Carly Tozer
PhD Candidate
Environmental and Climate Change Research Group
School of Environmental and Life Sciences
Faculty of Science and Information Technology
University of Newcastle
Callaghan NSW 2308 Australia
T: +61 2 4921 8656 F: +61 2 4921 6925
Email: carly.tozer@newcastle.edu.au

Wednesday 8th February 2012

Dear Prof. Remko Uijlenhoet (Handling Editor for this paper)

Thank you for providing the opportunity to address the comments made by the two anonymous reviewers of the paper "On the uncertainties associated with using gridded rainfall data as a proxy for observed" (Manuscript #: hess-2011-289).

The reviewers both acknowledge that the topic and content of the paper is suitable for publication subject to their questions and comments being addressed. Our responses to both reviewers' comments are included below along with details indicating how the paper has been revised.

We thank the reviewers for their comments which have improved the paper significantly and we hope that the paper is now suitable for publication in HESS.

If you require any further information please contact me.

Thank you
Carly Tozer

Anonymous Reviewer #1

Major Comment 1.1: Almost no information is provided about the spatial and temporal distribution of precipitation amounts within the analysis domain. For instance, the reader may be interested in the comparing the maps of Figure 4 with the maps of average yearly rainfall amounts from the 3 products.

Author Response: A map indicating the average annual rainfall (1900 – 2008) for the AWAP gridded data product has been included as Figure 4(a). AWAP has been included as it was used as the base dataset for the comparisons shown in Figure 4. This will give the reader an indication of the annual rainfall range across South Australia and provide context for the gridded data comparisons.

Major Comment 1.2: How do the authors interpret the systematic differences between datasets ("biases") found in section 4.1? Do the methodologies used to generate the 3 datasets explain these results? Are these differences significant at gauge locations? Would the authors expect the "unexplained microscale variance term" used in AWAP to produce systematic differences with gauge measurements?

Author Response: In regards to the question of *"systematic differences between the datasets"*, throughout the paper we suggest that the differences between the datasets are in fact not systematic. For example, in Section 4.1 (page 8409, lines 3-4) we state that "there does not seem to be any systematic pattern to the disagreement (i.e. the differences appear to be random)". In Section 6 (page 8413, lines 18-21) we state that "Gauges at different elevations and spatial scales were tested at different temporal scales (monthly, annual and seasonal) and the differences between the gridded datasets and between each gridded data set and gauged observations do not appear to be systematic". Hence, based on our analyses, we believe that the *"unexplained microscale variance term"* used in AWAP does not produce systematic differences with gauge measurements.

In regards to how the methods used to produce the datasets explain the differences both between the gridded datasets and between gridded and gauged data, we discuss throughout the paper that:

a) All three gridded datasets have been produced using different interpolation methods, and hence they will vary from each other (see Section 1, Section 2.1).
b) The SILO interpolation method is set to exactly interpolate gauged data and hence this is why it more closely matches gauged data, relative to the AWAP and BOM datasets (see Section 2.1, Section 4.2.2, Section 4.2.3, Section 6).
c) The methods used to develop the BOM and AWAP datasets aim to produce an accurate picture of the areal average and thus are not expected to exactly match gauged data (see Section 2.1, Section 4.2.2, Section 6).

We think that any further comment from us on whether or not the methods used to produce the gridded data sets explain the results is beyond the scope of this paper. Our aim is just to establish if/where/when the various gridded data products agree/disagree with each other and/or with gauged data. Questions on the reasons for any differences observed need to be answered by those responsible for producing the various gridded datasets.

Major Comment 1.3: What is the time-variability of the obtained results? Given the time extent of the 3 rainfall datasets, I encourage the authors to compare the 3 datasets for different periods. It would be of certain interest to relate the differences in gridded rainfall for different periods with the evolution of gauge density and distribution.

