Andrys, J., Lyons, T.J. and Kala, J. (2016) Evaluation of a WRF ensemble using GCM boundary conditions to quantify mean and extreme climate for the southwest of Western Australia (1970-1999). International Journal of Climatology, 36 (13). pp. 4406-4424.

Evaluation of a WRF Ensemble using GCM Boundary Conditions to Quantify Mean and Extreme Climate for the Southwest of Western Australia (1970-1999)

Julia Andrys*

State Centre of Excellence for Climate Change, Woodland and Forest Health, Murdoch University, Perth, Western Australia, Australia

Thomas J. Lyons

State Centre of Excellence for Climate Change, Woodland and Forest Health, Murdoch University, Perth, Western Australia, Australia

Jatin Kala

School of Veterinary and Life Sciences - Environmental and Conservation Sciences, Murdoch University, Perth, Western Australia, Australia and Australian Research Council Centre of Excellence for Climate Systems Science

**Corresponding author address:* State Centre of Excellence for Climate Change, Woodland and Forest Health, Murdoch University, Perth, Western Australia, Australia

E-mail: j.andrys@murdoch.edu.au

ABSTRACT

A high resolution (5 km), single initialisation, 30 year (1970-1999) Weather Research and Forecast regional climate model (RCM) ensemble for south-west Western Australia (SWWA) is evaluated. The paper focuses on the ability of the RCM to simulate winter cold fronts, which are the main source of rainfall for the region, and assesses the spatial and temporal characteristics of climate extremes within the region's cereal crop growing season. To explore uncertainty, a 4-member ensemble was run, using lateral boundary conditions from general circulation models (GCMs) of the Coupled Model Intercomparison Project Phase 3; ECHAM5, MIROC 3.2, CCSM3 and CSIRO mk3.5. Simulations are evaluated against gridded observations of temperature and precipitation and atmospheric conditions are compared to a simulation using ERA-Interim reanalysis boundary conditions, which is used as a surrogate truth. Results show that generally, the RCM simulations were able to represent the climatology of SWWA well however differences in the positioning of the subtropical high pressure belt were apparent which influenced the number of fronts traversing the region and hence winter precipitation biases. Systematic temperature biases were present in some ensemble members and the RCM was found to be colder than the driving GCM in all simulations. Biases impacted model skill in representing temperature extremes and this was particularly apparent in the MIROC forced simulation, which was the worst performing RCM for both temperature and precipitation. The dynamical causes of the biases are explored and findings show that nonetheless, the RCM provides added value, particularly in the spatio-temporal representation of wet season rainfall.

1. Introduction

General circulation models (GCMs) remain the primary source of information for projections of future climate change. While undeniably valuable, the coarse resolution of GCMs (100 to 250 km) limits their usefulness for assessing climate change at local and regional scales (1 to 10 km). Climate projections at this high resolution are important for assisting in the development of adaptation strategies; planning needed by industries such as agriculture and forestry to respond to the challenges faced by a changing climate.

In the context of climate information that is of value to agriculture, such as rain-fed cereal cropping, changes in extreme temperatures and precipitation patterns are of paramount importance. For example, screen temperatures of less than 2°C and greater than 34°C can have significant impacts on harvest yield and overall crop viability at the farm scale (Zheng et al., 2012), while shifting rainfall regimes may affect the future feasibility of marginal crop lands at a landscape scale (Ludwig et al., 2008). In southwest Western Australia (SWWA), cereal crops represent the majority (more than 60%) of the land use and contribute significantly to the regional economy (Varnas, 2014). The rain fed, winter growing croplands of SWWA have already experienced marked changes in climate, with an observed 30% decline in mean winter rainfall over the period 1970-2000 relative to the previous three decades (Bates et al., 2008). This reduction in rainfall has been attributed to a southward migration of storm tracks (Frederiksen and Frederiksen, 2007). Because this decline has predominantly impacted precipitation in July and August, when rainfall exceeds requirements, crop yields in SWWA have not deteriorated as a consequence (Turner and Asseng, 2005) although its negative impact on the region's native ecology is apparent (Brouwers et al., 2012). While agriculture has been able to adapt to the changes in the hydrological regime to date, marginal croplands in the east of SWWA face the prospect of becoming unviable if rainfall continues to decline. Therefore, future predictions of climate change are a critical component of adaptation strategies.

High resolution climate projections can be obtained through the use of regional climate models (RCMs), which account for regional influences on climate such as topography and land use, improving the modeling of mesoscale weather systems (Feser et al., 2011). Using GCMs or reanalysis as lateral boundary conditions, RCMs add value to these global models by improving the spatial representation of rainfall (Feldmann et al., 2008) and extreme events, such as heat waves (Gao et al., 2012). The ability of RCMs to add value to GCMs has been extensively evaluated. For example, Xue et al. (2007) showed that the choice of domain position and horizontal resolution had a significant impact on the utility of the RCM. This result was reinforced by Evans and McCabe (2013) who found that, in the southeast of Australia, increasing the resolution of the RCM improved model performance, particularly in coastal and mountainous regions. Song et al. (2008) undertook a RCM study for Australia at a 20 km resolution and were able to represent the seasonal distribution of rainfall, however at this scale, the influence of many topographical features was not represented. It is apparent that not all RCMs provide results of the same caliber and factors that have a substantial impact on the utility of a RCM include the dynamical core of the model itself, the choice of physical parameterisations, the capacity of the RCM to accurately represent the regional climatology when driven with reanalysis and finally, the performance of the RCM when GCMs are used as lateral boundary conditions (Xue et al., 2014).

Given the known sensitivity of RCMs to different physics and geographic regions, Kala et al. (2014) conducted an extensive sensitivity analysis using the Weather Research and Forecast Model Advanced Research core (WRF) to determine the most appropriate model physical parameterisations for SWWA. Following on from the work of Kala et al. (2014), the ability of WRF to simulate the historical climatology of SWWA using reanalysis data as lateral boundary conditions was evaluated by Andrys et al. (2015), who found that a 5 km horizontal resolution produced a skillful representation of the climate. This paper further extends on the work of Andrys et al. (2015) and Kala et al. (2014) by evaluating the capability of WRF to simulate the historical climate of SWWA using boundary conditions from four GCMs of the third Coupled Model Intercomparison Project (CMIP3). It is the final step in the validation of WRF for use in future climate projections in SWWA. In addition to examining the model's ability to represent the mean climatological conditions of the region, our analysis focuses on metrics that are of importance to cereal farming, including precipitation patterns and climatic extremes occurring in crop growth cycles.

2. Methods

a. The southwest of Western Australia (SWWA)

Typical of its mid-latitude location, the climate of SWWA is highly seasonal. The transition from cool wet winters to hot, dry summers is driven by the position of the subtropical high pressure belt, or subtropical ridge (SR), (Gentilli, 1971) which controls the passage of rain bearing cold fronts over the region in the winter. These frontal systems are the primary source of rain for much of SWWA and the region features a strong precipitation gradient, with rainfall declining from west to east. Summer rainfall is generally caused by surface convection however infrequent, large scale rain events do occur every 3 to 5 years when meriodonal troughs interact with tropical disturbances in the north of Western Australia (Wright, 1974). While SWWA is generally an area of low relief, topography still has a discernible influence on the region's climatology, particularly coastal precipitation. The Darling Scarp is an escarpment that produces a rapid change in elevation of approximately 300 m over 3 km and runs parallel to the coast, 25 km inland. The feature is apparent in the topographical map of the region shown in Figure 1(b) to the east of the city of

Perth. The escarpment results in a narrow band of elevated rainfall on the windward side which is challenging for mesoscale models to represent at moderate resolutions of approximately 10 km (Andrys et al., 2015), requiring instead a horizontal resolution closer to 0.5 km to comprehensively capture air flow across the escarpment and its associated turbulence (Pitts and Lyons, 1990). Most of the agricultural production in SWWA takes place inland of the Darling Scarp and the growing season for these croplands is in the cooler months of May to October.

b. Model Configuration

Employing the model configuration used by Andrys et al. (2015), a single initialisation, 30 year (with two month model spin up) regional climate simulation from 1970-1999 was conducted using WRF3.3 and lateral boundary conditions from four CMIP3 GCMs. The authors note that GCMs from CMIP5 (Taylor et al., 2012) represent the current state of the art for global climate models however the necessary 6-hourly fields required to run WRF were not available when simulations were commenced, hence the choice of CMIP3 GCMs. The GCMs; Max Planck Institute ECHAM5 model (Roeckner, 2003) (ECHAM), Center for Climate System Research Model for Interdisciplinary Research on Climate 3.2 (MIROC) (Hasumi and Emori, 2004), National Center for Atmospheric Research Community Climate System Model version 3 (CCSM) (Collins et al., 2006), Commonwealth Scientific and Industrial Research Organisation Mark 3.5 (CSIRO) (Gordon et al., 2002) were chosen based on the availability of data with 6-hourly fields. In choosing GCMs, consideration was given to the findings of Perkins et al. (2007) who evaluated the performance of CMIP3 GCMs for Australia and found that all of our chosen GCMs performed satisfactorily. Furthermore, in a subsequent study of the statistical independence of GCMs over Australia, Evans et al. (2014) found that both MIROC and ECHAM ranked highly in terms of model independence which warrants their use within a RCM ensemble. 6-hourly input data from the GCMs, which includes winds, geopotential height, temperature, humidity and pressure are ingested by the RCM at the lateral boundary of the outer domain only.

Our model utilises a three domain configuration (Fig.1(a)) with a 50:10:5 km horizontal resolution and 30 vertical levels. The choice of model physics was based on the findings of a prior sensitivity analysis of WRF to different physics and input data over SWWA (Kala et al., 2014). Parameterisation options include; the Single-Moment 5 class microphysics scheme (Hong et al., 2004), RRTM for long-wave radiation (Mlawer et al., 1997), Dudhia short-wave radiation (Dudhia, 1989), Yonsei University planetary boundary layer scheme, convective parameterisation on the first and second domains only from Kain Fritsch (Kain, 2004), the MM5 surface layer scheme (Grell et al., 2000) and Noah land surface model (Chen and Dudhia, 2001).

c. Observational Data

Observational data used for evaluation is from a daily gridded data set of maximum and minimum temperatures and rainfall provided by the Australian Bureau of Meteorology (Jones et al., 2009). The data, at a resolution of 5 km, is an interpolation from a network of weather stations across Australia and has been used as a validation tool for previous regional climate simulations in SWWA (Andrys et al., 2015; Kala et al., 2014) and other regions in Australia (Evans et al., 2011). King et al. (2013) established that, while this data set underestimates the contribution of extreme rainfall events, it is capable of reproducing trends and variability in extreme precipitation events for much of Australia, including SWWA.

The data was interpolated using simple inverse distance weighting to both domain two (10 km resolution) and three (5 km resolution) of the simulation. Andrys et al. (2015) found that the higher resolution, convection resolving 5 km domain was able to represent the overall climatology of SWWA generally better than the 10 km domain and as such this study will focus on the results of the 5 km domain. Data from the outer domain is used for examining large scale features such as mean sea level pressure (SLP) however our focus is on SWWA and so temperature and precipitation are not analysed for the the outer domain. To explore the source of temperature biases, monthly mean 2 m temperature data from each GCM used in the simulation was interpolated to the outer WRF grid and also compared with the observational data set.

The model configuration used in this study is identical to that of Andrys et al. (2015) who used ERA-Interim reanalysis (Dee et al., 2011) boundary conditions with WRF over the period 1981-2010. Because reanalysis data are constrained by observations we use the outputs from Andrys et al. (2015) as a “best-guess” of actual conditions to examine the validity of certain model diagnostics which are useful in depicting the synoptic meteorology, including mean SLP and 10 m wind vectors.

d. Evaluation Criteria

Daily rainfall and temperature distributions are assessed using probability density functions (PDFs). Simulated rainfall values less than 0.2 mm are excluded from the analysis as this falls below the detection level of the observations (Evans and McCabe, 2010). To examine the model's spatial performance at representing daily rainfall and temperatures we use a summary statistic known as relative entropy (R_E), which compares the observed and simulated distributions, and measures the difference between them. R_E has been used to compare GCM simulations with observations by Shukla et al. (2006) and Tippett et al. (2004) and also by Naveau et al. (2014) to detect changes in climate extremes. R_E is expressed by Cover and Thomas (2012) as;

$$R_E(p||q) = \sum_{x \in X} p(x) log \frac{p(x)}{q(x)} \tag{1}$$

where $p(x)$ and $q(x)$ are the observed and simulated PDFs respectively. As its name suggests, R_E is a relative measure rather than an absolute measure for examining model divergence. In cases were the model is showing perfect agreement with the observations, the R_E will be 0. A model with very poor agreement will have R_E approaching 1 however there is no absolute maximum value. Figure 2 illustrates examples of distributions with shifts in variance (a) and mean (c) that show good agreement having a corresponding R_E score of 0.01, while distributions with larger shifts in variance (b) and mean (d) show poor agreement, having a R_E score of 0.5. We acknowledge that there are a number of other metrics available for comparing observed and simulated PDFs; including the Kolmogorov-Smirnov (KS) test and the Perkins Skill Score (Perkins et al., 2007). We choose R_E because this method sums the log of the ratio between p(x) and q(x) at each bin, while both the Perkins Skill Score and KS test sum the difference in probabilities for each bin. By calculating the log of the ratio and not the difference, R_E ensures an equal weighting for changes in the tails of the distribution relative to changes at the distribution centre, where absolute changes are almost always the greatest.

