, 58(1), 29. URL http://web.maths.unsw.edu.au/~mss/index_files/speeretal_published.pdf.

Zappa, G., Shaffrey, L.C. and Hodges, K.I. 2013. The Ability of CMIP5 Models to Simulate North Atlantic Extratropical Cyclones*. *Journal of Climate*, 26(15), 5379–5396. ISSN 0894-8755, 1520-0442. URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00501.1.