INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2016)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.4653

The effect of bias correction and climate model resolution on wheat simulations forced with a regional climate model ensemble

Ian Macadam,[a,b]* Daniel Argüeso,[a] Jason P. Evans,[a] De Li Liu[c] and Andy J. Pitman[a]

[a] *Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, Australia*
[b] *Now at Met Office, Exeter, UK*
[c] *New South Wales Department of Industry, Skills and Regional Development, Wagga Wagga Agricultural Institute, Australia*

ABSTRACT: Regional climate model (RCM) simulations are often used with agricultural models to assess the impact of climate change on agriculture. This study assesses the suitability of climate data sets from an RCM ensemble for forcing simulations of wheat yields for New South Wales, Australia, performed using the Agricultural Production Systems sIMulator (APSIM). Differences in yields between APSIM simulations forced with RCM output for the 1990–2009 period and APSIM simulations forced with observations are examined, as are simulated changes in yields between 1990–2009 and 2060–2079. The RCM ensemble downscales four global climate models (GCMs) using three different RCMs, firstly to a horizontal resolution of approximately 50 km, and then to approximately 10 km. All of the RCM simulations, 50 and 10 km, are able to simulate the spatial pattern of yields across the study region. However, some simulations have biases in yields. The largest of these are due to biases in rainfall during the growing season inherited from the GCMs. If the RCM output is bias-corrected, the largest positive yield biases are reduced because the underlying biases in growing season rainfall are reduced. Simulated future changes in yields are affected because there is a non-linear relationship between simulated yields and growing season rainfall. Downscaling to 10 km, rather than 50 km, is only beneficial when bias correction is used. The bias-correction technique used does not eliminate all biases in growing season rainfall and increases some of the smaller biases. Nonetheless, because of the potentially large effect of the rainfall biases on simulated future yield changes, this study supports the use of bias correction in assessments of the impacts of climate change that use RCM output to force APSIM. Indeed, our findings suggest that bias correction may be necessary to obtain reliable future changes in other outputs of other biophysical models that respond non-linearly to climate inputs.

KEY WORDS downscaling; model evaluation; WRF; APSIM; agricultural model; GCM; New South Wales; Australia

Received 11 September 2015; Revised 16 December 2015; Accepted 20 December 2015

1. Introduction

Regional climate models (RCMs) are an increasingly important tool in the assessment of climate change impacts. They translate broad-scale representations of future climate conditions from global climate models (GCMs) to climate data on the fine spatial scales commonly required by impact studies. Many less computationally expensive methods of downscaling GCM output exist, including statistical downscaling methods and the scaling of observations by GCM-simulated climate changes (see Maraun *et al.*, 2010; Evans *et al.*, 2012b; Ekström *et al.*, 2015). However, dynamical downscaling using RCMs has the advantage that it produces internally consistent data sets that include many climate variables. RCM output is increasingly being used in studies of the impacts of climate change on agriculture (e.g. Holz *et al.*, 2010; Zhang *et al.*, 2015). We investigate the suitability of RCM output for this purpose. We use different data sets derived from an RCM ensemble designed to underpin climate adaptation efforts in New South Wales (NSW), Australia (Evans *et al.*, 2014) with an agricultural model commonly used to simulate wheat cropping in the region, the Agricultural Production Systems sIMulator (APSIM) (Keating *et al.*, 2003).

For an agricultural model to simulate realistic changes in crop yields in response to future climate changes, it must be forced with climate data that realistically represent relevant aspects of the recent climate (Hawkins *et al.*, 2013; Glotter *et al.*, 2014). This is because crops have non-linear responses to climate and other environmental factors, meaning that responses to climate change depend on some aspects of the initial climate state. For example, French and Shultz (1984) describe the non-linear response of wheat yields to rainfall in South Australia.

Like GCM output, RCM output contains errors relative to observations (e.g. Evans and McCabe, 2013; Bennett *et al.*, 2014). Therefore, because agricultural models require realistic climate data as input, many studies of the impacts of climate change on crop yields have used bias-corrected RCM output to force agricultural model

* Correspondence to: I. Macadam, Met Office, FitzRoy Road, Exeter, EX1 3PB, United Kingdom. E-mail ian.macadam@metoffice.gov.uk

simulations (e.g. Holz *et al.*, 2010; Zhang *et al.*, 2015). However, bias correction has limitations. Firstly, it relies on developing transformations between climate model output and corresponding observations and so cannot be applied to variables for which such observations are not available. Furthermore, it is not usually possible to correct all relevant errors for variables for which observations are available, such as the shape of distributions of daily values of climate variables (Hawkins *et al.*, 2013) or the temporal sequencing of values (Piani *et al.*, 2010). Bias-correction methods usually assume that errors do not vary with time throughout a climate model simulation (Bennett *et al.*, 2014), and any time-dependent components of errors are not corrected (Haerter *et al.*, 2011). Furthermore, bias correction can disrupt the physical consistency of the climate model output, changing, for example, the relationships between different climate variables (Piani *et al.*, 2010). Finally, some bias-correction methods modify simulated future climate changes in ways that may not be justifiable (Haerter *et al.*, 2011).

Recognizing the disadvantages of bias correction, Macadam *et al.* (2014) investigated the suitability of uncorrected RCM output for forcing APSIM simulations of wheat cropping in NSW. They considered nested 50 km resolution and 10 km resolution simulations of the recent climate from a single RCM forced with a single GCM. We extend the work of Macadam *et al.* (2014) by investigating APSIM wheat simulations forced with both uncorrected and bias-corrected output from an ensemble of 50 and 10 km RCM simulations. We ask whether there are benefits to bias correction and increased RCM resolution by analysing wheat yields simulated for a recent period and simulated future changes in yields.

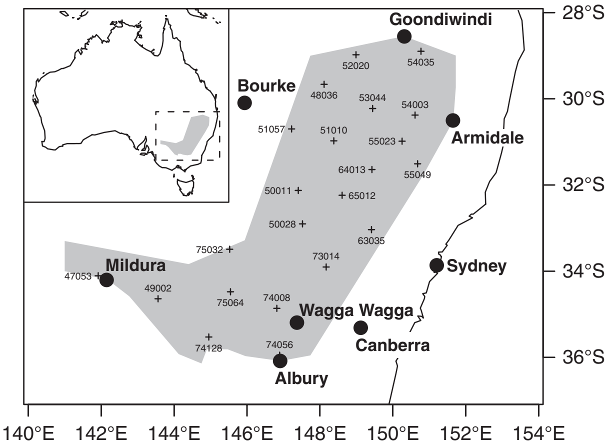

Figure 1. Maps showing the locations of the NSW wheat belt (inset) and the study region (main figure). The shaded region is the NSW wheat belt. The crosses show the 22 Australian Bureau of Meteorology observing sites considered by this study, with their station numbers. The black circles show selected towns and cities.

2. Method

2.1. Agricultural Production Systems sIMulator

APSIM is a well-developed biophysical model designed to simulate farming systems (Keating *et al.*, 2003). In Australia, the wheat module (Meinke *et al.*, 1998; Keating *et al.*, 2001; APSIM Initiative, 2014) has been widely used to examine the potential effects of increased atmospheric carbon dioxide (CO_2) concentrations and climate change on wheat yields (e.g. Howden *et al.*, 1999; Reyenga *et al.*, 1999; Howden and Crimp, 2005; Luo *et al.*, 2005; Wang *et al.*, 2009; Holz *et al.*, 2010; Luo *et al.*, 2010; Potgieter *et al.*, 2013).

APSIM is a paddock-scale model incorporating many empirically-derived algorithms representing relevant processes, such as those related to crop physiology. These algorithms have been extensively calibrated and evaluated against field observations by previous studies. Keating *et al.* (2003) list many of these studies. More recent studies include Chen *et al.* (2010), who calibrated and evaluated APSIM wheat and maize simulations against yield, biomass and leaf area observations from three field sites in China, and Liu *et al.* (2014), who evaluated the simulation of soil organic carbon storage for Wagga Wagga, NSW. As many APSIM-based studies of the impact of climate change on Australian wheat yields (e.g. Luo *et al.*, 2005; Luo *et al.*, 2009; Wang *et al.*, 2009; Luo *et al.*, 2010; Wang *et al.*, 2011) have done, this study relies on these previous studies to ensure plausible simulations of the processes governing wheat yields. Further calibration and evaluation of APSIM are beyond the scope of this study, which uses the model to assess RCM-derived climate data sets rather than to provide decision-relevant information on the impacts of climate change on wheat.

This study analyses annual wheat yields in units of mass per unit of area output by APSIM simulations. Data describing soil characteristics and crop management actions are simulation inputs, as are time series of daily rainfall, maximum air temperature (T_{Max}), minimum air temperature (T_{Min}) and solar radiation.

2.2. Observational data

We consider the same 22 Australian Bureau of Meteorology observing sites in the NSW wheat belt considered by Macadam *et al.* (2014) (Figure 1) and use the same point-based observational data set, the SILO patched point data set (Jeffrey *et al.*, 2001). In the SILO data set, gaps in meteorological records have been filled using spatial interpolation algorithms to produce complete records suitable as input to APSIM.

We also consider observational rainfall, T_{Max} and T_{Min} data from a gridded data set developed by Jones *et al.* (2009) for the Australian Water Availability Project (AWAP). The original AWAP data set, which is on a 5 km resolution grid, has been regridded by Evans and Argüeso (2014) onto the 50 and 10 km grids of the RCM simulations that we investigate. For each site, time series of daily rainfall, T_{Max} and T_{Min} for the 1990–2009 period were extracted from both the 50 and 10 km AWAP data sets. Following Macadam *et al.* (2014), data for the nearest gridpoint to each site were extracted.