We examine the average number of days that fronts traverse SWWA during winter using an automated front recognition technique, the thermal gradient recognition (TGR) algorithm. This method is based on thermal gradients at the 850 hPa level to detect the baroclinic zone within a cold front. Initially described by Mills (2005) and further validated by Hope et al. (2014) for applications in SWWA, parameters used for TGR in our study include a thermal gradient of greater than 2.5^{o}C 100 km^{-1} at 850 hPa that is accompanied by daily domain averaged rainfall greater than 0.5 mm. We note that Hope et al. (2014) employed a smaller temperature gradient of 1.3^{o}C 100 km^{-1} compared to our higher threshold. The former analyzed reanalysis data at a 250 km resolution which would have smoothed out the baroclinic zone within cold fronts, whereas our simulations at 10 and 5 km had more well defined baroclinicity, and hence, a larger threshold was warranted.

Given that the focus of this paper is on the ability of the RCM to provide climate information that is valuable to the agricultural sector of SWWA, our analysis of extreme indices is limited to the cereal crop growing season (May-October). Metrics are based on the core indices developed by the World Meteorological Organisation working group, the Expert Team on Climate Change Detection and Indices (ETCCDI) (Persson et al., 2007) which we modified to provide a better reflection of extreme conditions in SWWA. We redefine the summer days (SD) index to a count of days when the maximum temperature exceeds 34^{o}C based on findings by Asseng et al. (2011) who determined that temperatures in excess of this threshold can impact grain yield. The frost days (FD) index is modified to a count of days when minimum temperatures are lower than 2^{o}C following the work of Kala et al. (2009) who found that screen temperatures below 2^{o}C can result in foliage temperatures less than 0^{o}C.

Our choice of rainfall indices focus on rainfall intensity and distribution which are relevant for agriculture. We use the simple precipitation intensity index (SDII):

$$SDII_j = \frac{\sum_{w=1}^{w} RR_{wj}}{W} \tag{2}$$

where RRwj is the daily precipitation amount on wet days W when RR is greater than 1 mm in period j and W is the number of wet days in j. The total number of rain days (PRCPTOT) is a count of days where daily rainfall exceeds 1 mm. We also use the ETCCDI metrics, maximum length of dry spell (CDD) and maximum length of wet spell (CWD). These indices measure the longest span of days where rainfall is less than 1 mm for CDD and the longest span of days where rainfall is greater than 1 mm for CWD.

3. Results

The meriodonal movement of the SR underpins seasonality in SWWA, hence, it is important to evaluate the model's ability to capture this seasonal transition. Figure 3 shows the outer domain mean seasonal sea level pressure (SLP) for the ERA-Interim driven simulation of Andrys et al. (2015) (W-ERA), which is assumed to be a "best-guess" at reality, and the 4 GCM driven simulations. We compute seasonal means over the period 1981-1999, being the period when outputs are available for both W-ERA and the GCM driven simulations. The position of the SR to the south of Australia in summer and its northerly position over the continent in winter are apparent in W-ERA. While all of the simulations are able to represent this transition, there are distinct differences. The MIROC forced simulation (W-MIR) has the lowest SLP over SWWA in winter, suggesting a more northerly winter position of the SR, and hence a more northerly storm track than indicated by W-ERA. Conversely, the CSIRO driven model (W-CSI) is displaying a southerly position for the SR in winter which would lead to a more southerly storm track. Simulations using ECHAM (W-ECH) and CCSM (W-CCS) as boundary conditions are able to represent the winter SLP well, with W-CCS providing the closest match to W-ERA. The position of the SR during summer is generally well represented by all the ensemble members however both W-MIR and W-ECH have lower SLP relative to W-ERA, particularly in the region of Australia's mid south coast which suggests that the intensification of high pressure systems in this region, a major synoptic feature during the summer, is not fully captured by these simulations.

a. Seasonal Precipitation

Mean seasonal precipitation and simulation biases are shown in Figure 4. Observations highlight that most of the region's rainfall is in winter with a distinct west to east precipitation gradient. All of the simulations represent the seasonal transition of rainfall in SWWA.

Generally, model agreement with observations is satisfactory, with biases not exceeding +/- 20 mm month$_{-1}$*howeverbiasesofthismagnitudearemorenoteworthyinthesummerbecausemeanmonthlyrainfalldur*

CSIisabletorepresentthetemporaldistributionofthisrainfallwiththegreatestskill, *simulating*10*wetsummer*

CCSunderestimatesthetimingofregionalscalesummerrainfall, *withonly*2*wetsummerswhileW* − *MIRandW* − *ECHoverestimateevents*, *simulating*14*and*15*wetsummersrespectively.*

To investigate the impact of mean SLP on the number of frontal systems traversing SWWA, we examine the number of days that cold fronts are present over the region during winter using TGR and compare this with W-ERA for the period 1981-1999. The mean and standard deviation of winter front days are shown in Figure 5. W-MIR has the highest mean number of winter front days (30) followed by W-ECH (26). Both are higher than W-ERA (22). W-CCS, with an average of 24, represents front days well while W-CSI (16) is underestimating the number of winter fronts.

b. Daily Precipitation

Figure 6 shows the daily precipitation PDF of rainfall across all land points in the region. Observations show that rainfall less than 1 mm occurs 38% of the time. Daily rainfall exceeding 10 mm is uncommon, with a likelihood of 10% and rainfall greater than 25 mm day^{-1} has a probability of only 1%. W-MIR underestimates the likelihood of rain less than 1 mm day^{-1} by approximately 5% and overestimates the probability of days with rainfall greater than 4 mm. R_E for W-MIR (0.015) indicates that this simulation has the lowest agreement with observations. W-ECH follows a similar pattern to W-MIR, however the magnitude of the disparity for W-ECH is not as great, and this is reflected in an improved R_E of 0.010. Both W-CSI (R_E 0.007) and W-CCS (R_E 0.001) overestimate the chance of light rainfall and subsequently underestimate the likelihood of more intense rain, however this is small for the W-CCS simulation, hence the good R_E value.

The spatial distribution of R_E for daily rainfall is shown in Figure 7 which highlights that for all simulations, R_E is generally below 0.1. Inland areas show the best values for R_E and model deviation tends to increase in the north west corner of the domain for the W-CSI, W-MIR and W-ECH simulations. All simulations show poor agreement in the vicinity of the south west coastline, consistent with (Andrys et al., 2015).

c. Seasonal Temperatures

Observed seasonal mean maximum temperatures and simulation bias is shown in Figure 8. With the exception of W-CSI, simulations underestimate maximum temperatures. This is particularly apparent in the W-MIR simulation, where negative summer biases can exceed 5°C. W-ECH also displays a systematic cold bias up to 5°C. Both W-CCS and W-CSI show good agreement with observations, with biases generally not exceeding +/- 2°C. Seasonal mean observed minimum temperatures and model bias is shown in Figure 9. Overall, minimum temperatures show smaller biases than maximum temperatures, however the W-MIR cold bias persists, particularly in the summer. W-CCS performs well, with biases generally less than +/- 2°C. Likewise, W-ECH represents minimum temperatures with little bias. Contrary to its robust performance with respect to maximum temperatures, W-CSI displays a warm bias for summer minimum temperatures up to 5°C.

We consider the mean annual temperature bias of the GCM and the RCM between 1970 and 1999 in Figure 10. MIROC and ECHAM show very little bias in SWWA however CCSM and CSIRO are both displaying a warm bias, up to 5°C in the case of CSIRO. The RCM is able to eliminate much of this bias from W-CCS and W-CSI however WRF introduces a cold bias to W-MIR and W-ECH. In all cases the RCM is colder than its corresponding GCM.

W-MIR also displays a negative night time temperature bias which is strongest in summer. We investigate this by examining differences in air flow which are illustrated in Figure 11 showing the mean seasonal 10 m wind vectors between 1981-1999 for simulations and W-ERA. Summer winds in W-MIR are more meriodonal than W-ERA which displays a more zonal flow. Additionally, W-MIR winds in autumn and spring display a tendency towards onshore flow which is not found in W-ERA.

d. Daily Temperatures

Daily maximum temperature distributions for all land based grid points, including R_E scores, are shown in Figure 12 and summary statistics for these distributions are shown in Table 2. Observations show that the distribution has a short left tail suggesting that very cold maxima are rare, while the right tail is elongated, indicating that hot extremes are more likely than cold extremes. Summer days ($> 34^{o}$C) have an occurrence probability of 10%.

Ensemble members are able to simulate the shape of this distribution however, as expected by the high biases found in the seasonal analysis, there is a skew towards colder temperatures in W-MIR and W-ECH. W-MIR also displays decreased variability in maximum temperatures shown by a standard deviation 0.9^{o}C lower than observations (Table 2). Consequently, W-MIR has a poor R_E of 0.229. W-ECH represents the distribution of higher temperatures with more accuracy than W-MIR and has a R_E of 0.083. W-CSI (R_E 0.017) and W-CCS (R_E 0.035) show very good agreement with observations. W-CCS overestimates the likelihood of colder maxima and underestimates moderate maxima however its representation of temperatures above 34^{o}C is close to observations. Conversely, W-CSI overestimates the likelihood of higher temperatures but represents the distribution of colder maxima well.

The observed daily minimum temperature PDF, shown in Figure 13, follows a normal distribution, which indicates that warm extremes and cold extremes have an equal likelihood of occurrence. W-CCS simulates daily minima with the greatest accuracy, having a R_E of 0.018. W-ECH (R_E 0.026) also performs well however it is overestimating the variability of warmer minimum temperatures. Conversely, W-MIR (R_E 0.064) underestimates the overall variability of temperatures, particularly warm minima, and overestimates the likelihood of median temperatures. W-CSI displays significant warm bias for summer minimum temperatures and the impact of this bias is apparent in the PDF for W-CSI (R_E 0.087), which is skewed to the right.

Spatial R_E is considered for temperatures in Figure 14. The strong performance of W-CCS and W-CSI for maximum temperatures is apparent when compared with the much poorer R_E of W-MIR and W-ECH. Minimum temperature R_E corresponds with findings from the PDFs; W-CCS and W-ECH show generally strong performance throughout the domain while W-CSI and W-MIR do not perform as well.

e. Extreme Indices in the Growing Season

Indices related to extremes of temperature (FD and SU) and precipitation (PRCPTOT, CDD, CWD and SDII) as they occur during the SWWA growing season are shown in Figure 15. Observations show that the most intense rain in SWWA, indicated by SDII, is the orographically induced rainfall near the Darling Scarp (Fig.1(b)). The high resolution of the simulation means that all models can represent an increased SDII due to the Darling Scarp however none of the simulations are able to fully account for the magnitude of the SDII in this area. Rainfall intensity in the southern coastal region is significantly underestimated by all the simulations.

PRCPTOT observations show more than 100 days of rain each growing season in the south and as few as 30 in the north east. The spatial distribution of PRCPTOT is well represented and the magnitude is also generally well modeled however all simulations underestimate the number of rain days in the north east. W-MIR represents the high PRCPTOT values on the southern coast which are missed by the other simulations however it is overestimating in the domain interior. A tendency to overestimate CDD in the north east and underestimate CWD in the south west is common to all simulations.

Because the growing season occurs over the cooler months, the hot temperatures represented by SU are uncommon. The northern region experiences 2 SU each growing season and events do no generally occur to the south. On account of the strong cold bias displayed by W-MIR (Fig.8), this simulation does not represent SU at all. W-CCS and W-ECH both simulate SU well while W-CSI overestimates SU. Observations of FD show that frost does not commonly impact the coast. Inland areas are more susceptible to frost, experiencing between 8 to 30 FD in a growing season. Simulations represent the very low risk of frost along the coast however all simulations overestimate the occurrence of FD inland. This overestimation is the highest in W-MIR while W-CSI shows results closest to observations.

4. Discussion

Simulations are able to represent the topographically enhanced rainfall near the Darling Scarp and the strong west to east precipitation gradient which, due to the fine spatial scale of these features, are not well represented at the resolution of the driving GCMs. Some errors are systematic across all simulations, most notably the strong negative winter precipitation bias in the south west. We attribute this to the WRF model because a similar bias was also present in the 30-year ERA-Interim driven simulation of Andrys et al. (2015) which was caused by the south west boundary of domain 3 being too close to the coast.

Seasonal rainfall biases are smaller than those found in a regional climate study over Australia by Song et al. (2008) and comparable to the biases found by Evans and McCabe (2013) in south-east Australia. The wet summer bias in W-MIR and W-ECH and dry bias in W-CCS are caused by the poor representation of regional scale summer rainfall events. Because these summer rainfall events are associated with tropical disturbances in the north of Western Australia, we explore this region to attribute the bias. Summer SLP in Figure 3, displays apparent differences in the tropical regions of the outer domain, particularly off the northwest coast of Australia. W-CCS has higher pressure in this region compared to W-ERA whereas W-MIR and W-ECH display lower pressures. The ability of GCMs to represent tropical meteorology has been evaluated by Brown et al. (2013) who found that elements of the tropical climatology are poorly simulated by CMIP3 GCMs in the western tropical Pacific whereas Moise et al. (2012) identified uncertainties in the representation of the Australian tropical climate. Hence, the limitations of the GCMs in representing tropical meteorology is a likely source of error for these summer rainfall biases.

