Supplementary Information:
Historical trends of seasonal droughts in Australia

Matthew O. Grant[1,2], Anna M. Ukkola[1,2,3], Elisabeth Vogel[2,4,5], Sanaa Hobeichi[1,3], Andy J. Pitman[1,2], Alex Raymond Borowiak[2,5] and Keirnan Fowler[6]

[1]Climate Change Research Centre, UNSW Sydney, Kensington, NSW 2052, Australia
[2]Australian Research Council Centre of Excellence for Climate Extremes
[3]Australian Research Council Centre of Excellence for 21st Century Weather
[4]Water Research Centre, UNSW Sydney, Kensington, NSW 2052, Australia
[5]School of Geography, Earth and Atmospheric Sciences, University of Melbourne, Parkville, VIC 3052, Australia
[6]Department of Infrastructure Engineering, University of Melbourne, Parkville, VIC 3052, Australia

Correspondence to: Matthew O. Grant (matt.grant@unsw.edu.au)

S1 Methods

S1.1 Verification of hydrological drought trends

There are no available long-term observations of runoff or streamflow data which cover the whole Australian continent. As such we have used AWRA-L modelled runoff to quantify hydrological droughts. To ensure that hydrological drought trends based on AWRA-L runoff are reliable, we evaluated them against observed hydrological drought trends based on in-situ streamflow data from CAMELS-AUS v2 (Table 1). The CAMELS-AUS streamflow data was first filtered to remove stations with large data gaps. Catchments with missing data at over 5% of the timesteps were omitted. For the remaining catchments, gap-filling of missing data provided with the CAMELS-AUS dataset has been adopted (Fowler et al., 2024). Hydrological drought trends, based on observed streamflow and modelled runoff, were compared over two time periods: 1981-2020 and 1951-2020. 460 catchments have data spanning 1981 to 2020, which provides a large sample of catchments to compare with the modelled drought trends. Additionally, 34 of these catchments provided streamflow data back to 1951 and were used to evaluate how well AWRA-L captures drought trends over a longer time period.

Trends in time under drought were compared at the streamflow catchments for hydrological droughts calculated from AWRA-L runoff and the CAMELS-AUS streamflow. The observed streamflow time under drought trends were calculated at each catchment using the same method as described in Section 2.3 in the main text. To allow for direct comparison, the AWRA-L runoff was averaged over each catchment region and the drought metric and trend were calculated at each of these. The evaluation is presented in Figure S13.

S1.2 S/N ratio and KS test

The signal-to-noise (S/N) ratio and Kolmogorov-Smirnov (KS) test were used to determine if time and area under drought has emerged from their variability. For both methods the first 50 years (1911-1961) was used as a baseline to compare emergence to. The S/N ratio is one of the most widely used methods to test for emergence (Hawkins et al., 2020; Hawkins and Sutton, 2012). To calculate this, the relevant drought metric was annually averaged for area under drought and annually summed for time under drought. The mean of the baseline was subtracted from the time series to give the anomalies. To calculate the signal, a Locally Weighted Scatterplot Smoothing (LOWESS) model was fitted to these anomalies. LOWESS is commonly used to calculate climate signal due to its ability to fit to data of any shape (Hawkins et al., 2020). The timeseries of anomalies was then detrended, and another LOWESS model was fitted to the detrended timeseries. The noise was defined as the standard deviation of the residuals of the detrended LOWESS model. The signal is said to have emerged from the noise if the absolute value of the S/N ratio is above one (Frame et al., 2017).

Although not as common as the S/N ratio, the KS test can also be used to determine emergence (King et al., 2015; Mahlstein et al., 2011). The KS test compares the maximum of the difference between cumulative distribution functions (CDFs). Here we compare the CDFs of 20-year rolling windows across the entire timeseries to the CDF of the baseline period. Significance is assessed at the 95% level, and as such when the p-value is less than 0.05, the trend is said to have emerged. For both tests, emergence must remain until the end of the timeseries and must have emerged for at least 20 years (Hawkins et al., 2014). These criteria are applied to ensure the change is a substantial shift from the baseline variability, and not just a temporary deviation.

References

Fowler, K. J. A., Zhang, Z., and Hou, X.: CAMELS-AUS v2: updated hydrometeorological timeseries and landscape attributes for an enlarged set of catchments in Australia, Earth Syst. Sci. Data Discuss., 1–21, https://doi.org/10.5194/essd-2024-263, 2024.

Frame, D., Joshi, M., Hawkins, E., Harrington, L. J., and de Roiste, M.: Population-based emergence of unfamiliar climates, Nat. Clim. Change, 7, 407–411, https://doi.org/10.1038/nclimate3297, 2017.

Hawkins, E. and Sutton, R.: Time of emergence of climate signals, Geophys. Res. Lett., 39, https://doi.org/10.1029/2011GL050087, 2012.

Hawkins, E., Anderson, B., Diffenbaugh, N., Mahlstein, I., Betts, R., Hegerl, G., Joshi, M., Knutti, R., McNeall, D., Solomon, S., Sutton, R., Syktus, J., and Vecchi, G.: Uncertainties in the timing of unprecedented climates, Nature, 511, E3–E5, https://doi.org/10.1038/nature13523, 2014.

