Vol. 61: 215–229, 2014
doi: 10.3354/cr01258

CLIMATE RESEARCH
Clim Res

Published October 20

The use of uncorrected regional climate model output to force impact models: a case study for wheat simulations

Ian Macadam[1,*], Andy J. Pitman[1], Penny H. Whetton[2], De Li Liu[3], Jason P. Evans[1]

[1]Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, NSW 2052, Australia

[2]CSIRO Climate Adaptation National Research Flagship, Aspendale, VIC 3195, Australia

[3]NSW Department of Primary Industries, Wagga Wagga Agricultural Institute and EH Graham Centre for Agricultural Innovation (an alliance between NSW Department of Primary Industries and Charles Sturt University), Wagga Wagga, NSW 2650, Australia

ABSTRACT: Computationally-expensive regional climate models (RCM) are increasingly being used to generate local climate data for climate change impact studies. These studies usually process RCM output to remove errors in the simulated climate. However, this paper investigates the suitability of raw output from simulations of a single RCM as input to a biophysical impact model. Our study analyses errors in wheat yields simulated for New South Wales (NSW), Australia by the Agricultural Production Systems Simulator (APSIM) model forced with output from 2 RCM simulations with horizontal resolutions of approximately 50 and 10 km over NSW and with output from the global climate model (GCM) simulation that they downscale. Overall, across the NSW wheat belt, the ~50 km simulation has a better simulation of mean yields for the 1990–2010 period than the GCM simulation, and the ~10 km simulation has a better simulation than the ~50 km simulation. The average mean yield from APSIM simulations forced with observations is 3.5 t ha^{-1}. The average magnitudes of errors in mean yields for the GCM, ~50 km and ~10 km simulations are 1.2, 1.0 and 0.5 t ha^{-1} respectively. We suggest that the improvement in the simulation of mean yields with increasing climate model resolution is largely due to an improvement in the simulation of mean rainfall totals for the growing season. However, for a given value of mean growing season total rainfall, all 3 climate model simulations have a climate that is more conducive to high yields than the observed climate. This difference must be due to errors in other aspects of the simulated climates.

KEY WORDS: Downscaling · Model evaluation · Weather Research and Forecasting · WRF · Agricultural Production Systems Simulator · APSIM · Agricultural model · GCM · New South Wales · Australia

1. INTRODUCTION

Many studies of the potential impacts of climate change (e.g. Luo et al. 2005, Chiew et al. 2009, Ashfaq et al. 2010, Holz et al. 2010, Potgieter et al. 2013) use biophysical impact models, representing, for example, agricultural or hydrological systems. These translate climate data into measures of climate impact, such as crop yields or runoff. They often take meteorological data applicable to a small geographical area or specific locations as input. For the output of an impact model to be a plausible representation of future conditions, the meteorological data used as input must be a plausible representation of future weather at the location under consideration, at least in terms of

*Corresponding author: i.macadam@unsw.edu.au

aspects to which the model is sensitive. Impact studies often source information about future climate changes from global climate model (GCM) simulations. However, it is widely recognised that raw GCM output does not give a sufficiently realistic representation of fine-scale meteorological data to give plausible output from impact models (e.g. Fowler et al. 2007, Chiew et al. 2009, Ashfaq et al. 2010, Maraun et al. 2010, Piani et al. 2010, Haerter et al. 2011,). Indeed, GCMs typically output grids of data values with each value representing average conditions over a grid cell at least 100 km across. Therefore, in impact studies it is common for coarse-resolution GCM output to be 'downscaled' to produce fine-scale data that is then used to force an impact model. Downscaling methods are numerous and diverse (Fowler et al. 2007, Maraun et al. 2010, Evans et al. 2012). They include the use of statistically corrected monthly GCM output in a weather generator (Liu & Zuo 2012), statistical correction of daily GCM output (Piani et al. 2010) and the perturbation of observations of a recent period with simulated future climate changes (e.g. Luo et al. 2005). However, these methods have disadvantages. Statistical correction can break the physical consistency between different aspects of the climate (Piani et al. 2010, Haerter et al. 2011) and may make modifications to simulated future climate changes that are not justified on physical grounds (Haerter et al. 2011). Perturbation methods may not represent all of the simulated future changes in climate to which an impact model may be sensitive (Chiew et al. 2009, Holz et al. 2010). In particular, some perturbation methods do not incorporate simulated changes in inter-annual climate variability (Chiew et al. 2009, Piani et al. 2010). A much more computationally-expensive approach, dynamical downscaling, involving the forcing of fine-resolution regional climate model (RCM) simulations with GCM output (e.g. Ashfaq et al. 2010, Holz et al. 2010), may have the potential to overcome these issues. Considerable resources are devoted to running RCM simulations to underpin assessments of climate change impacts. It is therefore important to assess these simulations.

Impact studies that use RCM simulations with an impact model do not usually force the impact model with the raw output of the RCM simulations. This is because the impact model output will be affected by errors in the simulated climate. It is more usual for RCM output to be statistically corrected before being input to the impact model (e.g. Ashfaq et al. 2010, Holz et al. 2010). As discussed above, statistical correction has disadvantages and it would be desirable not to have to use this approach if RCMs produced a sufficiently good representation of the real climate. This study presents a case study that tests whether this is so for a single RCM and impact model. This is done by examining errors in the output of the impact model when it is forced with RCM output. This is a stringent test, as RCM errors that appear insignificant in an analysis based on climate variables alone can result in significant errors in climate impacts due to the particular sensitivities of an impact model. These may include highly non-linear sensitivities, such as those involving critical thresholds. Furthermore, impact models may be sensitive to interactions between multiple climate variables that are not considered in climate-based assessments of RCM simulation performance.

This study investigates the suitability of uncorrected output from the pair of RCM simulations described by Evans & McCabe (2013) as input to simulations of wheat cropping in New South Wales (NSW), Australia, performed using the Agricultural Production Systems Simulator (APSIM) model (Keating et al. 2003). The RCM simulations have horizontal resolutions of approximately 50 and 10 km. These simulations were compared with the GCM simulation that they downscaled, in terms of their ability to simulate wheat yields for a recent period for 22 sites in the NSW wheat belt (Fig. 1).

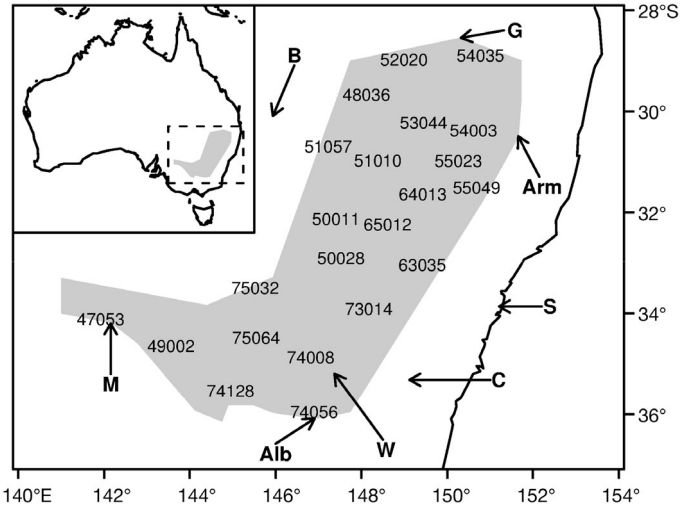

Fig. 1. New South Wales wheat belt (shading) and the Australian Bureau of Meteorology station numbers of the 22 analysis sites used in this study (main figure). Arrows point to the locations of Sydney (S), Canberra (C), Wagga Wagga (W), Albury (Alb), Mildura (M), Bourke (B), Goondiwindi (G) and Armidale (Arm)

2. METHOD

2.1. APSIM

The APSIM modelling framework (Keating et al. 2003) is designed to predict agricultural production by farming systems, and has modules representing crops, soil and farm management. The wheat module (Meinke et al. 1998, Keating et al. 2001, APSIM Initiative 2014) has been widely tested in Australia (Keating et al. 2003, their Table 2) and has been 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 & Crimp 2005, Luo et al. 2005, 2010, Wang et al. 2009, Holz et al. 2010, Potgieter et al. 2013).