Winter bias varies markedly between simulations and we attribute this to the position of the SR shown in Figure 3. Wet biases in W-MIR and W-ECH are caused by a northerly track of winter storms, resulting in more of these systems traversing SWWA. Conversely, the dry bias in W-CSI is attributed to the southerly winter position of the SR which forces a southerly storm track, reducing the number of fronts traversing the region. This attribution is in line with the findings of Argüeso et al. (2012) who, in a RCM study for Spain using WRF, established that model differences in wet season SLP contributed to precipitation biases as storm tracks were deviated from their observed position.

The high number of front days in W-MIR and W-ECH (Fig. 5) provides further evidence of a northerly storm track in both of these simulations. W-CCS had low winter rainfall bias because the simulation represented both the position of the SR and the number of winter front days well while

W-CSI underestimated the number of front days due to the simulation's southerly storm track. In an analysis of the position of the Austral jet stream, and hence storm tracks, Kidston and Gerber (2010) found a high degree of variability between CMIP3 GCMs. This spread is the likely cause of the differences in our simulations with respect to winter front days as these large scale features would be strongly influenced by the lateral boundary conditions used to drive WRF.

The performance of CMIP3 GCMs to simulate daily rainfall in regions of Australia, including SWWA, was evaluated by Perkins et al. (2007). They found that the GCMs, including the four used as boundary conditions in this study, overestimated the likelihood of low rainfall as much as two to three times. Our results show that this overestimation has been reduced by the RCM. Perkins et al. (2007) also found that ECHAM represented SWWA precipitation with greater skill than CCSM, CSIRO and MIROC. Based on our findings, W-CSI and W-CCS perform better than W-ECH. Furthermore, Perkins et al. (2007) found that CCSM was among the lower performing models for rainfall in SWWA however, our analysis shows that W-CCS displays the greatest skill. This suggests that a direct relationship cannot be assumed between the ranked performance of a GCM and the performance of the same GCM used as boundary conditions to drive an RCM. Determining why this is the case is outside the scope of this paper, however we can speculate that the higher resolution, or different dynamics and parameterisations, in the RCM are allowing the development of important local drivers in the W-CSI and W-CCS simulations. Alternatively, the higher resolution is realising some previously undetected issue with the lateral boundary conditions in the W-MIR simulation. Such an issue was found by Evans and McCabe (2013) in a RCM study over south-east Australia, who established that the GCM (in their case CSIRO) was transporting excessive moisture into the higher latitudes from the tropics however this excess moisture did not result in high precipitation biases until the resolution of the RCM was fine enough to fully resolve the topography of the region.

In terms of relative model performance, W-MIR has the highest overall bias and worst R_E while W-CCS provides the best representation of rainfall with very little bias and consistently good R_E. Because W-CCS is clearly the better simulation with respect to precipitation, we expect that W-CCS would also reproduce precipitation indices with the greatest skill however this is not always the case (Fig. 15). W-CCS and W-ECH generally provide the best representation of precipitation indices however they tend towards a dry bias; overestimating CDD and underestimating SDII, PRCPTOT and CWD. We find that W-CSI consistently underestimates rainfall indices due to the dry rainfall bias caused by the southerly storm track seen in this simulation. Conversely, W-MIR overestimates rainfall indices for PRCPTOT, SDII and CDD which is expected based on the northerly storm track found in this simulation. However, W-MIR provides the best simulation of rainfall around the Darling Scarp for SDII, PRCPTOT and CWD. None of the simulations can account for the full impact of the orography of the Darling Scarp which is in line with the findings of Pitts and Lyons (1990) who found that a resolution of 0.5 km was needed to fully represent the turbulent air flow initiated by the Scarp.

Argüeso et al. (2012) used WRF to downscale ECHAM and CCSM GCM data in a regional climate study over Spain and included extreme precipitation metrics in their evaluation criteria. They found that their simulation using ECHAM boundary conditions represented the CWD and CDD with greater skill than the simulation driven by CCSM however we find very little difference between W-ECH and W-CCS. Indeed, we find very little difference between all of the simulations with respect to these precipitation indices compared with seasonal differences in rainfall bias. However, while Argüeso et al. (2012) examined a number of different rainfall regions and different seasons in their analysis of indices, we consider only the growing season where rainfall is almost exclusively from southwesterly frontal systems. The relative homogeneity of our results indicate that all of our simulations show skill in representing the spatio temporal characteristics of rainfall in SWWA during the growing season.

High resolution simulations of precipitation are important for agriculture in SWWA because of the region's large east-west precipitation gradient during autumn, winter and spring, which all of the simulations are able to represent. Furthermore, to be of use to agriculture, accurate simulation of indices such as the SDII and CDD are vital because these cannot be derived from monthly rainfall values alone. The spatial variability observed in these indices has been well represented by all simulations in the inland region, especially by W-CCS.

Some simulations display strong biases for seasonal temperatures. With the exception of a warm bias in W-CSI, which we attribute to bias in the CSIRO GCM (Fig. 10), simulations tend to be cold, particularly for daytime temperatures. Andrys et al. (2015) previously found that WRF produced a cold bias for daytime temperatures in SWWA while WRF was also shown to introduce a cold bias for south east Asia (Chotamonsak et al., 2011) and Norway (Heikkilä et al., 2011). These studies are in line with our finding that the RCM is always colder than its corresponding GCM which accounts for the daytime cold bias in W-ECH and W-MIR, however this does not explain why the magnitude of the cooling between the GCM and the RCM is different for each simulation. For example, while areas of the W-CCS simulation are up to 4^{o}C colder than CCSM, the difference is only 1-2^{o}C between W-ECH and ECHAM. However, as we have highlighted in this paper with respect to precipitation and as has been demonstrated by other regional climate studies (Evans and McCabe, 2013), dynamical downscaling of a GCM does not produce a linear response in the corresponding RCM.

MIROC has been shown by Perkins et al. (2007) to represent Australian temperatures well however its performance globally tends to be low when compared with other CMIP3 GCMs (Randall et al., 2007). Connolley and Bracegirdle (2007) performed an assessment of CMIP3 GCMs over the Antarctic region and found that MIROC was one of the lowest performing simulations in this region. Furthermore, Irving et al. (2011) conducted a similar analysis of CMIP3 GCMs in the Pacific Islands region and also determined that MIROC performed poorly with respect to temperature. The extent of our simulation's outer domain means that boundary conditions are drawn from regions where MIROC has shown poor performance and as such it is likely that the GCM is also contributing to the negative temperature bias found in W-MIR. This suggests that, when choosing boundary conditions for RCM, the performance of the GCM in the vicinity of the outer domain lateral boundary is more important than the performance of the GCM over the specific area of study.

W-MIR also displays a cold nighttime bias which is strongest in summer. This bias can be attributed to anomalies in air flow in the W-MIR simulation (Fig. 11). With the exception of the coastal sea breeze circulation, mean summer 10 m wind direction in the SWWA is predominantly easterly. This flow is caused by high pressure systems in the Great Australian Bight which result in hot, dry winds from the continental interior dominating the SWWA wind field. The persistence of high pressure in this region is evident from the summer pattern of mean SLP in the W-ERA simulation in Figure 3. The summer meriodonal flow apparent in W-MIR suggests that the simulation is not advecting hot air from the continental interior which is contributing to the cold bias.

In their analysis of CMIP3 GCM daily temperatures for Australia, Perkins et al. (2007) found that GCMs produced temperature distributions that were too broad. We find that, based on the observed and simulated daily temperature standard deviations (Table 2), this distribution spread appears to be somewhat reduced by the RCM. When spatial performance is considered, coastal regions consistently display the poorest R_E scores in each simulation. This indicates that the difficulty representing the distribution of coastal maximum temperatures is a function of the WRF model configuration rather than the lateral boundary conditions themselves. This is consistent with the findings of Andrys et al. (2015) who also found issues in representing the daily distribution of temperatures in the coastal region. With the exception of W-MIR, all the simulations show an inferior R_E over the Perth metropolitan area. This also follows the findings of Andrys et al. (2015) who suggested that a reduction in simulation skill for night time temperatures over the metropolitan area was a result of the lack of representation of urban land use.

Because W-CCS represents temperatures over 34^{o}C well, it is also expected to simulate SU well and this is demonstrated in Figure 15. Maximum temperature skewness in W-MIR and W-ECH and the overestimation of high maxima in W-CSI mean that these simulations are unable to represent SU with the same skill as W-CCS. However, while W-CSI demonstrated the lowest performance for nighttime temperatures, minimum temperature distributions (Fig. 13) show that W-CSI has the best overall agreement for temperatures below 2^{o}C. This suggests that W-CSI, despite its warm minimum temperature bias, will represent FD with the greatest skill and this is shown in Figure 15. All simulations overestimate FD somewhat and share a common spatial pattern, which indicates that the distribution of FD in the region is heavily influenced by the RCM.

It is apparent that W-MIR is not representing seasonal or daily temperatures with the same level of skill as the other simulations, indicated by the high negative temperature biases and the poor values of R_E for daily temperatures shown for both minimum and maximum temperatures. W-CCS is the only simulation which shows low bias for both minimum and maximum temperatures and consistently strong R_E scores for daily temperature distributions. While W-CSI is able to represent daytime temperatures well and also very cold minimum temperatures, the simulation shows a high bias and overall low skill for nighttime temperatures. Conversely, W-ECH represented nighttime temperatures well but performed poorly for daytime temperatures.

Minimum and maximum temperatures are of interest to agriculture, however the growing season distribution of temperature extremes, including FD and SU, are more relevant. While simulation bias has introduced errors in representing the extent and magnitude of temperature indices in SWWA, W-CCS provides a good representation of SU, and both W-CCS and W-CSI are able to represent the spatial distribution of FD to a reasonable degree.

These findings demonstrate the merits of our RCM ensemble however there are limitations with our experimental design which warrant consideration. While Kala et al. (2014) established the most appropriate model physics options for WRF in SWWA, our study is limited in that our simulations used a single RCM only. WRF has a known sensitivity to parameterisation schemes, for example precipitation is sensitive to the choice of convective scheme while temperatures are sensitive to the PBL scheme (Argueso et al., 2011). Other regional climate simulations have employed an ensemble of RCMs to reduce the uncertainty from using a single model; either through the use of different dynamical cores (Solman et al., 2013) or by imposing different physical parameters within the same modeling framework (Evans et al., 2014). The use of additional RCMs was not computationally feasible for this project and as such our results are constrained by the uncertainty inherent in using a single RCM.

5. Conclusion

We present an evaluation of the RCM, WRF 3.3, for SWWA between 1970-1999 using four CMIP 3 GCMs; CCSM3, CSIRO mk3.5, ECHAM5 and MIROC3.2 (med-res) as lateral boundary conditions. Our analysis focused on the ability of the downscaled GCMs to represent the climate of the cereal crop growing season in SWWA, which runs from May to October. The growing season is of particular interest because dryland cereal crops are a major contributor to the economy of the region and are at a high risk of being impacted by changing hydrological regimes in the future.

Simulation performance was varied. Seasonal rainfall bias was generally low however there are elements of bias related to systematic errors from the WRF model itself and from errors in the lateral boundary conditions. For example, the dry winter rainfall bias in the south west corner can be attributed to model error because the domain boundary was located too close to the SWWA coastline (Andrys et al., 2015). Conversely, the wet inland winter biases shown by W-MIR and W-ECH are caused by a northerly storm track allowing too many cold fronts to traverse the region, which we attribute to the lateral boundary conditions. Dry summer rainfall biases in W-CCS and wet biases in W-MIR and W-ECH can also be attributed to the lateral boundary conditions because of GCM limitations in modeling tropical meteorology (Brown et al., 2013; Moise et al., 2012).

WRF demonstrated a tendency to simulate colder temperatures than those found in the GCMs and maximum temperature biases were considerable in some simulations. For example, W-MIR showed summer daytime cold biases exceeding 5^{o}C in some areas and this bias impacted the representation of extreme indices. While a portion of this bias was due to the cooling tendency seen in WRF, we also attribute this bias to the lateral boundary conditions that have been shown to demonstrate poor performance in the vicinity of our simulation outer domain.

We find that GCMs which rank highly when evaluated using PDFs of rainfall and temperature will not necessarily perform the best when used to provide lateral boundary conditions to a RCM. For example, Perkins et al. (2007) found CCSM to be among the worst performing GCMs for SWWA however, after downscaling, we find that W-CCS provided the best representation of rainfall distribution, exceeding the performance of W-CSI, W-ECH and W-MIR, whose corresponding GCMs all provided a closer approximation of daily rainfall than CCSM. This indicates that the suitability of a GCM for dynamical downscaling cannot necessarily be determined by how well it represents temperature and precipitation in a region. This finding is supported by Evans and McCabe (2013) who found that surface variables may not be sufficient to fully assess the capability of a GCM for regional climate modeling. Furthermore, the poor performance of the W-MIR simulation, which contrasts with the strong performance of MIROC over Australia, suggests that the performance of the GCM in the vicinity of the RCM lateral boundary may be a better indicator for how the GCM will perform when it has been downscaled.