Hawkins, E., Frame, D., Harrington, L., Joshi, M., King, A., Rojas, M., and Sutton, R.: Observed Emergence of the Climate Change Signal: From the Familiar to the Unknown, Geophys. Res. Lett., 47, e2019GL086259, https://doi.org/10.1029/2019GL086259, 2020.

King, A. D., Donat, M. G., Fischer, E. M., Hawkins, E., Alexander, L. V., Karoly, D. J., Dittus, A. J., Lewis, S. C., and Perkins, S. E.: The timing of anthropogenic emergence in simulated climate extremes, Environ. Res. Lett., 10, 094015, https://doi.org/10.1088/1748-9326/10/9/094015, 2015.

Mahlstein, I., Knutti, R., Solomon, S., and Portmann, R. W.: Early onset of significant local warming in low latitude countries, Environ. Res. Lett., 6, 034009, https://doi.org/10.1088/1748-9326/6/3/034009, 2011.

Supplementary Tables

Table S1: The R^2 scores between the predicted drought trend and true drought trend for each of the random forest models trained for the quantification of contribution of the key hydrometeorological variables to agricultural and hydrological drought trends

	Agricultural Drought				Hydrological Drought			
	DJF	**MAM**	**JJA**	**SON**	**DJF**	**MAM**	**JJA**	**SON**
Central Slopes	0.64	0.64	0.55	0.58	0.61	0.60	0.56	0.57
East Coast	0.68	0.74	0.60	0.63	0.56	0.54	0.47	0.48
Monsoonal North	0.75	0.84	0.74	0.73	0.61	0.67	0.50	0.55
Murray Basin	0.63	0.58	0.63	0.63	0.54	0.46	0.48	0.55
Rangelands	0.76	0.79	0.82	0.78	0.55	0.53	0.64	0.55
S/SW Flatlands	0.62	0.73	0.81	0.78	0.38	0.45	0.58	0.65
Southern Slopes	0.46	0.74	0.65	0.64	0.50	0.59	0.56	0.58
Wet Tropics	0.69	0.67	0.76	0.68	0.63	0.70	0.67	0.57

Supplementary Figures

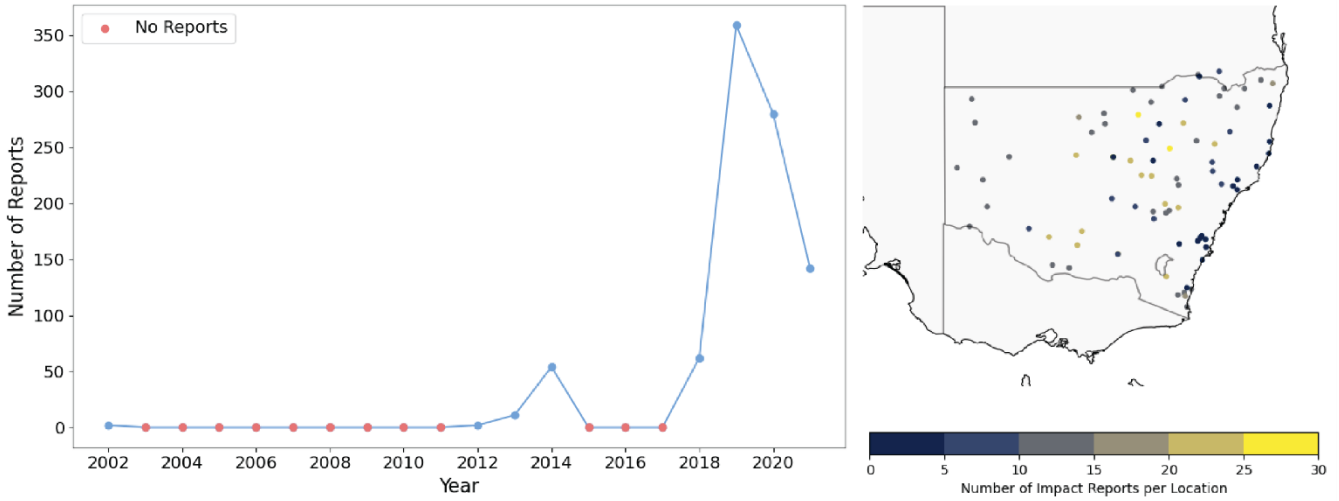


Figure S1: The years (left) and locations (right) for which there were drought impact reports in the data used to create the impacts-based drought metric.

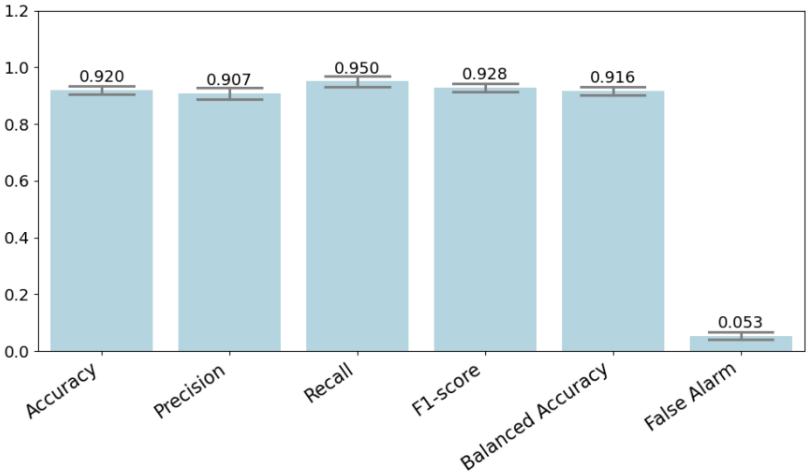


Figure S2: The skill scores for various metrics used to test the performance of the random forest model when trained on the drought impact reports. 100 models were trained, changing the random seed and train/test split each time. The bars indicate the mean skill score between the 100 models. The error bars indicate the variance of the skill scores between the 100 models. All metrics take values between 0 and 1, with 1 indicating the best possible performance for all metrics aside from False Alarm (where 0 indicates the best possible performance).