APSIM is a numerical model incorporating many empirically-derived equations representing relevant processes, such as those related to crop physiology. It is a point model, and outputs crop yields in units of mass per unit of area. Soil and atmospheric conditions are assumed to be constant over the spatial scale represented by a simulation. Data describing soil characteristics and crop management actions are simulation inputs.

APSIM also requires time series of daily rainfall, maximum air temperature (Tmax), minimum air temperature (Tmin) and solar radiation as input. The provision of these data is commonly termed 'meteorological forcing' or simply 'forcing', irrespective of whether the data are sourced from observations or a climate model. To avoid confusion we use the term 'forcing' in this way.

2.2. Observational data

This study considers a selection of sites in the NSW wheat belt for which data are available from the SILO patched point dataset (Jeffrey et al. 2001), hereafter the 'SILO dataset'. An updated version of the dataset comprises complete time series of 15 meteorological variables for the period since 1889 for approximately 4700 locations across Australia (see Queensland Government 2012). The dataset is based on records from the network of meteorological observing stations maintained by the Australian Bureau of Meteorology. These records contain data gaps, making them unsuitable as a direct input to APSIM. In the SILO dataset these data gaps have been 'patched' using spatial interpolation algorithms to produce complete records that can be used as input to APSIM.

We only considered sites for which the SILO dataset contains relatively little patched data. A total of 22 sites were selected for analysis from the 63 SILO sites in the NSW wheat belt that are based on meteorological records with at least 120 yr of data with no Australian Bureau of Meteorology quality control flags. The 22 sites are distributed reasonably evenly across the wheat belt (Fig. 1). For each site, time series of daily rainfall, Tmax, Tmin and solar radiation for the 1891–2010 period were extracted. For rainfall, Tmax and Tmin, these time series are based on direct measurements. Direct solar radiation measurements are only available for ~25 sites across Australia and so much of the solar radiation data extracted from the SILO dataset were derived from records of sunshine duration and cloud cover (Jeffrey et al. 2001).

2.3. Climate model output

This study examines the RCM simulations described in detail by Evans & McCabe (2013). They used an RCM to downscale data from a simulation of the CSIRO Mk3.5 GCM (Gordon et al. 2002), a coupled atmosphere–ocean GCM that contributed to the World Climate Research Programme (WCRP)'s Coupled Model Intercomparison Project Phase 3 (CMIP3) multi-model dataset (Meehl et al. 2007a). The CSIRO Mk3.5 GCM has a horizontal resolution of ~220 km over Australia. The simulation considered used historical forcing data before 2001 and forcing data consistent with the Intergovernmental Panel on Climate Change (IPCC) SRES A2 scenario for emissions of greenhouse gases and sulphate aerosols (Nakićenović & Swart 2000) after 2001.

Evans & McCabe (2013) performed 2 nested simulations of the Advanced Research WRF (Weather Research and Forecasting) version 3 RCM maintained by the National Center for Atmospheric Research (NCAR) (Skamarock et al. 2008). The first simulation, covering Australia and the surrounding ocean at a horizontal resolution of ~50 km, provided boundary conditions for the second simulation, covering southeast Australia at a horizontal resolution of ~10 km. This RCM setup has been found to perform well for the 1985–2008 period when driven by reanalysis data (Evans & McCabe 2010, Evans & Westra 2012). Evans & McCabe (2013) performed simulations of the 1985–2100 period. The ~50 km simulation was forced with boundary conditions derived from the CSIRO Mk3.5 GCM simulation and the time-varying atmospheric CO_2 concentrations used to force the GCM simulation were used as input to both RCM simulations.

Evans & McCabe (2013) provide full details on the setup of the RCM simulations, including a figure showing the model representation of topography at the resolutions of the RCM, and GCM.

2.4. SILO-forced APSIM simulations

This study examines errors in wheat yields simulated by APSIM when it is forced with climate model output. Since the extent to which APSIM is capable of reproducing wheat yields in the real world is beyond the scope of the study, these errors are not calculated relative to recorded wheat yields but relative to yields generated by forcing APSIM with climate data from the SILO dataset. The issue of how well the SILO dataset resembles the real climate is not addressed.

For each of the 22 analysis sites, an APSIM simulation spanning the 1985–2010 period was performed forced with SILO observational data. An identical model setup was used for all 22 analysis sites with the wheat variety set to Ventura. The soil type was set to 'Brown Clay (Bellata No 121)' from the APSoil database (see Liu et al. 2014), from which the model obtains soil properties. In each simulation year, APSIM was set up to simulate the sowing of the wheat in the austral autumn or early winter (between 1 May and 1 July). The wheat was sown as soon 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. For consistency with the climate model simulations, for all of the APSIM simulations performed, atmospheric CO_2 concentrations were set approximately equal to the multi-model mean mid-range carbon cycle projection for the SRES A2 emissions scenario (Meehl et al. 2007b, their Fig. 10.26). For the 1985–2010 period, these approximate observed concentrations. Their use in the SILO-forced APSIM simulations means that any differences between wheat yields from these simulations and those from APSIM simulations forced with climate model output are not due to differences in CO_2 concentrations, and can be attributed to differences between the meteorological forcing datasets.

Trends and other forms of long-period variability in wheat yields simulated by APSIM can arise from trends and long-period variability in simulated surface organic matter and soil nitrogen content. These can mask, or even be misinterpreted as, effects due to changes in climate. To avoid this problem, the amount of surface organic matter was reset to its initial value, 1 t ha^{-1}, on 1 January of every simulation year. The nitrate and ammonium contents of each soil layer were also reset at the start of every simulation year to levels that ensured that the growth of the wheat was not limited by nitrogen. This is a typical feature of APSIM simulations performed for climate change impact studies (e.g. Luo et al. 2005, 2013) and ensures that simulations reflect variability in the climate forcing data and not variability in soil nitrogen content.

For each simulation, it was not possible to set the initial water content of each soil layer to a value consistent with the meteorological forcing data for the simulation. Therefore, soil water content values were not reset to their initial values at the start of every simulation year but were allowed to evolve from year to year. Data for the 5-yr 1985–1989 period were excluded from the analysis of the output of the simulations 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. The use of a 5-yr spin-up is probably longer than strictly necessary (see Yang et al. 1995) but we wanted to ensure our results were insensitive to initial conditions.

2.5. APSIM simulations forced with climate model output

Time series of daily rainfall, Tmax, Tmin and solar radiation were extracted from the gridded output of each of the 3 climate model simulations for each of the 22 analysis sites for the 1985–2100 period. For each site, the data extracted were data for the nearest climate model gridpoint to the site. For each site, the spatial displacement of the site from the nearest gridpoint may contribute to differences between the extracted climate model data and the SILO data for the site. Interpolating data from multiple nearby gridpoints would have resolved this issue but would have introduced an alternative source of error by generating data for the site with less temporal variability than the data for any of the nearby gridpoints.

The extracted data were used to force an APSIM simulation of the 1985–2100 period for each site for each climate model simulation. Aside from the longer simulation period, the model setup was effectively identical to the 1985–2010 SILO-forced simulations.

2.6. Calculation of errors in simulated wheat yields

Errors in wheat yields from the APSIM simulations forced with climate model output were evaluated by

comparing them with yields from the APSIM simulations forced with SILO data. Yields for the 21 yr 1990–2010 period were used for the comparisons. Data for the 5 yr 1985–1989 period were excluded, as explained in Section 2.4.

Unforced inter-annual variability in the climate system means one would not expect exact correspondence between data output from a climate model for any given year and observational climate data for that year. This lack of exact correspondence would exist even if the climate model and observational datasets were perfect representations of the real climate system. Therefore, one would not expect the wheat yield for a given year from an APSIM simulation forced with climate model data to be the same as the wheat yield from an APSIM simulation forced with SILO data for that year. However, if the comparison period is long enough for the inter-annual climate variability to be well sampled, then 2 perfect representations of the real climate would produce similar distributions of wheat yields. Numerous statistics can be used to summarise distributions but the analysis presented here confines itself to using two. The mean, $\overline{Y}$, of the 21 yield values for the 1990–2010 period, Y_y, is used to characterise average yields for each climate model simulation. The characterisation of the inter-annual variability in yields, a property that is critical to the viability of farms, is based on distributions of yield anomalies, $\acute{Y}_y$, where

$$\acute{Y}_y = Y_y - \overline{Y} \quad (1)$$

The inter-annual variability in yields is characterised by the Mean Absolute Deviation (MAD) of the 21 yield values, where:

$$\mathrm{MAD} = \frac{1}{21} \sum_{y=1990\,\mathrm{to}\,2010} |\acute{Y}_y| \quad (2)$$

The consideration of the inter-annual variability in yields is consistent with recent studies of the potential impact of climate change on crop yields (e.g. Luo et al. 2010, Potgieter et al. 2013), though the use of the MAD statistic is uncommon. The MAD statistic is used here in preference to the standard deviation statistic and derivatives thereof (e.g. the Coefficient of Variability statistic used by Luo et al. 2010). This is because the basis of the standard deviation on squared deviations from the mean make it less robust to outliers and, therefore, to sampling, than the MAD.

