# KNOWN ISSUES IN THE BARRA DATASETS

Last updated: 14 August 2020

Chun-Hsu Su, Nathan Eizenberg, Bureau of Meteorology

## Contents



## 1. Data completeness

| Release date | Data released |
|---|---|
| July 2019 | 1990-2019-Feb (29-year) BARRA-R<br>1990-2019-Feb (29-year) BARRA-AD<br>1990-2019-Feb (29-year) BARRA-PH<br>1990-2019-Feb (29-year) BARRA-SY<br>1990-2019-Feb (29-year) BARRA-TA |
| January 2019 | 2003-2016 (14-year) BARRA-R<br>2006-2016 (11-year) BARRA-AD<br>2007-2016 (10-year) BARRA-PH<br>2007-2015 (8-year) BARRA-SY<br>2007-2016 (10-year) BARRA-TA |
| January 2018 | 2010-2015 (6-year) BARRA-R<br>2010-2015 (6-year) BARRA-SY<br>2010-2015 (6-year) BARRA-TA |

There are some parameters that have partial completeness over the 29-years. For more information see the parameter master list via the webpage.

**Commented [NE1]:** Need to make this link to the webpage version of the parameter master list

## 2. Missing Digital Object identifiers (DOI)

Ignore the DOI found in the netcdf files. DOI will be assigned during the second half of 2019 when the BARRA v1 is completed.

## 3. Field discontinuity

Daily reanalyses are produced by 4 cycles of data assimilation and model forecasts. The 4 forecasts can be stitched together to provide 24 hourly data, but the forecast fields are discontinuous at the crossover between consecutive cycles.

## 4. Grid-point storms

The Unified Model (UM) sometimes produces unrealistic extreme precipitation and exceedingly large vertical wind speed in a single isolated grid-point, known as "grid-point storms". They can also affect other variables, such as temperature and humidity. Grid-point storms affect only a very small percentages of the BARRA data. We advise that these model limitations be considered when using extreme precipitation values and vertical wind speed values in this dataset.

Grid point storms are a model artefact and higher resolution models are more predisposed to it due to model instability from higher numerical sensitivity. Numerical noise during computations can accumulate to trigger a fictitious storm by the explicit grid-scale convection used in BARRA regional (BARRA-R) and downscaling (BARRA-xx) models. The condensation heat release at the saturated grid box leads to a strong uplift (large vertical wind). The model then removes this excess moisture in the column by generating very large precipitation localised at that cell. This is more likely to occur over land in the tropics and sub-tropics during the warm seasons, when there is sufficient warm moisture supply at the surface. This phenomenon can be reduced by running the UM with very short integration time steps, allowing for more frequent model adjustments to numerical noises. However, this is computationally expensive – at present, BARRA-R uses 5-minute time steps, whereas BARRA-xx uses 1-minute time steps. Only when the UM runs crashes, shorter time steps are used, namely 2.5 minutes for BARRA-R and 30 seconds for BARRA-xx. Our choice of grid-box dependent CAPE (convective available potential energy) closure also aims to reduce the model stability and grid-point storms [Zhu & Dietachmayer, 2015].

To demonstrate the frequency of grid point storms in terms of precipitation, BARRA-R, BARRA-SY and BARRA-TA data were analysed based on their total precipitation fields. These fields correspond to the 12-hour accumulation field (named as 'accum_prcp') for BARRA_R and 9-hour accumulation field ('accum_ls_prcp') for BARRA-SY and BARRA-TA. They are chosen for the convenience of being able to extract the data from a single forecast run.

We describe our analysis to estimate the prevalence of the grid-point storms in the data. Two criteria are used to detect a grid point storm in BARRA data and they are described as follows. In a first step, we use ERA-Interim data [Dee et al., 2011] to provide latitude-dependent rainfall threshold values for identifying instances of exceeding high rainfall in BARRA. The daily rainfall accumulation from ERA-Interim between 1990 and 2017 are calculated. For identifying grid-point storms in BARRA-R, the ERA-Interim daily values between 65 and 180 °E in latitude are pooled according to their latitude. The maximum of the ERA-Interim's daily rainfall amounts across the time period and across this longitude domain is used as a threshold. This is illustrated in Figure 1 for BARRA-R model. The maximum of daily amounts is used as there is considerable smoothing in the ERA-Interim precipitation fields due to its coarse resolution.