Author Response: Additional figures have been added to the paper (Fig. 4d and Fig. 4e) that show the differences between the 3 gridded datasets in different years (i.e. 1900, 1930, 1960 and 1990). These are complemented by Fig. 5a, which indicates the spatial distribution of the rainfall gauges in 1900, 1930, 1960 and 1990 and Fig. 5b, which shows the evolution of the number of rain gauges in SA from 1900 to 2009. Further discussion has also been added to Section 4.1.

Major Comment 1.4: How was the "random location in SA" selected? A more systematic analysis would be necessary in order to generalize any conclusion obtained from the analysis relative to Figure 5. Similarly as for previous comment, this analysis would benefit from information on the gridded rainfall amounts.

Author Response: Several ungauged locations were selected across South Australia. All showed similar results (i.e. the gridded datasets do not agree) therefore only one point was selected to illustrate our point. We understand that a more systematic selection criteria would be required to generalise the findings however this is not the main focus of the paper. The analysis was used as a lead in to the comparison between gauged and gridded rainfall data.

Major Comment 1.5: The hydrological model is calibrated using the observations of a single rain gauge.What would be the effect of calibrating the hydrological model with the gridded rainfall datasets on the hydrological simulations?

Author Response: The model has now also been calibrated to AWAP gridded rainfall (extracted for the grid within which the rainfall gauge is located) to provide a comparison with the gauged rainfall calibration. The results are presented in an updated figure (now Fig. 9) and discussed in Section 5.

Major Comment 1.6: In Figure 8 it seems that in the period 1996-2009 the use of AWAP rainfall inputs produces better flow simulations than using gauged rainfall (which was used for model calibration in the period 1970-1986). In particular, the use of gauge measurements systematically overestimates observed runoff. How do the authors explain such a behavior? Flow simulations with BOM and SILO rainfall inputs should also be included in Figure 8.

Author Response: As requested, the figure (now Fig. 9) has been updated to include flow simulated using SILO rainfall (for the model calibrated using gauged rainfall). BOM data was not available on a daily timestep and thus is not included in the rainfall runoff modelling (this is now mentioned in Section 3.3). Note that flow simulated using AWAP rainfall (for the model calibrated using AWAP rainfall) is also included in the updated figure (see response to Major Comment 1.5).

The point made about AWAP rainfall inputs producing better flow simulations post 1996 than that produced using gauged inputs is a good observation and appears true – especially 1997-1999. We looked into this and found that the rainfall gauge used (23808) actually had a significant change in location in 2002 such that the post-2002 gauged data was not actually consistent with the pre-2002 data that the model was calibrated and validated on. We have revised the Figure (now Fig 9) and caption to reflect this and this partially addresses the comment (i.e. at least for the period post 2002).

In regards to why AWAP inputs perform better than gauged inputs for 1997 – 1999 we have no concrete explanation other than to speculate that this period, which coincided with extreme drought conditions in south east Australia, may have been associated with diversions or extractions within the catchment that were not properly accounted for in the 'naturalised' flow record or were not adequately represented or parameterised in the calibration period. Some new text has been inserted to acknowledge the point made by the reviewer.

Major Comment 1.7: As reported by the authors (page 8404, last paragraph), the SILO dataset is based on an exact interpolation technique. This implies that the SILO dataset exactly reproduces the observed rainfall values at rain gauge locations. This is responsible for the very good agreement between SILO datasets and rain gauge observations presented

in Tables 2, 3 and 4 and Figures 6 and 7. Although the authors state such an extent in several parts of the paper (for example in page 8413, lines 21-25: "SILO is a much better fit to the gauged data but this is to be expected as the method used to develop the SILO database involves a step that directly fits the gridded data to the gauged observation."), little discussion is made beyond the good correspondence between SILO and gauge measurements (for instance in Sections 4.2.1 or 4.2.3). For instance, further discussion on how the authors explain the differences between the 3 gridded datasets could be interesting for the reader.