The statistical significance of the errors in yields was assessed using the 2-sample Kolmogorov-Smirnov test (Siegel 1956). This non-parametric test assesses whether 2 samples are from the same distribution, and makes no assumptions about the form of this distribution. It was used to compare each 21-value sample of actual yield values from each climate model simulation with yields derived from the SILO dataset. In addition, it was applied to corresponding samples of yield anomalies to assess the significance of errors in inter-annual variability.

2.7. Calculation of errors in simulated growing season rainfall

Rainfall has been widely recognised as the dominant factor in determining the productivity of dryland farming in Australia, including in APSIM-based studies of wheat yields (e.g. Wang et al. 2009, Luo et al. 2013). Wang et al. (2009) noted the sensitivity of wheat yields simulated by APSIM for Wagga Wagga, NSW (Fig. 1) to annual total rainfall. They reported that decreasing rainfall had negative effects on wheat productivity due to reduced crop transpiration and water use efficiency. Rain falling before crop planting can affect crop processes during the growing season due to storage of water in the soil. However, Australia has some of the highest soil evaporation rates in the world (Haverd et al. 2013). APSIM's simulation of wheat yields are more sensitive to rainfall during the growing season, although this is not always true of wheat yield in the real world (see the results of Sadras et al. 2012 based on field measurements from South Australia). To investigate links between the errors in yields derived from the climate model output and errors in simulated rainfall, rainfall totals for the May–December season (the growing season) were calculated for each of the years 1990 to 2010 for each of the APSIM simulations. Mean and MAD values for the rainfall totals for the 1990–2010 period, and errors in these values, were then calculated as for wheat yields. The statistical significance of errors in rainfall totals was also assessed using the method used for wheat yields.

3. RESULTS

3.1. Wheat yield errors

Fig. 2 shows mean and MAD values for wheat yields for the SILO dataset and CSIRO Mk3.5 GCM, ~50 km WRF RCM and ~10 km WRF RCM simulations for the 1990–2010 period. Fig. 3 shows errors in mean and MAD values for yields for the 3 different climate model simulations relative to the values for the SILO dataset. To facilitate the subsequent discus-

sion of errors in simulated yields the wheat belt is divided into 4 regions: southwest (SW), northwest (NW), northeast (NE) and southeast (SE).

Fig. 2a shows that the mean yield averaged across the 22 APSIM simulations forced with SILO data is 3.5 t ha^{-1}. The largest mean yields occur for the 9 sites in the NE and SE regions, and exceed 4 t ha^{-1} for all but 1 of these sites. The smallest mean yields, < 2 t ha^{-1}, are for 3 sites in the SW region. Mean yields vary between 2 and 4 t ha^{-1} for the remaining sites in the SW and NW regions.

Fig. 2b shows that the mean yield averaged across the 22 APSIM simulations forced with GCM data is 2.3 t ha^{-1}, 1.2 t ha^{-1} less than for the SILO-forced simulations. Fig. 3a shows that the APSIM simulations forced with data from the GCM underestimate mean yields for all sites except one. However, with the exception of one site in the SW region, the differences in the distributions of yields derived from the GCM and SILO data are only statistically significant at the 5% level in the NE and SE regions. Errors in mean yields exceed 1 t ha^{-1} for all sites in these regions and 2 t ha^{-1} for 3 sites in the SE region. Errors in mean yields are <1 t ha^{-1} for all sites in the NW and SW regions. A comparison of Figs. 2a and 2b reveals that this pattern of errors for the GCM simulation results from an underestimation of the spatial variability in mean yields.

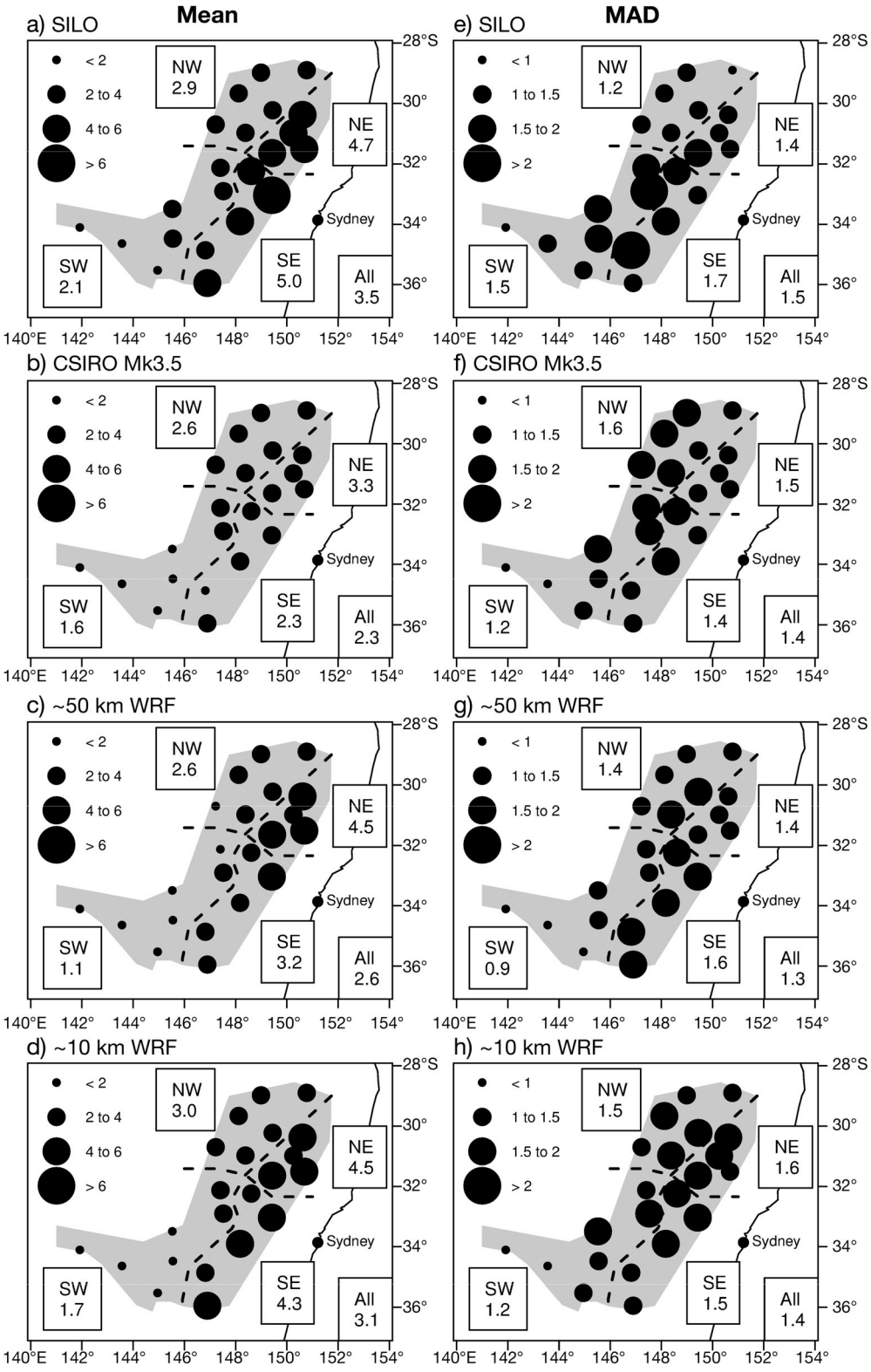

Fig. 2. (a–d) Mean and (e–h) mean absolute deviation (MAD) values (in t ha^{-1}) for wheat yields for the SILO dataset and CSIRO Mk3.5 GCM, ~50 km WRF and ~10 km WRF simulations for the 1990–2010 period. Dashed lines demarcate the 4 regions of the wheat belt considered in this study: southwest (SW), northwest (NW), northeast (NE) and southeast (SE). Mean values for the regions and across all sites are shown in boxes

