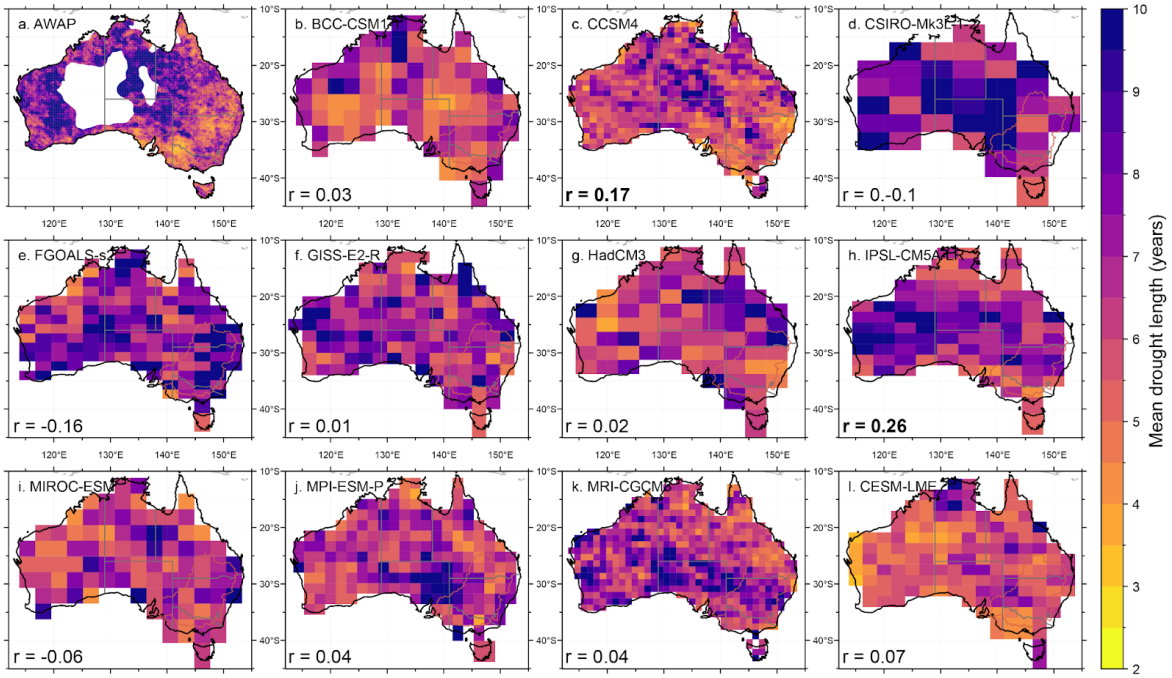
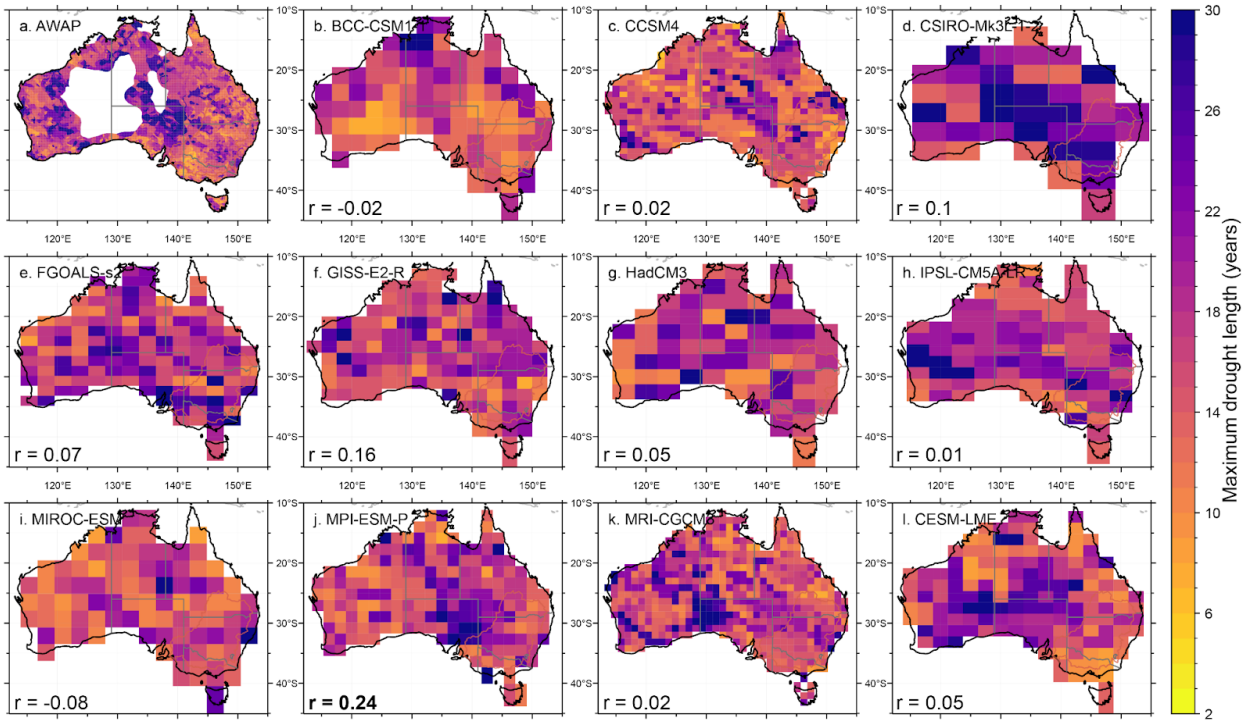
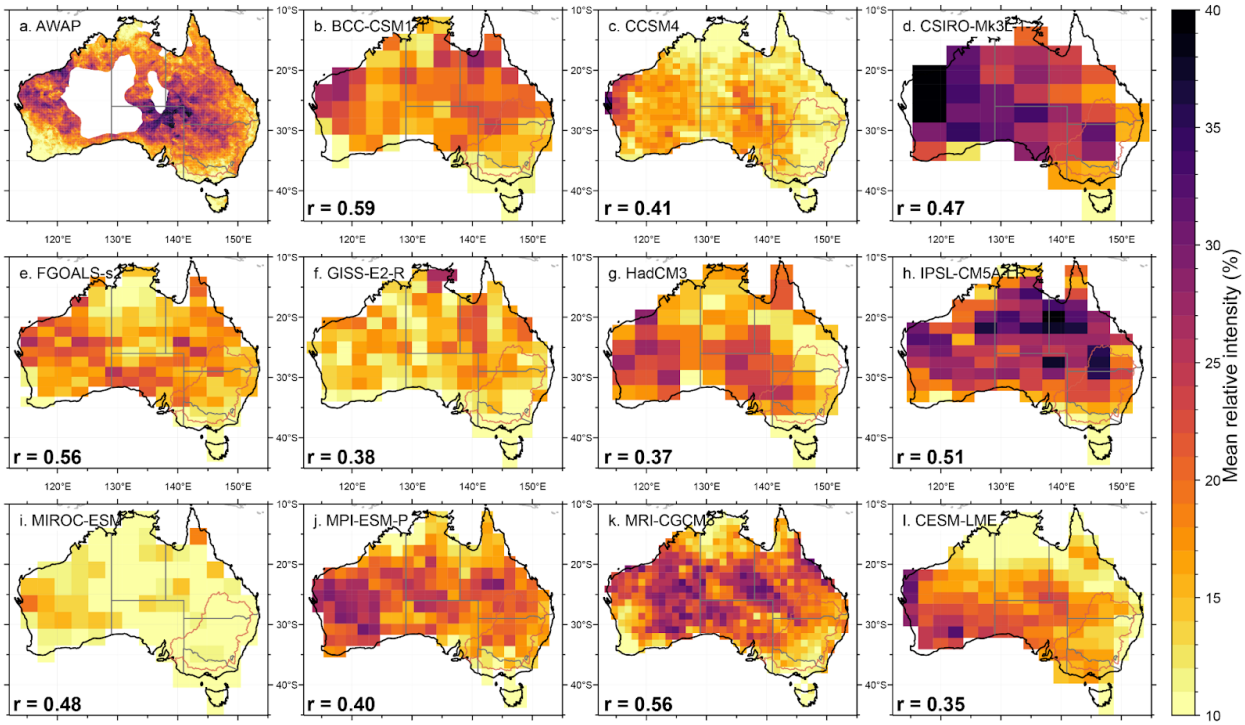