If possible, I suggest comparison against an independent reference (e.g. rain gauge records not used to produce the gridded datasets). This would allow the authors to assess how the 3 gridded datasets reproduce independent observations. Otherwise, the lack of an independent reference limits the interest of the comparison between gridded datasets and rain gauges records.

Otherwise, given that "in reality, assessing the fit of the AWAP and BOM datasets to a gauged point location is not a fair comparison: : :" I would suggest balancing the paper by enhancing the material and discussion on the comparison of section 4.1 and on the time variability and significance of the results, which would make the paper more interesting (see also major comment 1 and minor comment 8).

Author Response: The Bureau of Meteorology holds all good quality rainfall data in Australia. This data has been used in the production of all three gridded datasets. Independent rain gauges do exist (e.g. data may be collected manually by a farmer in regional SA) but are very difficult to obtain and are of questionable quality, consistency and comparability with the BOM high-quality dataset – as would any results be that are based on this "independent" data. Further, the reviewers comment that "*lack of an independent reference limits the interest of the comparison between gridded datasets and rain gauges records"* is not seen by us as a bad thing – our intent here is to see if/where/when the various gridded data products agree/disagree with each other and/or with gauged data so as to determine the "best" data source for hydroclimatic impact analysis in SA. As such, existing and accessible (and quality controlled) gridded and gauged records are all we are interested in and no independent reference was sought for this study.

As suggested by the reviewer, additional discussion of the time variability of the differences between the gridded datasets has been included in Section 4.1.

Major Comment 1.8: The paper states that "the intention here is to quantify the differences between various gridded data sources, and how they each compare with observed point data, such that these differences can be considered and accounted for in (: : :) studies that utilize gridded data". However, very little is said about the representativeness of point observations to estimate mean areal rainfall (which is, in many cases, the variable of interest) and on how this representativeness relates to the spatial variability of the rainfall field. This is especially relevant in the discussion of the results obtained with the hydrological model in the Finniss basin (193 km2) presented in section 5.

Author Response: We agree that point rainfall data is not necessarily representative of catchment average rainfall but in practice point data is still used in hydrological modelling. Many hydrological models used in Australia use point data as an input (e.g. SIMHYD, IQQM, AWBM) and these models have been recently used in major studies undertaken in Australian catchments. Additional references and discussion around this important point have been included in Section 5.

Minor Comment 1.9: Abstract: Should contain a description of the main findings and conclusions of the paper.

Author Response: An extra sentence has been added – line 20-27, page 1 of revised paper.

Minor Comment 1.10: Caption of Fig. 1. Please, add that the "Random ungauged point" is marked with a green dot. Also, a more visible symbol could help the reader to find the location of such a point in the map.

Author Response: The random ungauged point is now shown as a green star. The figure caption has been updated to reflect this change.

Minor Comment 1.11: Reference to Figure 3 appears in the text before Figure 2. Please, change the numeration.

Author Response: This issue has been addressed.

Minor Comment 1.12: Page 8406, line 5: ": : : the gridded datasets are intended to represent the same observed (or real) situation: : :". Given that observations also suffer from errors, something like "the gridded datasets are intended to represent the real situation" would be more strictly correct.

Author Response: The sentence has been revised as suggested.

Minor Comment 1.13: Page 8407, lines 17-22 and elsewhere in the text. Perhaps something like "grid cell" or "grid unit" would be more appropriated than "grid" to describe one element of the grid.

Author Response: "Grid cell" has been used to describe one element of the grid as suggested.

Minor Comment 1.14: Page 8408, line 18: The notation BOM/AWAP and SILO/AWAP may be confusing.

Author Response: Agreed. Reference to BOM/AWAP or SILO/AWAP or similar has been removed.

Minor Comment 1.15: Figure 4. How do the areas with smaller differences between rainfall datasets match the spatial distribution of rain gauges?

Author Response: This question has been discussed in an updated and extended Section 4.1. Note that Fig. 4d, 4e and 5a have been included to assist with this discussion.