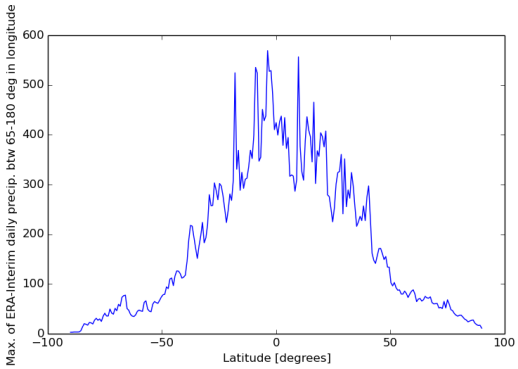


*Figure 1 Maximum daily rainfall (mm) in ERA-Interim during the period 1990-2017. The maximum is taken over the longitude domain from 65 to 180° East.*

After identifying high rainfall instances based on the ERA-Interim-derived thresholds, we analyse the spatial variability of precipitation fields in the vicinity of each of these instances in a second step. In particular, we examine the 12-hour accumulation fields for BARRA-R, and 9-hour accumulation fields for BARRA-x. The spatial extent of each instance analysed is a 9×9 grid area centered at the location of the high-rainfall cell. An instance is regarded as a grid point storm if the sum of the values in the 80 surrounding grid cells is only 2% of the value at the centre grid cell (Figure 2). The size of the grid area (9×9) and the fraction (2%) are chosen subjectively based on testing them against a small training set verified by visual inspections, and should not be considered as optimal values. Increasing the fraction identifies more grid point storms, and vice versa. Further investigations are needed to ascertain the presence of grid-point storms at regions where they persist.

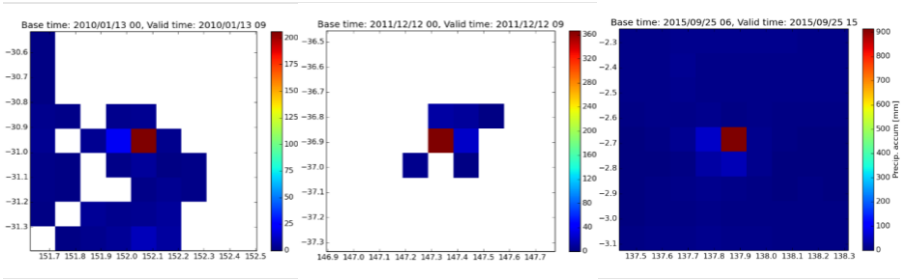


*Figure 2 12-hour rainfall accumulation for three detected grid point storms in BARRA-R (13 January 2010, 12 December 2011, and 25 September 2015).*

Figure 3 shows the counts of identified grid point storms in the 6-year (2010-2015) BARRA-R data. They are identified close to the boundary of the domain and ice sea of Antarctica and are mostly in the tropics between +20 to -20 degrees in latitude and in topographically complex areas of New Zealand. They are mostly concentrated over land points; 99.23% of the identified instances are over land points. We identified 93164 instances, which is 0.001 % of the number of (land and sea) grid cells in 8763 forecast runs over the 6-year period; total number of land and sea grid points for BARRA-R is 768 x 1200 x 8763 = 8075980800. This is 0.01% of the number of land-only grid points. Within the Australian domain (110.8° to 158°E, and 45.6° to 10.5°S), there are 1097 instances, which is about 0.0001% of the total grid cells.

For BARRA-AD, we have identified 40 possible instances of grid point storms, mostly over the ocean, in 25 forecast runs. These are out of 14608 forecast runs over the 10-year period (2006-2015). 14 instances of grid point storms, mostly over the ocean, in 10 forecast runs for BARRA-PH during 2007-2016. 157 instances are identified in 127 forecast runs for BARRA-SY during 2010-2015. Finally, for BARRA-TA, we found 25 instances of grid point storms in 22 forecast runs over the 2007-2016 period. These locations are plotted in Figure 4.