In common with the GCM simulation, the RCM simulations have negative biases in mean yield across most of the wheat belt (Fig. 3b,c). Fig. 2c,d shows that, whereas the mean yield averaged across the 22 sites is 3.5 t ha^{-1} for the SILO dataset, it is 2.6 and 3.1 t ha^{-1} for the ~50 km and ~10 km WRF simulations, respectively. Fig. 4a,b shows the magnitudes of the errors in mean yields for the ~50 km and ~10 km WRF simulations plotted against corresponding values for the GCM simulation, for sites in the SW, NW, NE and SE regions of the wheat belt. For most sites, the performance of the ~50 km WRF simulation in terms of mean yields is comparable to that of the CSIRO Mk3.5 GCM simulation (Fig. 4a). However, the distributions of yields derived from the 2 simulations are significantly different for 3 of the 4 sites in the NE region and 1 site in the SE region. The errors in mean yields for the ~50 km WRF simulation are smaller than for the GCM simulation for these sites. The overall performance of the ~10 km WRF simulation is superior to that of the GCM and the ~50 km WRF simulation in terms of errors in mean yields. Although the performance of the ~10 km WRF simulation is comparable to that of the GCM simulation in the NW and SW regions (Fig. 4b), the distributions of yields derived from the 2 simulations are significantly different for all but 2 of the sites in the NE and SE regions. The errors in mean yields for the ~10 km WRF simulation are smaller than for the GCM simulation for these sites. This is because the spatial variability in mean yields is well-captured by the ~10 km WRF simulation (Fig. 2d). The result is that the ~10 km WRF simulation negative biases in mean yield are less than 1 t ha^{-1} for all but 1 site (Fig. 3c).

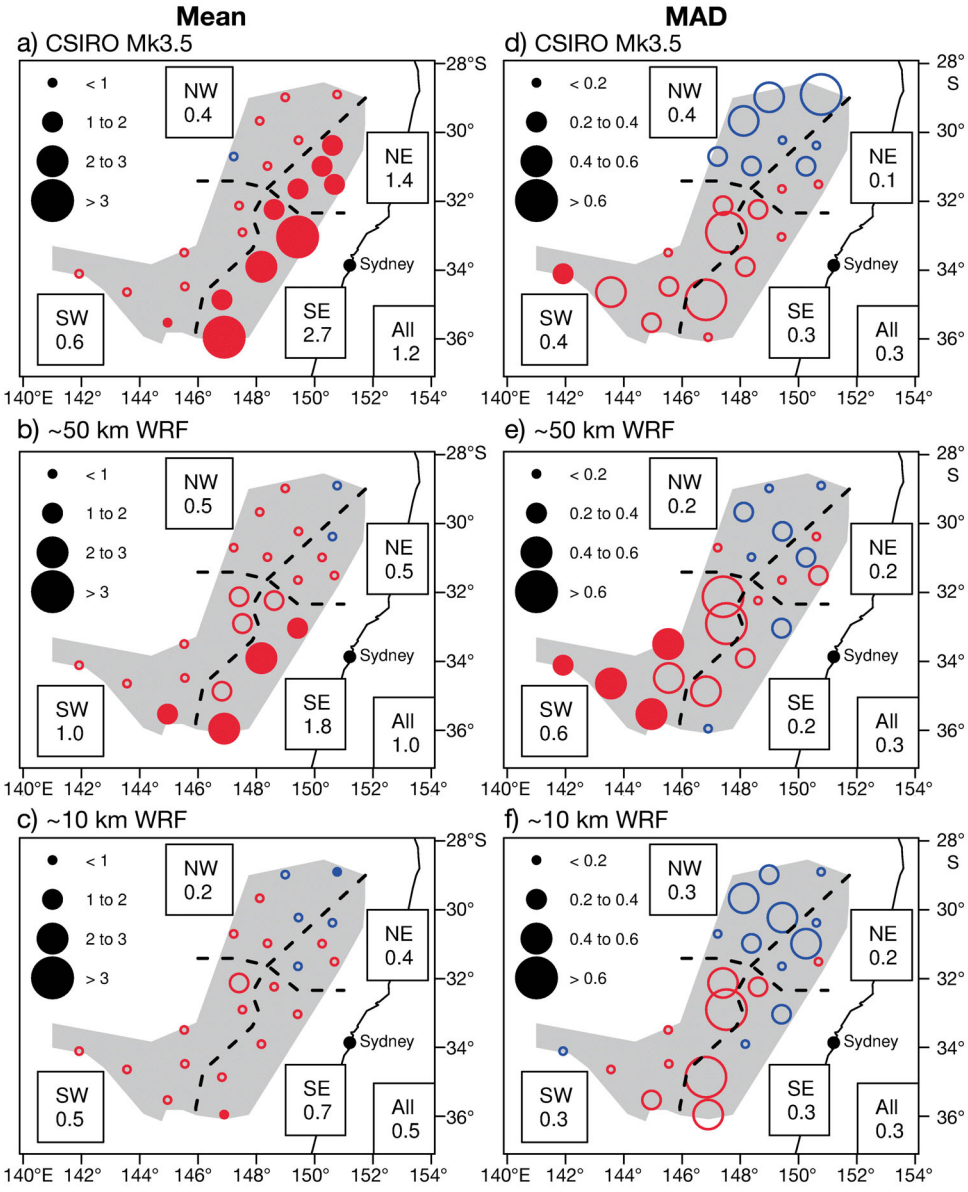

Fig. 3. Errors in (a–c) mean and (d–f) MAD values (in t ha^{-1}) for wheat yields for the CSIRO Mk3.5 GCM, ~50 km WRF and ~10 km WRF simulations for the 1990–2010 period. Red and blue symbols: negative and positive errors, respectively. Filled circles: errors that are significantly different from corresponding SILO-derived distributions at the 5% level. Values in boxes: mean magnitude of errors for the regions and across all sites. Abbreviations as in Fig. 2

Fig. 2e shows that the average MAD value for all sites is 1.5 t ha^{-1} for the SILO dataset. However, there is a tendency for variability to be smaller in the NW and NE regions, in which MAD values exceed 1.5 t ha^{-1} for only 1 site, than in the SW and SE regions, in which 2 sites have MAD values that exceed 2 t ha^{-1}.

Fig. 2f shows that the mean MAD value averaged across the 22 APSIM simulations forced with GCM data is 1.4 t ha^{-1}, similar to the value of 1.5 t ha^{-1} for

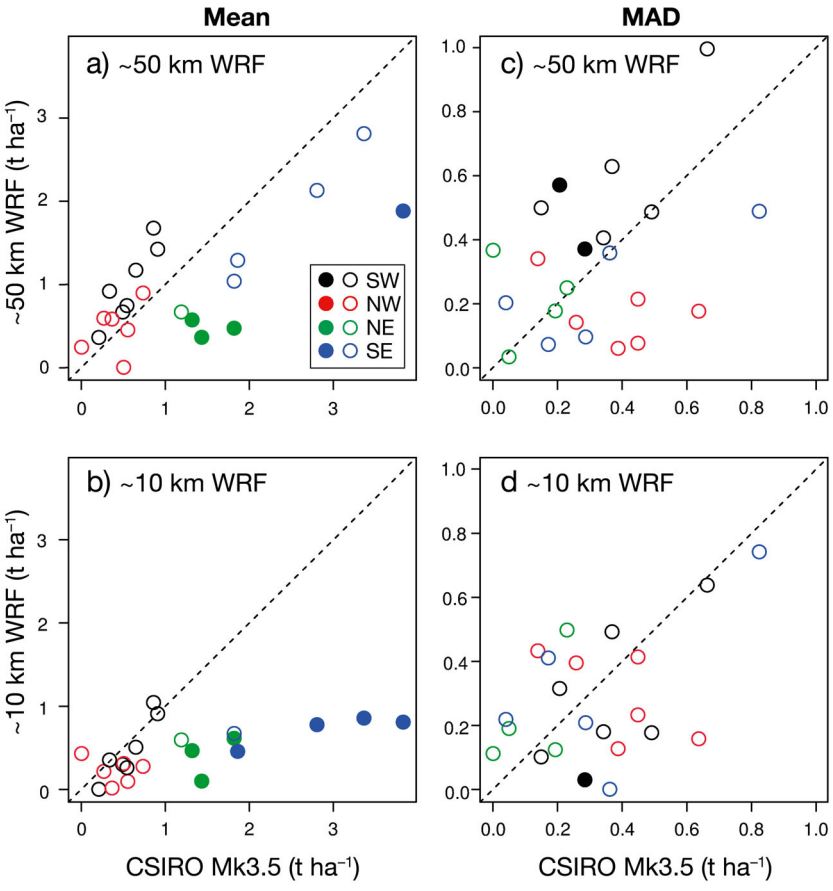

Fig. 4. Magnitudes of errors in (a,b) mean and (c,d) MAD values for wheat yields for the ~50 km WRF and ~10 km WRF simulations for the 1990–2010 period plotted against corresponding values for the CSIRO Mk3.5 GCM simulation. Black, red, green and blue symbols denote data for sites in the SW, NW, NE and SE regions (see Fig. 2) of the wheat belt, respectively. Filled circles: errors that are significantly different from corresponding GCM-derived distributions at the 5% level. Abbreviations as in Fig. 2

