

Ensemble one-kilometre forecasts for the South Esk Hydrological Sensor Web

J. Katzfey^a and M. Thatcher^a

^a CSIRO Marine and Atmospheric Research, PMB 1, Aspendale, Victoria 3195 Australia
Email: jack.katzfey@csiro.au

Abstract: An ensemble-based forecast system (60 km, 8 km, and 1 km resolution) using the Conformal Cubic Atmospheric Model (CCAM) has been set up on the Tasmanian Partnership for Advanced Computing (TPAC) cluster to provide probabilistic rainfall forecasts for the South Esk catchment in Tasmania. The 60 km forecasts are for five days, while the 8 km and 1 km forecasts extend for three days. Ensemble forecasting is typically employed to assess various sources of uncertainty (e.g., breeding perturbations for initial conditions). Here, the ensemble members are based upon different initial analyses (generated at different centres), different initial times, alternative downscaling approaches (such as far-field nudging, spectral nudging and using alternative 60 km forecasts), and varying model physics/dynamics settings. This approach to ensemble forecasting is sometimes called a "poor man's" ensemble, although it has the advantage of being reasonably computationally efficient and considers a wider range of sources of forecast uncertainty. The forecasts are then used to provide improved stream flow forecasts for the South Esk River. It is expected that the ensemble-based high-resolution forecasts will have a positive impact on the South Esk River stream flow predictions. Additional improvements in stream flow forecasting may result from merging of the model-based ensemble forecasts with available station data.

The forecasts have now been running for more than a year and are being validated, both with gridded Australian Bureau of Meteorology analyses of rainfall observations and available station data. As the forecast resolution increased, the bias decreased, while the mean absolute error and root mean square error increased. The index of agreement shows little change with resolution. The time correlation is highest for the 8 km forecasts, as is the distribution, where the slope is closest to one. More comprehensive validation of different ensemble members, different lead times and different regions, as well as comparison against station data, is underway.

Keywords: *Ensembles, rainfall, high-resolution predictions, numerical weather prediction.*

1. INTRODUCTION

Validation of high-resolution forecasts is becoming more important as the resolution of atmospheric models becomes finer. It is possible to use data from the observational weather station network to demonstrate similarity of model results with the observed large-scale weather patterns, but often the detail of the observations is insufficient to capture the fine-scale features generated by the models, which may have particular importance to some applications, especially those related to precipitation.

In this study, high-resolution ensemble atmospheric model predictions for Tasmania, Australia are described. These forecasts were generated in order to evaluate the utility of high-resolution forecasts in stream flow forecasting for the South Esk River in northeast Tasmania. The catchment area is shown in Figure 1, along with locations of the rain gauge network. This work is based upon a similar system used to forecast the wind for the America's Cup sailing event (see Katzfey et al., 2008). A description of the model and forecasts are given in the next section. This is followed by some preliminary verification results, with a summary and conclusions in the final section.

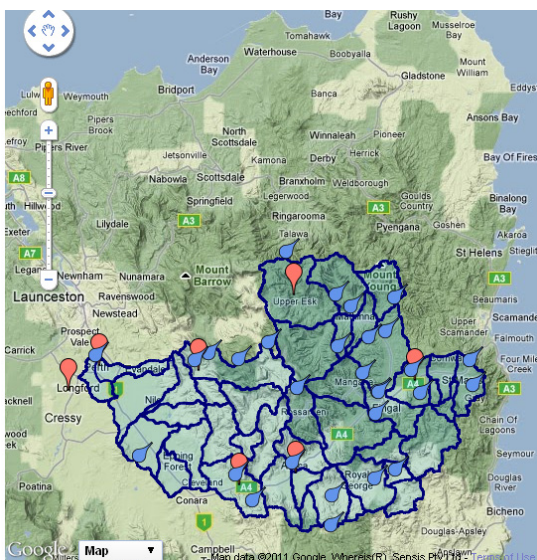


Figure 1: Google map of Esk catchment in northeast Tasmania. Catchment is outlined in dark blue, and rain gauges are light blue markers. Red markers are locations of stream flow measurements.