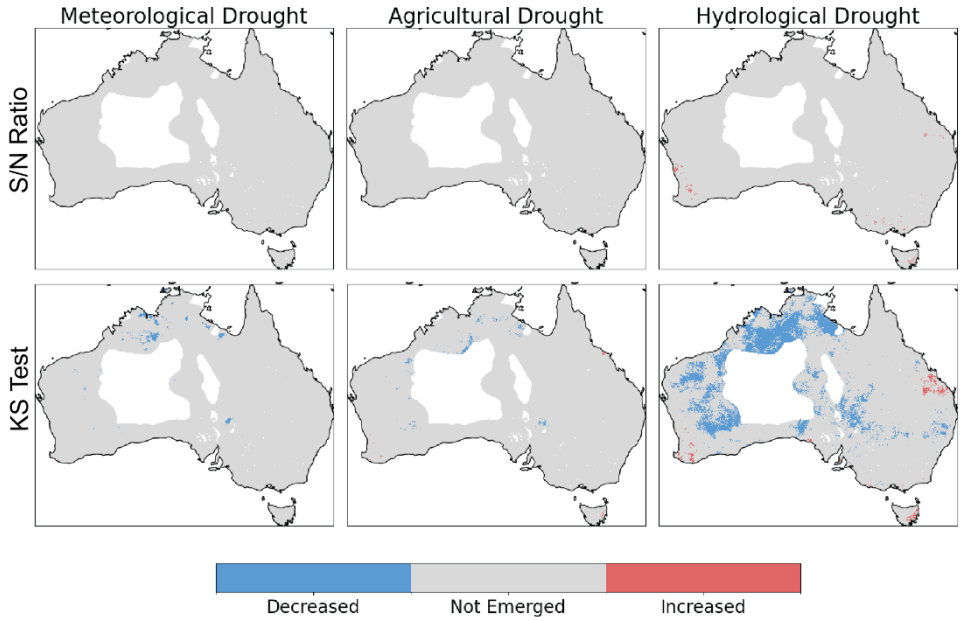


Figure S3: Trend emergence for time under drought characteristic, shown for each of the traditional drought types. This is shown for both the signal-to-noise (S/N) ratio (top row) and the Kolmogorov-Smirnoff (KS) test (bottom row).

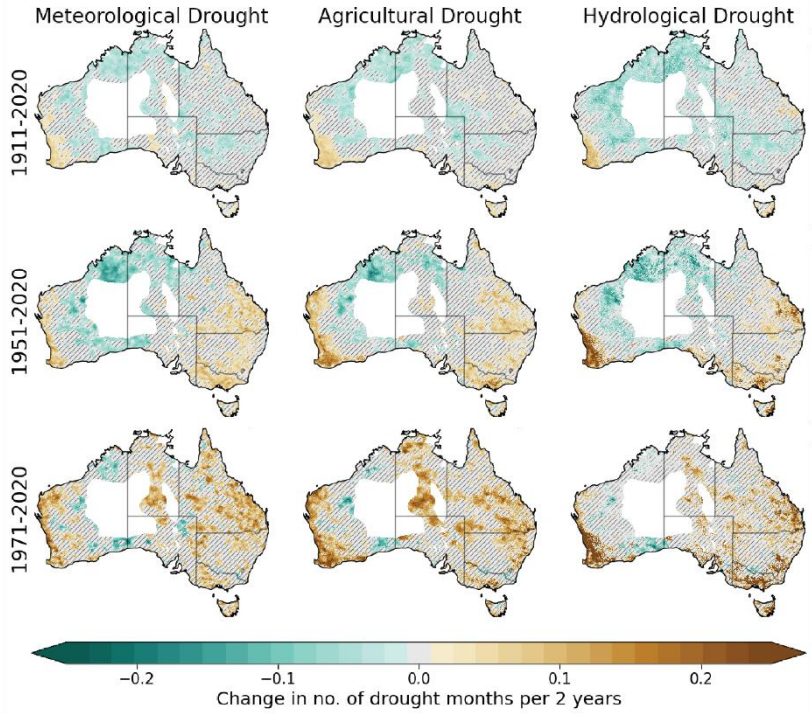


Figure S4: Trends in time under drought for the three traditional drought types and three time periods. The maps show the change in the number of drought months, equivalent to Figure 2 in the main text but using 2-year blocks for resampling. The hatching indicates where the trend is not significant ($p > 0.05$). The white spaces indicate the area masked out due to sparse observation network.