2.3. RCM output

We use output from an ensemble of RCM simulations performed by the NSW and Australian Capital Territory Regional Climate Modelling (NARCliM) project (Evans *et al.*, 2014). The simulations use the Intergovernmental Panel on Climate Change's SRES A2 scenario for future emissions of greenhouse gases and sulphate aerosols (Nakićenović and Swart, 2000). SRES A2 simulations of four GCMs that contributed to the World Climate Research Programme's Coupled Model Intercomparison Project phase 3 (CMIP3) (Meehl *et al.*, 2007) were downscaled with three different RCMs. The GCMs used were the CCCMA CGCM3.1(T74), CSIRO-MK3.0, ECHAM5/MPI-OM and MIROC3.2(medres) GCMs, hereafter the 'CCCMA3.1', 'CSIROMK3.0', 'ECHAM5' and 'MIROC3.2' GCMs. This set of GCMs was selected based on ability to reproduce aspects of the observed climate, independence of errors in the reproduction of the observed climate and maximization of sampling of uncertainty in future changes in annual mean temperature and rainfall (see Evans *et al.*, 2014 for further details). The RCMs were selected from 36 different configurations of the Advanced Research WRF (Weather Research and Forecasting) version 3 modelling system (Skamarock *et al.*, 2008) that used different combinations of parametrization schemes for cumulus convection, cloud microphysics, radiation and the boundary layer. The set of three RCMs used, designated 'R1', 'R2' and 'R3', was selected based on ability to reproduce aspects of the observed climate and independence of errors in the reproduction of the observed climate (Evans *et al.*, 2012a).

For each of the 12 different GCM-RCM combinations, downscaling was done in two stages. Firstly, the RCM was used to simulate a domain covering Australia and the surrounding ocean at a horizontal resolution of 50 km using boundary forcing from the GCM. Secondly, the RCM was used to simulate a domain covering south-east Australia at a horizontal resolution of 10 km using boundary forcing from the 50 km simulation. Hence, 24 RCM simulations were performed for each of the periods simulated, 1990–2009, 2020–2039 and 2060–2079. We discuss analysis based on just the 1990–2009 and 2060–2079 simulations. This is because climate changes due to greenhouse gas forcing are larger between 1990–2009 and 2060–2079 than between 1990–2009 and 2020–2039. Analysis of the 2060–2079 simulations was therefore likely to provide the best opportunity to consider non-linear responses of APSIM to substantial future climate changes.

As well as uncorrected RCM output, the NARCliM project generated bias-corrected data. Daily rainfall, T_{Max} and T_{Min} data were corrected towards observations from the AWAP data set using a quantile-mapping technique (Argüeso *et al.*, 2013; Evans and Argüeso, 2014). The underlying idea of the method was to correct the probability distribution function of each variable using observations of the NARCliM recent period (1990–2009). At each model grid point, a theoretical cumulative probability distribution function is fitted to the data for each NARCliM simulation and for the corresponding AWAP data set (50 or 10 km). A gamma function is used for rainfall and a Gaussian function is used for T_{Max} and T_{Min}. The cumulative probability of a given simulated event is estimated using the theoretical function for the model and it is replaced with the event with equal cumulative probability from the theoretical function fitted to observations. All simulated values for each variable are corrected individually and, as a result, the simulated probability function is corrected towards the observed one for each location independently. Residual biases may still remain in the corrected model outputs because the fitting is performed towards theoretical functions, which ensures that a cumulative probability can be assigned to all simulated values.

Time series of daily rainfall, T_{Max}, T_{Min} and solar radiation were extracted from the gridded output of all of the 1990–2009 and 2060–2079 NARCliM RCM simulations for each of the 22 sites. For rainfall, T_{Max} and T_{Min}, both uncorrected and bias-corrected data were extracted. As for the AWAP data, data for the nearest gridpoint to each site were extracted.

2.4. APSIM simulations

For each of the 22 sites, an APSIM simulation was performed forced with uncorrected and bias-corrected rainfall, T_{Max} and T_{Min} data from each of the 1990–2009 and 2060–2079 NARCliM RCM simulations. Because bias-corrected solar radiation data were not available, both types of simulations used uncorrected solar radiation from the NARCliM simulations. For each site, 1990–2009 simulations were also performed forced with SILO observational data and a combination of rainfall, T_{Max} and T_{Min} data from the 50 and 10 km AWAP observational data sets and SILO solar radiation data.

All of the APSIM simulations simulated Ventura wheat growing in 'Brown Clay (Bellata No 121)' soil using a setup almost identical to that used by Macadam *et al.* (2014). In each simulation year, the wheat was sown as soon after 30 April as the quantity of water in the top layer of soil exceeded a prescribed threshold. If this threshold was not exceeded by 1 July, then the wheat was sown on that date. Atmospheric CO_2 concentrations followed the SRES A2 emissions scenario.

The simulations differed slightly from those of Macadam *et al.* (2014) in their treatment of initial conditions. Macadam *et al.* (2014) reset the amount of surface organic matter and the nitrate and ammonium contents of each soil layer to their initial values on 1 January every simulation year. Because the growing period of wheat in our simulations sometimes extended beyond 1 January, we reset these variables on 1 February.

Macadam *et al.* (2014) excluded output from the first 5 years of their simulations from their analysis to allow for the simulated soil water conditions to spin-up from a state close to their initial conditions to a state consistent with the meteorological forcing data. This limited their analysis to the 21-year 1990–2010 period. We did not wish to exclude any data from our 20-year 1990–2009 and 2060–2079 analysis periods. To spin-up our simulations, we duplicated the 20 years of forcing data, ran 40-year simulations and then disregarded the first 20 years of output.

2.5. Analysis of simulated wheat yields

We judge the realism of relevant aspects of the RCM-simulated recent climate by comparing wheat yields from APSIM simulations forced with RCM output with wheat yields from simulations forced with climate observations. Observations of yields are not used. This approach allows an assessment of the NARCliM climate simulations, distinct from any assessment of the crop model, imperfections in which are not addressed. This approach was used recently by Glotter *et al.* (2014) to evaluate the utility of dynamical downscaling in agricultural impacts projections for the United States. We calculated the mean value of the 20 annual wheat yield values from each of the 1990–2009 APSIM simulations forced with NARCliM output and subtracted the mean yield from the corresponding 1990–2009 APSIM simulation forced with SILO data. We regard the resulting differences as yield biases, and refer to them as such.

We evaluated future changes in wheat yields for the APSIM simulations forced with NARCliM output by subtracting the 1990–2009 mean yields from corresponding 2060–2079 mean yields. We refer to the resulting data as 'future yield changes'.

3. Results

3.1. APSIM simulations forced with SILO data

Figure 2 shows mean wheat yields for the 1990–2009 period from the APSIM simulations forced with SILO data. These are used as a reference data set to which yields from the simulations forced with RCM output are compared. The mean simulated yield averaged across the 22 sites is 3.4 t ha^{-1}. This is comparable to the actual yields of between 2.5 and 3.5 t ha^{-1} reported for most of the region for 1999–2000 by Wang *et al.* (2015). Note, however, that we expect the simulated yields to differ from observed yields due to imperfections in APSIM and unrealistic aspects of the APSIM setup that we have used.

To facilitate the subsequent discussion, the wheat belt is divided into western, central and eastern regions. The sites in the eastern region have the highest mean yields, exceeding 4 t ha^{-1}. The sites with the lowest mean yields, less than 3 t ha^{-1}, are in the western region. The sites in the central region have intermediate mean yields, between 3 and 4 t ha^{-1}.

3.2. Fifty kilometre NARCliM simulations

Figures 3(a) and 4(a) show mean wheat yields for the 1990–2009 period and future yield changes from the APSIM simulations forced with the 50 km resolution NARCliM data. Figure 3(a) also shows mean yields for 1990–2009 derived from the SILO and AWAP 50 km data

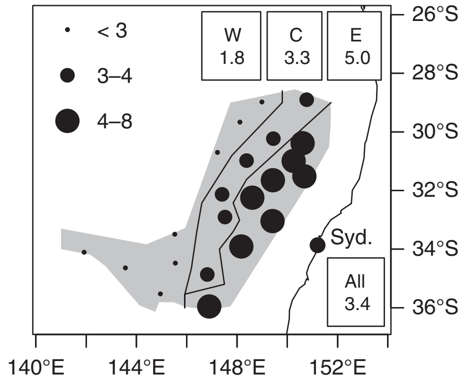

Figure 2. Map of mean wheat yields (in t ha^{-1}) for the 1990–2009 period from APSIM simulations forced with SILO data. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values for the regions and across all sites are shown in boxes.

sets. Figures S1 and S2, Supporting Information, show the NARCliM-derived data in map form.

Figure 3(a) shows that, generally, the 50 km simulations forced with different GCMs give different yields from each other, and from the SILO data set. For example, the mean yields averaged across all 22 sites for the CCCMA3.1, CSIROMK3.0, ECHAM5 and MIROC3.2 GCMs downscaled with the R1 RCM, 2.6, 5.4, 4.3 and 6.9 t ha^{-1}, respectively (Figure S1), differ from each other, and from the mean yield for the SILO data set, 3.4 t ha^{-1} (Figure 2). In contrast, mean yields for the three different RCMs are similar when they are forced with the same GCM. For example, wheat-belt-wide mean yields differ by only 0.2 t ha^{-1} across the three RCM simulations forced with the CCCMA3.1 GCM (Figure S1).

Figure 3(a) shows that there are negative yield biases for all or almost all sites for each of the CCCMA3.1-forced simulations. The wheat-belt-wide mean yield is 3.4 t ha^{-1} for the SILO data set (Figure 2) but only between 2.4 and 2.6 t ha^{-1} for the CCCMA3.1-forced simulations (Figure S1). There are positive biases for all or almost all sites for each of the ECHAM5-, CSIROMK3.0- and MIROC3.2-forced simulations. Respectively, wheat-belt-wide mean yields are between 4.1 and 4.7 t ha^{-1}, between 4.4 and 5.4 t ha^{-1} and between 6.4 and 7.8 t ha^{-1} for these sets of simulations (Figure S1).