2. MODELLING AND FORECAST SYSTEM

2.1. Conformal Cubic Atmospheric Model

CSIRO Marine and Atmospheric Research have been undertaking regional climate modelling for nearly two decades. For the past decade the Conformal Cubic Atmospheric Model (CCAM) has been the mainstay of CSIRO dynamical downscaling (McGregor, 1996, 2005a, 2005b; McGregor and Dix, 2001, 2008). CCAM is a full atmospheric GCM formulated on the conformal-cubic grid. CCAM includes a quite comprehensive set of physical parameterizations for climate modelling, including the GFDL parameterizations for long-wave and short-wave radiation (Schwarzkopf and Fels, 1991; Lacis and Hansen, 1974), with interactive cloud distributions determined by the liquid and ice-water scheme of Rotstavn (1997). The model employs a stability-dependent boundary layer scheme based on Monin-Obukhov similarity theory (McGregor et al., 1993). A canopy scheme is included, as described by Kowalczyk et al. (1994), having six layers for soil temperatures, six layers for soil moisture (solving Richard's equation) and three layers for snow. The cumulus convection scheme uses mass-flux closure, as described by McGregor (2003), and includes downdrafts, entrainment and detrainment. CCAM is run in stretched-grid mode utilizing the Schmidt (1977) transformation. The model can downscale to higher resolutions using either far-field nudging or a scale-selective digital filter (Thatcher and McGregor, 2009) to force the coarser resolution information into the finer resolution simulation. With far-field nudging (FFN), the surface pressure and the winds and temperatures above about 850 hPa and away from the high resolution region were nudged to the larger scale forcing data with a one-hour relaxation timescale. The scale-selective digital filter (SSDF) replaces the same fields above a certain scale with the large-scale forcing data (see details of this method in Thatcher and McGregor, 2009).

2.2. Forecast set-up

The forecast system was set up to download two freely available global analyses: one from the U.S. National Centers for Environmental Prediction (GFS) and the other from the Canadian Meteorological Centre (CMC). The GFS analysis is available at 1.0° and 0.5° degree longitude/latitude resolution, while the CMC analysis is available at 0.7° longitude/latitude resolution. The forecasts were initialized by interpolating the analyses to the CCAM variable resolution grid with 60 km resolution over an area of approximately 3000 km by 3000 km. For all except the GFSf run (see discussion later and Table 1), the 60 km forecast was run for 5 days. The 60 km forecast was then interpolated to a CCAM variable-resolution grid with 8 km resolution over an area of approximately 500 km by 500 km, and run for 3 days. Finally, the 8 km forecast was interpolated to a CCAM variable-resolution grid with 1 km resolution over an area of approximately 50 km by 50 km, and run for 3 days. The output domains and topography of the various resolution runs are shown in Figure 1. At 60 km resolution, Tasmania is one hill with highest elevation about 800 m. With 8 km resolution, the main topographical features of Tasmania are evident, including the western ranges, the Tamar Valley running southeastward from northcentral Tasmania, and higher elevations in northeast Tasmania, which is the area of focus of this study. The 1 km resolution domain better resolves the detailed topography of northeastern Tasmania, capturing the highest peak (Ben Lomond) with its elevation of more than 1500 m.