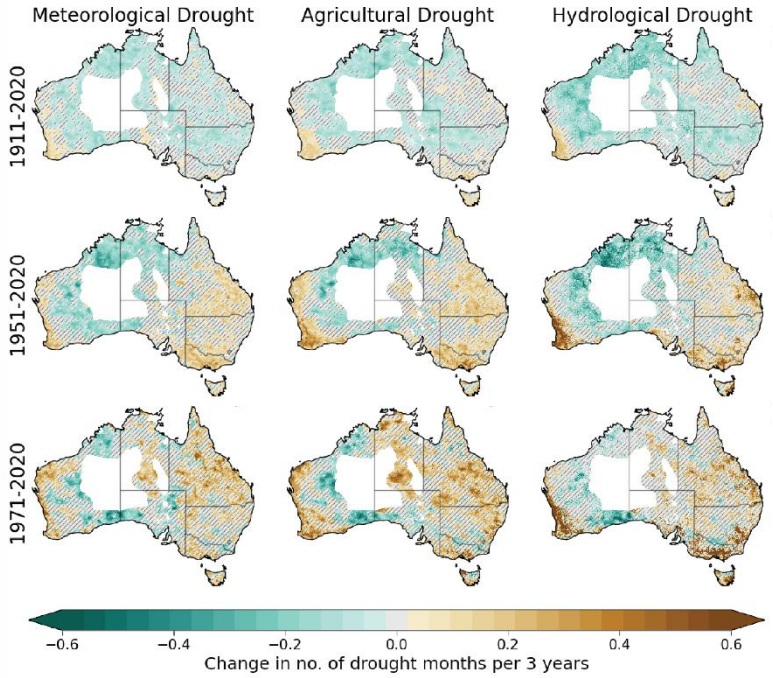


Figure S5: Trends in time under drought for the three traditional drought types and three time periods. These maps show the change in number of drought months, equivalent to Figure 2 in the main text but using 3-year blocks for resampling. The hatching indicates where the trend is not significant (p > 0.05). The white spaces indicate the area masked out due to sparse observation network.

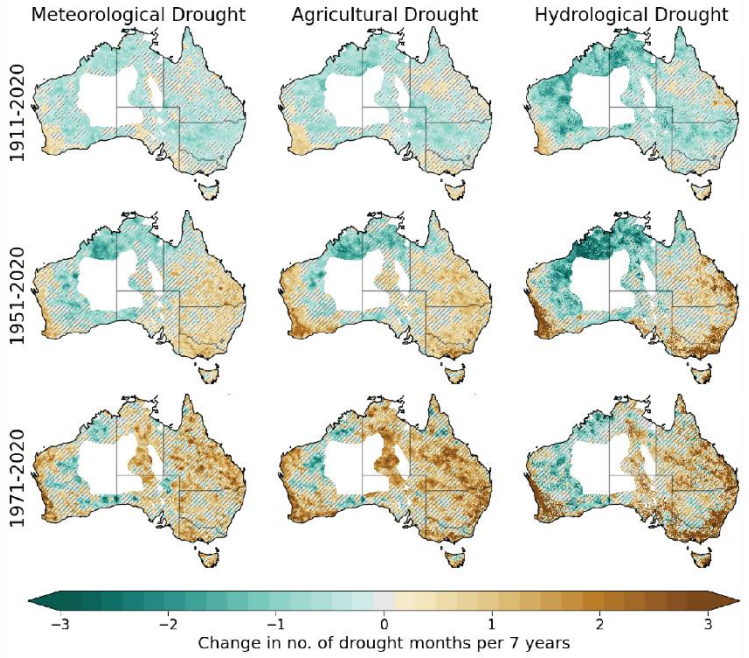


Figure S6: Trends in time under drought for the three traditional drought types and three time periods. The maps show the change in number of drought months, equivalent to Figure 2 in the main text but using 7- year blocks for resampling. The hatching indicates where the trend is not significant (p > 0.05). The white spaces indicate the area masked out due to sparse observation network.

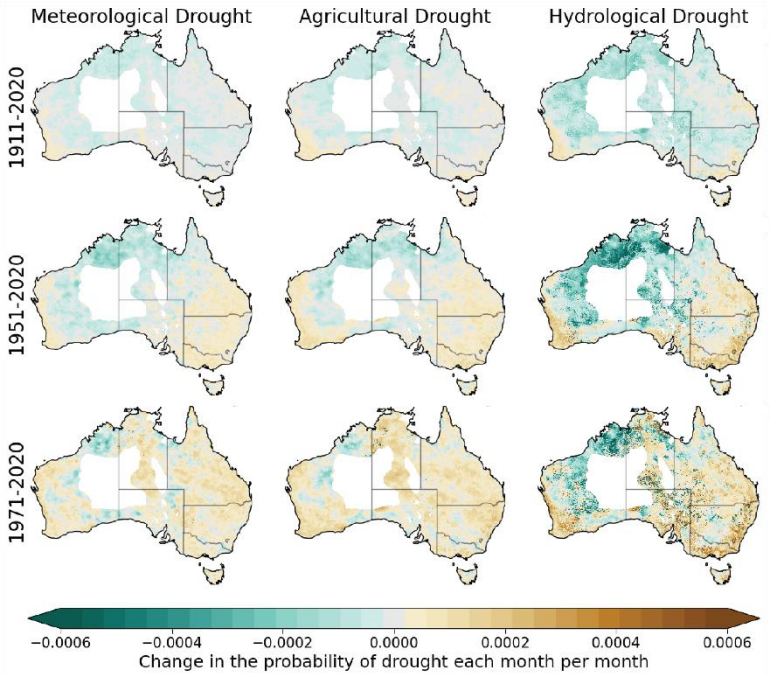


Figure S7: Trends in time under drought determined using a logistic regression model for the three traditional drought types and three time periods. The maps show the change in probability of a drought month. The white spaces indicate the area masked out due to sparse observation network.