the SILO-forced simulations. However, the figure also shows that the spatial variability in MAD values seen in the SILO-derived yield dataset is not well captured by the dataset derived from the GCM simulation. Fig. 3d reveals that this is due to positive errors in the NW region and negative errors in the SW and SE regions. However, except for 1 site in the SW region, these biases do not cause significant differences in yield anomalies between the APSIM simulations forced with GCM data and those forced with SILO data.

Fig. 2g,h shows that, in common with the GCM simulation, the 2 RCM simulations have mean average MAD values that are similar to that of the SILO-derived yield dataset, 1.3 t ha^{-1} and 1.4 t ha^{-1} for the ~50 km and ~10 km WRF simulations, respectively, but fail to reproduce the pattern of spatial variability in MAD values seen for the SILO-derived data. Fig. 4c,d shows the magnitudes of the errors in MAD values for wheat yields for the ~50 km and ~10 km WRF simulations plotted against corresponding values for the GCM simulation. The correlation between errors in MAD values between each of the WRF simulations and the GCM is not as strong as for errors in mean yields. However, overall, across the NSW wheat belt, there is some similarity between the MAD errors for both RCM simulations and the MAD errors for the GCM. Like the MAD errors for the GCM, the MAD errors for the RCMs are generally not great enough to cause significant differences in yield anomalies between the APSIM simulations forced with RCM data and those forced with SILO data (Fig. 3d–f). The results for the ~50 km WRF simulation for the SW region are an exception. This is the only case in which there are significant differences in yield anomalies between the APSIM simulations forced with climate model data and those forced with SILO data for multiple sites. This indicates that the ~50 km simulation has poorer performance than both the ~10 km simulation and the GCM simulation in terms of errors in the inter-annual variability in yields.

3.2. Links between wheat yields and growing season rainfall totals

Fig. 5 shows mean and MAD values for wheat yields for the 1990–2010 period plotted against corresponding values for May–December rainfall totals. Fig. 5a shows an overall association between mean wheat yields and mean May–December rainfall totals, with high-yield sites generally having relatively large mean rainfall totals and low-yield sites having smaller mean rainfall totals. It is notable that the crosses in Fig. 5a, representing data derived from the SILO dataset, are generally located to the right of the other symbols, representing data derived from the climate model simulations. This indicates that a given mean May–December rainfall total corresponds to a greater mean wheat yield in the data derived from the climate model simulations than in

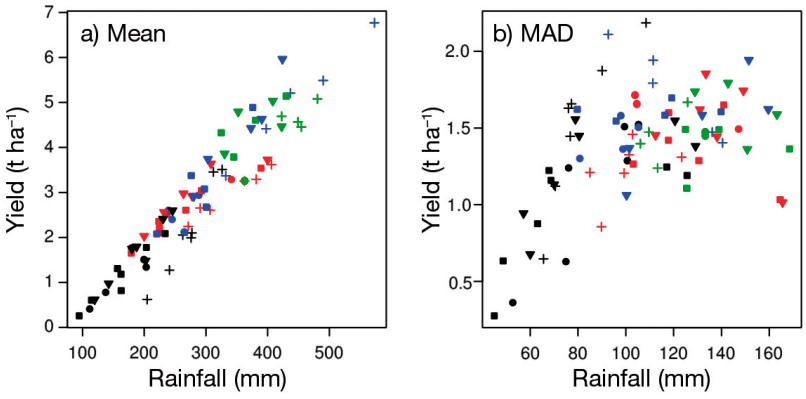

Fig. 5. (a) Mean and (b) mean absolute deviation (MAD) values for wheat yields plotted against mean and MAD values for May–December rainfall totals for the 1990–2010 period, showing data for (+) the SILO dataset, and for the (●) CSIRO Mk3.5 GCM, (■) ~50 km WRF and (▲) ~10 km WRF simulations. Black, red, green and blue symbols: data for sites in the SW, NW, NE and SE regions (see Fig. 2) of the wheat belt, respectively

the SILO-derived data. Possible reasons for this are discussed in the final section of this paper.

There is no obvious overall correspondence between MAD values for May–December rainfall totals and MAD values for mean wheat yield. Fig. 5b, which shows the MAD values for wheat yields for each site plotted against the MAD values for May–December rainfall totals, shows an association between MAD values for wheat yields and MAD values for May–December rainfall totals only for the sites in the SW region, which are denoted by black symbols. In this region, mean yields and mean May–December rainfall totals are relatively small (Fig. 5a). It is therefore likely that MAD values for both of these variables are limited by the sizes of values for individual years, and therefore vary with mean values.

3.3. Errors in growing season rainfall totals

The data for the SILO dataset plotted in Fig. 5a show that observed mean May–December rainfall totals for the 1990–2010 period are between 200 and 500 mm for 21 of the 22 sites. All 3 climate model simulations have negative biases in mean May–December rainfall total for this period for at least 21 of the 22 sites (Fig. 6a–c). The differences between the simulated and observed distribution of rainfall totals are statistically significant at the 5% level for at least 13 sites. The mean magnitude of error averaged across all sites is 106 mm for the GCM simulation (Fig. 6a). However, the underestimation of mean May–December rainfall totals is particularly pronounced in the SE region. Errors exceed 150 mm for 4 out of 5 sites in this region and the mean magnitude of error for sites in the region is almost 200 mm. The underestimation of rainfall in the SE region decreases with increasing climate model resolution (Fig. 6a–c). The mean magnitude of error for sites in the region is 152 mm for the ~50 km WRF simulation and 92 mm for the ~10 km WRF simulation. The distributions of rainfall totals simulated by the GCM and the ~50 km simulation are significantly different from those of the SILO dataset for all sites in the region. However, the distributions simulated by the ~10 km simulation are significantly different from those of the SILO dataset for only 3 of the 5 sites in the region. Considering all sites, the ~10 km WRF simulation has the smallest errors in mean May–December rainfall totals of the 3 climate model simulations (Fig. 6c). The mean magnitude of error averaged across all sites is 78 mm for this simulation. The corresponding value for the ~50 km WRF simulation is 105 mm (Fig. 6b), almost as large as that for the GCM simulation. This is due to the ~50 km WRF simulation having larger errors for the SW region and, to a lesser extent, the NW region than the GCM simulation. Whereas the mean magnitudes of error for the SW and NW regions are 84 and 66 mm respectively for the GCM simulation (Fig. 6a), they are 110 and 76 mm respectively for the ~50 km WRF simulation (Fig. 6b).

Fig. 6d–f reveals that none of the simulated distributions of anomalies in May–December total rainfall are significantly different from the observed distributions of anomalies. This is similar to the corresponding results for wheat yields, though the distributions of yield anomalies for the ~50 km WRF simulation are significantly different from the SILO-derived distributions for 4 sites in the SW region. There are no statistically significant differences between the distributions of rainfall anomalies for either of the RCM simulations and the distributions for the GCM simulation (not shown).