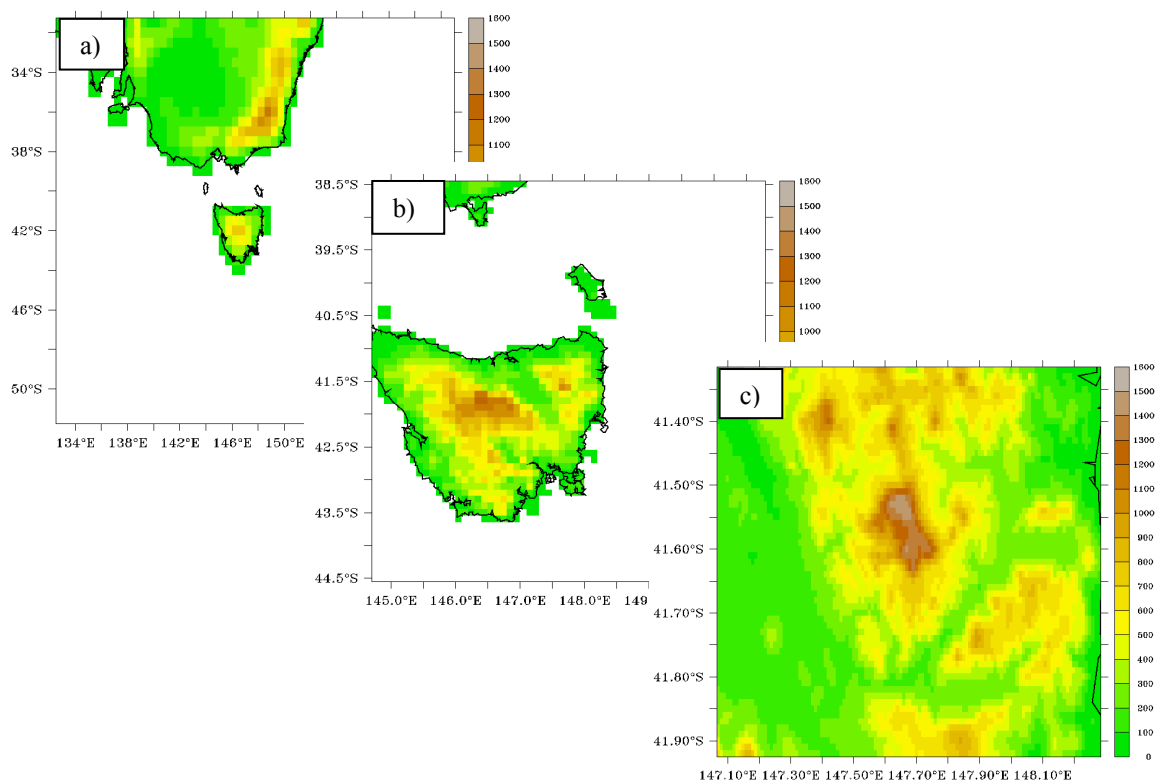


Figure 2: Topography (m) of CCAM a) 60 km, b) 8km and c) 1 km grids.

A suite of high-resolution forecasts was set up in order to assess some of the uncertainty related to initial conditions, downscaling method used and model settings. Table 1 provides a summary of the 32 forecasts completed each day. First, to assess the impact of different initial conditions, both the GFS and CMC analyses were used to initialise the forecasts. In addition, the GFS analyses were used at both 1.0° and 0.5° resolution to test the impact of resolution of the initial conditions. In another test, the GFS forecast was applied as large-scale forcing in the 8 km CCAM forecast (GFSf) to test the impact of this procedure on the high-resolution forecast. Different downscaling methodologies were tested, including FFN (far-field nudging) and SSDF (scale-selective digital filter). Boundary-layer parameterisations (t_{ke}) and number of vertical levels (27 vs 18) were varied to see if this affected the forecasts. Finally, the GFS forecasts were started every 6 hours, while the CMC forecasts were started every 12 hours, because these were the available analyses times. This allows testing of the impact of different initial times combined with varying amounts of observational data, since more data is available at 00 and 12 UTC.