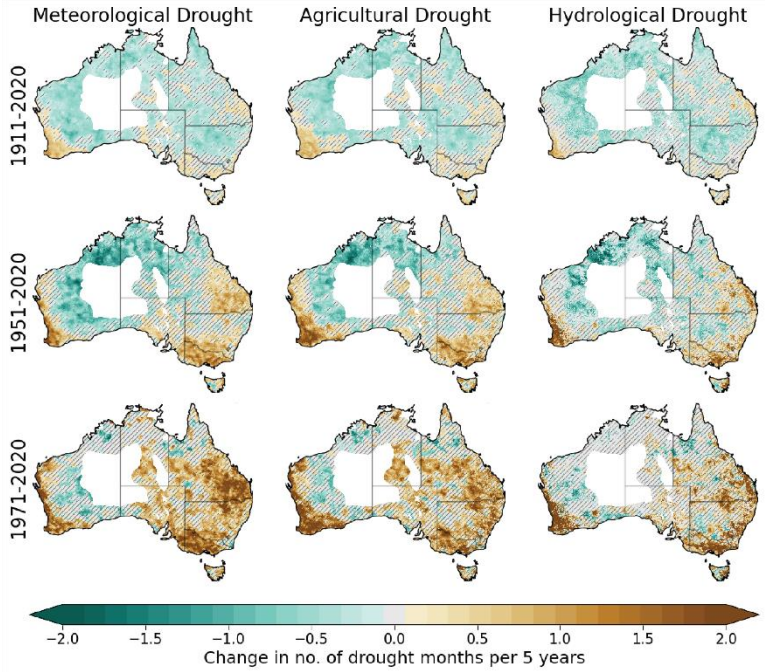


Figure S8: Trends in time under drought for annual-scale drought months determined using 12-month running means. The maps show the change in number of drought months per 5 years for the three traditional drought types and three time periods. The hatching indicates where the trend is not significant (p > 0.05). The white spaces indicate the area masked out due to sparse observation network.

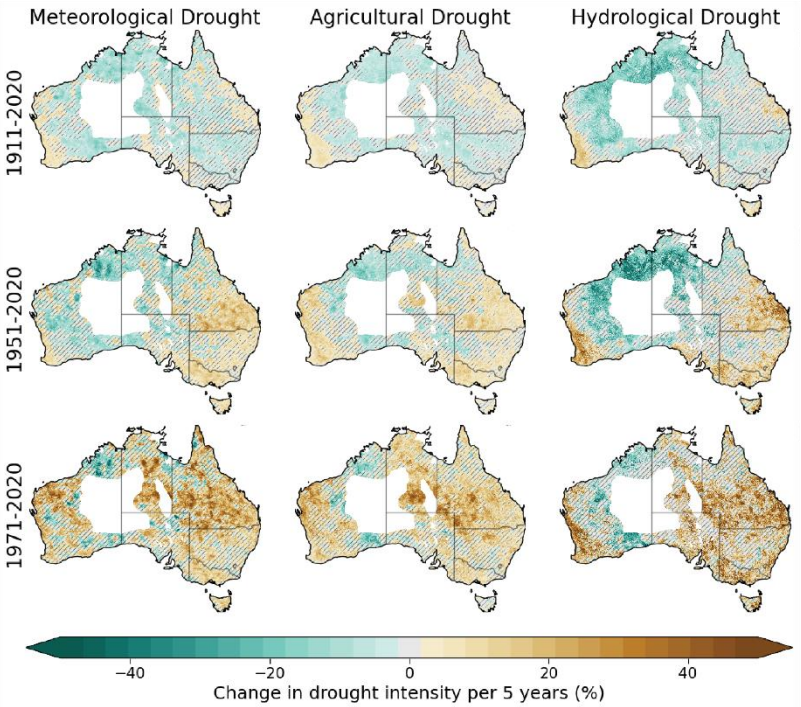


Figure S9: Trends in drought intensity for the three traditional drought types and three time periods. The hatching indicates where the trend is not significant (p > 0.05). The white spaces indicate the area masked out due to sparse observation network.