Minor Comment 1.16: Figure 4. The mean annual rainfall (at least for the AWAP product) in the domain would be useful for the reader. Also, how systematic are the results of Fig. 4? It would be interesting to add some Figures with the percentiles (e.g. 15% and 85%) of the relative differences in yearly rainfall to quantify the variability of the errors.

Author Response: A map indicating the average annual rainfall (1900–2008) for the AWAP gridded data product has been included as Fig. 4a. To investigate how systematic the differences between the 3 gridded data products are we have included Fig. 4d and Fig. 4e (also mentioned in our response to Major Comment 1.3). These figures show the differences between the 3 gridded datasets in different years (i.e. 1900, 1930, 1960 and 1990).

We have also looked at the 'tails' of the gridded and gauged datasets (i.e. the low and high rainfall periods) by assessing both the number of 'no rainfall' months and the 99th percentile monthly rainfall recorded for each dataset (see Table 5 and Table 6). Further discussion has consequently been added to Section 4.

Minor Comment 1.17: Section 4.1: the use of the term "error" implicitly assumes that AWAP is more trustable than SILO and BOM. Instead, I would suggest using the term "difference".

Author Response: The term 'error' has only been used in reference to the differences between gridded and gauged datasets, in particular the RMSE between gridded and gauged rainfall. It has not been used in reference to the differences between the AWAP, SILO and BOM gridded datasets.

Minor Comment 1.18: All throughout the text: I suggest emphasizing on the time accumulation windows (monthly vs yearly) for which the analyses are made. For instance, in the last paragraph of page 8408 and figure 5, it should be specified what the results are presented for yearly accumulation products (which is only specified in the title of the y-axis of Figure 5).

Author Response: This request has been fulfilled throughout the paper. For example, the Fig. 5 (now Fig. 6) caption has been updated to clearly indicate that annual rainfall data has been used in the analysis.

Minor Comment 1.19: Similarly as for Figure 4 (see comment 8), the time series of yearly accumulated rainfall in Figure 5 would be useful for the reader. Also, Information about the measured average yearly accumulation is necessary in Table 2.

Author Response: Note that Fig. 5 is now Fig. 6 and includes both the difference in annual rainfall totals between the gridded datasets (Fig. 6a) and the difference between the annual rainfall totals as a percentage of AWAP annual average rainfall (Fig. 6b). Whilst we have not specifically included the annual rainfall totals for each grid, we feel that Fig. 6a provides context to Fig. 6b (formerly Fig. 5). Table 2 (now Table 3) has also been updated to include the annual average rainfall for each gauge.

Minor Comment 1.20: The discussion on the large RMSE values at the gauge at the highest location is rather speculative. Perhaps further justification should be provided.

Author Response: We have included two references to this discussion (see Section 4.2.1) and have reworded the text such that it is less speculative.

Minor Comment 1.21: Page 8409, lines 24-25: "Figure 2 shows the location of the four SA grids investigated in the annual rainfall extremes assessment and the stations within each grid". This sentence may need some rephrasing.

Author Response: This analysis this text was referring to has now been altered and hence this sentence has been removed from the paper.

Minor Comment 1.22: Section 4.2.2: How do the techniques to produce 3 gridded datasets tackle the presence of multiple rain gauges in a grid cell? The discussion in this section should consider this.

Author Response: This is a good question and something we have also considered, however an extensive literature review was undertaken and no specific explanation of how the

interpolation techniques used to create the 3 datasets deal with the presence of more than one rain gauge in a grid cell were provided. As per our response to Major Comment 1.2, questions related to the development of the gridded data sets (not covered in publicly available literature) should be referred to the developers of the data sets.

Minor Comment 1.23: Page 8410, lines 1-9 and Figure 6. The term "events" may be confusing when it refers to yearly accumulations.

Author Response: We have changed Fig 6 (now Fig. 7) and have consequently removed any reference to 'events' in the figure and corresponding discussion.

Minor Comment 1.24: All Figure and Table captions should be more descriptive of all the elements in the Figures and Tables.