Although there are wheat-belt-wide biases in the NARCliM-derived yields relative to the SILO-derived data, all 12 50 km NARCliM simulations broadly capture the overall spatial pattern of mean yields for the SILO-derived data set. However, the size of the west-to-east 'gradient' in mean yields for some of the NARCliM simulations is different from that for the SILO data. For the SILO data, Figure 2 shows that the mean yield averaged across the sites in the eastern region is 3.2 t ha^{-1} higher than the mean yield averaged across the sites in the western region. The gradients for the CSIROMK3.0-forced simulations are larger, with the mean yield averaged across the sites in the eastern region being 4.1–4.6 t ha^{-1} higher than the mean yield

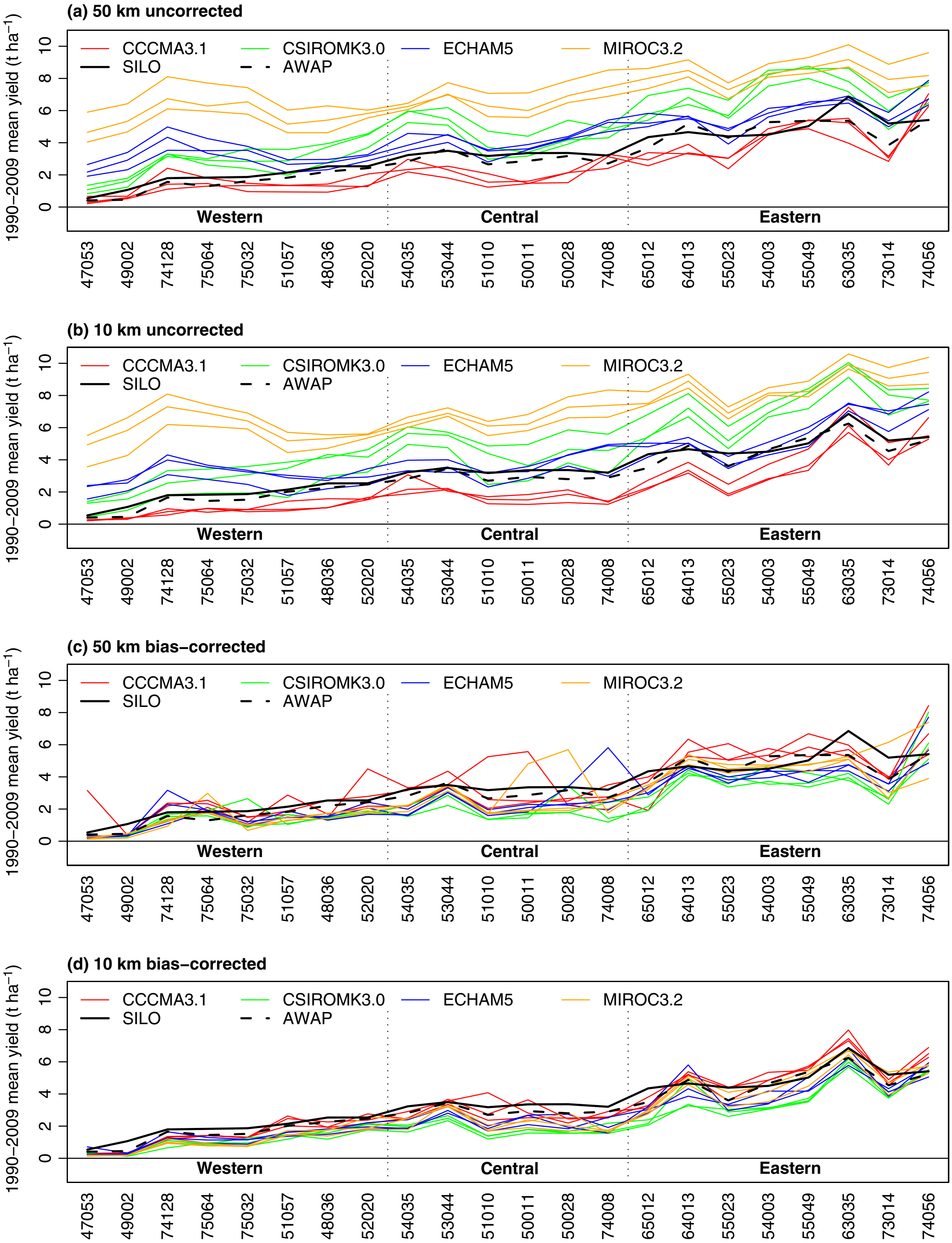

Figure 3. Graphs of mean wheat yields for the 1990–2009 period from APSIM simulations forced with observational data and 50 km resolution (a), 10 km resolution (b), bias-corrected 50 km resolution (c) and bias-corrected 10 km resolution (d) NARCliM data. The 22 analysis sites are listed along the *x*-axis, with low-yield sites on the left and high-yield sites on the right. The vertical dotted lines show the division of the wheat belt into western, central and eastern regions.

averaged across the sites in the western region (Figure S1). Figure 3(a) shows that this is due to greater positive yield biases in the east than in the west. The gradients for the MIROC3.2-forced simulations are smaller than for the SILO data, with the mean yield averaged across the sites in the eastern region being only 2.3–2.6 $\mathrm{t\,ha^{-1}}$ higher than the mean yield averaged across the sites in the western region (Figure S1). This is due to greater positive yield biases in the west than in the east (Figure 3(a)).

Figure 3(a) shows that, generally, the mean yields for the AWAP 50 km data are slightly smaller than for the SILO data. The mean yield averaged across all sites is

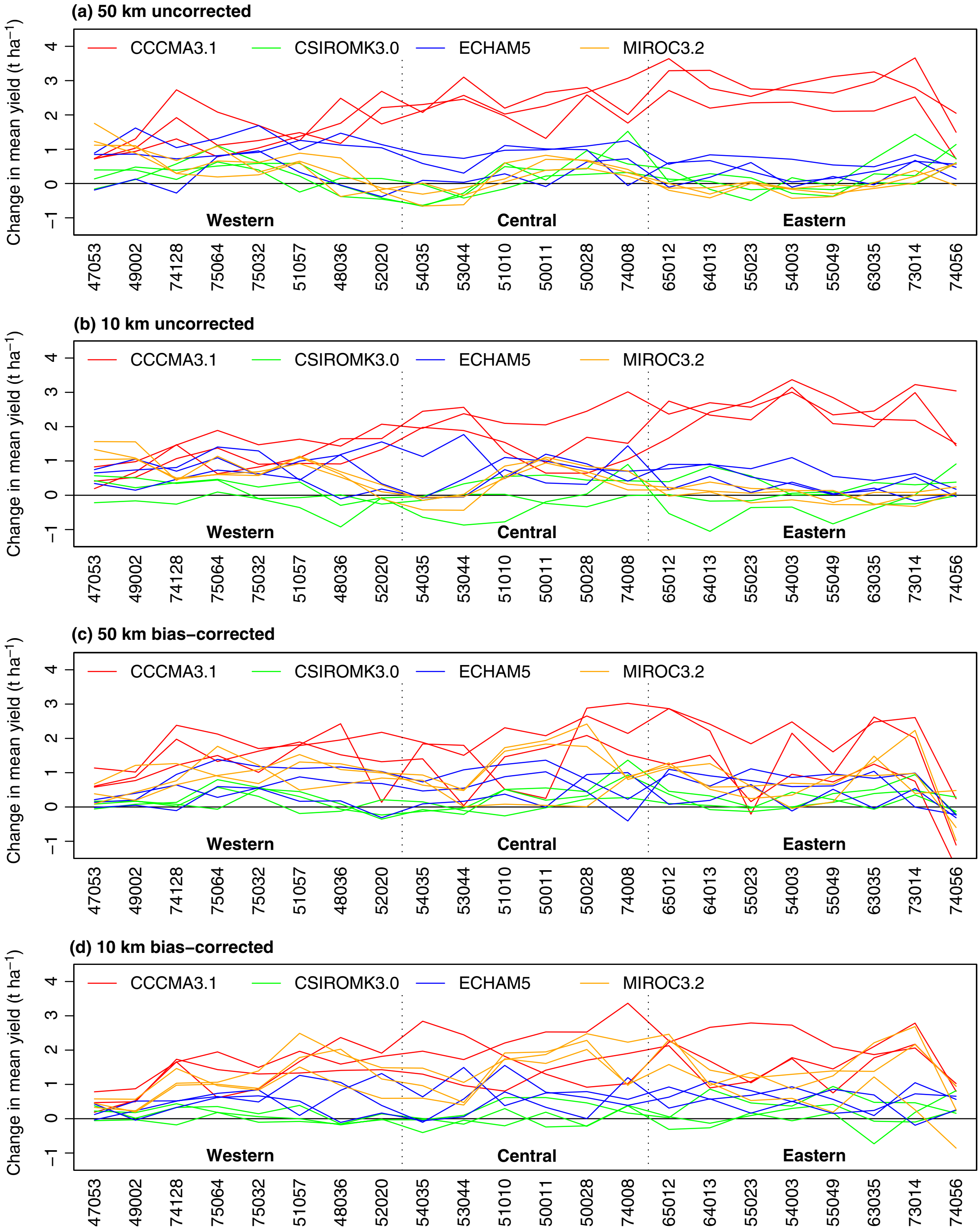

Figure 4. Graphs of changes in mean wheat yields between 1990–2009 and 2060–2079 from APSIM simulations forced with 50 km resolution (a), 10 km resolution (b), bias-corrected 50 km resolution (c) and bias-corrected 10 km resolution (d) NARCliM data. The 22 analysis sites are listed along the *x*-axis, with low-yield sites on the left and high-yield sites on the right. The vertical dotted lines show the division of the wheat belt into western, central and eastern regions.

3.1 t ha^{-1} for the AWAP 50 km data (not shown), whereas it is 3.4 t ha^{-1} for the SILO data. Except for sites 63035 and 73014 in the east, the difference in mean yields for individual sites is generally smaller than the differences between the NARCliM-derived yields and the SILO-derived yields. Therefore, the findings described above for yield biases relative to the SILO data are also true of biases relative to the AWAP 50 km data.