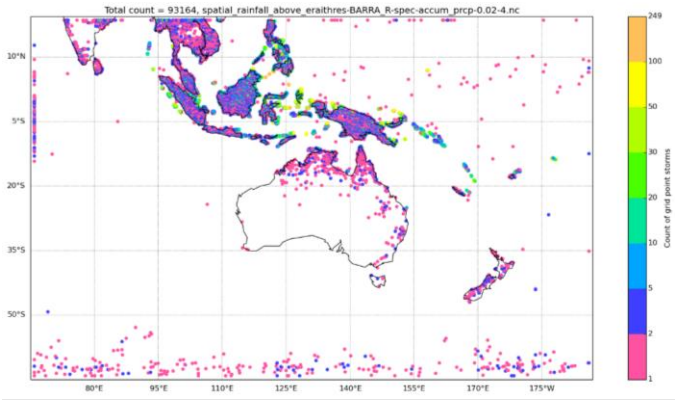


*Figure 3 Counts of "identified" grid point storms in BARRA-R 2010-2015 data. The counts are recorded for each 12 km grid cell.*

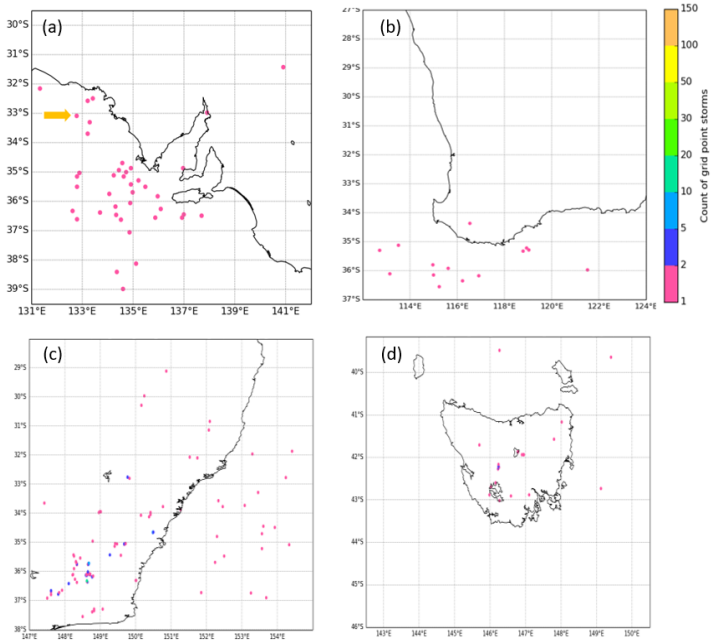


*Figure 4 Counts of identified grid point storms in (a) BARRA-AD 2006-2015, (b) BARRA-PH 2007-2016, (c) BARRA-SY 2010-2015 and (d) BARRA-TA 2007-2016 data. The counts are recorded for each 1.5 km grid cell.*

## 5. Spurious change in screen-level dewpoint temperature

There is a known problem in the screen-level (2 m) dewpoint temperature diagnostic (namely, 'dewpt_scrn') from UM over Australia. This stems from the JULES treatment of evaporation and the assumed humidity of the air in the soil's pores. Most land surface models including JULES assume that the air in the soil pores is saturated (i.e., over-estimated). This leads to screen-level dewpoint temperature to show spurious, abrupt increases. This problem is most striking around 09 to 11 UTC (sunset-early evening), i.e., most apparent in the forecast files produced by 06 UTC cycles. In particular when the surface winds drop in the evening, the aerodynamic wind resistance increases sharply and the term for the humidity of the air in the soil pores dominates the calculation of 2 m specific humidity (qsair_scrn).

The diagnostic can be recalculated using the lowest model-level fields; in particular, the water vapour pressure can be calculated from model-level specific humidity (diagnostic is 'spec_hum' in 'spec' or 'mdl') and air pressure (pressure in 'spec' or 'mdl'), which can in turn be used to approximate dew point temperature via the Clausius-Clapeyron relation. Figure 5 illustrates an example. Note that 'spec_hum' and 'pressure' parameters for the lowest model level are available to users.

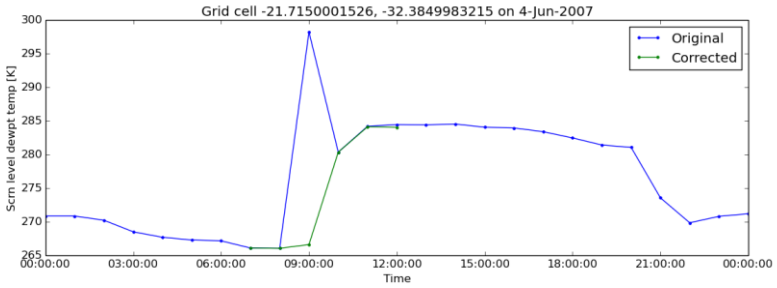


*Figure 5 Timeseries of the screen-level dewpoint temperature diagnostic (dewpt_scrn) at a grid point from BARRA-R, showing a spurious jump in dewpoint. Blue is the original UM diagnostic, and green corresponds to estimates calculated from the lowest model-level fields using the Clausius-Clapeyron relation.*