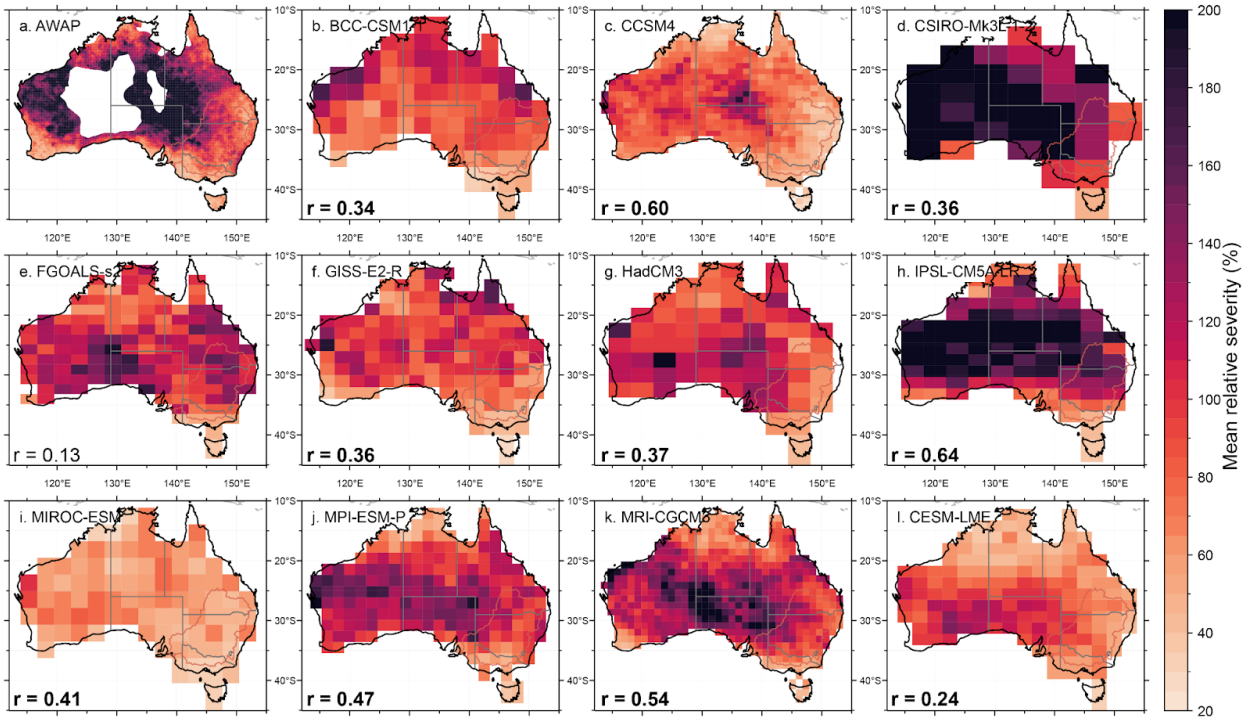
Supporting Fig 1: Mean multi-year drought length in (a) observations (1900-2000) and (b-l) model simulations of the same period. Showing the CESM LME ensemble mean. Spatial correlations of each model with observations are shown in the lower left corner in b-l. Bold text denotes a significant ($p < 0.05$) correlation.



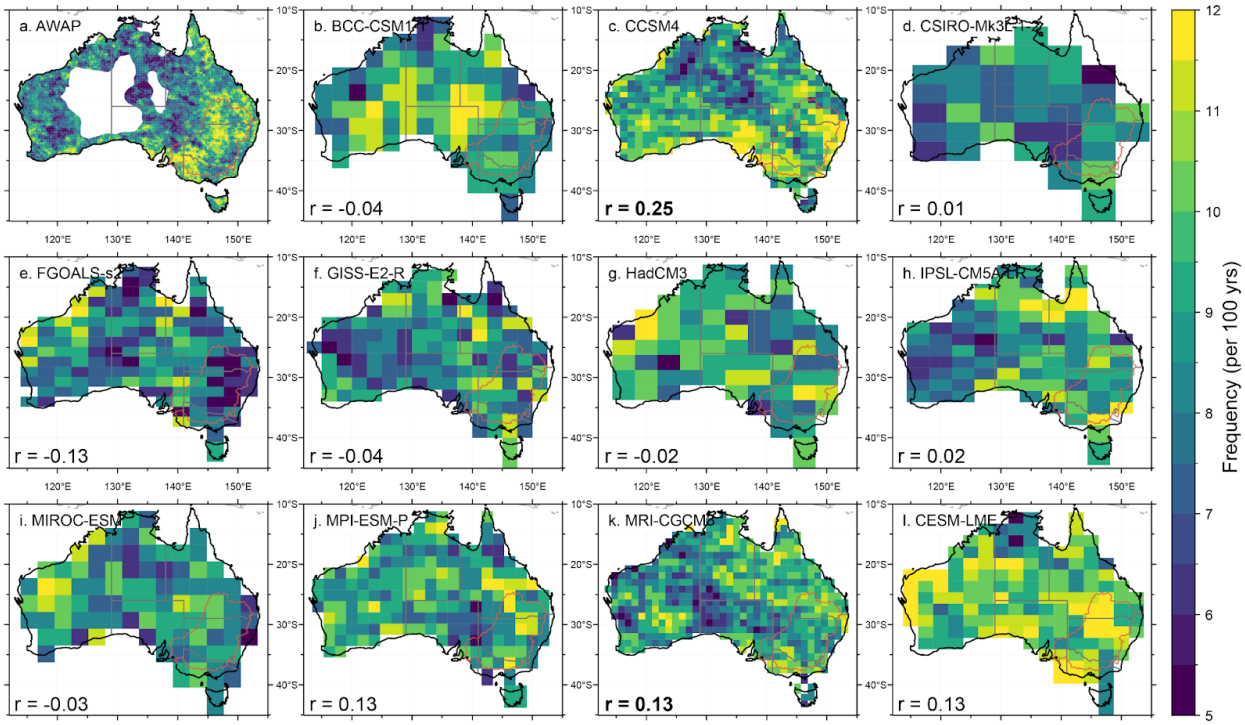
Supporting Fig 2: Maximum multi-year drought length in (a) observations (1900-2000) and (b-l) model simulations of the same period. Showing the CESM LME ensemble mean. Spatial correlations of each model with observations are shown in the lower left corner in b-l. Bold text denotes a significant ($p < 0.05$) correlation.