4. DISCUSSION

The APSIM simulations forced with output from all 3 climate model simulations described by Evans & McCabe (2013) have a tendency to underestimate the SILO-derived wheat yields for the 1990–2010 period (Fig. 3a–c). For the GCM simulation, the underestimation of wheat yields is particularly pro-

nounced, and statistically significant, along the eastern edge of the wheat belt. This underestimation along the eastern edge of the wheat belt decreases with increasing climate model resolution. The ~10 km WRF simulation has smaller errors in mean yields for 1990 to 2010 than the ~50 km WRF simulation, which has smaller errors than the ~220 km GCM simulation. The yields for the ~10 km WRF simulation are only significantly different from the SILO-derived yields for 2 sites. In general, differences in the inter-annual variability in yields for the 1990–2010 period between all 3 climate model simulations and the SILO dataset are not statistically significant (Fig. 3d–f).

The general underestimation of the 1990–2010 SILO-derived wheat yields by the climate model simulations is consistent with a general underestimation of May–December rainfall across the NSW wheat belt by the 3 simulations (Fig. 6a–c). For some, but not all, of the sites in the east of the wheat belt, the reduced underestimation of yields with increasing climate model resolution coincides with a reduced underestimation of May–December rainfall. In terms of errors in mean May–December rainfall across the entire wheat belt, only the ~10 km WRF simulation is superior to the CSIRO Mk3.5 GCM simulation, as the ~50 km WRF simulation has errors greater than the GCM simulation for some sites in the west of the wheat belt. However, the effect of these errors on simulated wheat yields is generally not great enough to cause significant differences in yields between the APSIM simulations forced with data from the ~50 km simulation and those forced with SILO data (Fig. 3b).

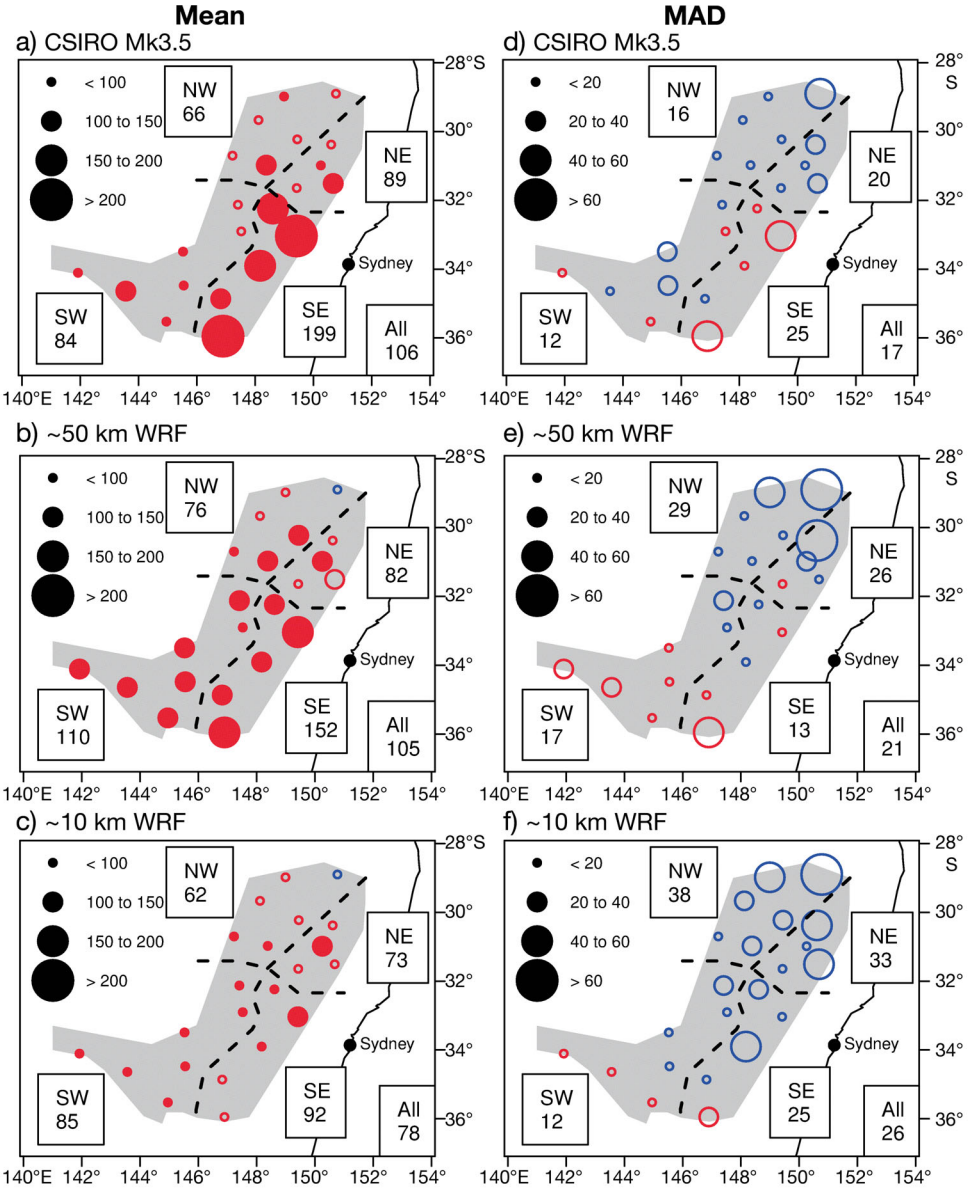

Fig. 6. Errors in (a–c) mean and (d–f) MAD values (in mm) for May–December rainfall totals for the CSIRO Mk3.5 GCM, ~50 km WRF and ~10 km WRF simulations for the 1990–2010 period. Red and blue symbols: negative and positive errors, respectively. Filled circles: errors that are significantly different from corresponding SILO-derived distributions at the 5% level. Values in boxes: mean magnitude of errors for the regions and across all sites. Abbreviations as in Fig. 2

Fig. 5a shows an overall association between mean wheat yields and mean May–December rainfall totals. We suggest that the main reason that the RCM simulations have a better reproduction of the SILO-derived wheat yields in the east of the wheat belt than the GCM simulation is that the RCM simulations have smaller errors in mean values of May–December total rainfall in this region. Similarly, we suggest that the ~10 km WRF simulation has a better reproduction of the SILO-derived yields in the east

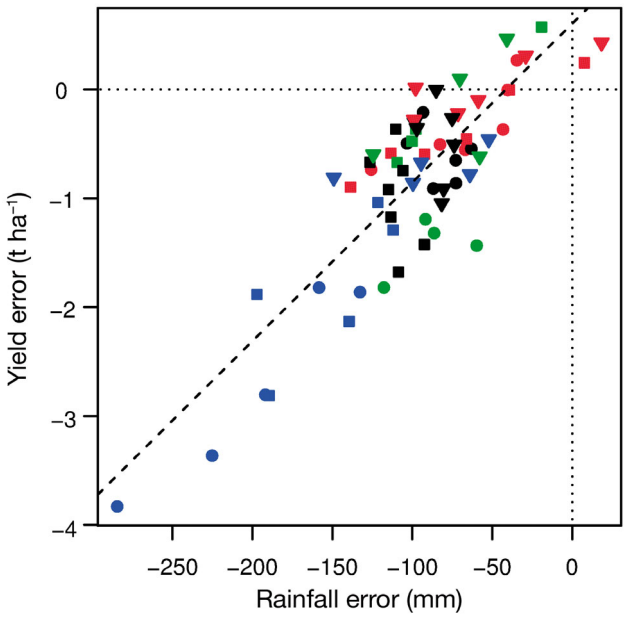

Fig. 7. Errors in mean wheat yields plotted against errors in mean May–December rainfall totals for the 1990–2010 period, showing data for the (●) CSIRO Mk3.5 GCM, (■) ~50 km WRF and (▲) ~10 km WRF simulations. Black, red, green and blue symbols: data for sites in the SW, NW, NE and SE regions (see Fig. 2) of the wheat belt, respectively. Dashed line: best-fit line calculated using ordinary least squares regression. Abbreviations as in Fig. 2

than the ~50 km WRF simulation is because it has smaller errors in mean values of May–December total rainfall in this region. Fig. 7 supports this. It shows an ordinary least squares regression of errors in mean yields against errors in mean May–December rainfall totals. The correlation coefficient between the yield errors and rainfall errors is 0.85.