## 6. Domain boundary artefacts

The BARRA-XX sub domains are nested within BARRA-R, which provides their lateral boundary conditions. The differences in science configuration and spatial resolutions between BARRA-R and the BARRA-XX can provide anomalous dynamics in the boundary of BARRA-XX. These boundary artefacts persist across a forecast run, as hourly BARRA-R fields are used across the entire forecast run. For instance, the boundary of BARRA-TA shows far less precipitation. We advise discarding data near the boundary.

We have examined the discontinuities of the precipitation and 10 m wind speed near the boundary and observed smoothed fields within a margin of 0.34 degrees, along both the latitude and longitude axes. This should be treated as a minimum margin, which is related to the rim width (0.324 degrees) of the lateral boundary data from BARRA-R. The impacts of the boundary can propagate further into the domain.

Figure 6 shows such artefacts in modelled precipitation and screen temperature fields for BARRA-SY. Outside the margin, there is a significant reduction in precipitation and smoothed temperature, wind

speed and dewpoint fields (see figure below). We use the mean daily values across the 2010-2015 period to identify a minimum margin of 0.34 degrees, along both the latitude and longitude axes.

There are also boundary artefacts in the BARRA-R dataset with a minimum margin of 0.888 degrees recommended.

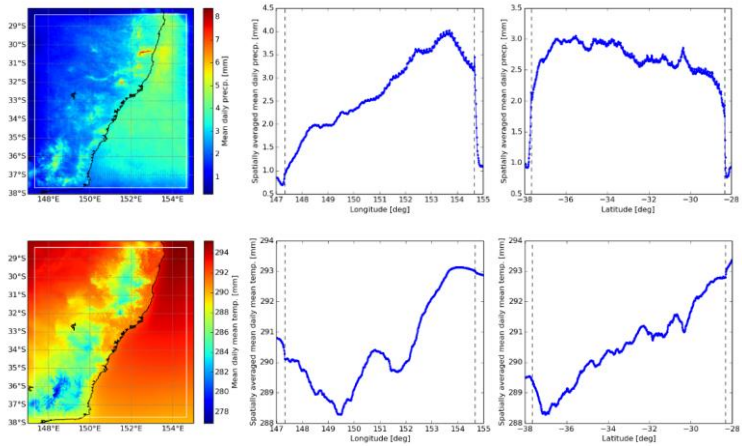


*Figure 6 Boundary artefacts in BARRA-SY modelled precipitation (first row), and modelled screen temperature (second row). The mean daily values over the 2010-2015 period are mapped in the first column. A transect along longitude (second column), and along latitude (third column) is shown to identify a minimum margin of 0.34 degrees (white box and dashed lines) to omit the artefacts.*

## 7. Extreme wind gust spike values in high-resolution datasets

The pre-10.8 versions of the UM with higher resolution (< 60 km) model configurations can produce spurious, extreme wind gust values at singular grid cells. These values present as spatially and temporally isolated gusts of over 100 m/s in both calm and highly unstable conditions over very rough surfaces (Figure 7). Investigations found that the inflated gust values are caused by inconsistencies between the explicit and implicit surface friction velocities used to calculate the diagnosed wind gust (Ma et al. 2018). A fix has been implemented in newer versions of the UM to always use implicitly derived friction velocities which removes the spikes completely.

BARRA-R, TA, SY and PH were produced with UM versions 10.2 and 10.6, so they exhibit some wind gust spikes. The BARRA-AD dataset was produced with a patched model that included the implicit wind gust fix and so does not exhibit this issue. Users are advised to take care when examining extreme values of the 'wndgust10m' parameter including the replacing parameters 'max_wndgust10m' and 'min_wndgust10m' in the BARRA-R, TA, SY and PH datasets. Values exceeding 40 m/s should be inspected and disregarded if they are clearly isolated in space and time.

