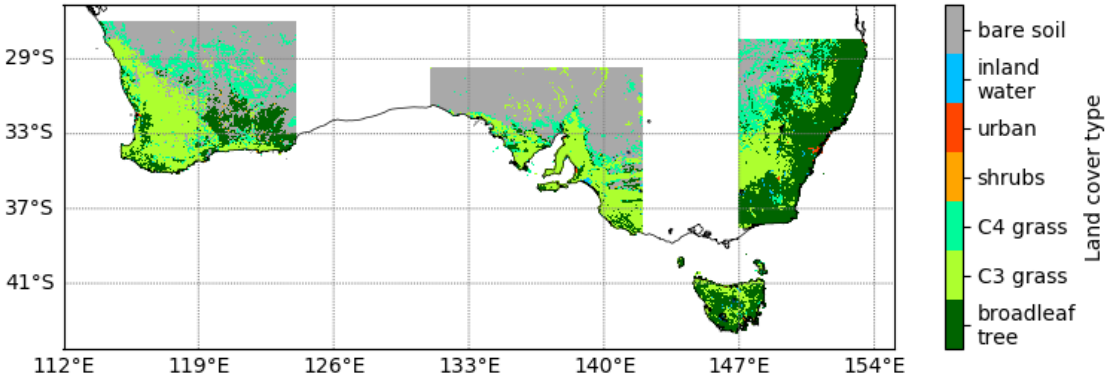

Figure S1: Most dominant surface types at each grid cell mapped on the BARRA-C 1.5 km grid.

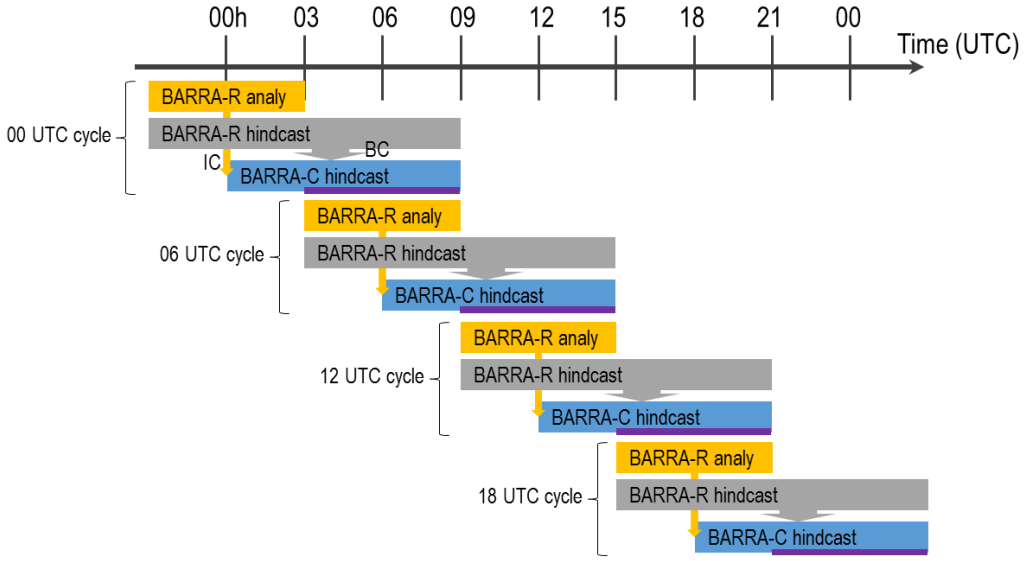

Figure S2: Cycling set-up of BARRA-C at base time t_0 = 00:00, 06:00, 12:00, and 18:00 UTC. Each UM hindcast is initialized at t_0 with BARRA-R analysis centred at t_0 and subjects to boundary conditions from BARRA-R hindcast from t_0, t_0+1h, …, t_0+9h. The purple bars indicate the time steps of the model states that have been archived.