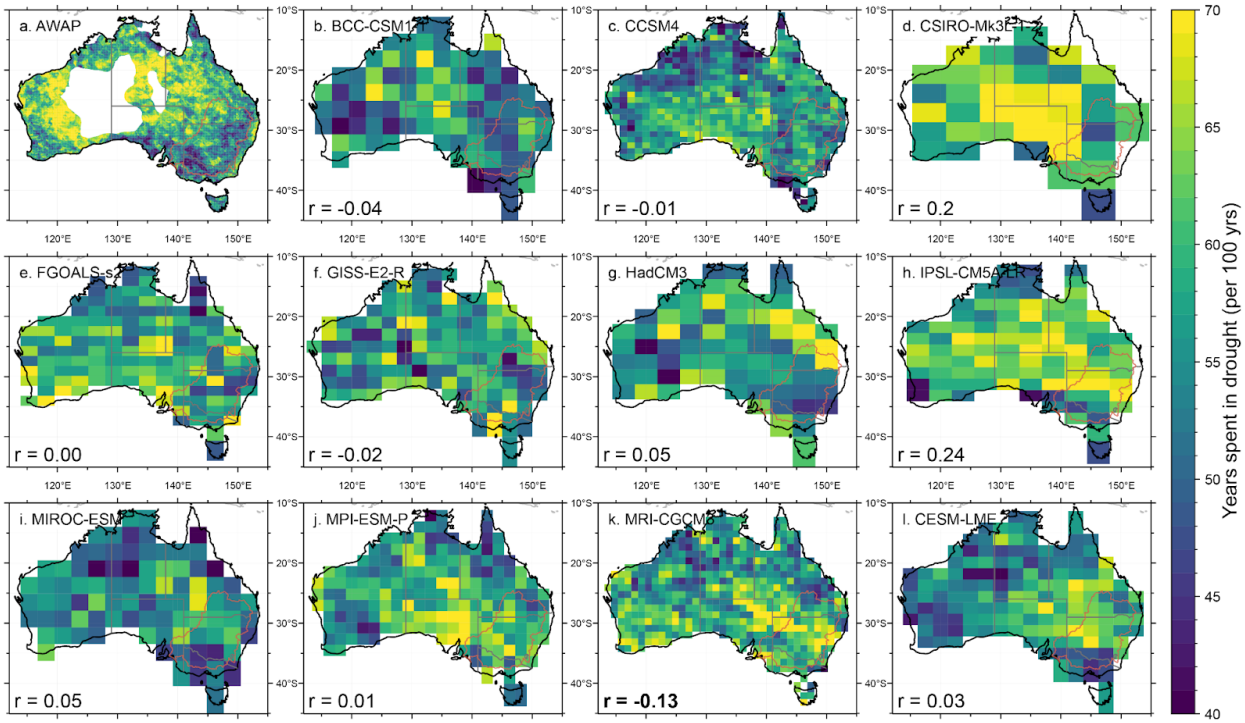
Supporting Fig 3: Relative intensity of multi-year droughts in (a) observations (1900-2000) and (b-l) model simulations of the same period. Showing the CESM LME ensemble mean. Spatial correlations of each model with observations are shown in the lower left corner in b-l. Bold text denotes a significant ($p < 0.05$) correlation.



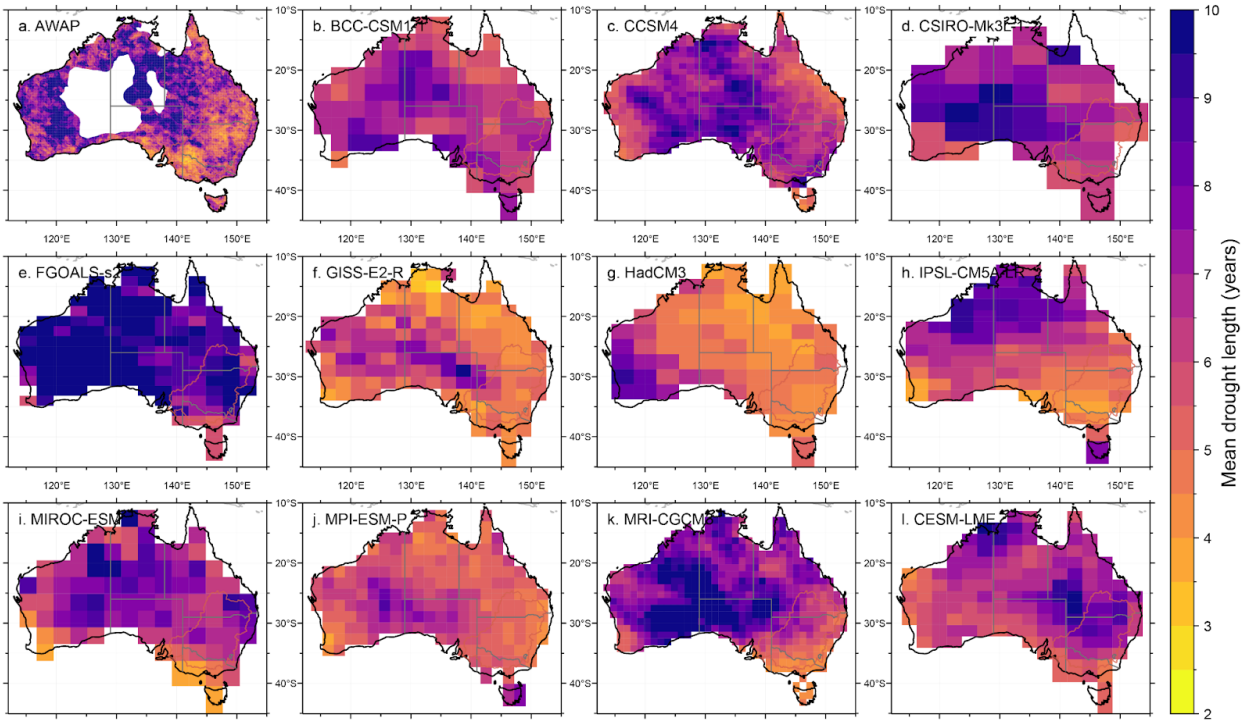
Supporting Fig 4: Relative severity of multi-year droughts in (a) observations (1900-2000) and (b-l) model simulations of the same period. Showing the CESM LME ensemble mean. Spatial correlations of each model with observations are shown in the lower left corner in b-l. Bold text denotes a significant ($p < 0.05$) correlation.