Table 1. Range of simulations completed each day, with model options, initial conditions data (IC) and various resolutions listed. FFN: far-field nudging; SSDF: scale-selective digital filter; tke: turbulent kinetic energy boundary layer scheme; 27 levels: number of vertical levels used in model (normally 18). “GFSf” in resolution column indicates that the GFS global forecast was used as large-scale forcing for the 8 km CCAM forecast. The 60 km forecasts were for 5 days, while the 8 km and 1 km forecasts were for 3 days, except for the GFSf, which was only for 24hrs.

Name	Initial UTC	Model options	IC	Resolutions		
				60 km	8 km	1 km
Std.GFS	00/06/12/18	FFN	GFS(1.0°)	60 km	8 km	1 km
Std.GFS _h	00/06/12/18	FFN	GFS(.5°)	60 km	8 km	1 km
Std.GFS _f	00/06/12/18	FFN, only 24 hour forecast	GFS	GFS _f	8 km	1 km
Stds.GFS	00/06/12/18	SSDF	GFS	60 km	8 km	1 km
Tke.GDF	00/06/12/18	SSDF w/tke	GFS	60 km	8 km	1 km
L27.GFS	00/06/12/18	SSDF w/27 levels	GFS	60 km	8 km	1 km
Std.CMC	00/12	FFN	CMC	60 km	8 km	1 km
Stds.CMC	00/12	SSDF	CMC	60 km	8 km	1 km
Tke.CMC	00/12	SSDF w/tke	CMC	60 km	8 km	1 km
L27.CMC	00/12	SSDF w/27 levels	CMC	60 km	8 km	1 km

Figure 3 shows an example of the ensemble predictions of accumulated rainfall forecasts for the location 147°E and 41.42°S at 60 km, 8 km and 1 km resolution for the period 00 UTC 28 September to 00 UTC 30 September 2011. As can be seen from these plots, a large spread of rainfall forecasts is evident for any given location, due to uncertainty of the timing, location and amount of rainfall. However, the general timing of rainfall is similar, with the main rainfall event occurring around 12 UTC 28 September and a small event around 06 UTC 29 September 2011. The next section presents a brief summary of some of the verification of the forecasts.

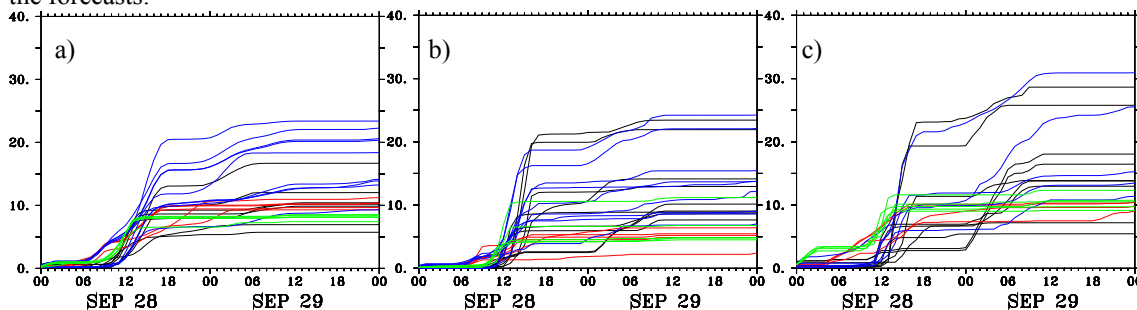


Figure 3: Plot of forecast accumulated rainfall (mm) from ensemble of forecasts from Table 1, at a) 60 km resolution, b) 8 km resolution, c) 1 km resolution. Black are forecasts started 00 UTC, red are forecasts started 06 UTC, blue are forecasts started 12 UTC and green are forecasts started 18 UTC, all started on 27 September 2011.

3. VERIFICATION

In this report, only verification of the Std.GFS 60 km, 8 km and 1 km forecasts for the period 23 July 2010 to 19 July 2011 is presented. The rainfall is verified against the gridded rainfall analyses at .25° resolution produced by the Bureau of Meteorology (BoM), based upon interpolation of the observations. The analysis was generated by using an optimised Barnes successive correction technique that applies a weighted averaging process to the station data. Topographical information was included by the use of rainfall ratio (actual rainfall divided by monthly average) in the analysis process (see details in Jones et al., 2009).