In a recent review of regional climate modeling, and the conditions under which they add value to GCM data, Xue et al. (2014) highlighted the importance of the appropriate choice of GCM data and a robust model set up. We have identified some issues with both the lateral boundary conditions and the model itself in this study. However one simulation, W-CCS, represents the climate of SWWA remarkably well and two further simulations (W-CSI and W-ECH) provide a satisfactory representation. We note issues with W-CSI representing minimum temperatures and W-ECH with maximum temperatures and suggest caution when using their results for those variables. The W-MIR simulation consistently performed with the lowest skill; cold temperature biases resulted in large errors when extreme temperature indices were examined and errors in mean SLP resulted in wet summer and winter precipitation biases. Based on these findings, we do not recommend that the W-MIR simulation be used for future climate analysis for SWWA. Notwithstanding, when compared with the findings of Perkins et al. (2007), the RCM has significantly improved upon the daily distribution of precipitation and allowed for the development of more intense rainfall events. The strong performance of the RCM is particularly apparent in representing the spatiotemporal distribution of wet season rainfall, which is significant for future applications of this data in agricultural adaptation planning. Based on these findings, we have validated the capability of the individual ensemble members W-ECH, W-CCS and W-CSI to represent the historical climate of SWWA and have confidence in the use of the RCM for analysis of future climate scenarios.

Acknowledgments. This research was supported by an Australian Grains Research and Development Corporation (GRDC) Grant (MCV0013). Julia Andrys is supported by an Australian Postgraduate Award and a GRDC Top Up Scholarship. Jatin Kala is supported by the Australian Research Council Centre of Excellence for Climate Systems Science (CE110001028). The research group lead by Associate Professor Jason Evans at the University of New South Wales, Australia, provided the modifed version of WRFv3.3 used in this study, and assisted in the pre-processing of the input data. Computational modeling was supported by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. The project was funded under the National Computational Merit Allocation Scheme and the Pawsey Partner Allocation Scheme. All of this support is gratefully acknowledged.

References

Andrys, J., T. J. Lyons, and J. Kala, 2015: Multi-decadal Evaluation of WRF Downscaling Capabilities Over Western Australia in Simulating Rainfall and Temperature Extremes. *Journal of Applied Meteorology and Climatology*, **54**, 370–394, doi:http://dx.doi.org/10.1175/JAMC-D-14-0212.1.

Argüeso, D., J. M. Hidalgo-Muñoz, S. R. Gámiz-Fortis, M. J. Esteban-Parra, and Y. Castro-Díez, 2012: Evaluation of WRF Mean and Extreme Precipitation over Spain: Present Climate (197099). *Journal of Climate*, **25**, 4883–4897, doi:10.1175/JCLI-D-11-00276.1.

Argueso, D., J. M. Hidalgo-Mutildenoz, S. R. Gacuteamiz-Fortis, M. J. Esteban-Parra, J. Dudhia, and Y. Castro-Diez, 2011: Evaluation of WRF parameterizations for climate sutdies over Southern Spain using a multi-step regionalization. *J. Clim.*, **24**, 5633–5651.

Asseng, S., I. Foster, and N. C. Turner, 2011: The impact of temperature variability on wheat yields. *Global Change Biology*, **17 (2)**, 997–1012, doi:10.1111/j.1365-2486.2010.02262.x.

Bates, B. C., P. Hope, B. Ryan, I. Smith, and S. Charles, 2008: Key findings from the Indian Ocean Climate Initiative and their impact on policy development in Australia. *Climatic Change*, **89**, 339–354, doi:10.1007/s10584-007-9390-9.

Brouwers, N. C., J. Mercer, T. Lyons, P. Poot, E. Veneklaas, and G. Hardy, 2012: Climate and landscape drivers of tree decline in a Mediterranean ecoregion. *Ecology and Evolution*, **3**, 67–79.

Brown, J. N., and Coauthors, 2013: Implications of CMIP3 model biases and uncertainties for climate projections in the western tropical Pacific. *Climatic Change*, **119**, 147–161, doi:10.1007/s10584-012-0603-5.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, **129**, 569–585.

Chotamonsak, C., E. P. Salathé Jr, J. Kreasuwan, S. Chantara, and K. Siriwitayakorn, 2011: Projected climate change over Southeast Asia simulated using a WRF regional climate model. *Atmospheric Science Letters*, **12**, 213–219.

Collins, W. D., and Coauthors, 2006: The formulation and atmospheric simulation of the Community Atmosphere Model version 3 (CAM3). *Journal of Climate*, **19**, 2144–2161.

Connolley, W. M., and T. J. Bracegirdle, 2007: An Antarctic assessment of IPCC AR4 coupled models. *Geophysical Research Letters*, **34**, 1–6, doi:10.1029/2007GL031648.

Cover, T. M., and J. A. Thomas, 2012: *Elements of information theory*. 2nd ed., John Wiley & Sons, Hoboken.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, doi:10.1002/qj.828.

Dudhia, J., 1989: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences*, **46**, 3077–3107.

Evans, J. P., M. Ekström, and F. Ji, 2011: Evaluating the performance of a WRF physics ensemble over South-East Australia. *Climate Dynamics*, **39**, 1241–1258, doi:10.1007/s00382-011-1244-5.

Evans, J. P., F. Ji, C. Lee, P. Smith, D. Argüeso, and L. Fita, 2014: Design of a regional climate modelling projection ensemble experiment - NARCliM. *Geoscientific Model Development*, **7**, 621–629, doi:10.5194/gmd-7-621-2014.

Evans, J. P., and M. F. McCabe, 2010: Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment. *Journal of Geophysical Research*, **115**, 1–15, doi:10.1029/2010JD013816.

Evans, J. P., and M. F. McCabe, 2013: Effect of Model Resolution on a Regional Climate Model Simulation Over Southeast Australia. *Climate Research*, **56**, 131–145.

Feldmann, H., B. Frueh, G. Schaedler, H.-J. Panitz, K. Keuler, D. Jacob, and P. Lorenz, 2008: Evaluation of the precipitation for South-western Germany from high resolution simulations with regional climate models. *Meteorologische Zeitschrift*, **17**, 455–465.

Feser, F., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn, 2011: Regional Climate Models Add Value to Global Model Data: A Review and Selected Examples. *Bulletin of the American Meteorological Society*, **92**, 1181–1192, doi:10.1175/2011BAMS3061.1.

Frederiksen, J. S., and C. S. Frederiksen, 2007: Interdecadal changes in southern hemisphere winter storm track modes. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **59**, 599–617.

Gao, Y., J. S. Fu, J. B. Drake, Y. Liu, and J. F. Lamarque, 2012: Projected changes of extreme weather events in the eastern United States based on a high resolution climate modeling system. *Environmental Research Letters*, **7**, 44 025.

Gentilli, J., 1971: *Climates of Australia and New Zealand*. Elsevier Pub. Co., 108–114 pp.

Gordon, H. B., and Coauthors, 2002: *The CSIRO Mk3 climate system model*, Vol. 130. CSIRO Atmospheric Research.

Grell, G. A., S. Emeis, W. R. Stockwell, T. Schoenemeyer, R. Forkel, J. Michalakes, R. Knoche, and W. Seidl, 2000: Application of a multiscale, coupled MM5/chemistry model to the complex terrain of the VOTALP valley campaign. *Atmospheric Environment*, **34**, 1435–1453.

Hasumi, H., and S. Emori, 2004: K-1 coupled model (MIROC) description. K-1 Technical Report 1. *Center for Climate System Research, University of Tokyo, Tokyo*.

Heikkilä, U., A. Sandvik, and A. Sorteberg, 2011: Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. *Climate dynamics*, **37**, 1551–1564.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. *Monthly Weather Review*, **132**, 103–120, doi:10.1175/1520-0493(2004)132⟨0103:ARATIM⟩2.0.CO;2.

Hope, P., and Coauthors, 2014: A Comparison of Automated Methods of Front Recognition for Climate Studies: A Case Study in Southwest Western Australia. *Monthly Weather Review*, **142 (1)**, 343–363, doi:10.1175/MWR-D-12-00252.1.

Irving, D. B., and Coauthors, 2011: Evaluating global climate models for the Pacific island region. *Climate Research*, **49**, 169–187, doi:10.3354/cr01028.

Jones, D. A., W. Wang, and R. Fawcett, 2009: High-quality spatial climate data-sets for Australia. *Aust. Meteorol. Oceanographic Journal*, **58**, 233–248.

Kain, J. S., 2004: The KainFritsch convective parameterization: an update. *Journal of Applied Meteorology*, **43**, 170–181.

Kala, J., J. Andrys, T. J. Lyons, I. J. Foster, and B. Evans, 2014: Sensitivity of WRF to driving data and physics options on a seasonal time-scale for the southwest of Western Australia. *Climate Dynamics*, doi:10.1007/s00382-014-2160-2.

Kala, J., T. J. Lyons, I. J. Foster, and U. S. Nair, 2009: Validation of a Simple Steady-State Forecast of Minimum Nocturnal Temperatures. *Journal of Applied Meteorology and Climatology*, **48**, 624–633, doi:10.1175/2008JAMC1956.1.

Kidston, J., and E. P. Gerber, 2010: Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology. *Geophysical Research Letters*, **37**, 1–5, doi:10.1029/2010GL042873.

King, A. D., L. V. Alexander, and M. G. Donat, 2013: The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia. *International Journal of Climatology*, **33**, 2376–2387.

Ludwig, F., S. P. Milroy, and S. Asseng, 2008: Impacts of recent climate change on wheat production systems in Western Australia. *Climatic Change*, **92**, 495–517, doi:10.1007/s10584-008-9479-9.

Mills, G. A., 2005: A re-examination of the synoptic and mesoscale meteorology of Ash Wednesday. *Aust. Met. Mag*, **54**, 35–55.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102 (D)**, 16 663–16 682, doi:10.1029/97JD00237.

Moise, A. F., R. A. Colman, and J. R. Brown, 2012: Behind uncertainties in projections of Australian tropical climate: Analysis of 19 CMIP3 models. *Journal of Geophysical Research: Atmospheres*, **117**, 1–16, doi:10.1029/2011JD017365.

Naveau, P., A. Guillou, and T. Rietsch, 2014: A non-parametric entropy-based approach to detect changes in climate extremes. *Journal of the Royal Statistical Society*, **76**, 861–884, doi:10.1111/rssb.12058.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate*, **20**, 4356–4376, doi:10.1175/JCLI4253.1.

Persson, G., L. Bärring, and E. Kjellström, 2007: *Climate indices for vulnerability assessments*. 111, SMHI, 1–80 pp.

Pitts, R. O., and T. J. Lyons, 1990: Airflow over a two-dimensional escarpment. II: Hydrostatic flow. *Quarterly Journal of the Royal Meteorological Society*, **116**, 363–378, doi:10.1002/qj.49711649207.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR).*, Cambridge University Press, Cambridge, chap. 8, 589–662.

Roeckner, E., 2003: The atmospheric general circulation model ECHAM 5. Part I: Model description, Rep. 349, Max Planck Inst. for Meteorol., Hamburg, Germany.

Shukla, J., T. DelSole, M. Fennessy, J. Kinter, and D. Paolino, 2006: Climate model fidelity and projections of climate change. *Geophysical Research Letters*, **33**, L07 702, doi:10.1029/2005GL025579.

Solman, S. A., and Coauthors, 2013: Evaluation of an ensemble of regional climate model simulations over South America driven by the ERA-Interim reanalysis : model performance and uncertainties. *Clim. Dyn.*, **41**, 1139–1157, doi:10.1007/s00382-013-1667-2.

Song, R., X. Gao, H. Zhang, and A. Moise, 2008: 20 km resolution regional climate model experiments over Australia : experimental design and simulations of current climate. *Aust. Met. Mag*, **57**, 175–193.

Taylor, K. E., R. J. Stouffer, and G. a. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, **93 (4)**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tippett, M. K., R. Kleeman, and Y. Tang, 2004: Measuring the potential utility of seasonal climate predictions. *Geophysical Research Letters*, **31**, 1–4, doi:10.1029/2004GL021575.

Turner, N. C., and S. Asseng, 2005: Productivity, sustainability, and rainfall-use efficiency in Australian rainfed Mediterranean agricultural systems. *Crop and Pasture Science*, **56**, 1123–1136.

Varnas, D., 2014: Western Australian Grains industry. URL https://www.agric.wa.gov.au/grains-research-development/western-australian-grains-industry.

Wright, P. B., 1974: Seasonal Rainfall in Southwestern Australia and the General Circulation. *Monthly Weather Review*, **102**, 219–232.

Xue, Y., Z. Janjic, J. Dudhia, R. Vasic, and F. De Sales, 2014: A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. *Atmospheric Research*, **147-148**, 68–85, doi:10.1016/j.atmosres.2014.05.001.

Xue, Y., R. Vasic, Z. Janjic, F. Mesinger, and K. E. Mitchell, 2007: Assessment of dynamic downscaling of the Continental U.S. Regional Climate using the Eta/SSiB regional climate model. *Journal of Climate*, **20**, 4172–4193, doi:10.1175/JCLI4239.1.

Zheng, B., K. Chenu, M. Fernanda Dreccer, and S. C. Chapman, 2012: Breeding for the future: what are the potential impacts of future frost and heat events on sowing and flowering time requirements for Australian bread wheat (Triticum aestivium) varieties? *Global Change Biology*, **18**, 2899–2914, doi:10.1111/j.1365-2486.2012.02724.x.