Author Response: Agreed. This has been done.

Minor Comment 1.25: Page 8410, line 17-20. NSE results show "that SILO is a better match to gauge data compared to AWAP and BOM (a result consistent with the RMSE analysis): : :". It is assumed that the authors refer to the results presented in section 4.2.1 for annual accumulations. This should be made explicit in the text. Otherwise, it could be confusing given that the NSE is biunivocally related with the RMSE (NSE=1-RMSE2/_2, where _2 is the variance of the reference).

Author Response: The sentence has been updated to read: *"Table 4 shows that SILO is a better match to gauged data compared to AWAP and BOM (a result consistent with the RMSE analysis of annual rainfall totals presented in Section 4.2.1)"*.

Minor Comment 1.26: Section 4.2.4 analyzes the number of zero monthly rainfall values for the different gridded products and compares them with observations. The minimum monthly accumulation detectable with the rain gauges should be reported. Also, is there any threshold applied to gridded rainfall data to distinguish rain from no-rain? I strongly suggest adding the values of gridded rainfall matching the cumulative probability of observed zero accumulations for the different gridded products (that is, for the first row of Table 4, the monthly rainfall for BOM, SILO and AWAP in station 16031 that is not exceeded for 162 months out of the total number of records in this station). This would allow better interpretation of the results.

Author Response: The following sentence has been included: **"**Note that months that have less than 1 mm of rainfall are recorded as having no rainfall (Bureau of Meteorology Climate Services 2011, pers. comm., 9 December). The relevant literature does not indicate a threshold that is applied to gridded data to distinguish rain from no-rain.
Based on the reviewer's comment, we have included two new tables (Table 5 and 6, in place of Table 4) that indicate:

- The number of months with less than 1 mm (i.e. 'no rainfall' months) recorded by each data product.
- The number of months (for each data product) that are greater than the gauged 99th percentile rainfall.
- The total rainfall (in mm) for each dataset that corresponds to the number of gauged 'no rain' months.
- The total rainfall (in mm) for each dataset that corresponds to the number of gauged months greater than the gauged 99th percentile rainfall.

Minor Comment 1.27: The SIMHYD model was calibrated in the Finniss River catchment (193 km2) with monthly records of rainfall and flows. Further description of the model and its calibration is necessary in the paper beyond the provided references. In particular, the type of model, rainfall inputs (aggregated or distributed) or number of parameters should be reported. Also, the fact that the records of a single rain gauge were used for the calibration of the model should be stated more clearly.

Author Response: Additional text describing the model and its calibration has been included in Section 3.3. Note that the model has now also been calibrated to AWAP data. The details of the two calibrations (i.e. using gauged and AWAP data) have been included in Section 3.3.

Minor Comment 1.28: Page 8413, line 17: "The results of this study have shown that the SILO, AWAP and BOM gridded datasets are not an exact match to gauged rainfall". According to the results presented in Tables 2-4 and Figures 6 and 7 SILO datasets match the observations of rain gauges used in the production of SILO remarkably well. Further support to that statement should thus be provided.

Author Response: Line 17, page 8413 has since been changed to read: "*The results of this study have shown that the SILO, AWAP and BOM gridded datasets vary, sometimes significantly, from gauged rainfall datasets, and importantly often do not capture gauged extreme events."*

Although it is a closer match to gauged relative to the AWAP and BOM datasets, throughout the paper we indicate that the SILO dataset does not exactly match the gauged rainfall. For example:

a) In Section 4.2.1 we clearly indicate that the annual RMSE for the SILO dataset at gauge 23736 is greater than 15%.
b) In Section 4.2.2 we indicate that despite being set to exactly interpolate gauged data, SILO cannot match gauged data at all gauges simultaneously for grids that encompass more than one gauge.
c) In Section 4.2.3 and in Table 4 and Fig. 8 (formerly Fig. 7), we indicate that although SILO is a better match to gauged compared with AWAP and BOM, it is not an exact match. This is most obvious at the high elevation gauged, 23736.
d) In Section 4.2.4 and in Tables 5 and 6 we clearly show that (along with AWAP and BOM), SILO does not capture the gauged extremes.