Figure 4(a) shows that 50 km simulations forced by the same GCM give similar future yield changes, though not to the same extent as for the 1990–2009 mean yields (Figure 3(a)). Most of the changes are increases in mean

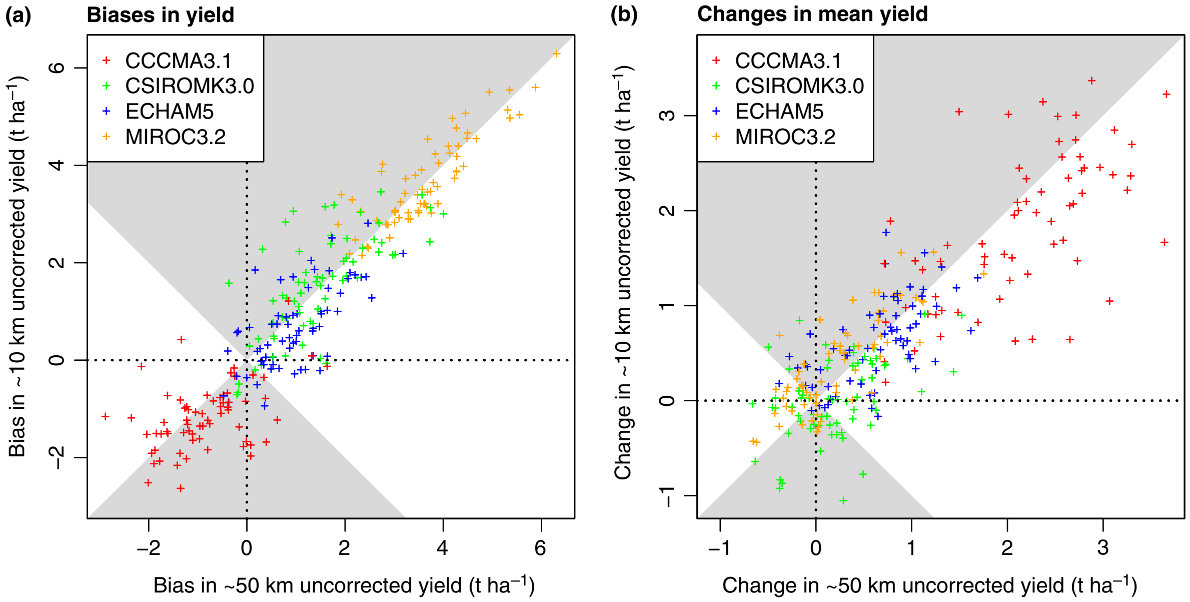

Figure 5. Biases in wheat yields (a) and changes in mean wheat yields between 1990–2009 and 2060–2079 (b) from APSIM simulations forced with 10 km resolution NARCliM data plotted against corresponding data for APSIM simulations forced with 50 km resolution NARCliM data. Data points in the shaded region have larger values for the 10 km data than for the 50 km data.

yield. The largest increases are for the CCCMA3.1-forced simulations. The mean yield increases averaged across all sites for these simulations are between 1.8 and 2.4 t ha^{-1} (Figure S2). The increases are largest for sites in the eastern region, where the average increases are between 2.1 and 2.9 t ha^{-1}, and smallest for the western region, where the average increases are between 1.3 and 1.8 t ha^{-1}. The CSIROMK3.0- and MIROC3.2-forced simulations give changes in yields of less than 1 t ha^{-1} for most sites. Some decreases of generally less than 0.5 t ha^{-1} are given for some sites, generally in the north. The magnitudes of future yield changes averaged across all sites for these simulations are between 0.2 and 0.5 t ha^{-1}. The ECHAM5 R2 simulation has similar yield changes to the CSIROMK3.0- and MIROC3.2-forced simulations. However, the ECHAM5 R1 and ECHAM5 R3 simulations have moderate increases in mean yield, with wheat-belt-wide mean changes of 0.7 and 1.0 t ha^{-1}. In contrast to the CCCMA3.1-forced simulations, these are smallest for the sites in the eastern region, where the average increases are 0.4 and 0.7 t ha^{-1}, and largest for the western region, where the average increases are 0.9 and 1.3 t ha^{-1}.

3.3. Ten kilometre NARCliM simulations

Figures 3(b) and 4(b) present the 1990–2009 mean wheat yields and future yield changes from the APSIM simulations forced with the 10 km resolution NARCliM data. Broadly, Figures 3(b) and 4(b) resemble Figures 3(a) and 4(a), indicating that many of the characteristics of the mean yields from these simulations are similar to those of the mean yields from the APSIM simulations forced with the 50 km resolution NARCliM data.

As for the 50 km NARCliM simulations, the mean yields and yield changes for the 10 km simulations are generally similar between simulations forced with the same GCM. Like the 50 km NARCliM simulations, the 10 km simulations broadly capture the overall spatial pattern of mean yields in the SILO-derived data set, but also have biases (Figure 3(b)). Because the yields for the AWAP 10 km data are similar to those of the SILO data, the biases relative to the AWAP 10 km data are similar to those relative to the SILO data. The spatial patterns of yield changes are similar between the 10 km simulations and the corresponding 50 km simulations, with increases for the CCCMA3.1-forced simulations that are larger towards the east and increases for the ECHAM5-forced simulations that are larger towards the west (Figure 4(b)).

Figure 5 shows the wheat yield biases and future yield changes from the APSIM simulations forced with the 10 km NARCliM data plotted against corresponding data for the APSIM simulations forced with the 50 km NARCliM data. Considering all 12 GCM-RCM combinations and all 22 sites, there are no clear overall systematic differences between the biases for the two different resolutions (Figure 5(a)). However, there are some notable differences in yield change for some of the simulations (Figure 5(b)). The 10 km CCCMA3.1 R2 and CCCMA3.1 R3 simulations have smaller increases in mean yield than the corresponding 50 km simulations. The magnitude of mean yield increase averaged across all sites is 1.6 t ha^{-1} for both the 10 km CCCMA3.1 R2 and CCCMA3.1 R3 simulations (Figure S4), whereas it is 1.8 and 2.4 t ha^{-1} for the corresponding 50 km simulations (Figure S2).

3.4. Bias-corrected NARCliM data

Figure 3(c) and (d) present mean wheat yields for the 1990–2009 period from the APSIM simulations forced with bias-corrected data from the 50 and 10 km resolution NARCliM simulations. These APSIM simulations capture the overall spatial pattern of mean yields in the

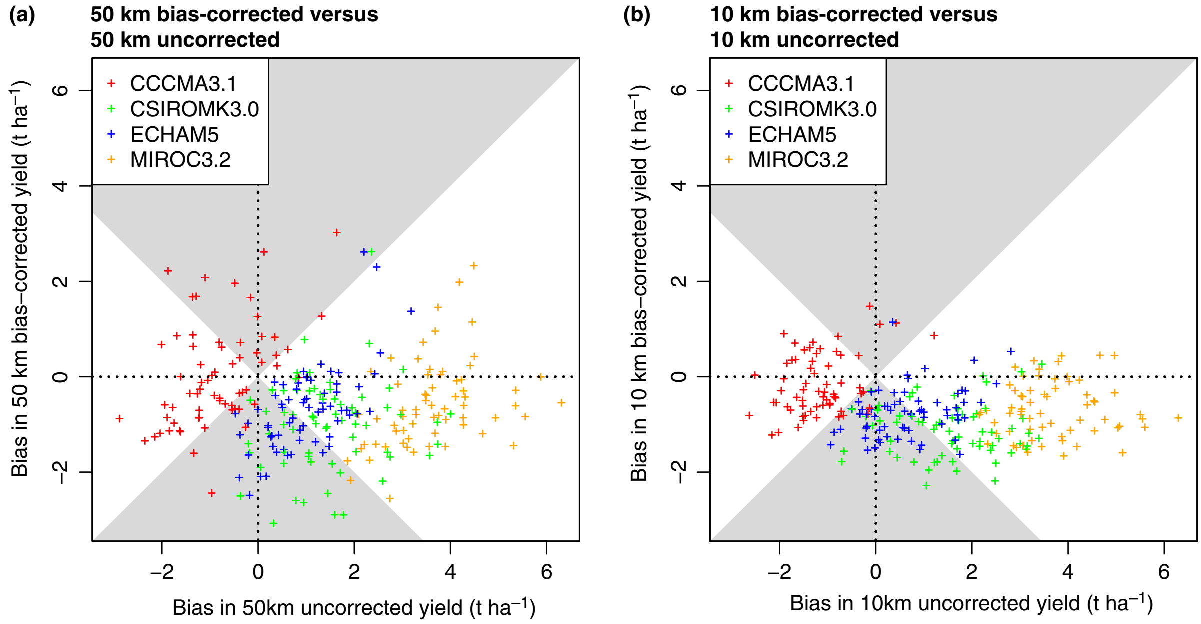

Figure 6. Biases in wheat yields from APSIM simulations forced with bias-corrected NARCliM data plotted against corresponding biases for APSIM simulations forced with uncorrected NARCliM data. Data points in the shaded region have larger biases for the bias-corrected data than for the uncorrected data.

SILO-derived data set and, unlike for the uncorrected NARCliM data, almost all of the NARCliM simulations have a yield 'gradient' between the low- and high-yield sites that is similar to that for the SILO data. For the uncorrected data, differences between the mean yields averaged across the sites in the eastern region and the mean yields averaged across the sites in the western region are 2.3–4.6 t ha^{-1} for the 50 km simulations (Figure S1) and 2.6–4.9 t ha^{-1} for the 10 km simulations (Figure S3). For the bias-corrected data, the corresponded ranges are 2.6–3.9 t ha^{-1} and 2.7–3.9 t ha^{-1} (Figures S5 and S7), closer to the value of 3.2 t ha^{-1} for the SILO data (Figure 2).

Figure 3 also shows that, for both RCM resolutions, the mean yields from the 12 different NARCliM simulations are much more similar to each other in terms of the magnitude of yields than for the corresponding uncorrected data. For the uncorrected data, the range of values for the wheat-belt-wide mean yield is 2.4–7.8 t ha^{-1} for the 50 km simulations (Figure S1) and 2.1–7.6 t ha^{-1} for the 10 km simulations (Figure S3). For the bias-corrected data, the corresponded ranges are 2.2–3.9 t ha^{-1} and 2.2–3.3 t ha^{-1} (Figures S5 and S7).

The bias-corrected data for both RCM resolutions give negative yield biases relative to the SILO data set for most of the 22 sites for all 12 simulations (Figure 3(c) and (d)). While the predominantly negative yield biases for the uncorrected CCCMA3.1-forced simulations are largely removed by bias correction, the predominantly positive biases for the uncorrected CSIROMK3.0-, ECHAM5- and MIROC3.2-forced simulations correspond to negative biases for the bias-corrected data. Because the NARCliM data are corrected towards the AWAP data set, the yields derived from the bias-corrected NARCliM data are generally closer to the yields derived from the AWAP data than to the SILO-derived yields. Nonetheless, the CSIROMK3.0-, ECHAM5- and MIROC3.2-forced simulations have negative biases relative to the AWAP data as well as relative to the SILO data.