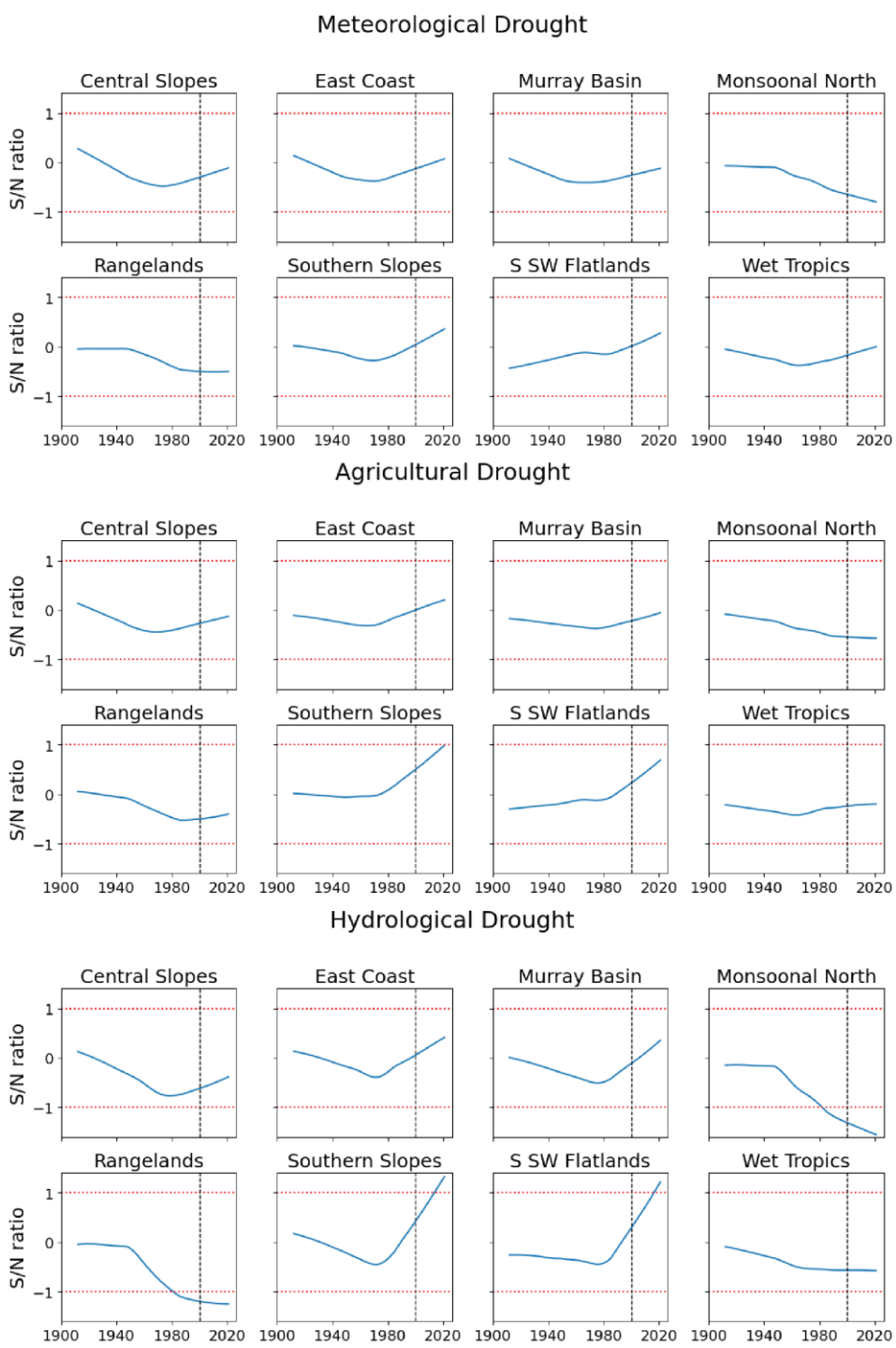
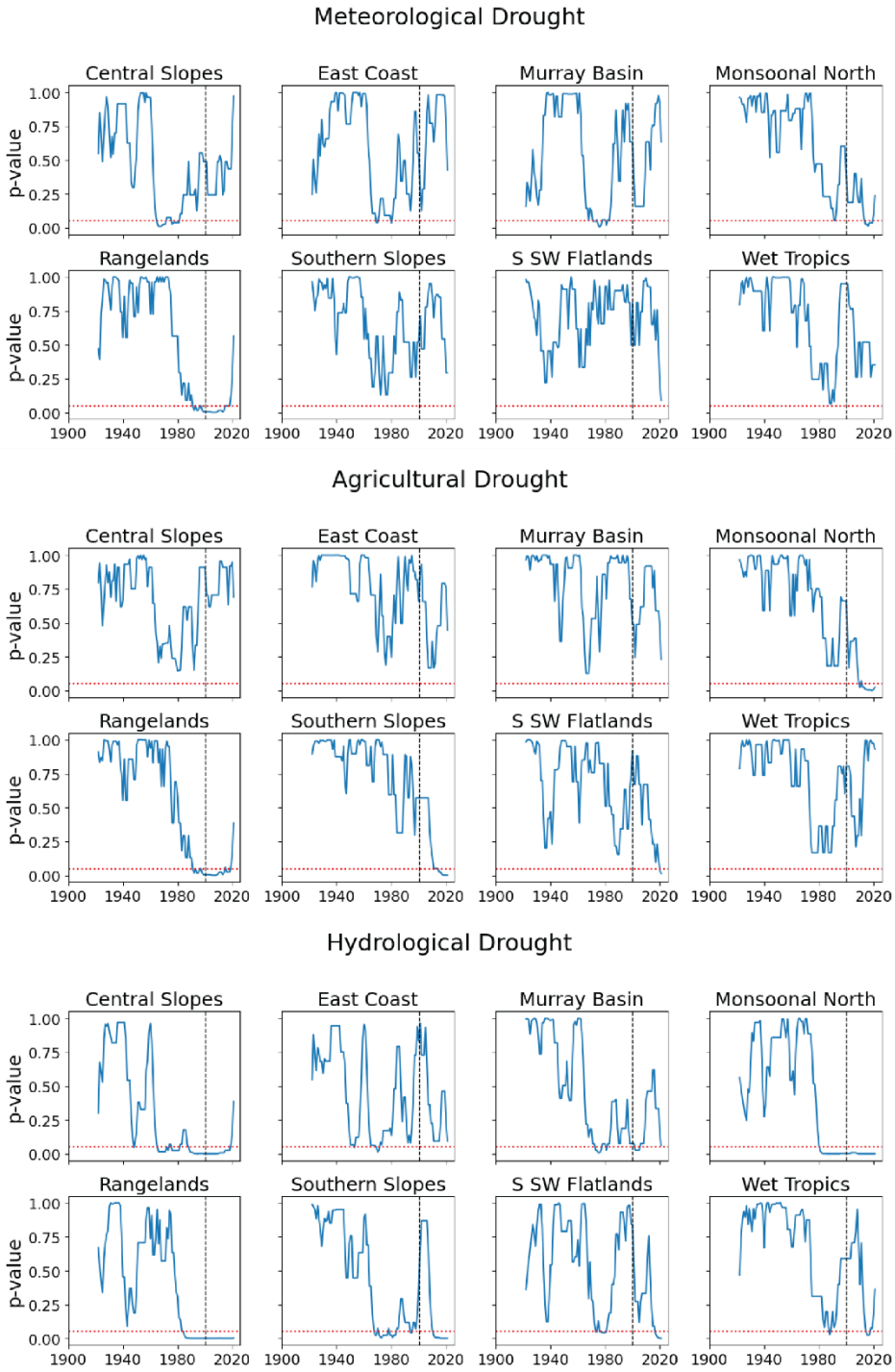