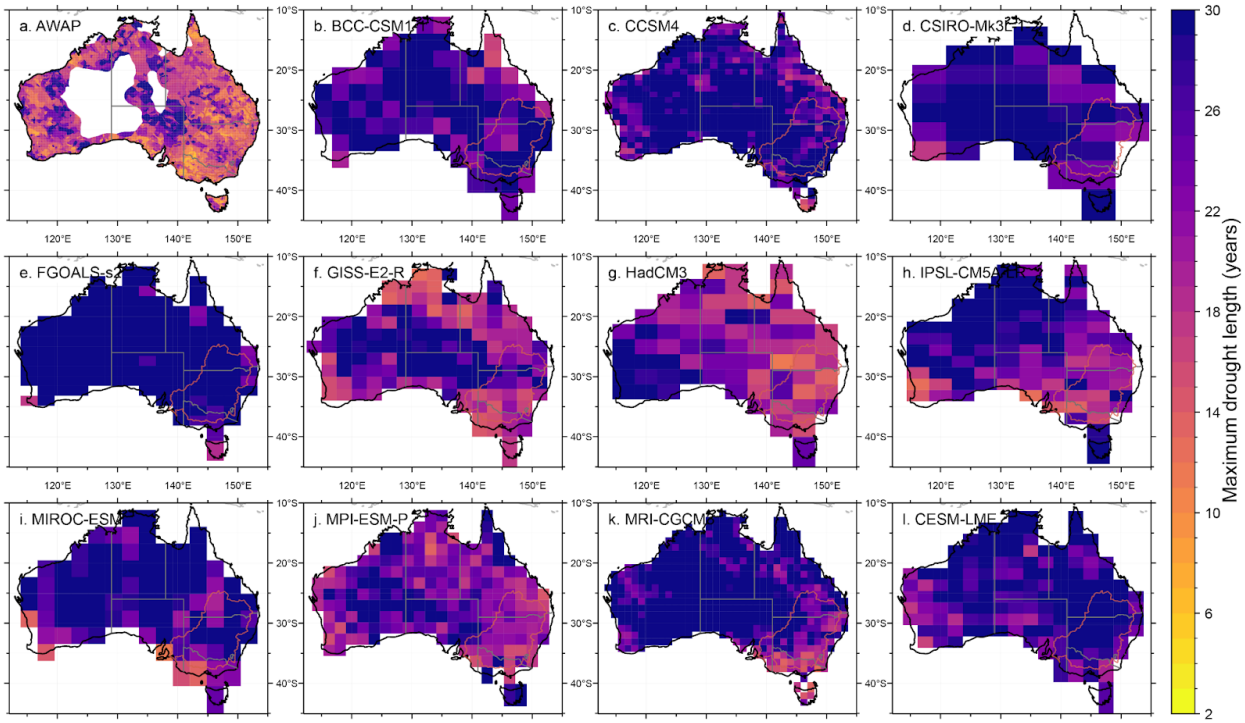
Supporting Fig 5: Frequency of multi-year droughts (per 100 years) in (a) observations (1900-2000) and (b-l) model simulations of the same period. Showing the CESM LME ensemble mean. Spatial correlations of each model with observations are shown in the lower left corner in b-l. Bold text denotes a significant ($p < 0.05$) correlation.



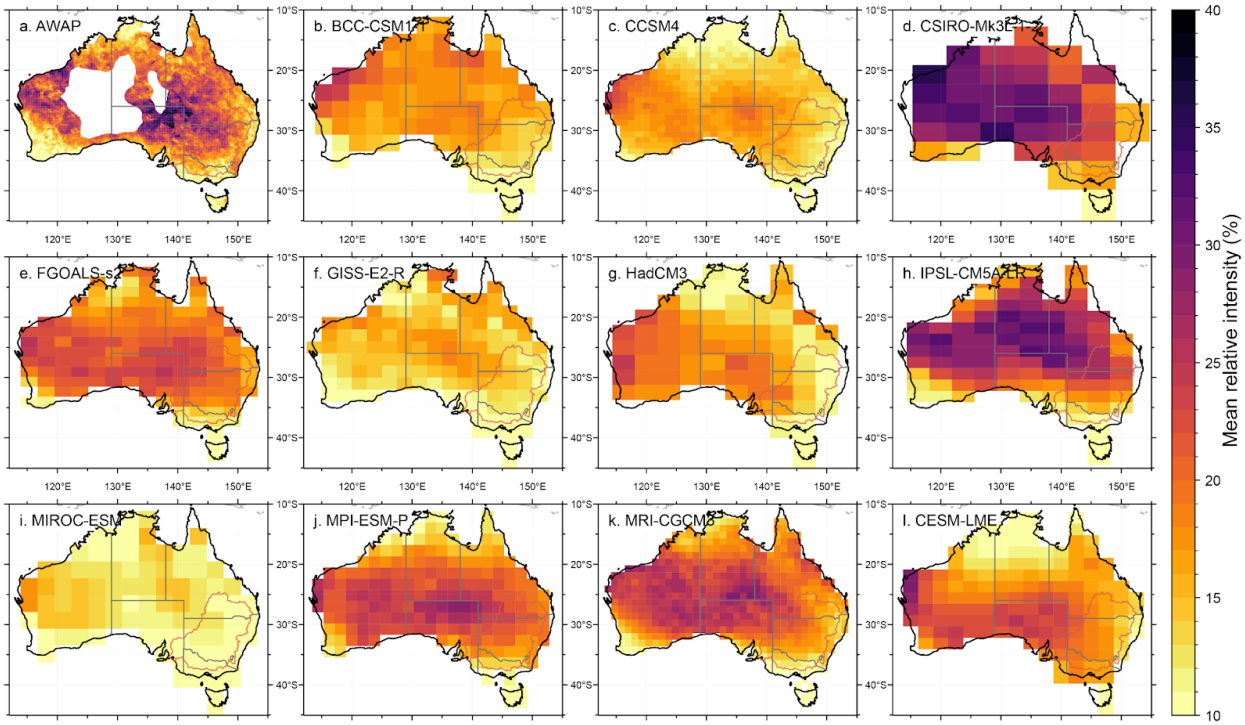
Supporting Fig 6: Total years spent in drought (per 100 years) in (a) observations (1900-2000) and (b-l) model simulations of the same period. Showing the CESM LME ensemble mean. Spatial correlations of each model with observations are shown in the lower left corner in b-l. Bold text denotes a significant ($p < 0.05$) correlation.



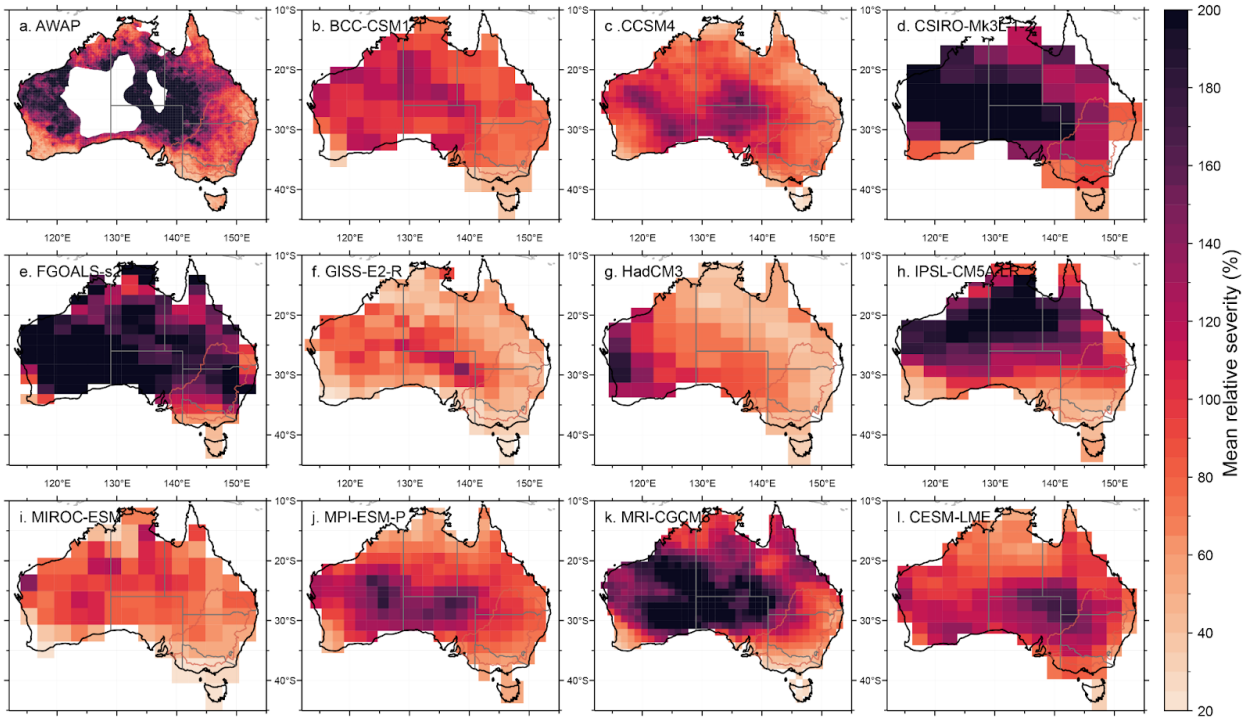
Supporting Fig 7: Mean multi-year drought length in (a) observations (1900-2000) and (b-l) model simulations of the pre-industrial last millennium (850-1849). Showing the CESM LME ensemble mean.