Figure 6 shows the wheat yield biases from the APSIM simulations forced with bias-corrected data from the 50 and 10 km NARCliM simulations plotted against corresponding data for the APSIM simulations forced with uncorrected NARCliM data. The bias correction is more effective at removing yield biases for the 10 km simulations than for the 50 km simulations. Most of the yield biases for the bias-corrected 10 km data (Figure 6(b)) are smaller than the corresponding yield biases for the 50 km data (Figure 6(a)) and the largest biases for the 10 km data, around 2 t ha^{-1}, are smaller than the largest biases for the 50 km data, around 3 t ha^{-1}. However, the relative performance of the bias correction between RCM simulations forced with different GCMs is similar for the two resolutions. For both resolutions, the bias correction is most effective at removing large positive yield biases. Biases exceeding 2 t ha^{-1} are reducing in magnitude, to less than 2 t ha^{-1} in almost all cases. This means that the magnitudes of the large biases for MIROC3.2-forced simulations are reduced to magnitudes similar to those for the other RCM simulations. The bias correction is least effective at removing the yield biases for the CSIROMK3.0- and ECHAM5-forced simulations. Although most of the largest biases for these simulations are reduced in magnitude, and some of the smaller biases are reduced, many of the smaller biases are actually increased in magnitude by the bias correction. Finally, the bias correction is moderately effective at removing the yield biases for the CCCMA3.1-forced simulations. Almost all of these

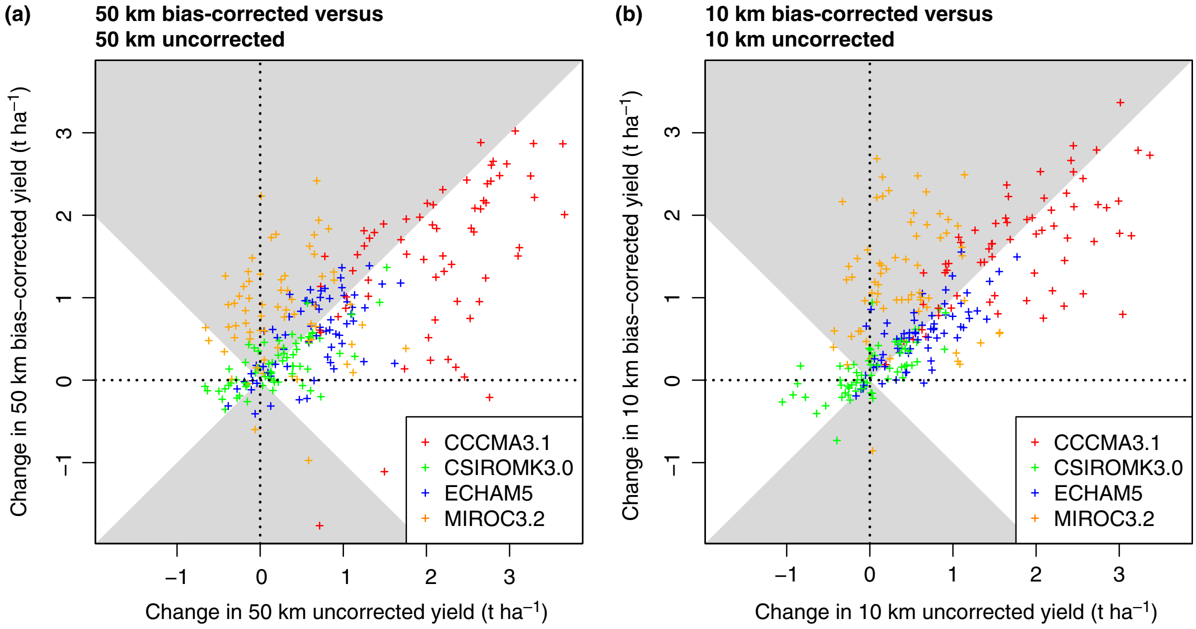

Figure 7. Changes in mean wheat yields between 1990–2009 and 2060–2079 from APSIM simulations forced with bias-corrected NARCliM data plotted against corresponding biases for APSIM simulations forced with uncorrected NARCliM data. Data points in the shaded region have larger changes for the bias-corrected data than for the uncorrected data.

biases are reduced in magnitude for the 10 km simulations, although a greater proportion of these biases are increased in magnitude for the 50 km simulations.

Figure 4(c) and (d) presents future yield changes from the APSIM simulations forced with bias-corrected data from the 50 and 10 km simulations. Figure 7 shows the yield changes for the bias-corrected data plotted against the changes for the uncorrected data. For both resolutions, bias correction makes little difference to most of the changes for the CCCMA3.1-, CSIROMK3.0- and ECHAM5-forced simulations. A notable exception is that many of the large increases in mean yield for the uncorrected CCCMA3.1-forced simulations for the eastern region and, for the 50 km simulations, the central region correspond to substantially smaller increases for the bias-corrected data. The magnitude of mean yield increase averaged across the sites in the eastern region is between 2.1 and 2.9 t ha^{-1} for the uncorrected 50 km CCCMA3.1-forced simulations (Figure S2), whereas it is between 1.0 and 2.1 t ha^{-1} for the corresponding bias-corrected data (Figure S6). Importantly, this removes almost all of the yield increases that exceed 3.0 t ha^{-1}, thus narrowing the range of yield changes simulated by the ensemble. In addition, the decreases in mean yield for the uncorrected data from the CSIRO-MK3.0-forced simulations correspond to smaller decreases or small increases for the bias-corrected data. However, the largest differences between the yield changes for the uncorrected and bias-corrected data are for the MIROC3.2-forced simulations. Most of the changes for these simulations are much larger for the bias-corrected data than for the uncorrected data. The wheat-belt-wide mean magnitude of change in yield is between 0.3 and 0.5 t ha^{-1} for the uncorrected 50 km MIROC3.2-forced simulations (Figure S2), whereas it is between 0.6 and 1.1 t ha^{-1} for the corresponding bias-corrected data (Figure S6).

4. Discussion

4.1. Yield–rainfall relationship

Rainfall during the growing season is a key determinant of dry-land wheat yields in southern Australia (e.g. French and Shultz, 1984) and this is reflected in studies that have used APSIM to simulate wheat cropping in the region (e.g. Wang *et al.*, 2009; Macadam *et al.*, 2014). Macadam *et al.* (2014) showed a strong association between wheat yields and May–December total rainfall in their APSIM simulations for the sites that we consider. Figure 8 shows mean wheat yields for the 1990–2009 period from the APSIM simulations forced with SILO observational data and uncorrected 50 and 10 km resolution NARCliM data plotted against corresponding mean May–December rainfall total data. The relationship between yields and rainfall is non-linear. Yields increase linearly with rainfall until mean May–December total rainfall exceeds around 500 mm. For wetter climates, other factors increasingly limit increases in yield per unit of rainfall. In APSIM, these include a limit on the rate of photosynthesis due to finite interception of solar radiation by leaves. Similar non-linear yield–rainfall relationships are observed in the real world due to this, and factors not represented by APSIM (e.g. crop pests and diseases) (French and Shultz, 1984).

4.2. Yield biases for uncorrected RCM data

Macadam *et al.* (2014) assessed whether biases in a range of different aspects of the simulated climate during the

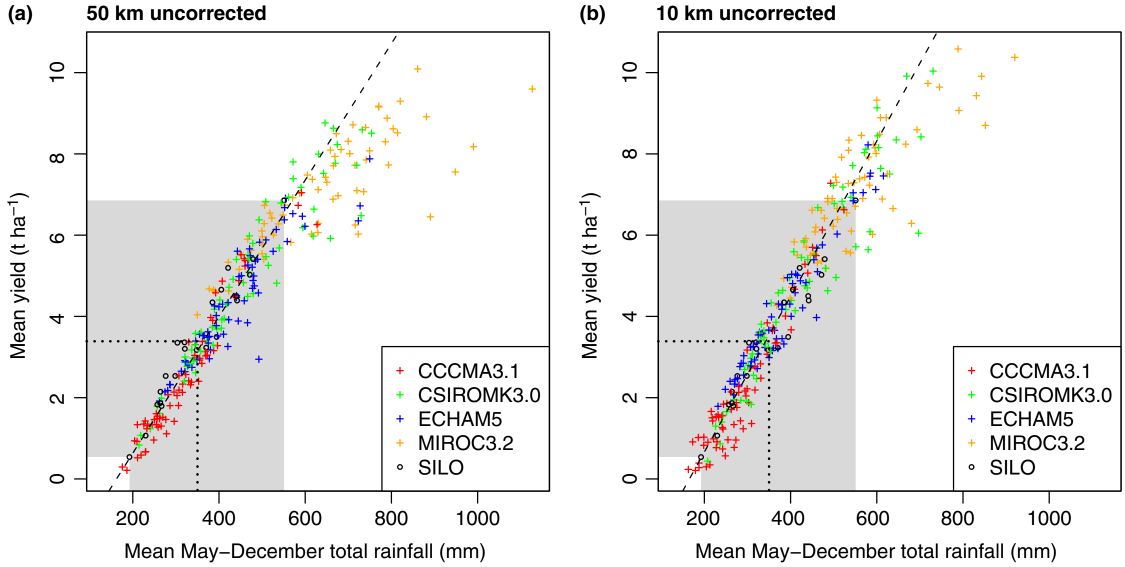

Figure 8. Mean wheat yields for the 1990–2009 period from APSIM simulations forced with NARCliM data and SILO observational data plotted against corresponding mean May–December rainfall total data. The shaded regions and the dotted lines show the range of values and the mean values for the SILO data. The dashed lines show linear ordinary least squares regression fits to the data for which the mean rainfall total is less than 500 mm.

growing season might be responsible for yield biases in their APSIM wheat simulations. These included the frequency and intensity of rainfall, mean and extreme temperatures, diurnal temperature range and solar radiation. However, for the single GCM-RCM combination that they investigated, they concluded that biases in simulated yields for the sites that we consider were driven by biases in May–December total rainfall. The strong association between mean yields and mean May–December total rainfall (Figure 8) suggests that the same is true for our APSIM simulations forced with data from 12 different GCM-RCM combinations.

NARCliM simulations forced with the same GCM have similar yield biases (Figure 3(a) and (b)). This suggests that the rainfall biases in an RCM simulation that are largely determining its yield biases are themselves largely determined by the forcing data from the driving GCM. Biases in the MIROC3.2 GCM are almost certainly the reason for the particularly poor correspondence between the yields for the APSIM simulations forced with data downscaled from this GCM and the yields for the APSIM simulations forced with observational data. Of the four GCMs chosen for downscaling in the NARCliM project, only MIROC3.2 has substantial positive biases in rainfall over the NSW wheat belt throughout the June–November season (Olson *et al.*, 2014). The RCM simulations forced with the MIROC3.2 GCM inherit these biases and are wetter, and more conducive to wheat growth, than the other RCM simulations (Figure 8), leading to large positive yield biases (Figure 3(a) and (b)). The enhancement of yields is especially large in the west of the wheat belt, where precipitation is least, and yields are most sensitive to precipitation. Therefore, the 'gradient' in yields across the wheat belt is smaller for the MIROC3.2-forced simulations than for the other simulations.