LIST OF TABLES

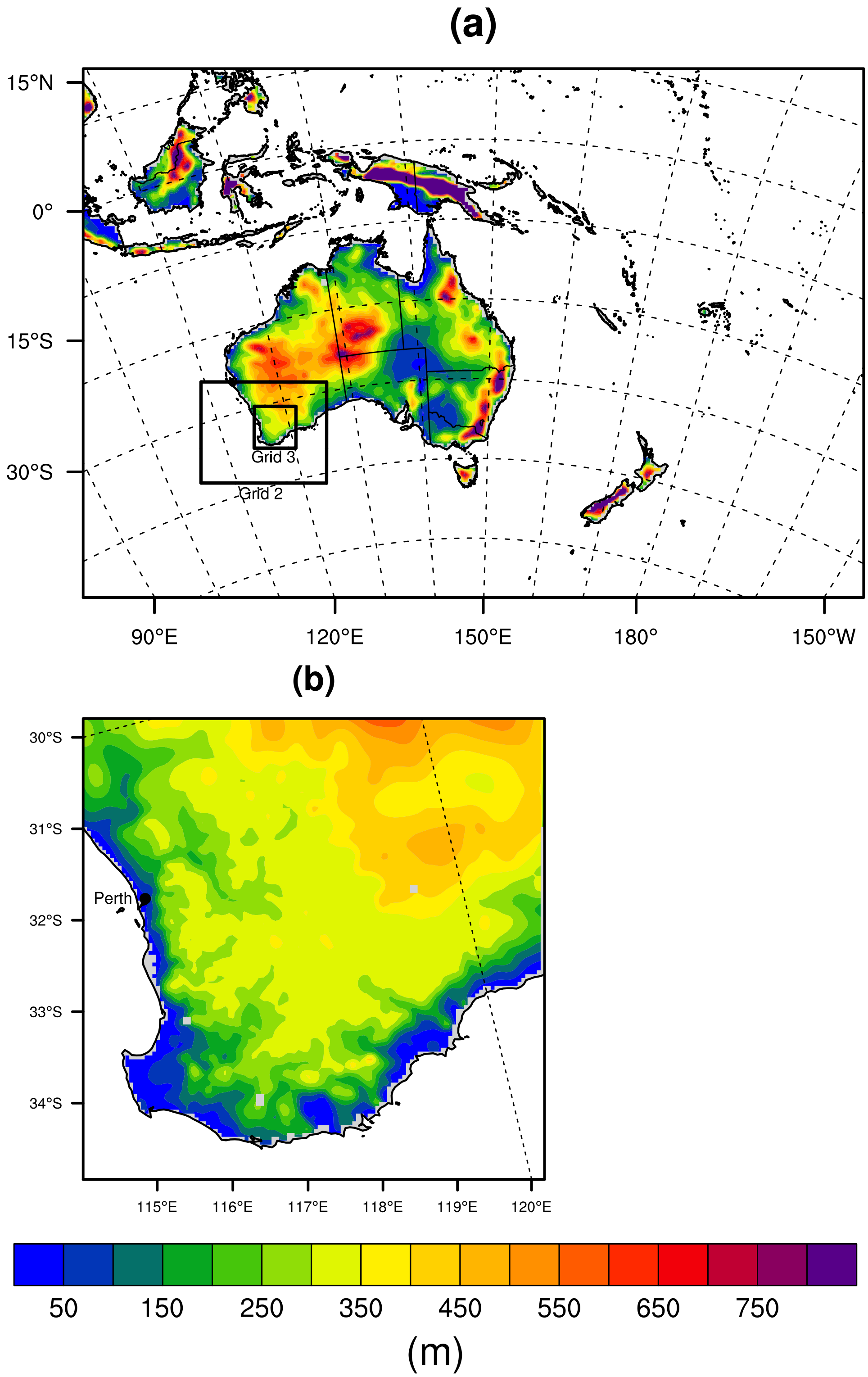
TABLE 1. Number of wet and dry summers from observations (OBS) and all simulations from 1970-1999. A wet summer has at least one month where domain averaged rainfall exceeds 20 mm

	OBS	W-MIR	W-CCS	W-ECH	W-CSI
Wet Summer	6	14	2	15	10
Dry Summer	24	16	28	15	20

TABLE 2. Domain averaged mean and standard deviation of observed and simulated daily minimum and maximum temperatures

	OBS	W-MIR	W-CCS	W-ECH	W-CSI
	Mean (Std)	Mean (Std)	Mean (Std)	Mean (Std)	Mean (Std)
Maximum Temperature (^{o}C)	23.2 (6.9)	19.4 (6.2)	22.3 (7.3)	20.7 (7.5)	23.2 (7.6)
Minimum Temperature (^{o}C)	10.4 (4.8)	9.1 (4.5)	10.5 (5.1)	10.4 (5.5)	12.1 (5.9)

LIST OF FIGURES



770 FIG. 1. Topographical map from Andrys et al. (2015) of (a) the model outer domain showing the extent of nested domains 2 (10 km resolution) and 3 (5 km resolution) used for simulations and (b) the location of Perth and the topography of the Darling Scarp within the 5 km domain.

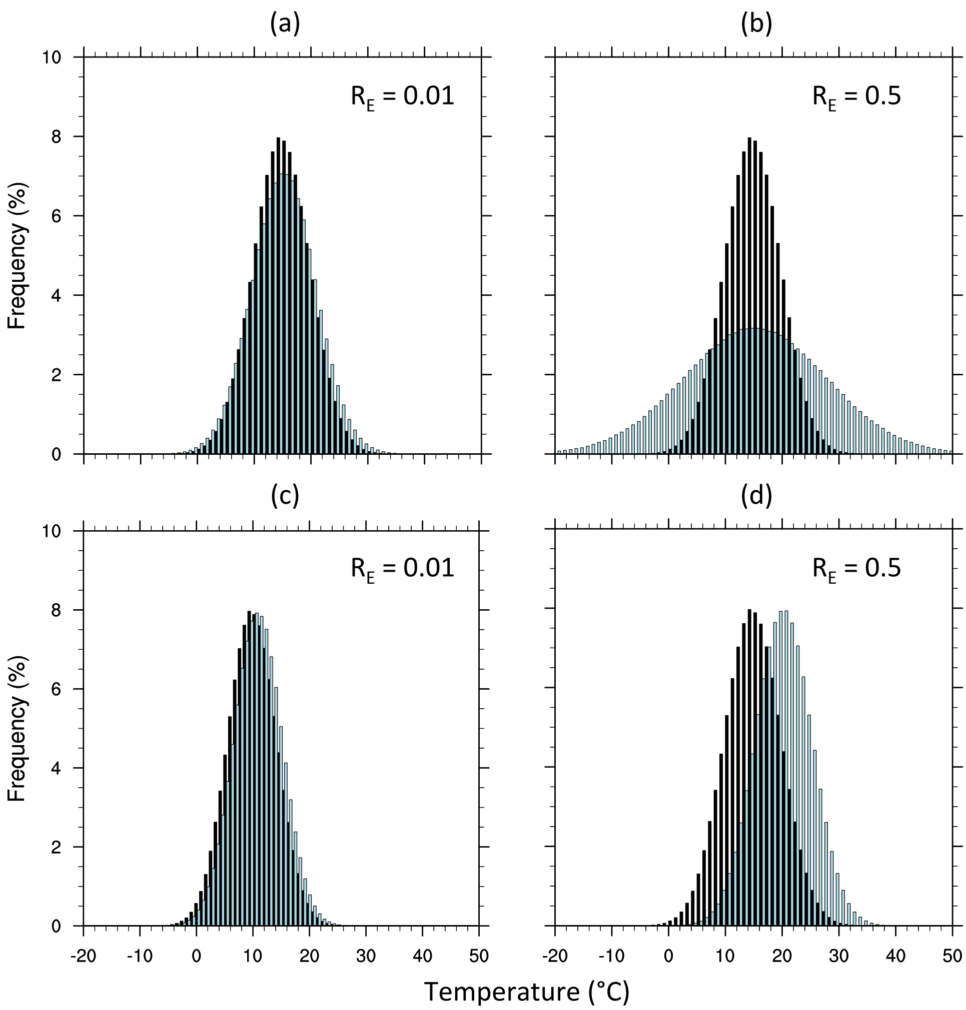


FIG. 2. Example PDF plots showing (a) distributions with equal means and a 10% variance shift having a R_E score of 0.01 representing good agreement, (b) distributions with the equal means and a 150% variance shift having a R_E score of 0.5 representing poor agreement, (c) distributions with a 5% mean shift and equal variance having a R_E score of 0.01 and (d) distributions with 33% mean shift and equal variance having a R_E score of 0.5.

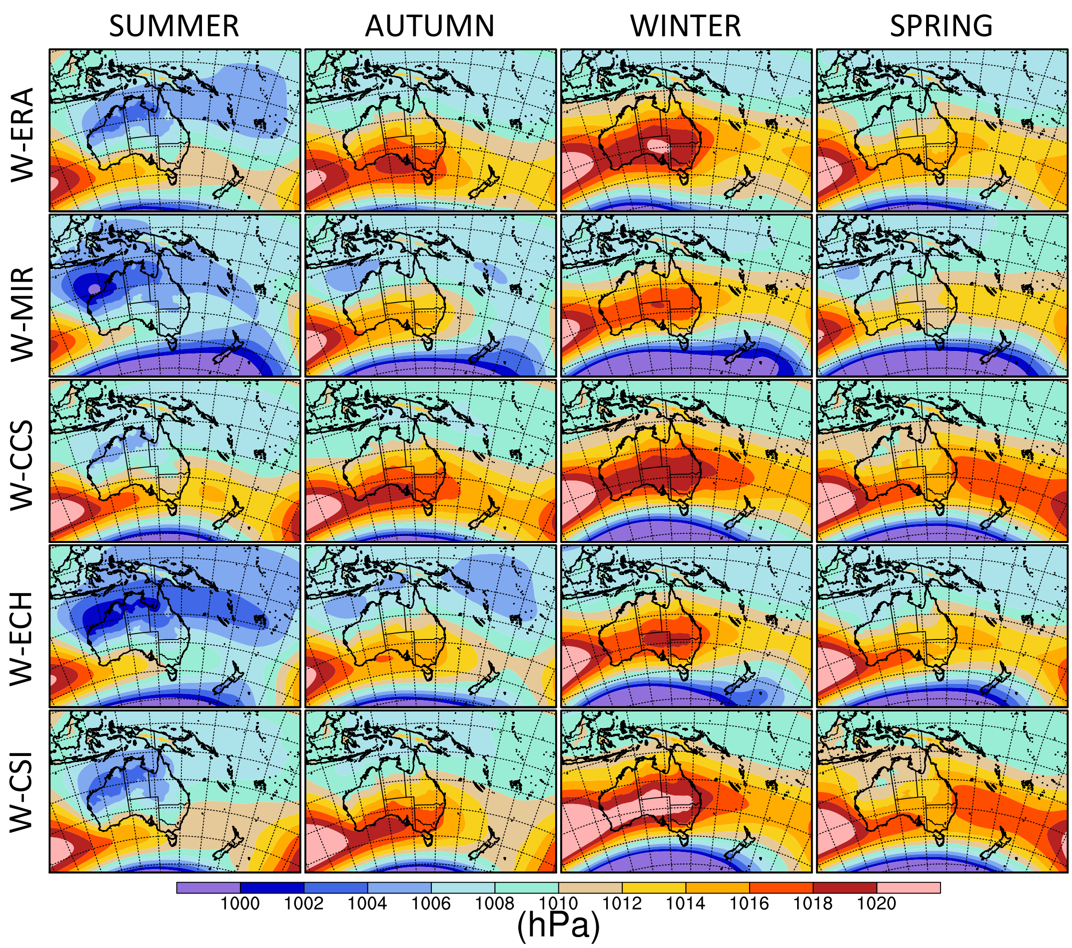
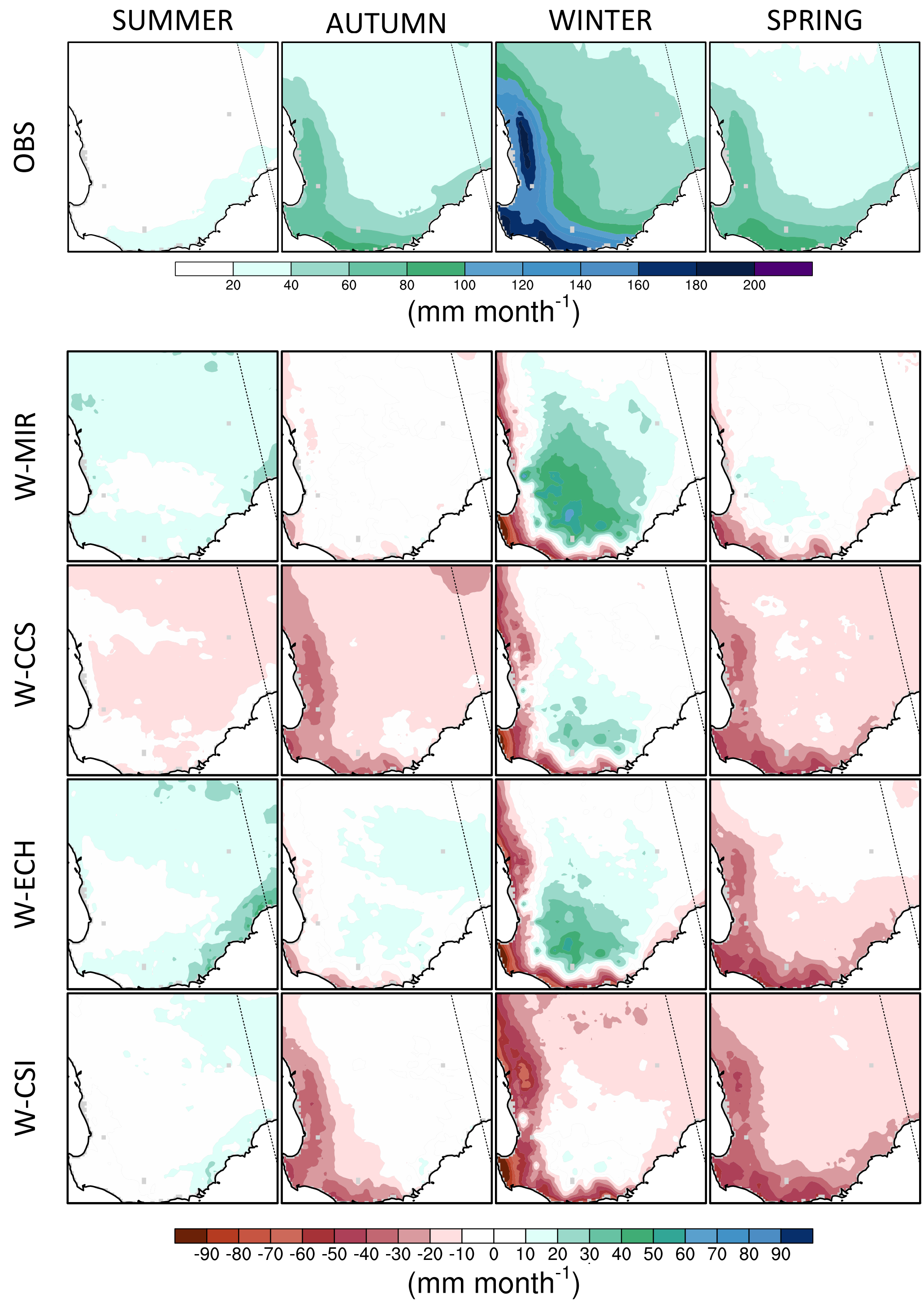
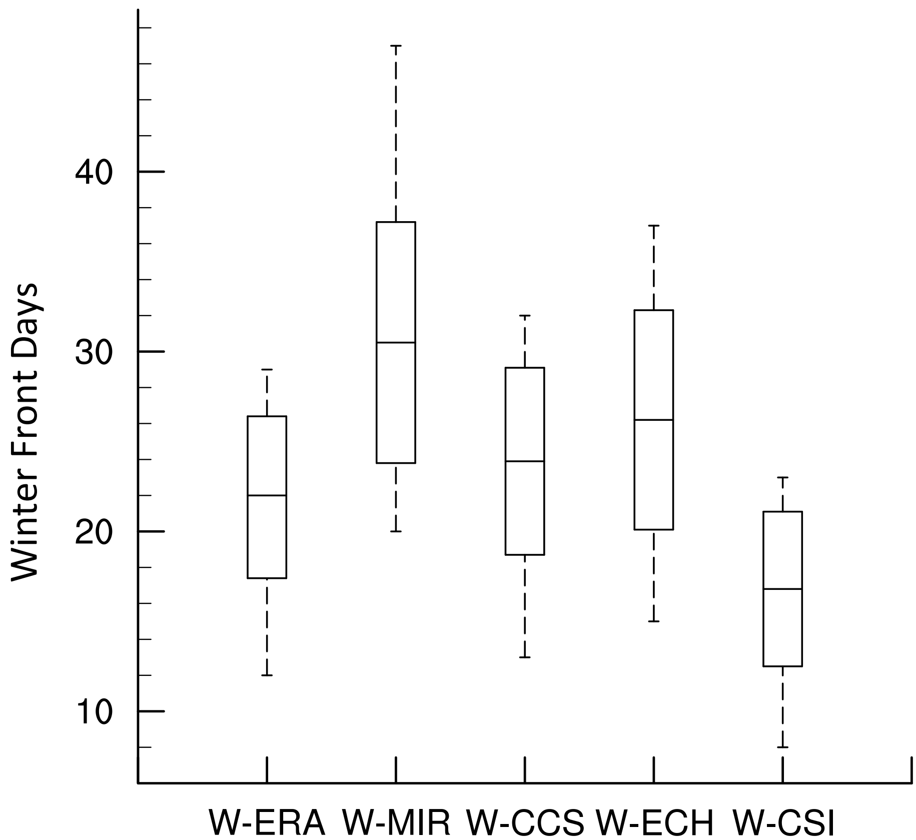