The averaged rainfall from the gridded analyses and from the various forecasts for this yearly period is shown in Figure 4. The observed rainfall (Figure 4a) shows little rainfall over central Australia, with higher amounts along the eastern coast. Higher rainfall is also evident over the Victorian mountains (southeastern Australia) and western Tasmania. The CCAM 60 km forecast (Figure 4b) captures this pattern reasonably well. The CCAM 8 km forecast rainfall (Figure 4c) begins to show the effect the high-resolution topography,

with a larger amount now evident over northeastern Tasmania, related to the high topography in this region. Additional detail in the forecasted rainfall, also related in part to the topography, is evident for the 1 km grid forecast (Figure 4d). For both the 8 km and 1 km resolution forecasts, the gridded dataset does not have enough spatial resolution to capture the details of the forecasts.

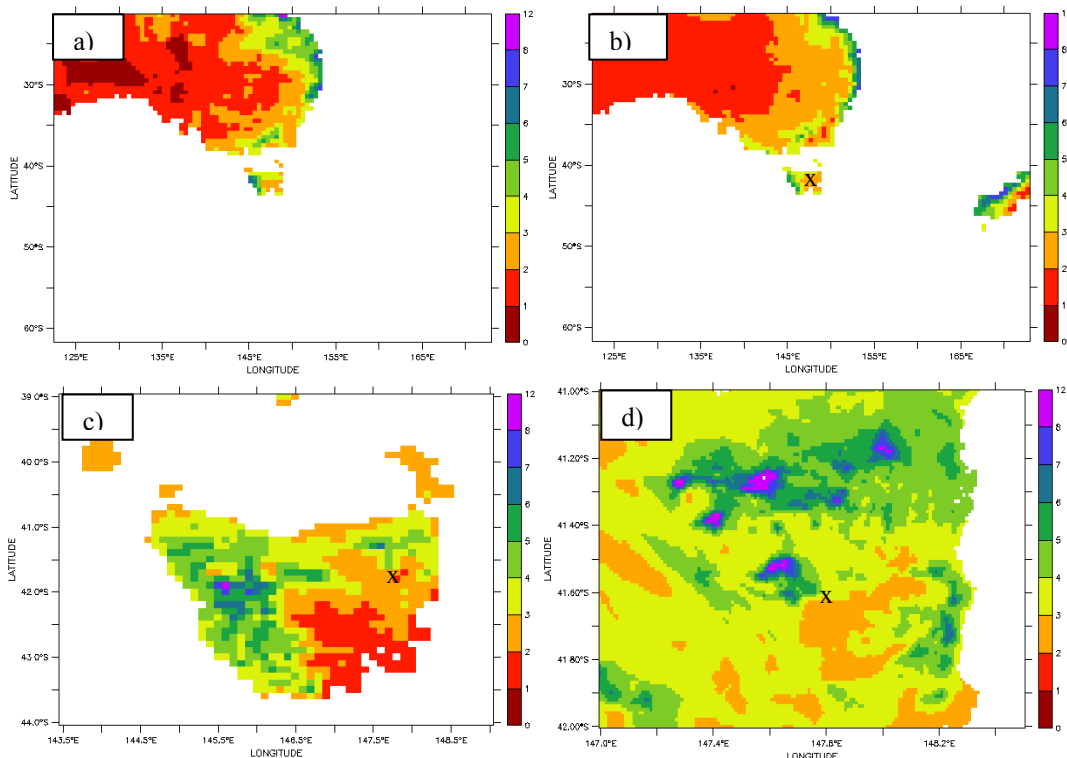


Figure 4: Averaged daily rainfall (mm/day) for the period 23 July 2010 to 19 July 2011: a) Observed gridded analysis, and CCAM Std.GFS 00-24 hour forecast at b) 60 km, c) 8 km and d) 1 km resolution. “x” marks location of the time series validation in figures 4b, 4c and 4d.