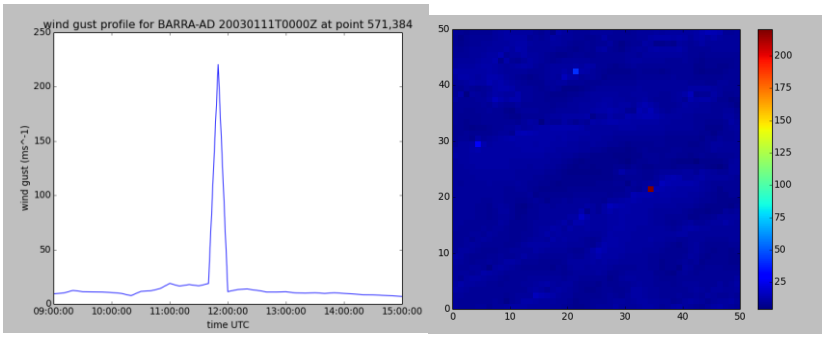


*Figure 7 An example of a wndgust10m spike in an experimental (pre-production) run of the BARRA-AD. Left panel shows the time profile of the wndgust10m over a 6-hour forecast, right shows a map snapshot at the time of the spike.*

A test was run to compare the different wind gust diagnostic schemes showing that the revised method eliminates the erroneous spikes. Figure 8 shows the results from Ma et al. (2018). In an experiment of 42x36 hour forecasts using UM 8.2 at 1.5km horizontal resolution over the Victoria and Tasmania region, the original scheme (denoted by CTL) produces wind gust spikes of 100-300 m/s, which are more frequent at night than during the day. The revised scheme (denoted by TST) had most of the maximum wind gust values between 30-40 m/s and none above 100 m/s.

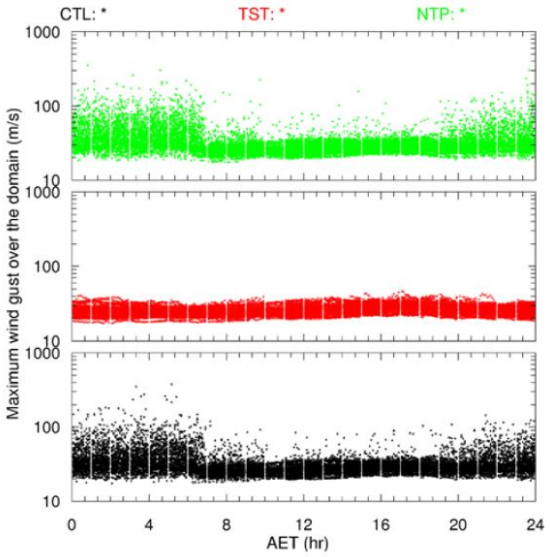


*Figure 8 Statistical verification of wind gust diagnostic scheme fix from Ma et al. 2018. Maximum wind gust values for 42x36 hour forecasts. CTL used the original scheme that uses effective roughness length parameter, NTP uses conventional roughness parameter, and TST refers to the revised scheme.*

We repeat the exercise of Ma et al. (2018) on the BARRA data and calculate the number of instances across the entire model domain, per forecast cycle, where hourly maximum wind gust values at a

grid cell exceed 100 m/s during each 6-hour period. In particular, January-June 2010 data are examined, and the results are shown in Figure 9. All models, except BARRA-AD, show the problem but the number of detected instances is limited to a maximum of 4 occurrences across the (BARRA-PH) model domain in a cycle.

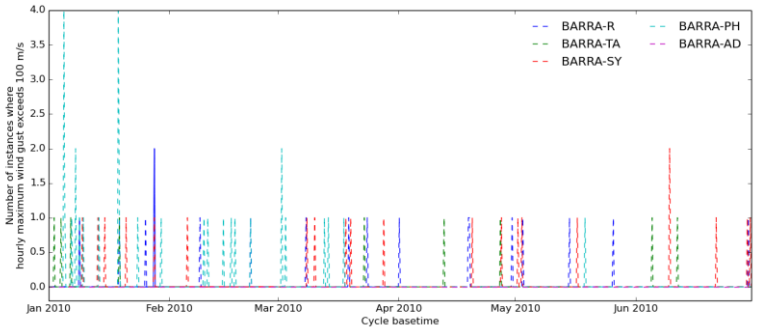


*Figure 9 Number of instances per forecast cycle where hourly maximum wind gust exceeds 100 m/s.*