4.3. Value of bias correction and higher climate model resolution

The strong association between mean wheat yields and mean May–December rainfall totals (Figure 8) suggests that the effect of the bias correction on yield biases could be explained by the effect of the bias correction on the rainfall totals. This is supported by the strong resemblance of Figure 9, showing biases, relative to the SILO data set, in May–December rainfall totals for the bias-corrected data plotted against corresponding biases for the uncorrected data, to Figure 6, the equivalent figure for yield biases.

It is noteworthy that the bias-correction method that we have used results in a data set that contains predominantly negative biases in May–December rainfall totals and, in some cases, results in an increase in the magnitude of small biases in May–December rainfall totals. In theory, this could be because we have defined rainfall biases relative to the point-based SILO observations. The bias correction corrects data towards the 50 and 10 km AWAP observational data sets. In theory, differences between the SILO and AWAP data sets could contribute to rainfall biases as we have defined them. However, in general, we found little difference between the data sets in terms of mean May–December rainfall totals for 1990–2009. We therefore suggest two alternative causes of the residual rainfall biases. Firstly, rainfall biases may remain in the bias-corrected NARCliM data due to mismatches between the actual daily rainfall data and the gamma cumulative probability distribution function fitted to them as part of the bias-correction procedure (Argüeso *et al.*, 2013). Secondly, bias correction of data for the daily timescale does not necessarily result in bias-free data at other timescales, such as the seasonal (May–December) timescale that we have analysed. Indeed, Haerter *et al.* (2011) showed that

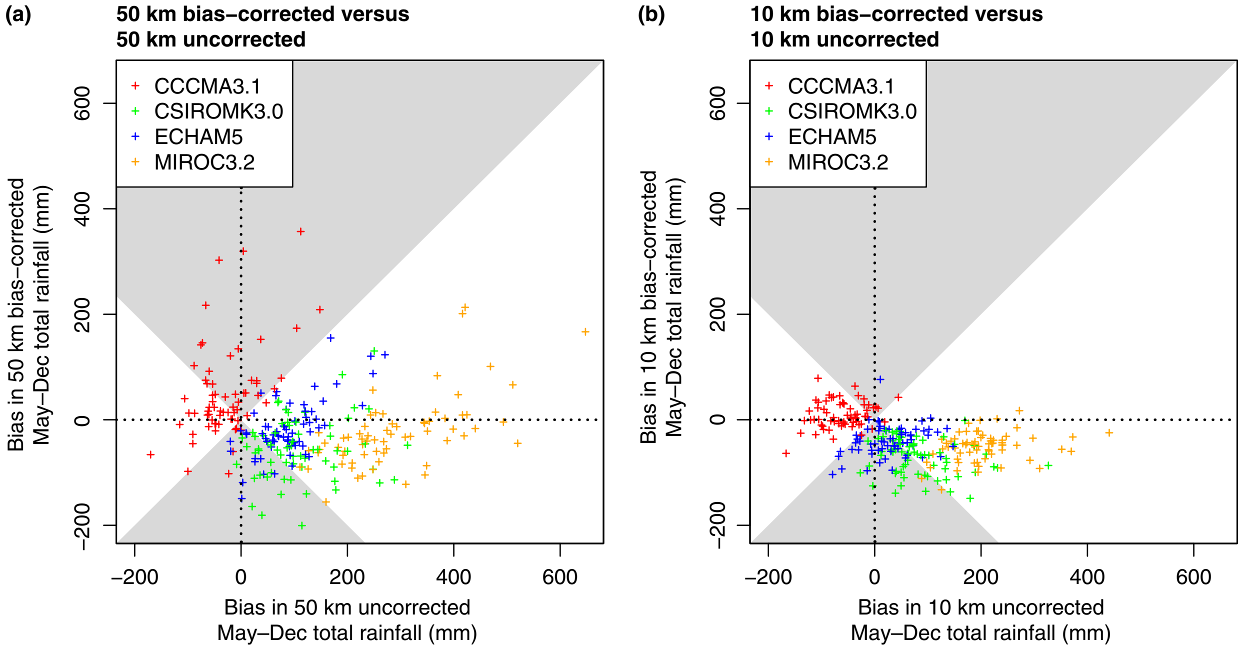

Figure 9. Biases in May–December rainfall totals for bias-corrected NARCliM data plotted against corresponding biases for uncorrected NARCliM data. Data points in the shaded region have larger biases for the bias-corrected data than for the uncorrected data.

bias correction of daily data can increase biases at other timescales.

Figure 10 shows changes in mean May–December rainfall totals between 1990–2009 and 2060–2079 for bias-corrected NARCliM data plotted against corresponding biases for uncorrected NARCliM data. The resemblance between Figures 10 and 7, which show changes in mean May–December rainfall totals and in wheat yields, is less than that between Figures 9 and 6, which show biases in mean May–December rainfall totals and in mean yields. This means that the association between future changes in yields and future changes in May–December rainfall totals is less than between biases in yields and biases in May–December rainfall totals. However, the changes in rainfall do have some influence on the changes in yields. The rainfall changes for both resolutions behave similarly. For all simulations, the sign of rainfall changes is the same between the corrected and uncorrected data for almost all sites and the magnitudes of rainfall changes are somewhat similar. The yield changes are more positive than one would expect if changes in mean May–December rainfall totals were their only driver. Therefore, future changes in atmospheric CO_2 concentrations, which act to increase yields, and/or other aspects of the climate are affecting the future yield changes.

Figure 10 shows that there are some systematic differences in the magnitude of the rainfall changes between the uncorrected and bias-corrected data. The rainfall increases for the CCCMA3.1-forced simulations are larger in the bias-corrected data than in the uncorrected data, while the rainfall increases for the MIROC3.2-forced simulations and the rainfall decreases for the CSIROMK3.0-forced simulations are smaller. It is possible that the reduction in the magnitude of the rainfall decreases for the CSIROMK3.0-forced simulations by the bias correction is the reason for the reduced magnitude of the yield decreases for these simulations for the bias-correction data. This effect is important because it narrows the range of yield changes simulated by the entire RCM ensemble. The modification of future changes in rainfall by the bias correction is a result of the correction of rainfall variability within the 1990–2009 period. The quantile mapping technique does not distinguish between different timescales of variability, so long-term climate trends are also modified by the method. Others (e.g. Haerter *et al.*, 2011; Kempel *et al.*, 2013; Bennett *et al.*, 2014) have questioned the validity of alteration of long-term climate trends by bias correction. It is unclear whether this feature of the bias-correction technique that we have used is desirable or not. Further work examining links between the reliability of simulated long-term climate changes and biases in the simulation of climate variability on shorter timescales would be needed to establish this.

Figure 8 shows a non-linear relationship between simulated yields and May–December total rainfall. Generally, bias correction has little effect on the simulated changes in mean yield between 1990–2009 and 2060–2079 where, for both the uncorrected and bias-corrected data, both the 1990–2009 and 2060–2079 mean yields are small enough to fall within the linear part of this relationship. Hence, bias correction makes little difference to most of the changes in mean yields for the CSIROMK3.0- and ECHAM5-forced simulations (Figure 7). Generally, the uncorrected data and bias-corrected data from these simulations give modest yields for the 1990–2009 period (Figure 3) and modest changes in yields between 1990–2009 and 2060–2079 (Figure 4). The same situation applies for the CCCMA3.1-forced simulations for the sites

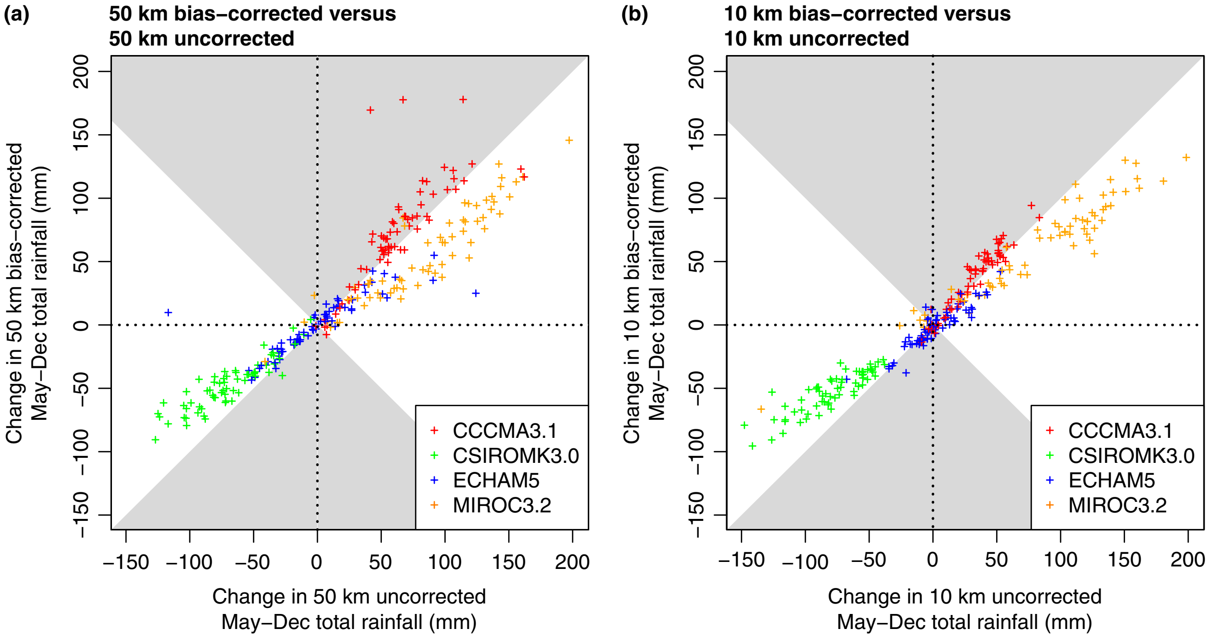

Figure 10. Changes in mean May–December rainfall totals between 1990–2009 and 2060–2079 for bias-corrected NARCliM data plotted against corresponding biases for uncorrected NARCliM data. Data points in the shaded region have larger changes for the bias-corrected data than for the uncorrected data.

in the west of the wheat belt. In contrast, the uncorrected data for the MIROC3.2-forced simulations give high yields for the 1990–2009 period due to positive rainfall biases and the non-linear yield–rainfall regime applies (Figure 8). Despite substantial increases in rainfall (Figure 10), yield changes for the uncorrected MIROC3.2-forced simulations are limited because the yields are relatively insensitive to rainfall changes. Yield changes are larger for the bias-corrected data from these simulations (Figure 4) as the bias correction removes the positive rainfall biases (Figure 9), making the yields more sensitive to rainfall changes. The non-linear yield–rainfall relationship also causes the CCCMA3.1-forced simulations to produce smaller yield changes for sites in the east of the wheat belt when negative rainfall biases are removed by the bias correction (Figures 4 and 9). In this case, however, the non-linear regime applies due to increases in rainfall due to bias correction and simulated future climate changes, rather than positive rainfall biases during the 1990–2009 period. The results for the MIROC3.2- and CCCMA3.1-forced simulations demonstrate the sensitivity of the future yield changes to rainfall biases. Given this, bias correction would seem to be a necessary step in using the NARCliM ensemble to force the APSIM wheat simulations. This is consistent with the views of others (e.g. Glotter *et al.*, 2014; Zhang *et al.*, 2015) who have used climate model output with non-linear agricultural models.