Figure S10: Timeseries of the signal-to-noise (S/N) ratio, calculated on the area under drought for each of the three traditional drought types and each of the NRM regions. The change in area under drought is said to have emerged if the S/N ratio remains outside the range of -1 to 1 (indicated by the red doted lines), for at least 20 years and until the end of the timeseries. The black dotted line indicates the year 2000 (20 years before the end of the timeseries).



115 **Figure S11: Timeseries of the Kolmogorov-Smirnoff (KS) test p-value, calculated on the area under drought for each of the three traditional drought types and each of the NRM regions. The change in area under drought is said to have emerged if the p-value remains below 0.05 (indicated by the red dotted line) for at least 20 years and until the end of the timeseries. The black dotted line indicates the year 2000 (20 years before the end of the timeseries).**

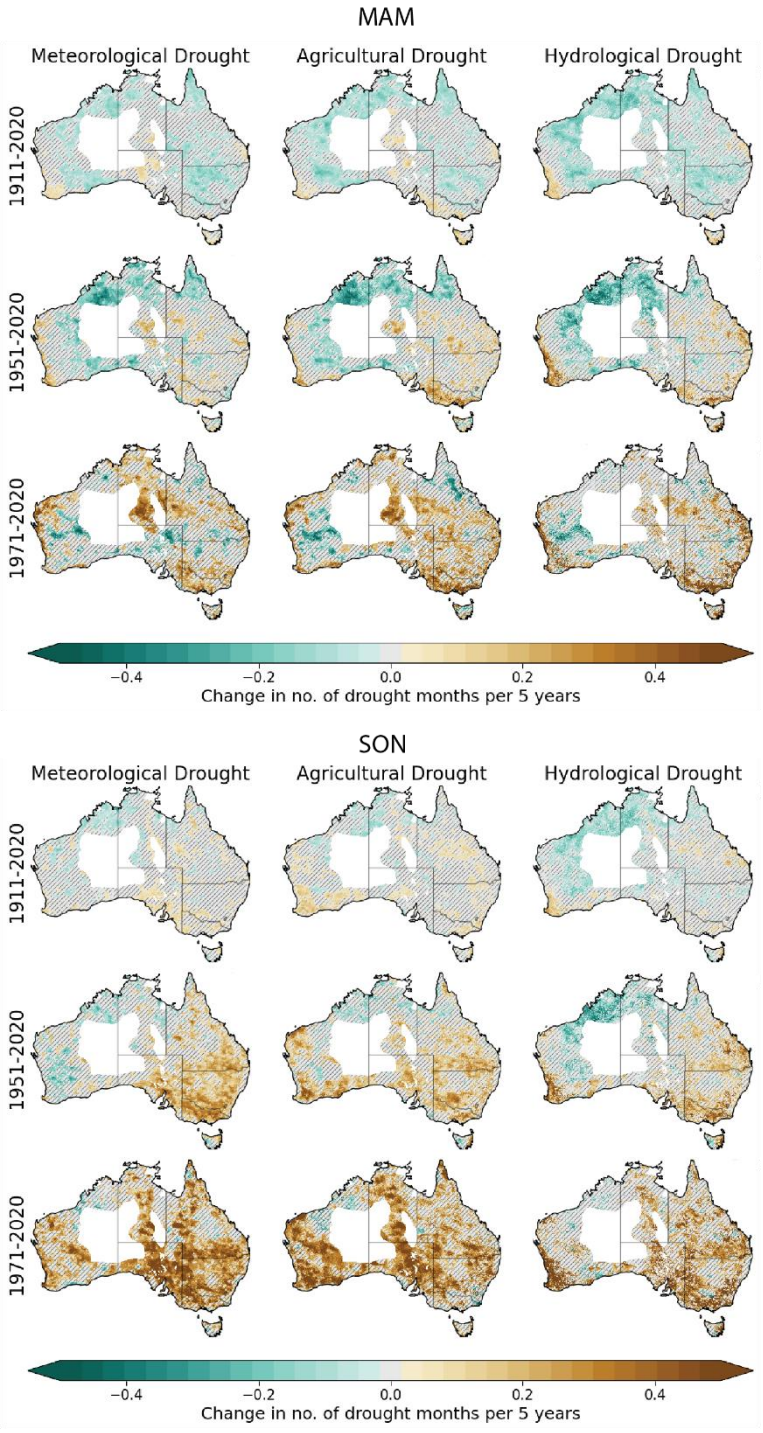