## 8. Biases in direct and diffused radiation

There exist biases in the UM's direct and diffuse radiation diagnostics (namely 'av_sfc_sw_dif', 'av_sfc_sw_dir' in 'slv'). These stem from the erroneous inclusion of photon scattering contribution in the calculation of direct flux. In particular, the use of delta-Eddington approximation in calculating the direct solar flux leads to over-estimation of the circumsolar radiation (photons being scattered into the region closely surrounding the solar disk by atmospheric agents such as aerosol and cloud particles). The resulting positive bias in direct flux leads to the negative bias in diffused flux. Further, the bias in irradiance is not constant across the irradiance values and can exceed 10%. The problem is described in Sun et al. (2016) and is fixed in the more recent versions of UM ≥ version 11.0 (BARRA-R uses UM version 10.2 while BARRA-xx uses UM version 10.6).

## 9. Potential evapotranspiration on tiles

The potential evapotranspiration (PET) diagnostic (namely, 'tiles_pot_et' in 'spec') is the ET that would occur if the soil and vegetation surfaces are saturated. The PET scales with the friction velocity for each tile type, which depends on the canopy heights etc. Therefore PET is provided for each tile type. The energy constraint is not imposed on the calculation of the diagnostic, which is different from using the Penman-Monteith equation (Harman, personal comm.). In other words, the PET diagnostic is simply driven by a surface to atmosphere humidity difference. Accordingly, it is expected that the PET diagnostic will be higher than the values in other PET products such as SILO PET (https://data.qld.gov.au/dataset/silo-climate-database).

Figure 10 shows that BARRA can provide PET estimates that are commensurate with SILO PET values by calculating PET from the Penman-Monteith equation with BARRA input fields.

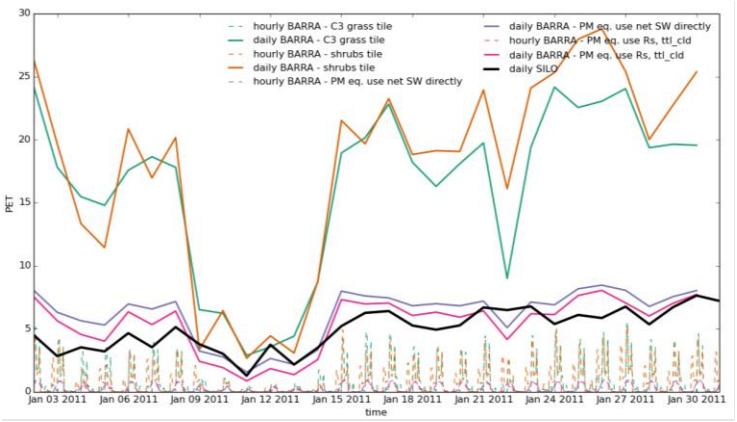


*Figure 10 Timeseries of PET from SILO climate database (black), and BARRA-R near location 35.76° south and 148.90° east. Different calculations of PET from BARRA-R are considered. The UM diagnostic for PET on grass and shrub tiles are shown in green and orange, and the PET calculated from the FAO (Food and Agriculture Organization of the United Nations, fao.org) Penman-Monteith equation using the BARRA-R forcing parameters are shown in purple and pink.*

*We thank Jie Jian (University of Melbourne) and Ian Harman (CSIRO) who provided this insight.*

## 10. Biases in 10 m wind speed

We found that the 10 m wind speed diagnostics ('uwnd10m', 'vwnd10m' in 'slv' and 'spec') are positively biased during low wind conditions and vice versa during strong wind speeds. There are many possible reasons for under-estimating strong winds: the inaccurate descriptions of boundary layer mixing and form drag for sub-grid orography, and of surface properties such as land cover and vegetation types. Changing the fractional area of the vegetation canopy modifies scalar roughness of the vegetated tiles, affecting the wind speed. The seemingly linear variation in wind speed is also known in the global reanalyses (e.g., Carvalho et al., 2014), and Rose and Apt (2016) attributed the problem of wind underestimation to inaccuracy in modelling wind speeds in unstable atmospheric conditions. The comparisons in Figure 11 are reported in Su et al. (2019).