Figure 3 shows that, for both NARCliM RCM resolutions, the wheat yields for the bias-corrected data for the 12 different NARCliM simulations are much more similar to each other than for the corresponding uncorrected data. This is due to the correction of rainfall towards observed values and we would expect similar behaviour for simulated yields for a future period. Furthermore, bias correction can reduce uncertainties in simulated future changes in yields where non-linearities in the yield–rainfall relationship are important. In our study, the correction of rainfall biases in the CCCMA3.1-forced simulations led to a reduction in the largest yield changes produced by the NARCliM ensemble. Hence, bias correction has the potential to reduce the apparent uncertainty in projections of future changes in wheat yields based on agricultural models forced with RCM ensembles. Ehret *et al.* (2012) caution that bias correction can 'hide' uncertainty in climate change impact studies without physical justification. For our case study, we argue that the reduction of uncertainty is valid if the mechanism is indeed the correction of rainfall biases that could result in unrealistic responses of simulated yields to simulated future rainfall changes. However, the reduction of uncertainty may not be valid for climate impacts that are more sensitive to relationships between different climate variables, and other relationships within climate data sets that, as pointed out by Ehret *et al.* (2012), can be made less realistic by bias correction.

If uncorrected RCM output is used to force APSIM, there is no apparent benefit in terms of bias reduction to using the 10 km simulations rather than the 50 km simulations (Figure 5). With bias correction, the largest yield biases for the 10 km simulations are smaller than those of the 50 km simulations (Figure 6) and there is a greater tendency for the bias correction to increase the magnitudes of the yield biases for the 50 km CCCMA3.1-forced simulations than for the corresponding 10 km simulations. Hence, the 10 km simulations are superior to the 50 km simulations only if bias correction is used. The bias correction is slightly more effective at removing yield biases relative to the SILO data for the 10 km simulations than for the 50 km simulations. This is likely because at some sites, the 10 km AWAP data used to correct the 10 km simulations more closely resemble the SILO data than the 50 km AWAP data used to correct the 50 km simulations. This certainly appears to be the case in the eastern region (Figure 3(c) and (d)), which has more complex topography than further westwards in the wheat belt.

4.4. Limitations of the study

This study does not incorporate bias-corrected solar radiation data as these were not available from the NARCliM project. Because solar radiation is related to rainfall, T_{Max} and T_{Min} (e.g. Liu and Scott, 2001), correcting these variables and not solar radiation may reduce the realism of the relationships between variables in the bias-corrected data sets that we have used to force APSIM. Further work could incorporate bias-corrected solar radiation data. Because there is a lack of direct observations of solar radiation, it may be necessary to correct simulated solar radiation towards values derived from satellite products and/or observations of sunshine hours or other variables. This may make the bias correction less reliable than for temperature and rainfall. The additional effect of bias correcting solar radiation data is only likely to be significant in APSIM simulations in which rainfall is plentiful and wheat growth is not always limited by water availability (e.g. simulations for the east of the wheat belt). For these simulations, altering solar radiation values could affect the relative importance of water and energy availability and thus alter the extent of the linear part of the yield–rainfall relationship shown in Figure 8.

Another limitation of this study is that the analysis is confined to mean yields over a 20-year period. The inter-annual variability in yields is important for the impacts of climate change on farming communities (e.g. Luo *et al.*, 2010; Potgieter *et al.*, 2013). It is debatable whether the 20-year data samples that were available to us would be sufficient for a robust analysis of inter-annual variability. Using 21-year data samples, Macadam *et al.* (2014) found few significant differences in the magnitude of yield variability between APSIM simulations forced with climate model output and simulations forced with observations. We do note, however, that yield variability may not be well-represented by GCM-RCM-APSIM setups like ours. In southeast Australia, inter-annual variability in rainfall is greatly affected by large-scale modes of climate variability, such as the El Nino Southern Oscillation (Risbey *et al.*, 2009), that are not well-simulated by GCMs. The RCM simulations are likely to be affected by errors in the representation of these modes in the driving GCM simulations, and such errors are difficult to correct using bias correction.

5. Conclusions

All of the RCM simulations that we examined, 50 and 10 km, are able to simulate the spatial pattern of high- and low-yield sites across the NSW wheat belt. However, the simulations have biases in yields. These are largely due to biases in May–December total rainfall, the largest of which are inherited from the forcing GCM. This is consistent with the results of Macadam *et al.* (2014). Using APSIM-forcing data from a single GCM downscaled with a single RCM, they found that biases in wheat yields in NSW from an APSIM setup similar to ours were largely determined by biases in May–December total rainfall. We find that this conclusion is robust to the use of all 12 GCM-RCM combinations in the NARCliM ensemble. However, this result may not apply to APSIM setups that are more sensitive to other climate variables or rainfall characteristics, such as rainfall intensity.

When a quantile-mapping bias-correction technique is applied to the daily NARCliM temperature and rainfall output, the largest positive wheat yield biases are reduced because the underlying biases in growing season rainfall are reduced. However, negative biases in growing season rainfall and, therefore, yield, remain after bias correction, and some smaller biases are actually increased. This could be due to the fitting of a theoretical cumulative probability distribution function to the daily data as part of the bias-correction procedure and/or the focus of the procedure on correcting biases in daily, rather than seasonal, data. Nonetheless, we believe that bias correction of the NARCliM data is appropriate for forcing APSIM simulations. This is because we have demonstrated that biases in rainfall can affect future changes in APSIM-simulated yields due to a non-linear relationship between simulated yields and growing season rainfall.

The value of finer resolution climate modelling is expected to be most apparent in areas of complex topography, such as near mountains and coasts (e.g. Evans and McCabe, 2013), and not in relatively flat inland areas, such as the NSW wheat belt. Consistent with this, when we use uncorrected RCM to force our APSIM simulations, there is no apparent benefit to downscaling to 10 km rather than 50 km. If the only purpose of the NARCliM simulations were to force APSIM wheat simulations, this result may appear to call into question the wisdom of incurring the additional computational cost of performing the 10 km simulations. However, the largest yield biases for the 10 km simulations are smaller than for the 50 km simulations when bias correction is used. Because we believe that bias correction is necessary, we conclude that there is additional value in the 10 km simulations for forcing our wheat simulations. It is unclear from this study whether there would be additional value in simulations finer than 10 km, the finest resolution of the NARCliM simulations. Because biases in rainfall are inherited from the forcing GCMs, bias correction would still be necessary at finer resolutions.

Although the APSIM model used by this study is widely used for climate change impact assessments, the study itself is not such an assessment. The focus of the study on climate data sets has necessarily meant that aspects of APSIM setup used are unrealistic. For example, the growth of the simulated wheat is never limited by soil nitrogen content. Nonetheless, this study provides evidence to support the use of bias correction in climate change impact assessments that use RCM output to force biophysical models. Although we have analysed wheat yields from APSIM simulations, our findings suggest that bias correction of RCM output may be necessary to obtain reliable future changes in other outputs of other biophysical models that respond non-linearly to climate inputs. Furthermore, although our study provides an example of where finer RCM resolution can be valuable when bias correction is used, the passing of GCM biases into downscaling RCM simulations suggests that finer RCM resolution will not negate the need for bias correction.

Acknowledgements

We thank Roman Olson of the University of New South Wales Climate Change Research Centre and the ARC Centre of Excellence for Climate System Science for useful discussions on the NARCliM Climatological Atlas (Olson *et al.*, 2014). This work was partly funded by the ARC Centre of Excellence for Climate System Science (grant CE110001028).

Supporting Information

The following supporting information is available as part of the online article:

Figure S1. Maps of mean wheat yields (in t ha^{-1}) for the 1990–2009 period from APSIM simulations forced with uncorrected 50 km resolution NARCliM data. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values for the regions and across all sites are shown in boxes.

Figure S2. Maps of changes in mean wheat yields (in t ha^{-1}) between 1990–2009 and 2060–2079 from APSIM simulations forced with uncorrected 50 km resolution NARCliM data. Blue and red circles denote increases and decreases, respectively. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values of the magnitudes of changes for the regions and across all sites are shown in boxes.

Figure S3. Maps of mean wheat yields (in t ha^{-1}) for the 1990–2009 period from APSIM simulations forced with uncorrected 10 km resolution NARCliM data. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values for the regions and across all sites are shown in boxes.

Figure S4. Maps of changes in mean wheat yields (in t ha^{-1}) between 1990–2009 and 2060–2079 from APSIM simulations forced with uncorrected 10 km resolution NARCliM data. Blue and red circles denote increases and decreases, respectively. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values of the magnitudes of changes for the regions and across all sites are shown in boxes.

Figure S5. Maps of mean wheat yields (in t ha^{-1}) for the 1990–2009 period from APSIM simulations forced with bias-corrected 50 km resolution NARCliM data. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values for the regions and across all sites are shown in boxes.

Figure S6. Maps of changes in mean wheat yields (in t ha^{-1}) between 1990–2009 and 2060–2079 from APSIM simulations forced with bias-corrected 50 km resolution NARCliM data. Blue and red circles denote increases and decreases, respectively. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values of the magnitudes of changes for the regions and across all sites are shown in boxes.

Figure S7. Maps of mean wheat yields (in t ha^{-1}) for the 1990–2009 period from APSIM simulations forced with bias-corrected 10 km resolution NARCliM data. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values for the regions and across all sites are shown in boxes.