The time series of the daily rainfall amounts at location 147.8°E and 41.6°S (see “x” in Figure 4) from the observed gridded analyses and the various 00-24 hour forecasts at 60 km, 8 km and 1 km resolution is shown in Figure 5. The model captures the timing and general amounts of the forecasted rainfall when compared with the observations. Comparison of the distribution of forecasted daily rainfall (averaged to 25 km for the 8 km and 1 km forecasts) with the gridded observations shows generally good agreement (Figure 6). The 60 km forecast tends to underpredict the heavier rainfall amounts, while the 1 km forecasts tend to overpredict them. This result can be verified by computing the statistics of the comparison, shown in Table 2. The results for the 8 km and 1 km forecasts were averaged to the same approximate resolution as the analyses. As the forecast resolution increased, the bias decreased, while the mean absolute error and root mean square error increased. The index of agreement shows little change with resolution. The correlations against the gridded analyses increase from 0.80 for the 60 km forecasts to 0.84 for the 8 km forecasts averaged to 25 km. The correlation decreases slightly, to 0.75, as the resolution increases to 1 km. This might be due to the accuracy of the gridded analyses, as there are probably insufficient observational data in this region. The linear regression slope confirms the distribution comparison. The 60 km forecasts have a regression slope of 0.80, the 8km slope is 1.10 and the 1 km slope is 1.27. Thus the distribution of the 8 km forecasts best matches that of the gridded observations (slope nearest one).

Examination of the spatial pattern of the correlations of the daily rainfall of the gridded analyses compared with the 60 km and 8 km forecasts (Figure 7) shows values of over 0.8 over the northeast region of Tasmania, the main area of interest. Note that the 1 km forecasts are not shown since the validation area is too small to validate against the 25 km gridded analyses. Correlations of over 0.7 are also evident along the western slopes of Tasmania. Regions of lower correlation are in the Tamar Valley in northcentral Tasmania and in southwest Tasmania for the 8 km forecasts. The reason for the lower correlations is a subject of future work.

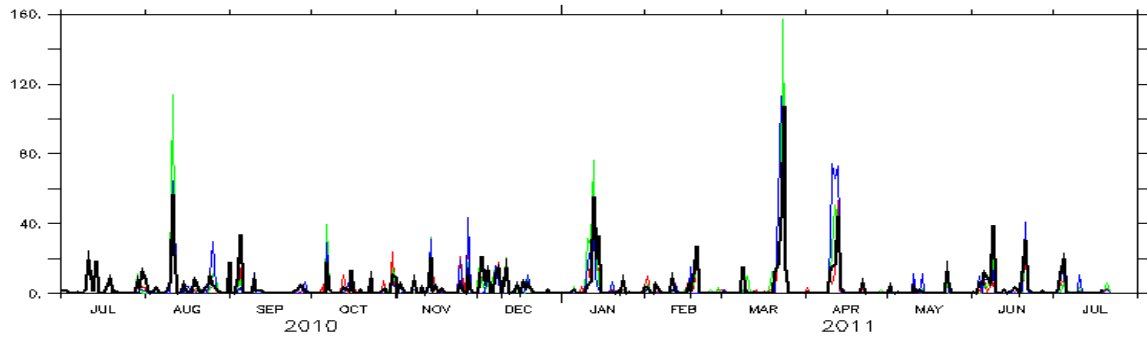


Figure 5: Time series of daily rainfall (mm) at 147.8°E and 41.6°S for gridded analyses (black) and CCAM 00-24 hour forecasts at 60 km (red), 8 km (green) and 1 km (blue) for 23 July 2010 to 19 July 2011.