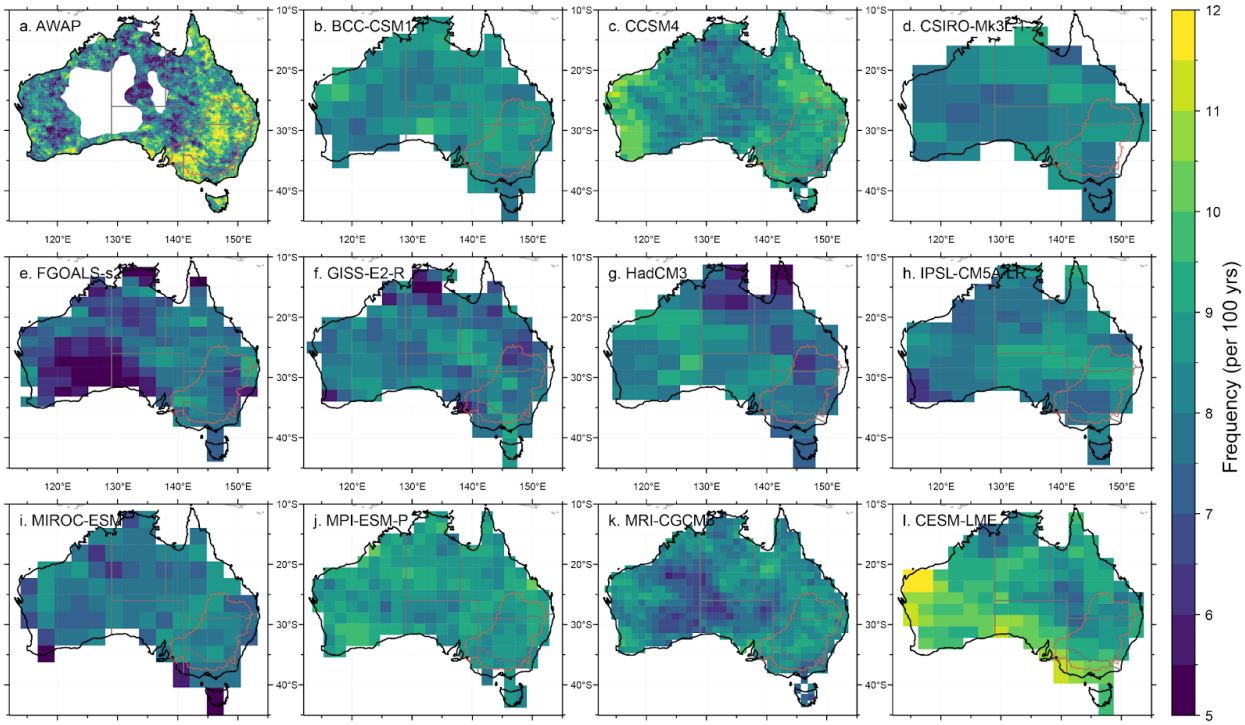
Supporting Fig 8: Maximum multi-year drought length in (a) observations (1900-2000) and (b-l) model simulations of the pre-industrial last millennium (850-1849). Showing the CESM LME ensemble mean.



35 **Supporting Fig 9**: Relative intensity of multi-year droughts in (a) observations (1900-2000) and (b-l) model simulations of the pre-industrial last millennium (850-1849). Showing the CESM LME ensemble mean.

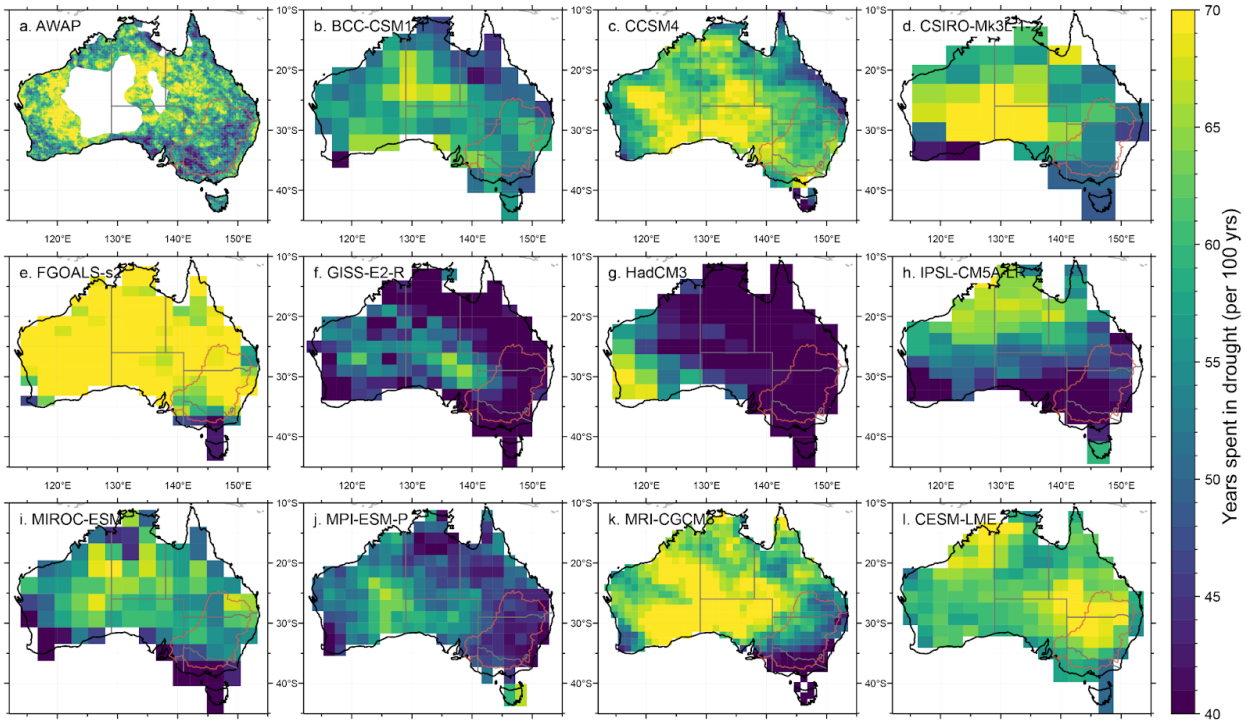


Supporting Fig 10: Relative severity of multi-year droughts in (a) observations (1900-2000) and (b-l) model simulations of the pre-industrial last millennium (850-1849). Showing the CESM LME ensemble mean.