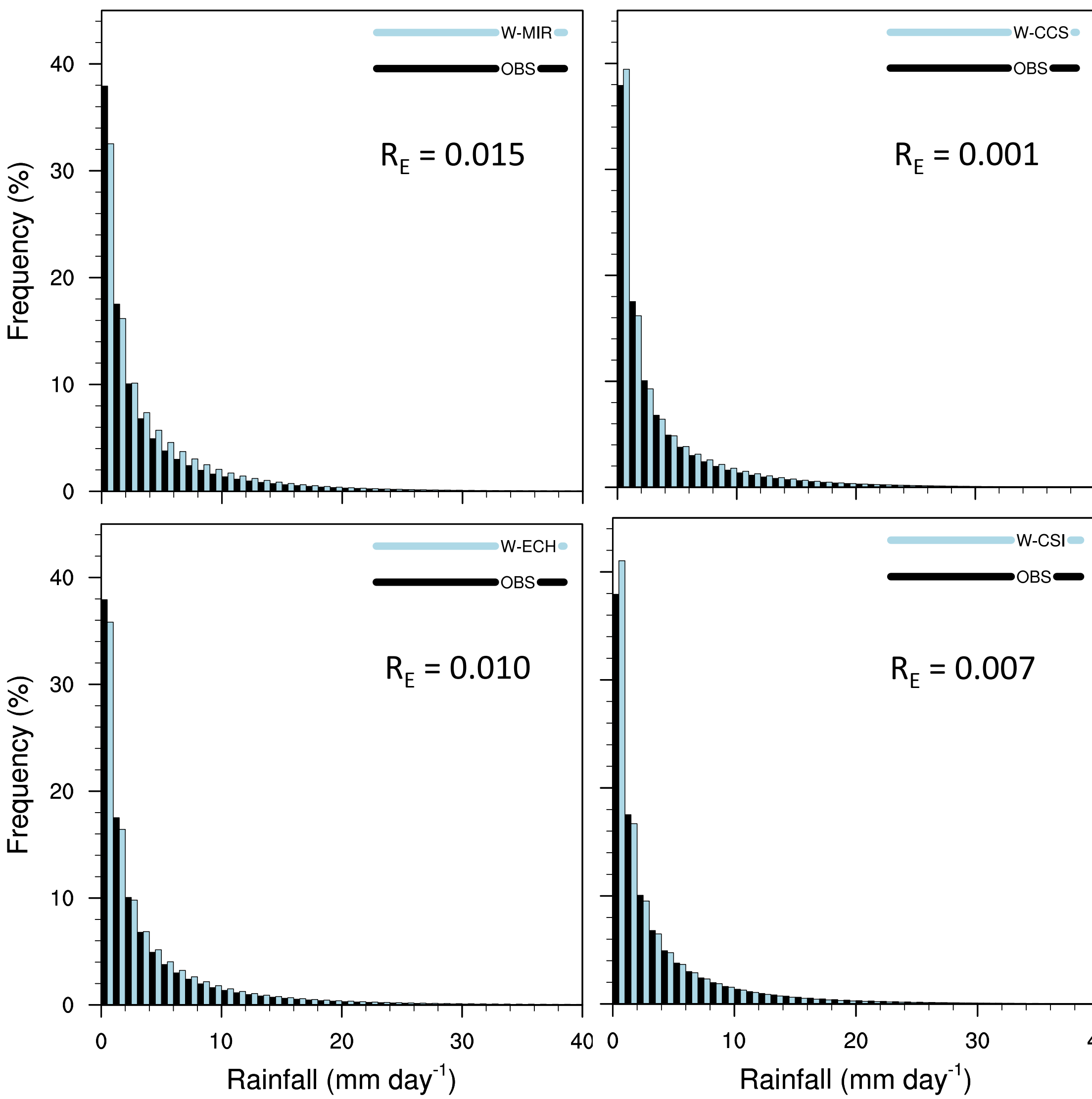
FIG. 3. Seasonal mean sea level pressure (1980-1999) for the WRF outer domain for simulations using ERA-Interim (W-ERA), MIROC3.2 (W-MIR), CCSM3 (W-CCS), ECHAM5 (W-ECH) and CSIRO Mk 3 (W-CSI) lateral boundary conditions.



780 FIG. 4. Observed (OBS) seasonal mean rainfall (top panel) and bias (bottom panels) for all simulations over
781 the period 1970-1999.



782 FIG. 5. Boxplot showing the range of winter front days by simulation. Centre line displays mean values, the box bounds one standard deviation from the mean and tails represent the range of values.



784 FIG. 6. Daily rainfall probability density functions for simulations and observations taken from all land based grid points from the 5 km domain. The R_E value comparing the similarity of the distributions is included for each plot.

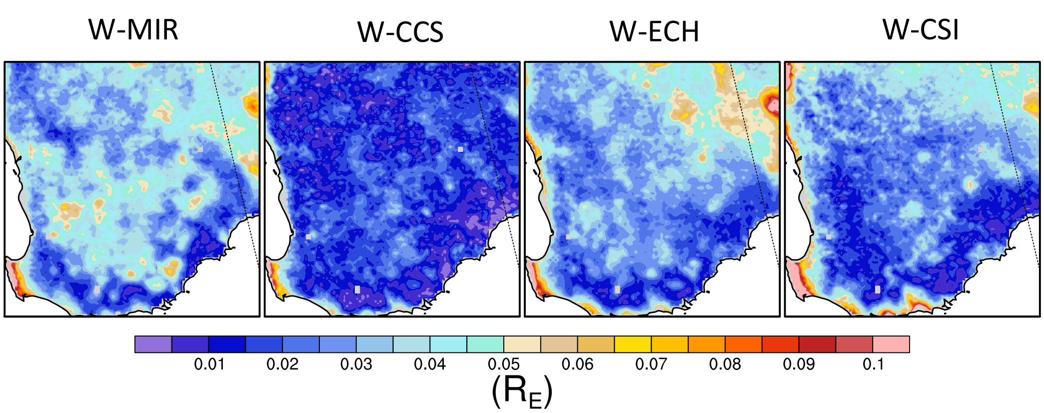


FIG. 7. Contour plot showing spatial distribution of daily rainfall R_E.