Fig. 5a shows that a given mean May–December rainfall total corresponds to a greater mean wheat yield in the data derived from the climate models than from the SILO data. The best-fit line in Fig. 7 shows a positive error in mean yields for zero error in mean May–December rainfall total. This indicates that mean yield is not only a function of mean rainfall and that the climate model simulations would be unable to reproduce the SILO-derived mean yields even if they reproduced the SILO mean rainfall data perfectly. Errors in other aspects of the climate are contributing to a systematic error in mean yields. For the 3 climate model simulations, Fig. 8 shows histograms of errors in other aspects of the climate that have the potential to affect the simulated yields.

Wang et al. (2009) investigated the sensitivities of APSIM wheat simulations for Wagga Wagga to meteorological forcing data and noted that yields were sensitive to the intensity of rainfall events. This suggests that climate model errors in the distribution of daily rainfall may be important, with the climate models distributing rainfall in a manner more conducive to growth of the simulated wheat than the SILO dataset. Fig. 8a,b shows that, in general, May–December rainfall simulated by the 3 Evans & McCabe (2013) climate model simulations is more frequent and less intense than in the SILO dataset. This suggests that for the climate model simulations, less water may be lost to the crop through surface runoff or drainage downward through the soil, both of which are favoured by more intense rainfall. However, rainfall that is more frequent and less intense than observed may not favour crop growth in the semi-arid climate of NSW, where a greater proportion of soil moisture is lost through evaporation than runoff and drainage. For NSW, a simulated climate with rainfall that is more frequent and less intense than observed rainfall is likely be less conducive to wheat growth due to increased evaporation of water that might otherwise have infiltrated the soil and moistened the root zone. It is therefore unlikely that errors in the frequency and intensity of rainfall contribute to the excessive wheat yields per millimetre of mean May–December total rainfall shown in Figs. 5a & 7.

APSIM wheat simulations are also sensitive to temperature. APSIM Initiative (2014) describes how temperature affects the simulation of many processes represented by the APSIM wheat module. To highlight the possible effects of some key sensitivities on the errors in simulated wheat yields, we have examined errors in the climate model simulations of some temperature variables. An in-depth analysis of the sensitivities of the APSIM simulations to temperature, and, indeed, other climate variables, is beyond the scope of this paper. Such an analysis would consider APSIM outputs related to individual climate-sensitive processes, such as the efficiency of photosynthesis.

One mechanism by which temperature can affect simulated wheat yields is by affecting the length of the growing season. Fig. 8c shows predominantly warm biases in the mean May–December temperatures simulated by the climate model simulations. These might be expected to decrease yields by shortening the growing season (Wang et al. 2009). Indeed, if the warm biases were not present, the mean yields for a given mean May–December rainfall total for the climate model simulations might be even greater. Although it is possible that the climate model errors in mean May–December temperatures are affecting

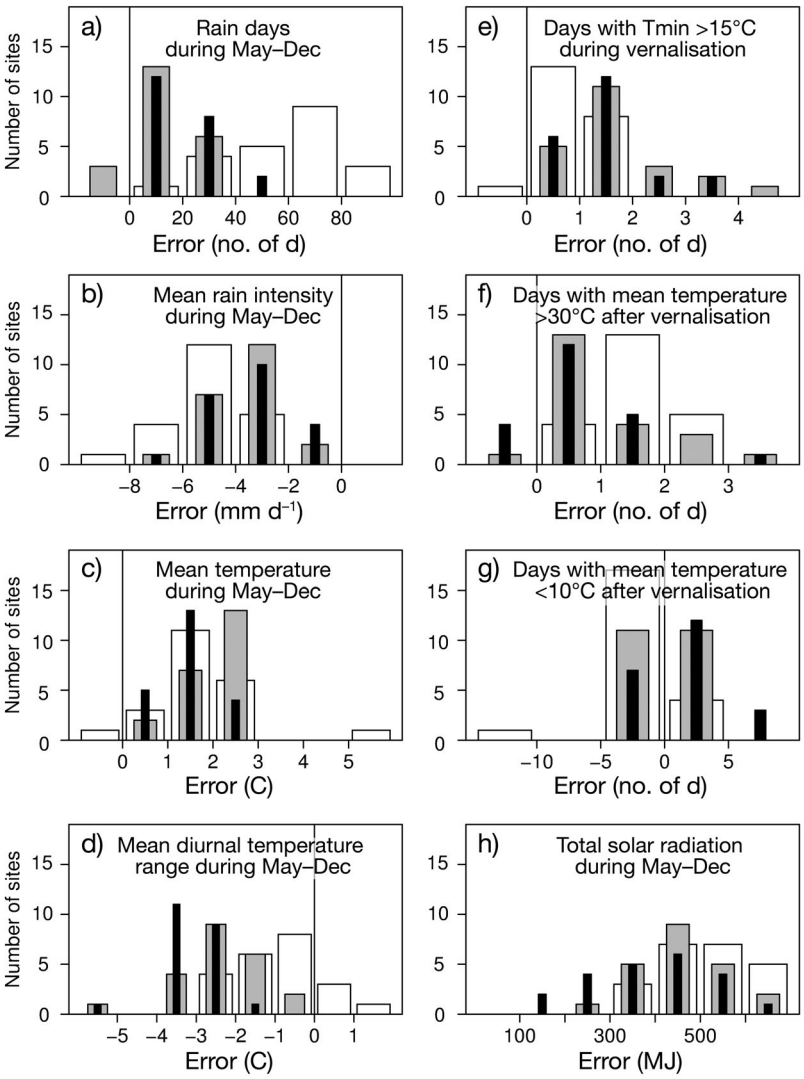

Fig. 8. Errors in mean values of various climate variables for the 1990–2010 period. White, grey and black bars: errors for the CSIRO Mk3.5 GCM, ~50 km WRF and ~10 km WRF simulations, respectively. Rain days are defined as days with non-zero rainfall. Rainfall intensity is defined as the total rainfall divided by the number of rain days

the simulated wheat yields, it is unlikely that the effect of mean temperature on the length of the growing season is responsible for the greater mean yields for a given mean May–December rainfall total for the climate model simulations than for the SILO data.

It is possible that conversion of available soil water to wheat yield is more efficient in the APSIM simulations forced with climate model output due to widespread negative biases in diurnal temperature range (i.e. Tmax minus Tmin) during the May–December season (Fig. 8d). In the APSIM configuration used, decreased values of diurnal temperature range affect yields by decreasing the humidity of the air, which increases the transpiration efficiency of the wheat.

Temperature extremes can also affect the simulated wheat. The sensitivity of the crop to temperature extremes depends on its phenological stage. High temperatures during vernalisation can have a detrimental effect on yields. Indeed, in APSIM, vernalisation will only occur on days on which Tmax is <30°C and Tmin is <15°C. Devernalisation can occur if Tmax exceeds 30°C. Tmax exceeded 30°C during vernalisation in only a few of our APSIM simulations and never on more than 3 d. Therefore, errors in the number of days on which Tmax exceeded 30°C were not calculated, as they would be unlikely to be robust. Fig. 8e shows errors in the mean number of days during vernalisation on which Tmin exceeded 15°C. For all 3 climate model simulations, for almost all sites, more such days are simulated than are observed. These errors should act to decrease yields relative to SILO-derived yields and so cannot be the reason for the excessive wheat yields per millimetre of mean May–December total rainfall for the climate model simulations.

Temperature extremes after the end of vernalisation can also affect the simulated wheat by reducing the rates of grain and root growth and the efficiency of photosynthesis. Fig. 8f shows errors in the mean number of days between vernalisation and the end of grain filling on which the mean temperature exceeded 30°C. On such hot days, root growth and the efficiency of photosynthesis can be reduced by >50%. For all 3 climate model simulations, for most sites, more hot days are simulated than are observed. Like the errors in Tmin extremes during vernalisation, these errors should act to decrease yields relative to SILO-derived yields, and so cannot be the reason for the excessive wheat yields per millimetre of mean May–December total rainfall for the climate model simulations.

Fig. 8g shows errors in the mean number of days between vernalisation and the end of grain filling on which the mean temperature was <10°C. On such cold days, grain and root growth can be reduced by

>50% and photosynthesis can be impaired. The figure shows that the CSIRO Mk3.5 GCM simulation has fewer cold days than the SILO dataset at most sites. For the ~50 km WRF simulation, half of the sites have more cold days than the SILO dataset and half have fewer. The ~10 km WRF simulation has more cold days than the SILO dataset at most sites. Since the errors in the frequency of cold days are different for the different climate model simulations, they cannot be the main reason for the excessive wheat yields per millimetre of mean May–December total rainfall as these apply to all 3 climate model simulations.