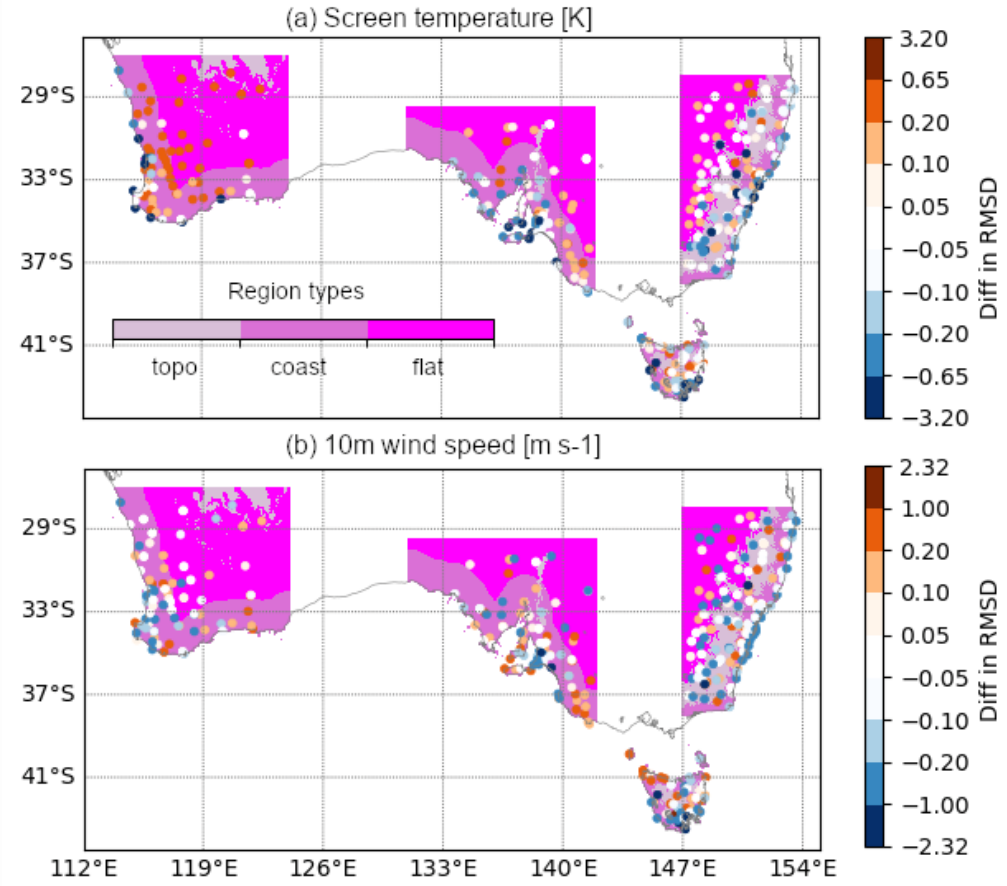

Figure S3: Difference in RMSD at each station between BARRA-C and BARRA-R for (a) screen temperature and (b) 10 m wind speed. The boxplots of the score from different models are plotted in Figure 2. The background shows three regions of analysis: complex topography ('topo'), coastal ('coast'), and flat.

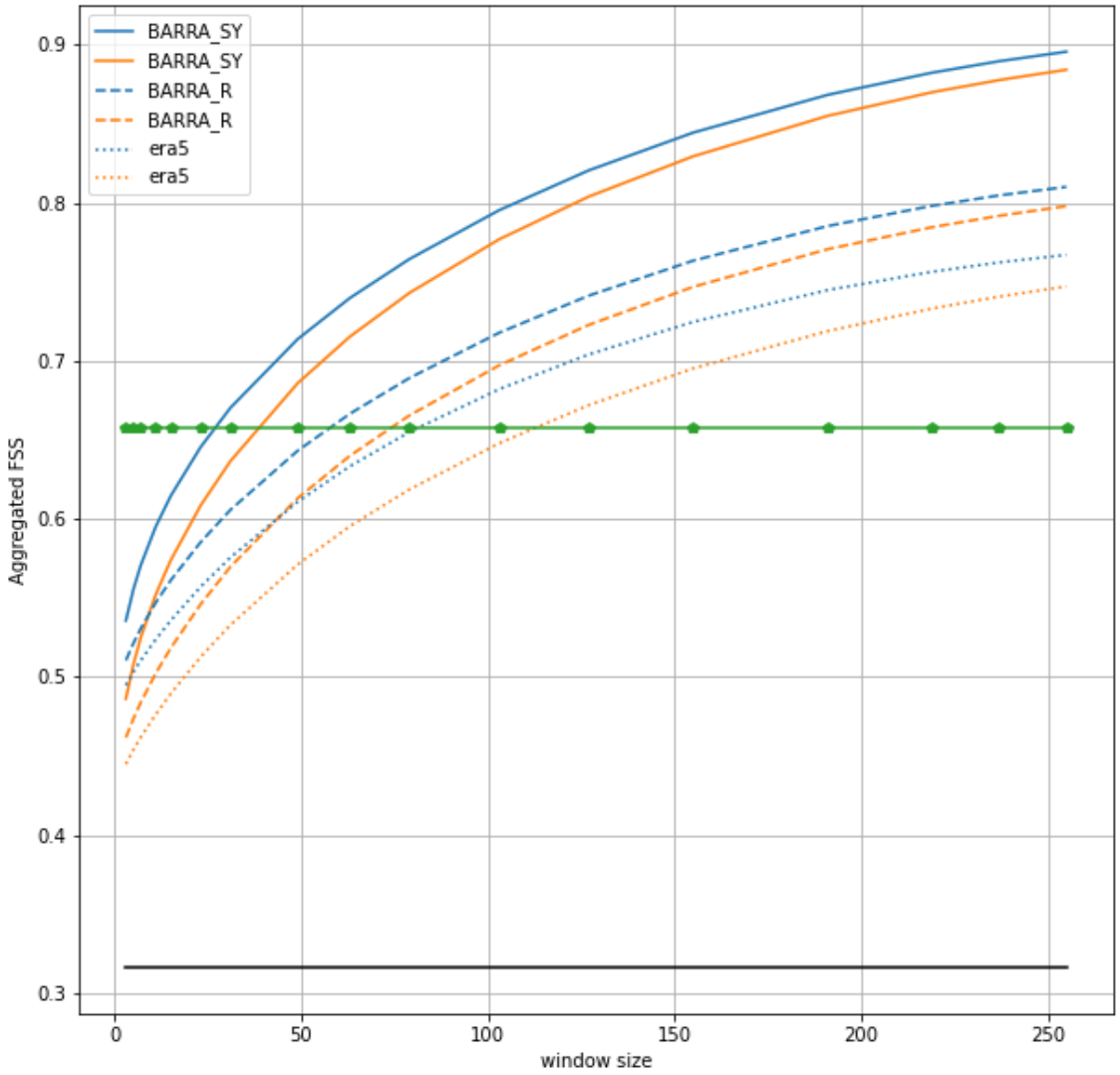

Figure S4: Aggregated FSS across 95 12-hour non-overlapping storm events as a function of neighbourhood distance (window length size i.e. number of 1.5km grid cells) for 12h rainfall above two percentile thresholds, 48th (blue) and 62nd (orange). Scores for three models are shown, BARRA-C (solid), BARRA-R (dashed) and ERA5 (dotted).