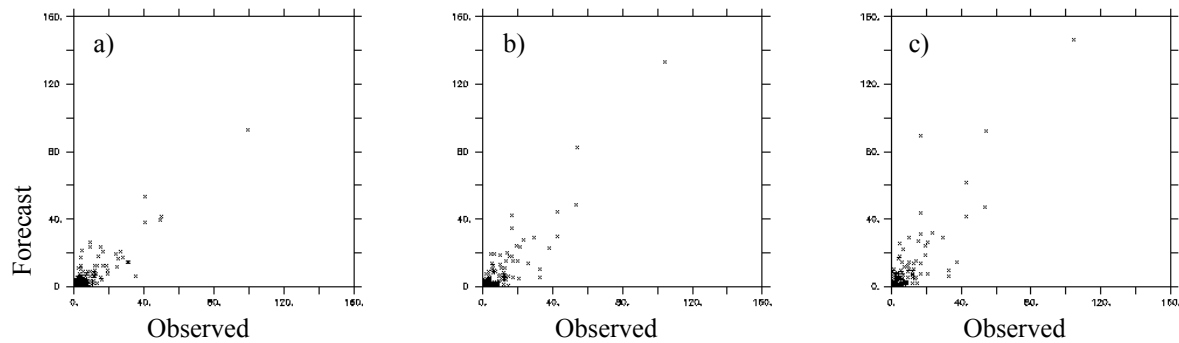


Figure 6: Plots of gridded observed (abscissa) and forecasted (ordinate) daily rainfall in mm for a) 60 km forecast, b) 8 km averaged to 25 km, and c) 1 km averaged to 25 km for location 147.8°E and 41.6°S and for 23 July 2010 to 19 July 2011.

Table 2: Validation statistics of daily rainfall forecasts (00-24 hour) for 147°E, 41.6°S and time period 23 July 2010 to 19 July 2011. The 8 km and 1 km forecasts were averaged to 25 km, like the gridded analyses. MAE is mean absolute error, RMSE is root mean square error, IOA is index of agreement and Corr. and Slope are from the linear regression of the observed and modelled rainfalls amounts.

Run	Mean obs	Mean fcst	Bias	MAE	RMSE	IOA	Corr.	Slope
60 km	3.60	2.76	-0.85	1.87	3.85	0.94	0.81	0.80
8 km	3.60	3.07	-0.54	1.98	4.50	0.94	0.84	1.10
1 km	3.52	3.59	0.07	2.30	6.48	0.90	0.75	1.27

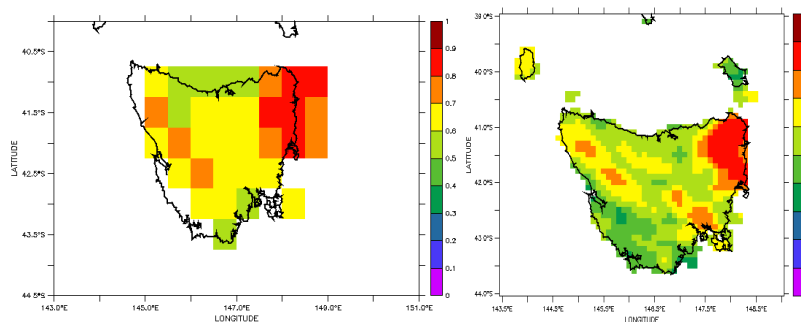


Figure 7: Correlations in time of the a) 60 km forecast and b) 8 km forecasts vs the gridded daily rainfall analyses for 23 July 2010 to 19 July 2011.