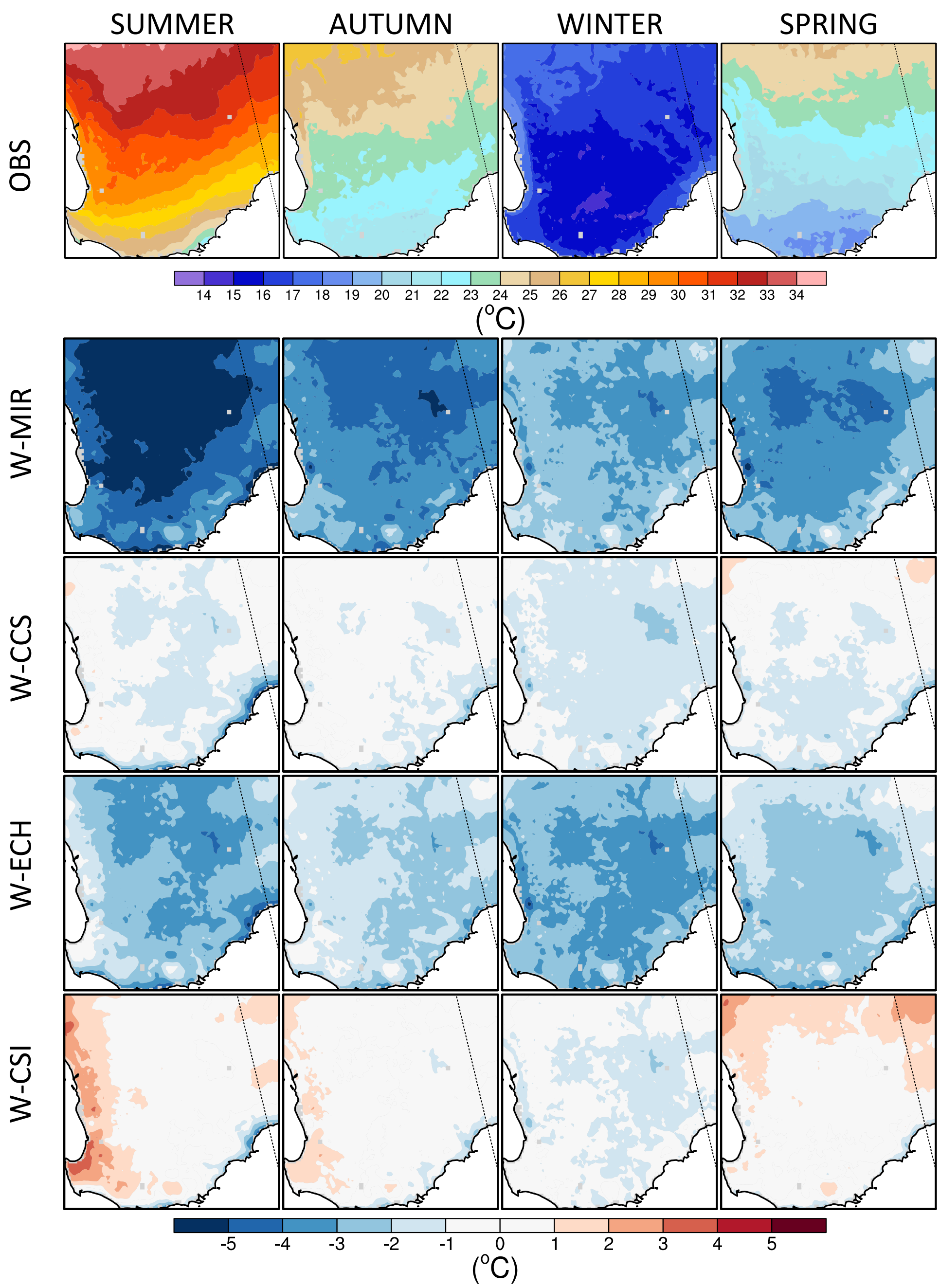
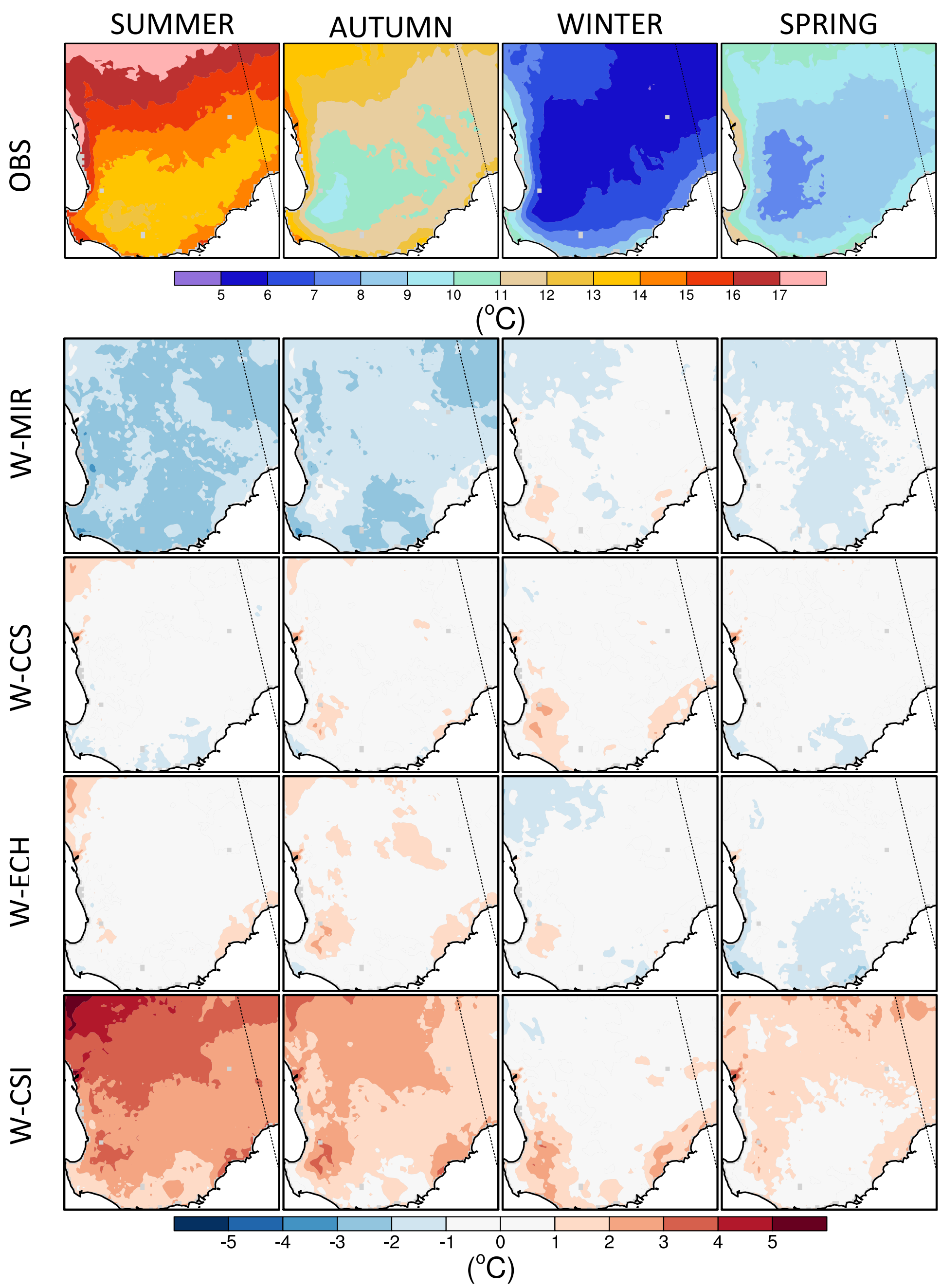
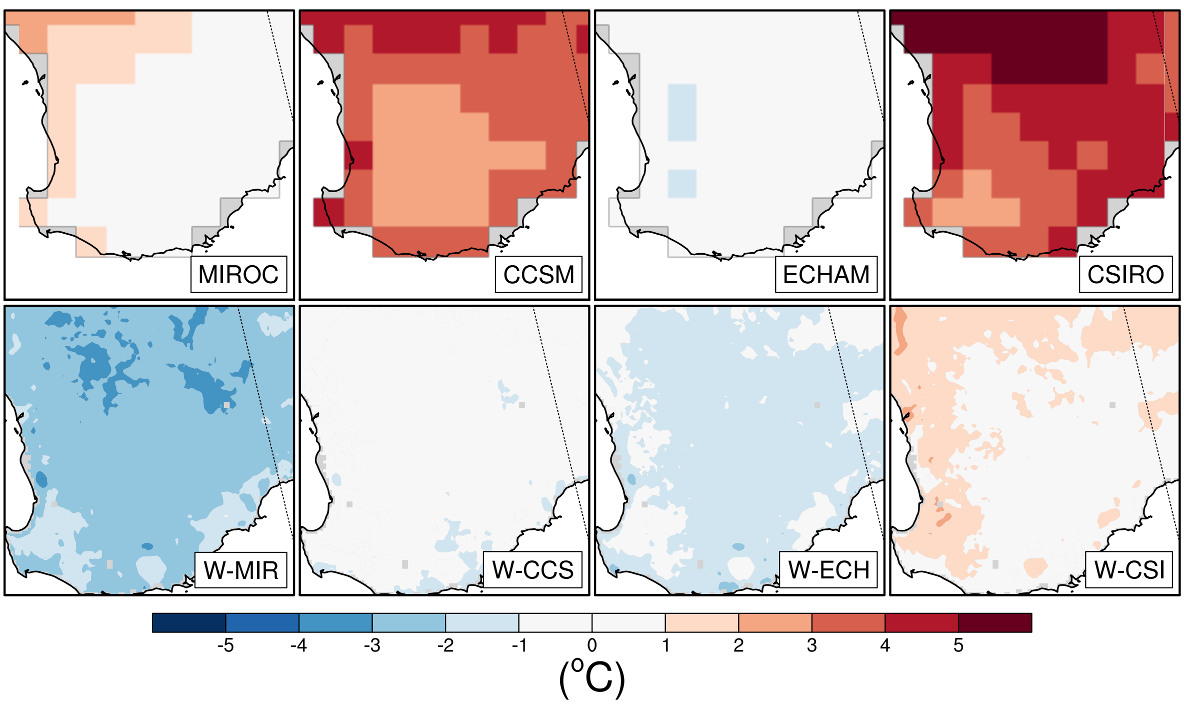


FIG. 8. Observed seasonal mean maximum temperatures (top panel) and bias (bottom panels) for all simulations over the period 1970-1999.



789 FIG. 9. Observed seasonal mean minimum temperatures (top panel) and bias (bottom panels) for all simula-
790 tions over the period 1970-1999.



791 FIG. 10. Mean annual temperature bias (1970-1999) for GCM model output and the corresponding RCM simulation.

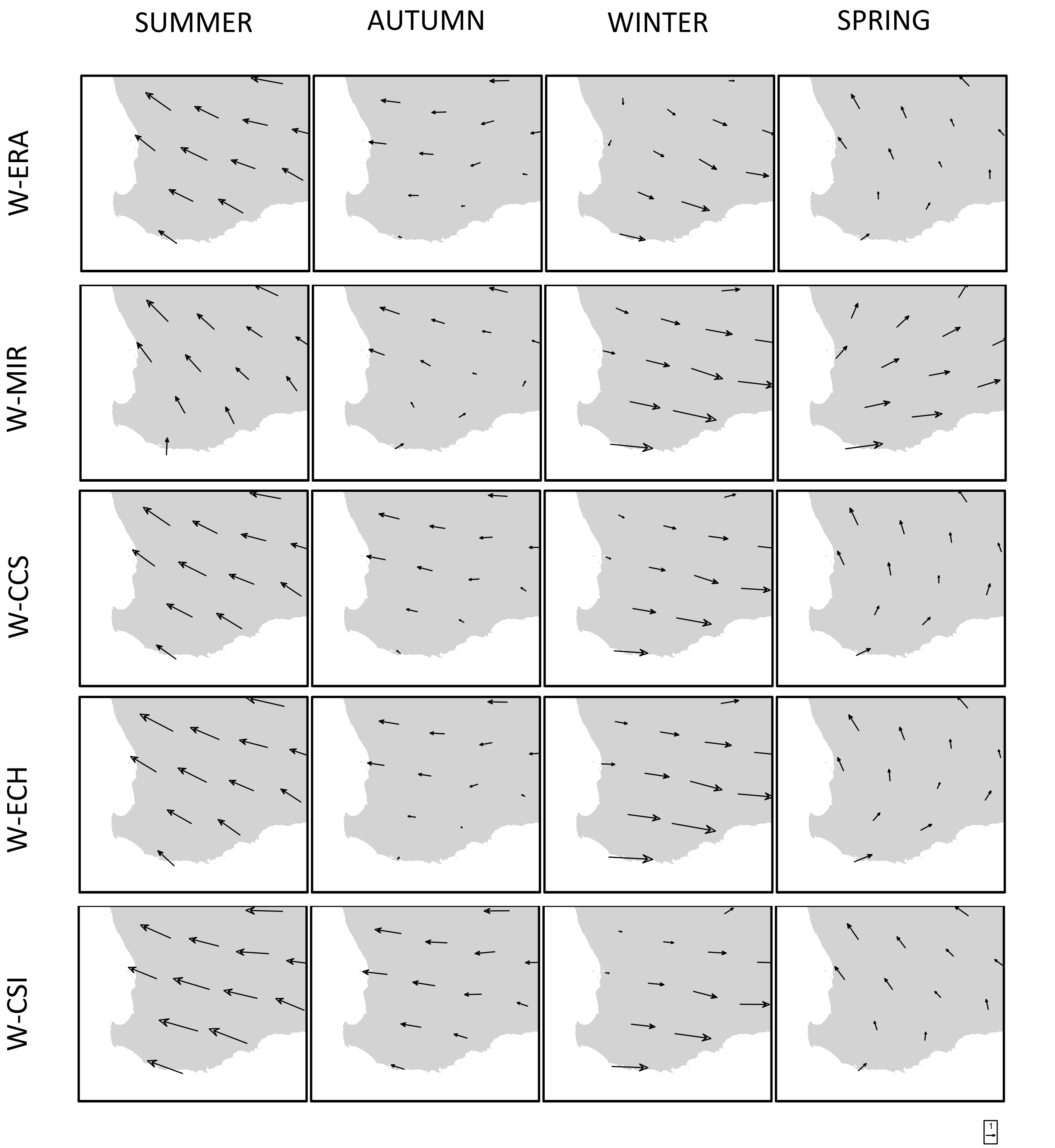


FIG. 11. Mean seasonal 10 m wind vectors for W-ERA and all simulations from 1980-1999. The reference vector represents a wind speed of 1 m s^{-1}