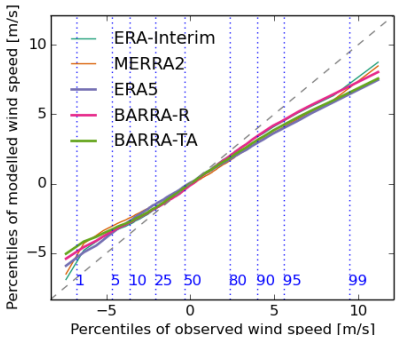


*Figure 11 Comparisons of percentile values between observations and reanalyses for 10 m wind speed during 2010-2013. The values from 1% to 99% percentiles are calculated using values deviated from monthly means. The vertical blue dashed lines indicate the corresponding percentiles of the observations. The comparisons include ERA-Interim, ERA5 and MERRA2 reanalyses.*

## 11. Soil moisture

We found that the BARRA's 4-layer soil moisture (namely 'soil_mois' parameter in 'slv') data set is problematic for several reasons. We advise users to use this data with caution and to refrain from using the soil moisture data in the bottom two layers (depths 0.35 to 3 m).

1. Recall that the daily re-initialization scheme in BARRA-R uses external/offline soil moisture data (see FAQ) to reinitialize the soil moisture state at 06 UTC cycle. We uncovered that the soil moisture state has not been updated consistently between the start dump file used by UM forecast task and the start dump file used by the EKF analysis. This led to analysis increments being added to inappropriate soil moisture states. Consequently, we find a superficial increase in soil moisture for all the 06 UTC cycle runs (Figure 12). The analysis and forecast are performed correctly for the other cycles.
2. The deeper soil moisture model layers takes a long time to spin up. While the soil moisture at the top layers stabilizes after one-month of runs, the bottom layers can take more than 1 year. BARRA-R was produced with many production streams with one-month overlapping runs (for practical reasons), and thus could not fully spin up the bottom soil moisture layers. This is also illustrated in Figure 12.
3. The external/offline soil moisture data from JULES and ACCESS-G are not homogeneous. This leads to discontinuities in soil moisture at changeover times between using different soil moisture data: first changeover on 31 December 2014, and 2 July 2016 for the second. The second changeover between ACCESS-G1 and ACCESS-G2 should be less drastic because both uses the same land surface scheme (MOSES2 in G1 and the newer JULES in G2 are scientifically equivalent) and soil parameter, albeit the two are scientifically distinctive forecast systems. The first changeover at end-2014 is more drastic because there are changes to the soil parameters. Further ACCESS-G (with a soil moisture nudging scheme) aims to improve the NWP forecasts, not to provide accurate soil moisture. Hence, ACCESS-G and offline JULES runs can provide different soil moisture. The discontinuities are found to be most striking for the soil moisture at the bottom two layers.

The impacts of points 2 and 3 in near-surface meteorology, if any, are more likely to be found in forested regions where trees extract water from the deep soil layers. The point 1 could influence the diurnal cycles of screen-level temperature and moisture. However, at this stage, we have not found concrete examples of these impacts.

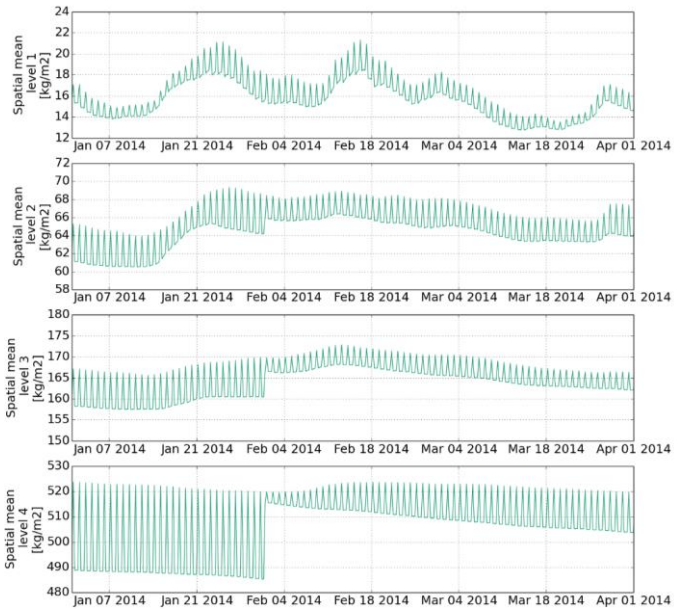


*Figure 12 6-hourly timeseries of soil moisture spatial averages from BARRA-R analysis valid at 00, 06, 12, and 18 UTC. This is plotted separately for each soil layer (top – level 1, bottom – level 4). This illustrates the issues described in point 1 and 2.*