4. SUMMARY AND CONCLUSIONS

In this study, a very high-resolution, “poor man’s” ensemble rainfall prediction system for the South Esk River catchment in Tasmania is described. The ensemble is based upon different analyses, different downscaling methodologies, and different model parameterisations. This report is a summary of the system and a preliminary verification of the Std.GFS 60 km, 8 km and 1 km forecasts over a one year period (23 July

2010 to 19 July 2011). The Australian Bureau of Meteorology gridded analyses of daily rainfall used as the verification data is provided on a 0.25° longitude/latitude grid, which is fine for the 60 km forecasts, but relatively coarse for the 8 km forecasts, and not at a suitable resolution to verify the 1 km forecasts. In general the forecasts appear to do a very reasonable job of capturing the timing, distribution and amounts of daily rainfall. The 60 km forecasts tend to underpredict heavy rainfall, while the 1 km possibly overpredicts, but this needs further validation against station data.

Future work will focus on validating the forecasts against high-quality observations more suitable for the 1 km forecasts as well as verification of more members of the ensemble in order to assess their accuracy, with particular focus on the impact of the different parameterisations. Further evaluation of other ensemble predictions for temporal and spatial rainfall distributions also needs to be done. The rainfall predictions are being used to improve stream flow forecasts in the South Esk River catchment. Validating these forecasts and exploring methods to improve these forecasts can have significant implications for the multiple usages of the river system.

ACKNOWLEDGMENTS

This work was co-funded by the CSIRO Water for a Healthy Country Flagship and the Intelligent Island Fund. The Intelligent Island Fund is administered by the Tasmanian Department of Economic Development, Tourism and the Arts

REFERENCES

- Ebert, E.E. 2001, 'Ability of a poor man's ensemble to predict the probability and distribution of precipitation', *Monthly Weather Review*, 129, 2461-2480.
- Jones, D.A., W. Wang and R. Fawcett (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal* 58, 233-248.
- Katzfey, J. J.; McGregor, J. L.; Thatcher, M. J., and Ebert, B. (2009). Verification of 1 km ensemble wind predictions. In Anderssen, R.S., R.D. Braddock and L.T.H. Newham (eds) *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009 proceedings; Cairns, Queensland; 2377-2383.
- Kowalczyk, E., J.R. Garratt, and P.B. Krummel. (1994). Implementation of a soil-canopy scheme into the CSIRO GCM - regional aspects of the model response. *CSIRO Div. Atmospheric Research Tech Paper No. 32*, 59 pp.
- Lacis, A. and J. Hansen (1974). A parameterization of the absorption of solar radiation in the Earth's atmosphere. *Journal of Atmospheric Science*, 31, 118-133.
- McGregor, J.L., H.B. Gordon, I.G. Watterson, M.R. Dix, and L.D. Rotstain (1993). The CSIRO 9-level atmospheric general circulation model. *CSIRO Div. Atmospheric Research Tech. Paper No. 26*, 89 pp.
- McGregor, J.L. (2003). A new convection scheme using a simple closure. In "Current issues in the parameterization of convection", *BMRC Research Report 93*, 33-36.
- McGregor, J.L. (2005). C-CAM: Geometric aspects and dynamical formulation [electronic publication]. *CSIRO Atmospheric Research Tech. Paper No. 70*, 43 pp.
- McGregor, J.L., and M.R. Dix (2001). The CSIRO conformal-cubic atmospheric GCM. In *IUTAM Symposium on Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics*, P. F. Hodnett (ed.), Kluwer, Dordrecht, 197-202.
- McGregor, J.L., and M.R. Dix (2008). An updated description of the Conformal-Cubic Atmospheric Model. In *High Resolution Simulation of the Atmosphere and Ocean*, K. Hamilton and W. Ohfuchi (eds.), Springer, 51-76.
- Rotstain, L.D. (1997). A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I: Description and evaluation of the microphysical processes. *Quarterly Journal of the Royal Meteorological Society*, 123, 1227-1282.
- Schwarzkopf, M.D. and S.B. Fels (1991). The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. *Journal of Geophysical Research*, 96, 9075-9096.
- Thatcher M. and J.L. McGregor (2009). Using a scale-selective filter for dynamical downscaling with the conformal cubic atmospheric model. *Monthly Weather Review*, 137, 1742-1752.