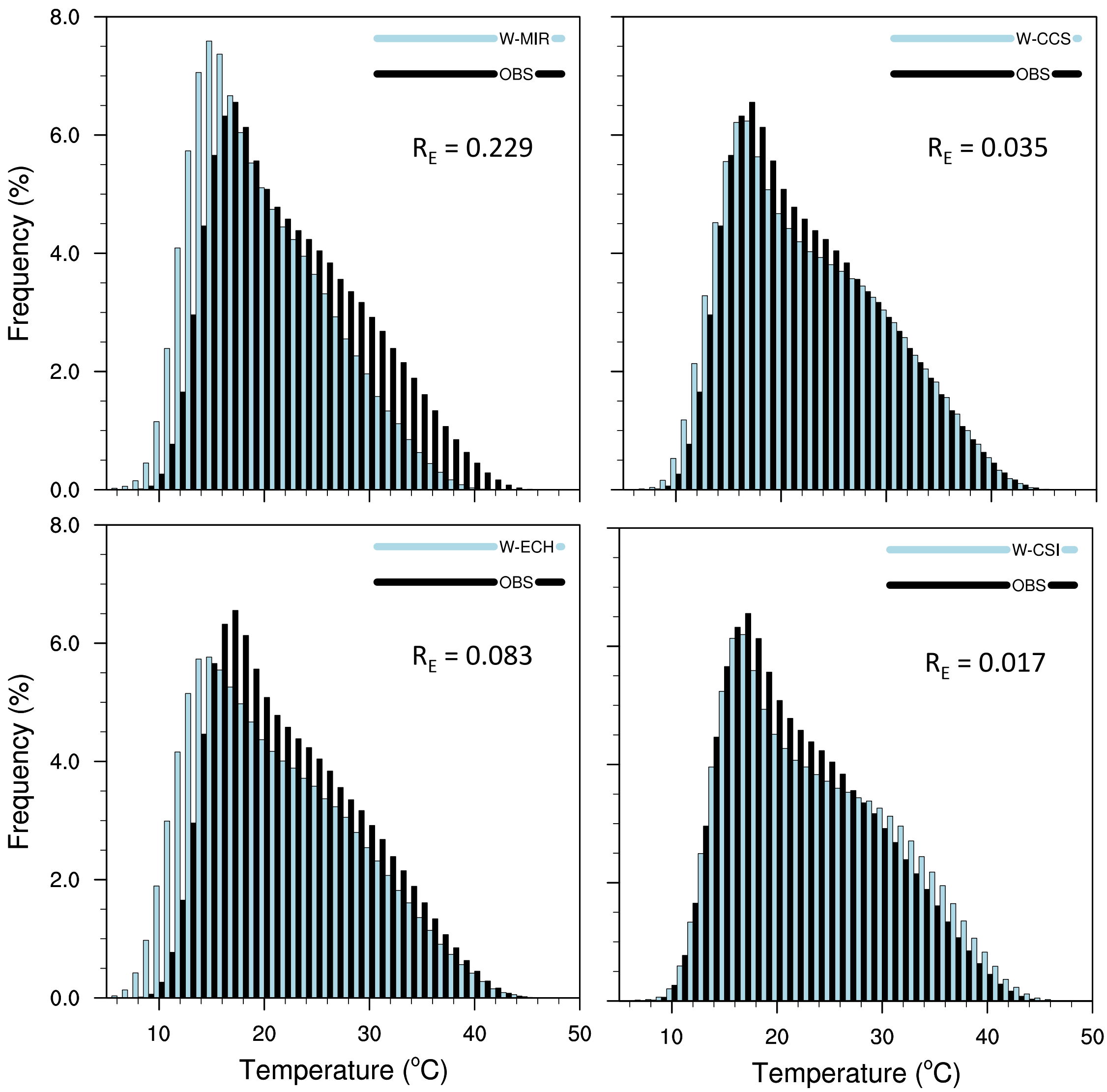
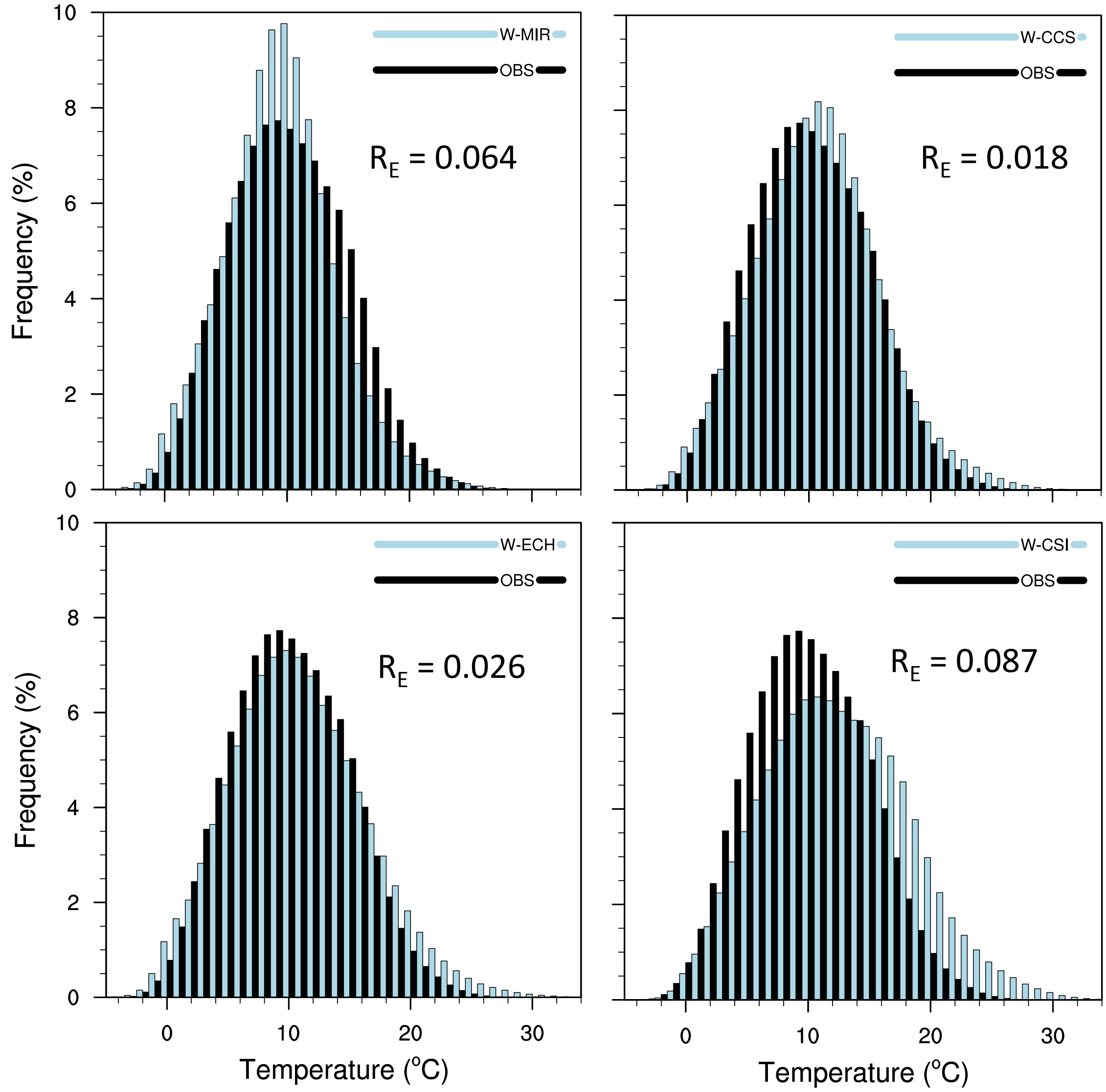


FIG. 12. Daily maximum temperature probability density functions for simulations and observations taken from all land based grid points from the 5 km domain. The R_E value comparing the similarity of the distributions is included for each plot.



798 FIG. 13. Daily minimum temperature probability density functions for simulations and observations taken
799 from all land based grid points from the 5 km domain. The R_E value comparing the similarity of the distributions
800 is included for each plot.

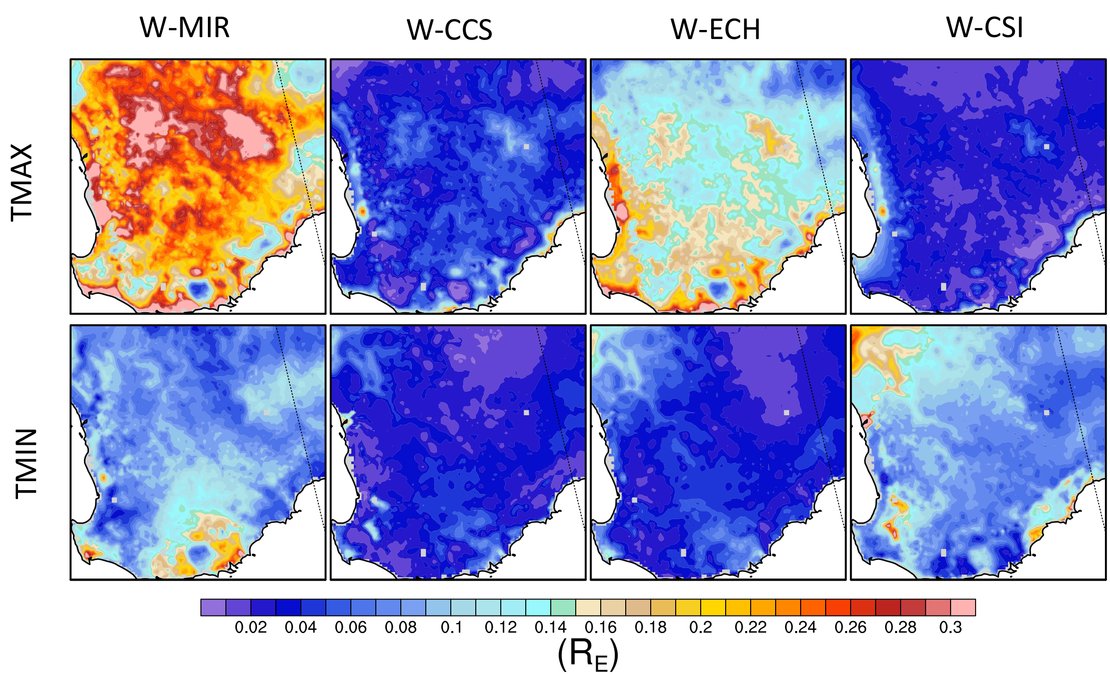
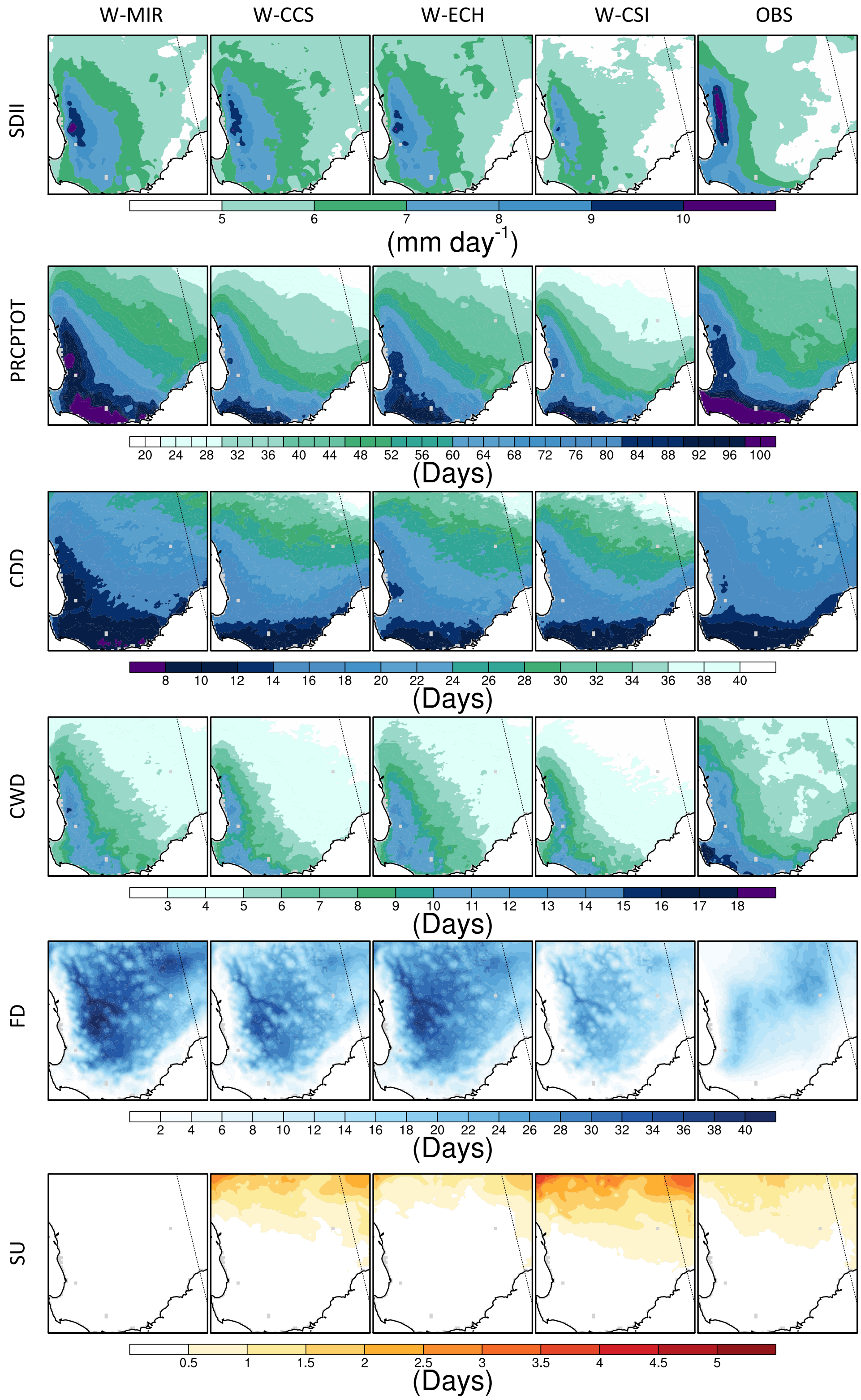


FIG. 14. Contour plots showing the spatial distribution of minimum and maximum temperature R_E.



801 FIG. 15. Contour plots showing the observed and simulated climatological mean of extreme indices calculated
802 over the SWWA growing season (May-October) only.