It is unlikely that the behaviour of solar radiation is responsible for the greater mean yields for a given mean May–December rainfall total for the climate model simulations than for the SILO data. Fig. 8h shows that all 3 climate model simulations have greater solar radiation values than the SILO dataset for all sites. This could be due to both climate model errors and errors in the SILO data arising from the derivation of solar radiation data from sparse direct measurements and records of sunshine duration and cloud cover (Jeffrey et al. 2001). However, it is likely that the yields derived from the climate model simulations and the SILO-derived yields are water-limited and insensitive to such differences in solar radiation. Evidence for this is the close association between mean wheat yields and mean May– December rainfall totals for each dataset, which holds for even the highest rainfall totals (Fig. 5a).

5. CONCLUSIONS

Many impact studies describe future changes in an impact relative to a recent baseline period. To simulate plausible future changes, an impact model must be forced with data that give a plausible representation of future changes in aspects of the climate to which it is sensitive. If the model has non-linear responses to changes in some aspects of the climate, then an accurate representation of those aspects in the baseline period is also necessary. For example, APSIM may require accurate baseline rainfall data to give plausible future changes in wheat yield as APSIM wheat simulations can have non-linear relationships between yield and rainfall (Wang et al. 2009, their Fig. 1i).

Relative to the use of uncorrected output from the CSIRO-Mk3.5 GCM, the use of uncorrected output from the Evans & McCabe (2013) RCM simulations as input to APSIM simulations representing wheat cropping in NSW results in smaller errors in yields for the 1990–2010 period. The average magnitudes of errors in mean yields for the GCM, ~50 km WRF and ~10 km WRF simulations are 1.2, 1.0 and 0.5 t ha^{-1} respectively. In the absence of compensating climate errors, this indicates a superior simulation of relevant aspects of the recent baseline climate by the RCM simulations, especially the ~10 km simulation. We suggest that the superior simulation of wheat yields by the RCM simulations is likely largely due to the superior simulation of mean values of May–December total rainfall. Further work involving ensembles of different GCMs and RCMs could contribute to establishing the generality of this result. However, if this study were to be repeated with the same APSIM setup but with different RCM simulations forced with different GCM simulations, we would expect that the RCM simulations would result in superior simulations of wheat yields so long as they had superior simulations of mean May–December total rainfall. The importance of mean May–December total rainfall might be less for other APSIM setups in which the simulated wheat may be more sensitive to other aspects of the climate, such as temperature.

As well as leading to a more reliable simulation of the response of yields to future changes in climate, the superior simulation of aspects of the recent baseline climate by the Evans & McCabe (2013) RCM simulations could also imply a more reliable simulation of future climate changes themselves. Further work to establish whether this is so is recommended, as other studies (e.g. Whetton et al. 2007, Giorgi & Coppola 2010) have found only weak links between simulated recent climate conditions and simulated future climate changes.

Acknowledgements. We thank Martix Dix of CSIRO for providing output from the CSIRO Mk3.5 GCM and Malte Meinshausen of the Potsdam Institute for Climate Impact Research for providing atmospheric CO_2 concentrations for the IPCC SRES A2 emissions scenario. Michael Grose of the CSIRO Climate Adaptation National Research Flagship and Annette Hirsch and Melissa Hart of the Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, University of New South Wales, and 3 anonymous reviewers provided comments that improved the paper. This work was partly funded by the ARC Centre of Excellence for Climate System Science (grant CE110001028), the CSIRO Climate Adaptation National Research Flagship, the NSW Environment Trust (RM08603) and the NSW State Leveraging Fund.

LITERATURE CITED

APSIM Initiative (2014) The APSIM-Wheat Module (7.5 R3008). www.apsim.info/Documentation/Model,Cropand

Soil/CropModuleDocumentation/Wheat.aspx (accessed 03 Jul 2014)
Ashfaq M, Bowling LC, Cherkauer K, Pal JS, Diffenbaugh NS (2010) Influence of climate model biases and daily scale temperature and precipitation events on hydrological impacts assessment: a case study of the United States. J Geophys Res 115:D14116, doi:10.1029/2009JD012965
Chiew FHS, Teng J, Vaze J, Post DA, Perraud JM, Kirono DGC, Viney NR (2009) Estimating climate change impact on runoff across southeast Australia: method, results, and implications of the modeling method. Water Resour Res 45:W10414, doi:10.1029/2008WR007338
Evans JP, McCabe MF (2010) Regional climate simulation over Australia's Murray-Darling basin: a multitemporal assessment. J Geophys Res 115:D14114, doi:10.1029/2010JD013816
Evans JP, McCabe MF (2013) Effect of model resolution on a regional climate model simulation over southeast Australia. Clim Res 56:131–145
Evans JP, Westra S (2012) Investigating the mechanisms of diurnal rainfall variability using a Regional Climate Model. J Clim 25:7232–7247
Evans JP, McGregor JL, McGuffie K (2012) Future regional climates. In: Henderson-Sellers A, McGuffie K (eds) The future of the world's climate. Elsevier, Amsterdam
Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int J Climatol 27:1547–1578
Giorgi F, Coppola E (2010) Does the model regional bias affect the projected regional climate change? An analysis of global model projections. Clim Change 100:787–795
Gordon HB, Rotstayn LD, McGregor JL, Dix MR and others (2002) The CSIRO Mk3 climate system model. Technical paper, CSIRO Division of Atmospheric Research, Aspendale
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
Haverd V, Raupach MR, Briggs PR, Canadell JG and others (2013) Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles. Biogeosciences 10:2011–2040
Holz GK, Grose MR, Bennett JC, Corney SP and others (2010) Climate futures for Tasmania: impacts on agriculture. Technical report, Antarctic Climate and Ecosystems Cooperative Research Centre, Hobart
Howden SM, Crimp S (2005) Assessing dangerous climate change impacts on Australia's wheat industry. In: Zerger A, Argent RM (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Canberra, p 170–176
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
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
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
Keating BA, Carberry PS, Hammer GL, Probert ME and others (2003) An overview of APSIM, a model designed for farming systems simulation. Eur J Agron 18:267–288
Liu DL, Zuo H (2012) Statistical downscaling of daily climate variables for climate change impact assessment over New South Wales, Australia. Clim Change 115:629–666
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
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
Luo Q, Wen L, McGregor JL, Timbal B (2013) A comparison of downscaling techniques in the projection of local climate change and wheat yields. Clim Change 120: 249–261
Maraun D, Wetterhall F, Ireson AM, Chandler RE and others (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 and others (2007a) The WCRP CMIP3 multi-model dataset: a new era in climate change research. Bull Am Meteorol Soc 88: 1383–1394
Meehl GA, Stocker TF, Collins WD, Friedlingstein P and others (2007b) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z and others (eds) Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
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
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
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
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
Queensland Government (2012) What data are available from SILO? www.longpaddock.qld.gov.au/silo/data_available.html (accessed 26 Jul 2013)
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
Sadras VO, Lawson C, Hooper P, McDonald GK (2012) Con-

tribution of summer rainfall and nitrogen to the yield and water use efficiency of wheat in Mediterranean-type environments of South Australia. Eur J Agron 36:41–54

Siegel S (1956) Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York, NY

Skamarock WC, Klemp JB, Dudhia J, Gill DO and others (2008) A description of the Advanced Research WRF Version 3. Technical note, National Center for Atmospheric Research, 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

Whetton PH, Macadam I, Bathols JM, O'Grady JG (2007) Assessment of the use of current climate patterns to evaluate regional enhanced greenhouse response patterns of climate models. Geophys Res Lett 34:L14701, doi:10.1029/2007GL030025

Yang ZL, Dickinson RE, Henderson-Sellers A, Pitman AJ (1995) Preliminary study of spin-up processes in land surface models with the first stage of Project for Intercomparison of Land Surface Parameterisation Schemes Phase 1(a). J Geophys Res 100:16553–16578

Editorial responsibility: Gerrit Hoogenboom, Prosser, Washington, USA

Submitted: October 10, 2013 Accepted: July 31, 2014
Proofs received from author(s): October 2, 2014