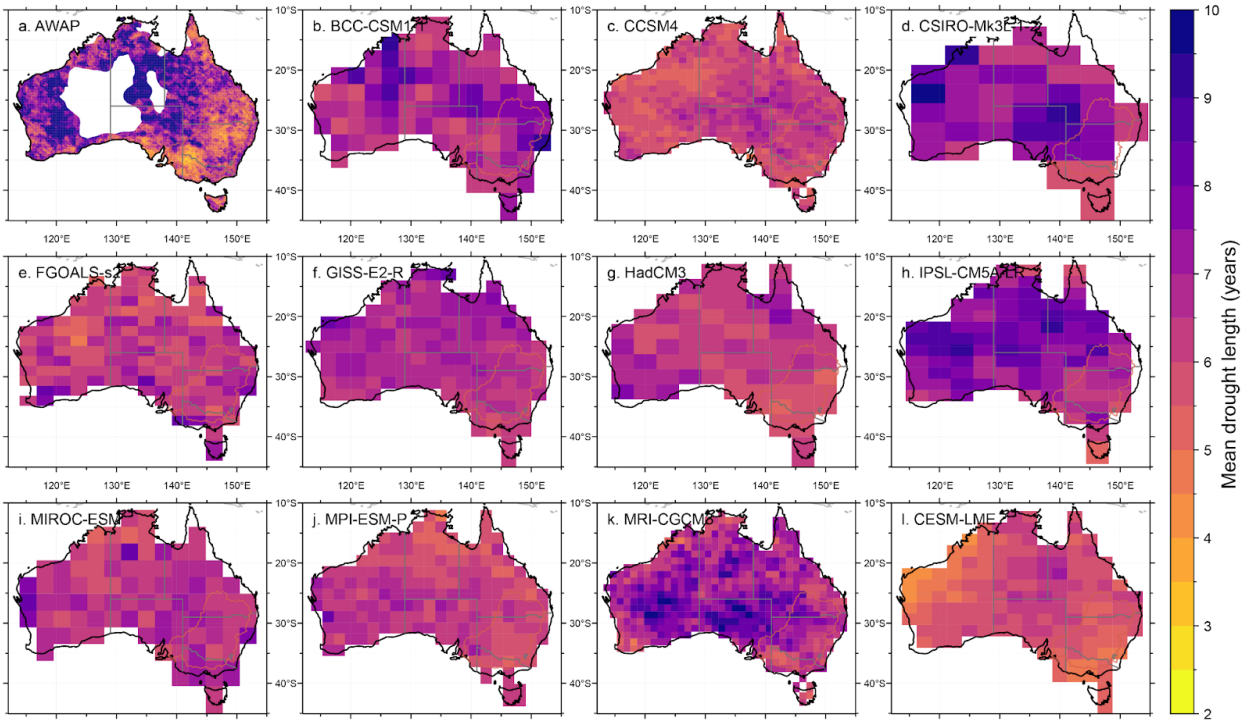


Supporting Fig 11: Frequency of multi-year droughts (per 100 years) in (a) observations (1900-2000) and (b-l) model simulations of the pre-industrial last millennium (850-1849). Showing the CESM LME ensemble mean.