Minor Comment 1.29: Page 8410 line 24: ": : :yet during summer BOM tends to record higher NSE values compared to AWAP". The text could be modified to something like ": : : yet during summer higher NSE values are obtained for BOM products".

Author Response: This sentence has been amended as suggested.

Minor Comment 1.30: Page 8414, lines 14-15: "Although the focus is of Fawcett et al. (2010) was on western Tasmania: : :" should be "Although the focus of Fawcett et al. (2010) was on western Tasmania: : :".

Author Response: This issue has been addressed

Anonymous Reviewer #2

Comment 2.1: A mention of the fundamental difference between the mean areal rainfall of the raster product and the point rainfall of the gauge observation is required in the text. I do not

expect that this is significant at the monthly and annual time scales that are used in the analysis, but it is worthwhile to discuss this in the paper.

Author Response: Agreed. This comment has been addressed in our response to Major Comment 1.8.

Comment 2.2: I missed a more formal evaluation of the probability distribution functions of the three products and the gauge data, a plot of the logarithm of the exceedence probability against the logarithm of rainfall would provide a clear evaluation of the differences between the tails of the four PDFs.

Author Response: We have looked into the 'tails' of the gridded and gauged datasets (i.e. the low and high rainfall periods) by assessing both the number of 'no rainfall' months and the 99th percentile monthly rainfall recorded for each dataset (see revised Table 4 (now Table 5) and new Table 6). Further discussion has also been added to Section 4. See also our response to Minor Comment 1.16.

Comment 2.3: The spatial and temporal structure of the differences between the products provides the key to understanding the impact of these differences in a hydrological model. If the residuals are essentially unbiased white noise then the hydrological model will smooth them out, but if the residuals have significant structure in space and time relative to the catchment scale then these biases will be manifest in the model output. Therefore a more formal evaluation of the spatial and temporal structure of the residuals would greatly assist one in understanding the likely impact of these residuals on a particular catchment scale. Variograms of the temporal and spatial differences would assist in evaluating this structure and a power spectrum of the spatial residuals will help with evaluating the structure at the larger scales.

Author Response: We agree with what the reviewer is suggesting but feel that it is outside the scope of our paper. The hydrological model is used as a simple case study to highlight our point that gridded data is not always a suitable proxy for gauged data and that the 'errors' in gridded data (relative to gauged data) will propagate and magnify during the modelling process.

Comment 2.4: Figure 5 used the percentage difference in annual totals, but perhaps the difference should be used instead since there are likely to be some extremely dry years where a modest absolute difference results in a large percentage difference.

Author Response: Agreed. Fig. 6 (formerly Fig. 5) has been updated to include Fig. 6a, which features the difference between SILO and AWAP and BOM and AWAP annual rainfall totals. This figure complements Fig. 6b, which features the differences between annual rainfall totals as a percentage of AWAP annual average rainfall for SILO and AWAP and BOM and AWAP datasets at the random location in SA.

Comment 2.5: Figure 6 was difficult to read and understand, perhaps a plot more like a box and whisker plot could be used rather than representing the individual gauges. I was impressed at the range in annual totals within a single pixel, perhaps this is worth a comment in the paper. I assume that there are orographic influences at play.

Author Response: As suggested by the reviewer, we have updated the figure (now Fig. 7) to include box and whisker plots of the gridded and gauged data for the four selected grid cells. The corresponding discussion in Section 4.2.2 has been updated and includes a brief comment on the large range in annual rainfall totals present in the one grid cell.

Comment 2.6: I think that the term "pixel" would be better than "grid", eg the last sentence in the discussion.

Author Response: As per response to Minor Comment 1.13 the term 'grid cell' has been included to describe one element of the grid.