Figure S12: Seasonal trends in time under drought for autumn (MAM) and spring (SON). The maps show the change in number of drought months per 5 years for the three traditional drought types and three time periods. The hatching indicates where the trend is not significant (p > 0.05). The white spaces indicate the area masked out due to sparse observation network.

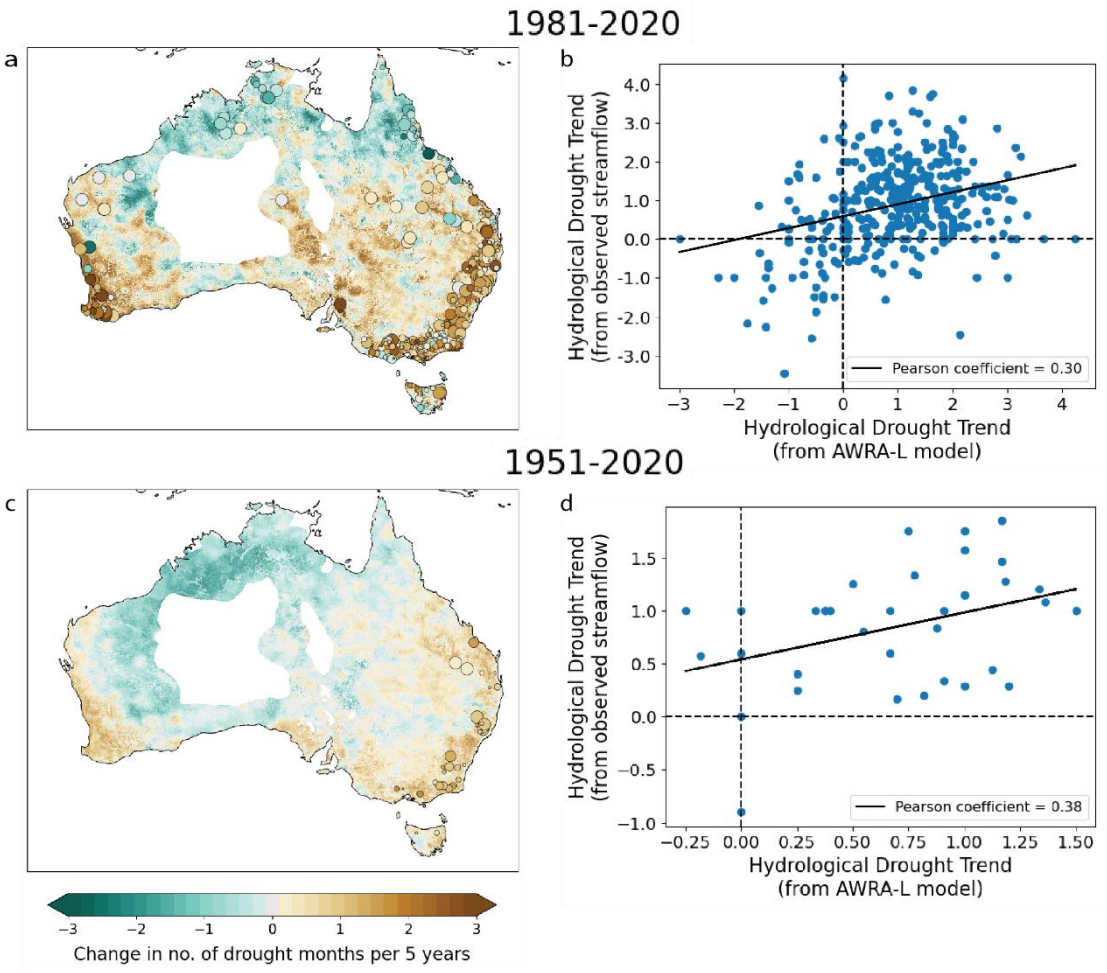


Figure S13: Evaluation of AWRA-L runoff against observed streamflow time under drought trends. Panels a and c show the observed streamflow time under drought trends at the catchments overlayed onto the AWRA-L runoff time under drought trends. Panels b and d show scatterplots of the AWRA-L runoff time under drought trends against the observed streamflow time under drought trends. Both types of plots are shown for 1981-2020 (a-b) and 1951-2020 (c-d) trends. The white spaces on a and c indicate the area masked out due to sparse observation network.