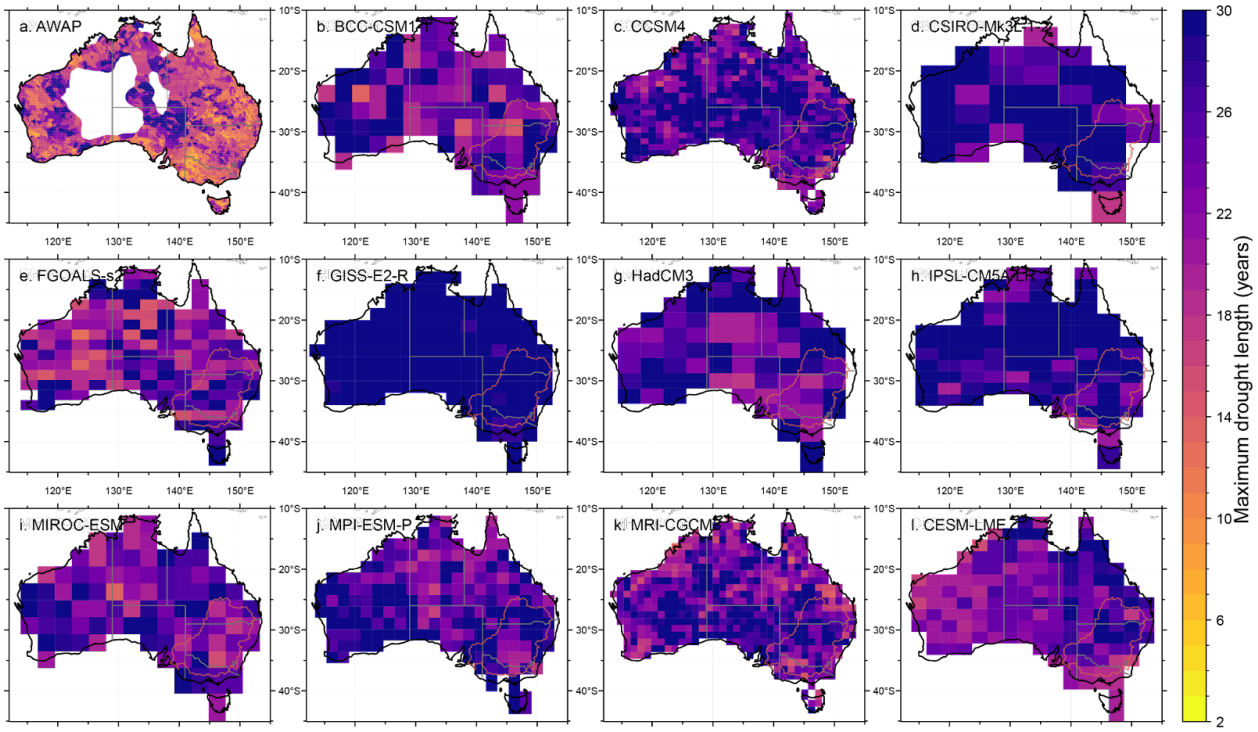
45



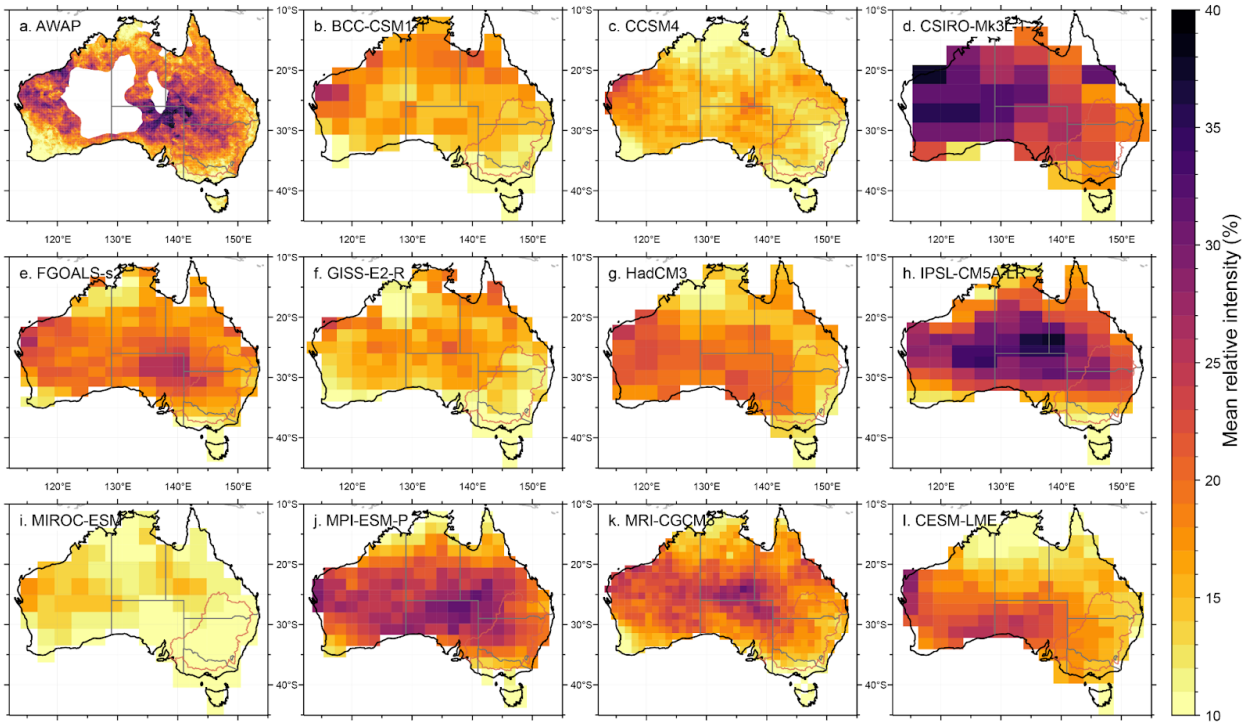
Supporting Fig 12: Total years spent in drought (per 100 years) in (a) observations (1900-2000) and (b-l) model simulations of the pre-industrial last millennium (850-1849). Showing the CESM LME ensemble mean.



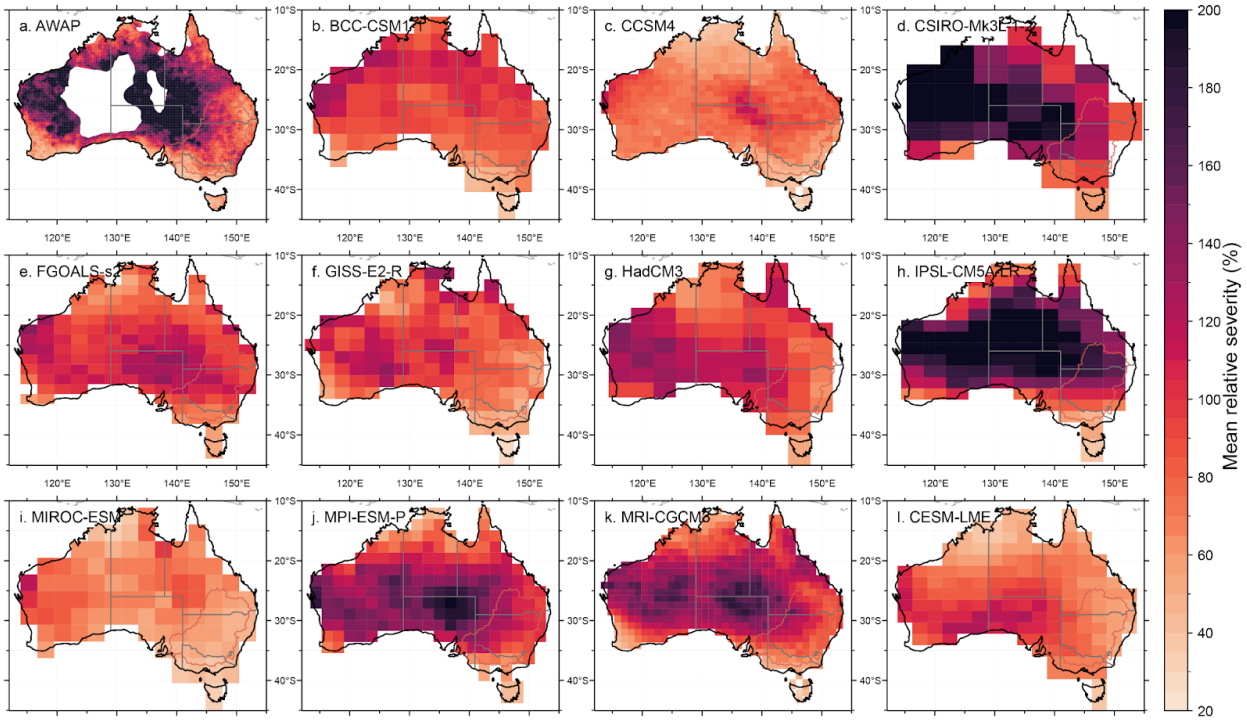
Supporting Fig 13: Mean multi-year drought length in (a) observations (1900-2000) and (b-l) piControl simulations . Showing the CESM LME ensemble mean.



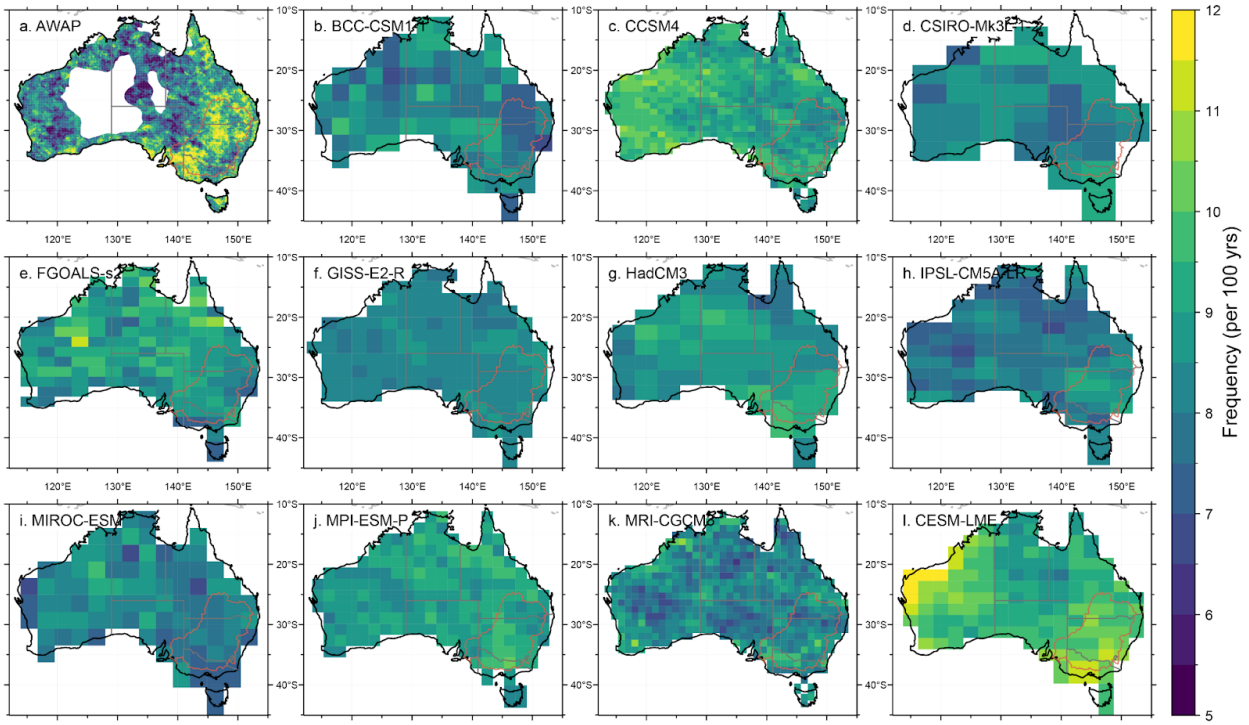
Supporting Fig 14: Maximum multi-year drought length in (a) observations (1900-2000) and (b-l) piControl simulations. Showing the CESM LME ensemble mean.