*We thank Ziliang Tan from University of Tasmania for bringing the above point 3 to our attention.*

## 12. Reduced pressure level set available on ma05

Due to space limitations on NCI's Raijin we have *temporarily* reduced the pressure level set for a subset of BARRA-xx pressure level output files available for ma05 members. The reduced set contains 21 unique pressure levels where the full set has 37 levels. The reduced set only affects the pressure level parameters `air_temp`, `geopt_ht`, `relhum`, `wnd_ucmp`, and `wnd_vcmp`. For BARRA-R, users can find the full 37 pressure-level sets from 14 August 2020.

In general, the two BARRA model configurations, BARRA-R and BARRA-xx, produce 'pressure level' or 'prs' fields on a set of 37 isobaric levels. The standard 37 pressure levels as well as the reduced 21 levels are shown below in Table 1. Note that in the reduced set the lowest pressure level is 10hPa (c.f. 0.1hPa in the full set) but that every available level is included between 1000-800hPa.

We plan to have the full set of pressure levels for all parameters and years when more resources have been allocated the project. Limited support is available for making these data available before then. If your research requires access to these data please let us know at helpdesk.reanalysis@bom.gov.au.

*Table 1 BARRA-xx output pressure level set. Atmospheric parameters in the 'prs' stream are available as output on each of the pressure levels. Levels in bold (and asterisked) are included in the reduced 21 level set*

| Pressure level index | Pressure level (hPa) |
|---|---|
| **37*** | **1000** |
| **36*** | **975** |
| **35*** | **950** |
| **34*** | **925** |
| **33*** | **900** |
| **32*** | **850** |
| **31*** | **800** |
| 30 | 750 |
| **29*** | **700** |
| **28*** | **600** |
| **27*** | **500** |
| 26 | 450 |
| **25*** | **400** |
| 24 | 350 |
| **23*** | **300** |
| 22 | 275 |
| **21*** | **250** |
| 20 | 225 |
| **19*** | **200** |
| 18 | 175 |
| **17*** | **150** |
| **16*** | **100** |
| **15*** | **70** |
| **14*** | **50** |
| **13*** | **30** |
| **12*** | **20** |
| **11*** | **10** |
| 10 | 7 |
| 9 | 5 |
| 8 | 3 |
| 7 | 2 |
| 6 | 1 |
| 5 | 0.7 |
| 4 | 0.5 |
| 3 | 0.3 |
| 2 | 0.2 |
| 1 | 0.1 |

## References

Carvalho, D., Rocha, A., Gomez-Gesteira, M., and Santos, C. S.: WRF wind simulation and wind energy production estimates forced by different reanalyses: comparison with observed data for Portugal, Appl. Energy, 117, 116-126, doi: 10.1016/j.apenergy.2013.12.001, 2014.

Dee, D. P., et al., 2011, The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quart. J. Roy. Meteorol. Soc., 137, 553-597.

Ma Y., Dietachmayer G., Steinle P., Lu W., Rikus L. & Sgarbossa D., 2018, Diagnose Wind Gusts from High Resolution NWP Modelling over Mountainous Regions, Bureau Research Report No. 029, April 2018, http://www.bom.gov.au/research/publications/researchreports/BRR-029.pdf

Rose, S., and Apt, J.: Quantifying sources of uncertainty in reanalysis derived wind speed, Renewable Energy, 94, 157-165, doi: 10.1016/j.renene.2016.03.028, 2016.

Su, C.-H., Eizenberg, N., Steinle, P., Jakob, D., Fox-Hughes, P., White, C. J., Rennie, S., Franklin, C., Dharssi, I., and Zhu, H.: BARRA v1.0: the Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis for Australia, Geosci. Model Dev., 12, 2049-2068, https://doi.org/10.5194/gmd-12-2049-2019, 2019.

Sun, Z., Li, J., He, Y., Li, J., Liu, A., and Zhang, F., 2016, Determination of direct normal irradiance including circumsolar radiation in climate/NWP models, Q. J. R. Meteorol. Soc., 142, 2591-2598, doi: 10.1002/qj.2848.

Zhu, H., & Dietachmayer, G., 2015, Improving ACCESS-C convention settings, Bureau Research Report No. 008, December 2015, http://www.bom.gov.au/research/publications/researchreports/BRR-008.pdf