Figure S8. Maps of changes in mean wheat yields (in t ha^{-1}) between 1990–2009 and 2060–2079 from APSIM simulations forced with bias-corrected 10 km resolution NARCliM data. Blue and red circles denote increases and decreases, respectively. The black lines show the division of the wheat belt into western (W), central (C) and eastern (E) regions. Mean values of the magnitudes of changes for the regions and across all sites are shown in boxes.

References

APSIM Initiative. 2014. *The APSIM-Wheat Module (7.5 R3008)*. http://www.apsim.info/Documentation/Model,CropandSoil/CropModuleDocumentation/Wheat.aspx (accessed 03 July 2014).

Argüeso D, Evans JP, Fita L. 2013. Precipitation bias correction of very high resolution regional climate models. *Hydrol. Earth Syst. Sci.* **17**: 4379–4388, doi: 10.5194/hess-17-4379-2013.

Bennett JC, Grose MR, Corney SP, White CJ, Holz GK, Katzfey JJ, Post DA, Bindoff NL. 2014. Performance of an empirical bias-correction of a high-resolution climate dataset. *Int. J. Climatol.* **34**: 2189–2204, doi: 10.1002/joc.3830.

Chen C, Wang E, Yu Q. 2010. Modeling wheat and maize productivity as affected by climate variation and irrigation supply in North China Plain. *Agron. J.* **102**: 1037–1049, doi: 10.2134/agronj2009.0505.

Ehret U, Zehe E, Wulfmeyer W, Warrach-Sagi K, Liebert J. 2012. Should we apply bias correction to global and regional climate model data? *Hydrol. Earth Syst. Sci.* **16**: 3391–3404, doi: 10.5194/hess-16-3391-2012.

Ekström M, Grose MR, Whetton PH. 2015. An appraisal of downscaling methods used in climate change research. *WIREs Clim. Change* **6**: 301–319, doi: 10.1002/wcc.339.

Evans JP, Argüeso D. 2014. Guidance on the use of bias corrected data. *NARCliM Technical Note 3*. NARCliM Consortium: Sydney, Australia.

Evans JP, McCabe MF. 2013. Effect of model resolution on a regional climate model simulation over southeast Australia. *Clim. Res.* **56**: 131–145, doi: 10.3354/cr01151.

Evans JP, Ekström M, Ji F. 2012a. Evaluating the performance of a WRF physics ensemble over South-East Australia. *Clim. Dyn.* **39**: 1241–1258, doi: 10.1007/s00382-011-1244-5.

Evans JP, McGregor JL, McGuffie K. 2012b. Future regional climates. In *The Future of the World's Climate*, Henderson-Sellers A, McGuffie K (eds). Elsevier: Waltham, MA, Oxford, UK and Amsterdam, The Netherlands.

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

French RJ, Shultz JE. 1984. Water use efficiency of wheat in a Mediterranean-type environment: II some limitations to efficiency. *Aust. J. Agric. Res.* **35**: 765–775, doi: 10.1071/AR9840765.

Glotter M, Elliott J, McInerney D, Best N, Foster I, Moyer EJ. 2014. Evaluating the utility of dynamical downscaling in agricultural impacts projections. *Proc. Natl. Acad. Sci. USA* **111**: 8776–8781, doi: 10.1073/pnas.1314787111.

Haerter JO, Hagemann S, Moseley C, Piani C. 2011. Climate model bias correction and the role of timescales. *Hydrol. Earth Syst. Sci.* **15**: 1065–1079, doi: 10.5194/hess-15-1065-2011.

Hawkins E, Osborne TM, Ho CK, Challinor AJ. 2013. Calibration and bias correction of climate projections for crop modelling: an idealised case study over Europe. *Agric. For. Meteorol.* **170**: 19–31, doi: 10.1016/j.agrformet.2012.04.007.

Holz GK, Grose MR, Bennett JC, Corney SP, White CJ, Phelan D, Potter K, Kriticos D, Rawnsley R, Parsons D, Lisson S, Gaynor SM, Bindoff NL. 2010. Climate futures for Tasmania: impacts on agriculture. Technical Report. Antarctic Climate and Ecosystems Cooperative Research Centre: Hobart, Australia

Howden SM, Crimp S. 2005. Assessing dangerous climate change impacts on Australia's wheat industry. In *MODSIM 2005 International Congress on Modelling and Simulation*, Zerger A, Argent RM (eds). Modelling and Simulations Society of Australia and New Zealand: Melbourne, Australia.

Howden SM, Reyenga PJ, Meinke H. 1999. Global change impacts on Australian wheat cropping. Report to the Australian Greenhouse Office. CSIRO Division of Wildlife and Ecology: Canberra, Australia.

Jeffrey SJ, Carter JO, Moodie KB, Beswick AR. 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environ. Model. Softw.* **16**: 309–330, doi: 10.1016/S1364-8152(01)00008-1.

Jones DA, Wang W, Fawcett R. 2009. High-quality spatial climate data-sets for Australia. *Aust. Meteorol. Oceanogr. J.* **58**: 233–248.

Keating BA, Meinke H, Probert ME, Huth NI, Hills IG. 2001. NWheat: documentation and performance of a wheat module for APSIM. Tropical Agronomy Memorandum. CSIRO Division of Tropical Agriculture: St Lucia, Australia.

Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, Holzworth D, Huth NI, Hargreaves JNG, Meinke H, Hochman Z, McLean G, Verburg K, Snow V, Dimes JP, Silburn M, Wang E, Brown S, Bristow KL, Asseng S, Chapman S, McCown RL, Freebairn DM, Smith CJ. 2003. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* **18**: 267–288, doi: 10.1016/S1161-0301(02)00108-9.

Kempel S, Frieler K, Warszawski L, Schewe J, Piontek F. 2013. A trend-preserving bias correction – the ISI-MIP approach. *Earth Syst. Dyn.* **4**: 219–236, doi: 10.5194/esd-4-219-2013.

Liu DL, Scott BJ. 2001. Estimation of solar radiation in Australia from rainfall and temperature observations. *Agric. For. Meteorol.* **106**: 41–59, doi: 10.1016/S0168-1923(00)00173-8.

Liu DL, Anwar MR, O'Leary G, Conyers MK. 2014. Managing wheat stubble as an effective approach to sequester soil carbon in a semi-arid environment: spatial modelling. *Geoderma* **214–215**: 50–61, doi: 10.1016/j.geoderma.2013.10.003.

Luo Q, Jones RN, Williams M, Bryan B, Belotti W. 2005. Probabilistic distributions of regional climate change and their application in risk analysis of wheat production. *Clim. Res.* **29**: 41–52, doi: 10.3354/cr029041.

Luo Q, Bellotti W, Williams M, Wang E. 2009. Adaptation to climate change of wheat growing in South Australia: Analysis of management and breeding strategies. *Agric. Ecosyst. Environ.* **129**: 261–267, doi: 10.1016/j.agee.2008.09.010.

Luo Q, Bellotti W, Hayman P, Williams M, Devoil P. 2010. Effects of changes in climatic variability on agricultural production. *Clim. Res.* **42**: 111–117, doi: 10.3354/cr00868.

Macadam I, Pitman AJ, Whetton PH, Liu DL, Evans JP. 2014. The use of uncorrected regional climate model output to force impact models: a case study for wheat simulations. *Clim. Res.* **61**: 215–229, doi: 10.3354/cr01258.

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I. 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* **48**: RG3003, doi: 10.1029/2009RG000314.

Meehl GA, Covey C, Delworth T, Latif M, McAvaney B, Mitchell JFB, Stouffer RJ, Taylor KE. 2007. The WCRP CMIP3 multi-model dataset: a new era in climate change research. *Bull. Am. Meteorol. Soc.* **88**: 1383–1394, doi: 10.1175/BAMS-88-9-1383.

Meinke H, Hammer GL, van Keulen H, Rabbinge R. 1998. Improving wheat simulation capabilities in Australia from a cropping systems perspective. III. The integrated wheat model (I_WHEAT). *Eur. J. Agron.* **8**: 101–116, doi: 10.1016/S1161-0301(97)00015-4.

Nakićenović N, Swart R (eds). 2000. Special Report on Emissions Scenarios. Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK.

Olson R, Evans JP, Argüeso D, Di Luca A. 2014. NARCliM climatological atlas. *NARCliM Technical Note 4*. NARCliM Consortium: Sydney, Australia.

Piani C, Weedon GP, Best M, Gomes SM, Viterbo P, Hagemann S, Haerter JO. 2010. Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol.* **395**: 199–215, doi: 10.1016/j.jhydrol.2010.10.024.

Potgieter A, Meinke H, Doherty A, Sadras VO, Hammer G, Crimp S, Rodriguez D. 2013. Spatial impact of projected changes in rainfall and temperature on wheat yields in Australia. *Clim. Change* **117**: 163–179, doi: 10.1007/s10584-012-0543-0.

Reyenga PJ, Howden SM, Meinke H, McKeon GM. 1999. Modelling global change impacts on wheat cropping in south-east Queensland, Australia. *Environ. Model. Softw.* **14**: 297–306, doi: 10.1016/S1364-8152(98)00081-4.

Risbey JS, Pook MJ, McIntosh P, Wheeler MC, Hendon HH. 2009. On the remote drivers of rainfall variability in Australia. *Mon. Weather Rev.* **137**: 3233–3253, doi: 10.1175/2009MWR2861.1.

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X-Y, Wang W, Powers JG. 2008. A description of the Advanced Research WRF Version 3. Technical Note. NCAR: Boulder, CO.

Wang J, Wang E, Luo Q, Kirby M. 2009. Modelling the sensitivity of wheat growth and water balance to climate change in Southeast Australia. *Clim. Change* **96**: 79–96, doi: 10.1007/s10584-009-9599-x.

Wang J, Wang E, Liu DL. 2011. Modelling the impacts of climate change on wheat yield and field water balance over the Murray-Darling Basin in Australia. *Theor. Appl. Climatol.* **104**: 285–300, doi: 10.1007/s00704-010-0343-2.

Wang B, Chen C, Liu DL, Asseng S, Yu Q, Yang X. 2015. Effects of climate trends and variability on wheat yield variability in eastern Australia. *Clim. Res.* **64**: 173–186, doi: 10.3354/cr01307.

Zhang Y, Zhao Y, Chen S, Guo J, Wang E. 2015. Prediction of maize yield response to climate change with climate and crop model uncertainties. *J. Appl. Meteorol. Climatol.* **54**: 785–794, doi: 10.1175/JAMC-D-14-0147.1.