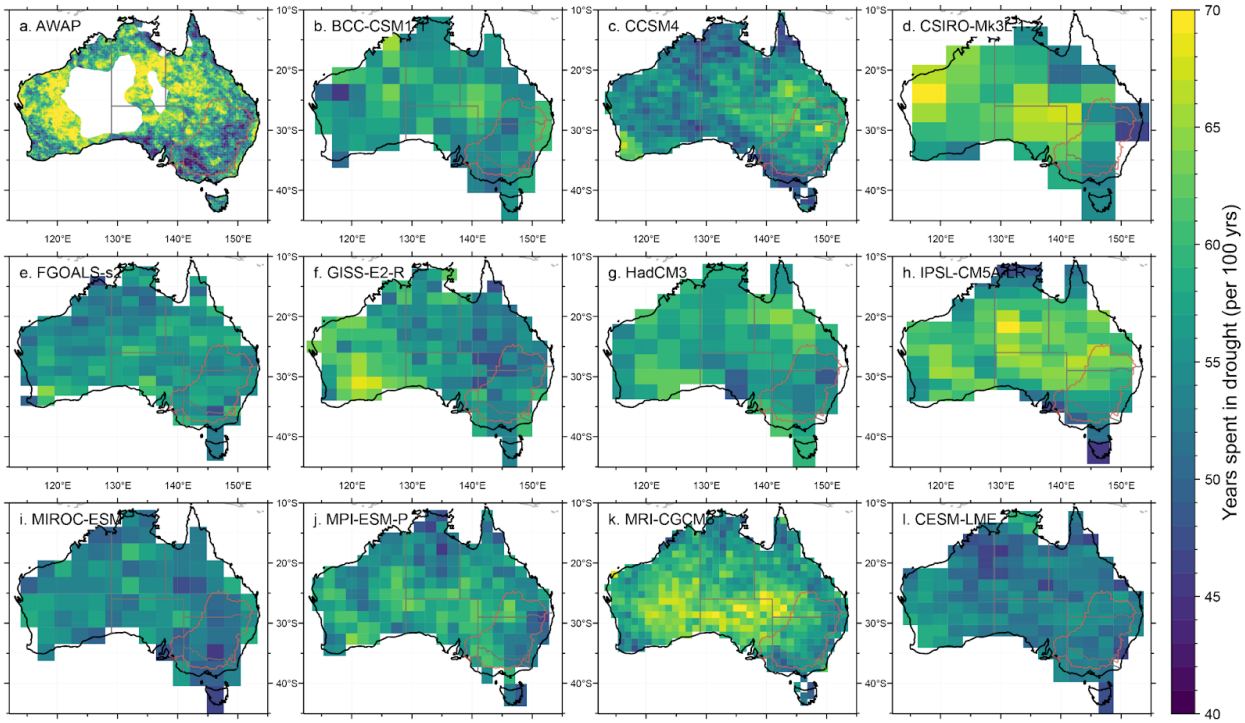
Supporting Fig 15: Relative intensity of multi-year droughts in (a) observations (1900-2000) and (b-l) piControl simulations. Showing the CESM LME ensemble mean.



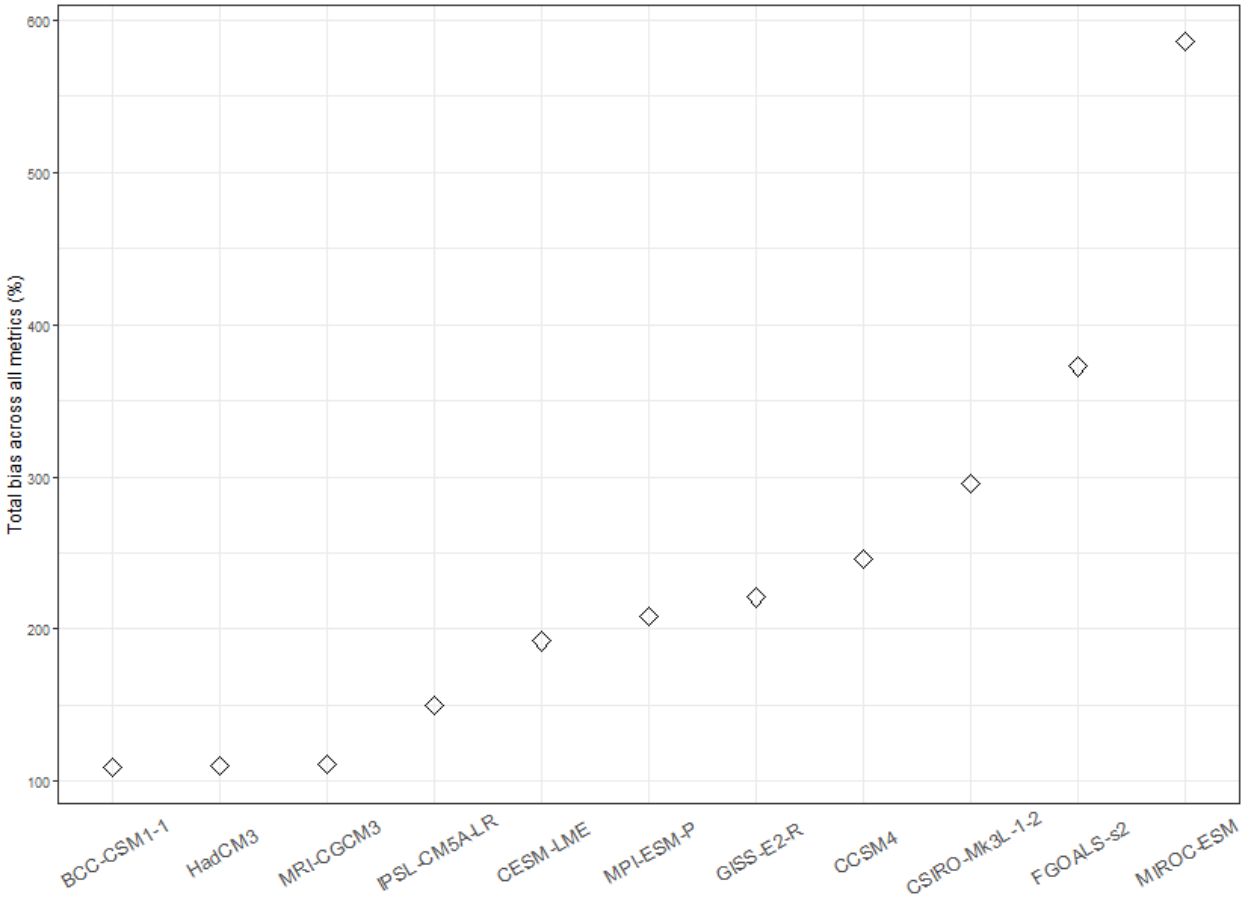
Supporting Fig 16: Relative severity of multi-year droughts in (a) observations (1900-2000) and (b-l) piControl simulations. Showing the CESM LME ensemble mean.



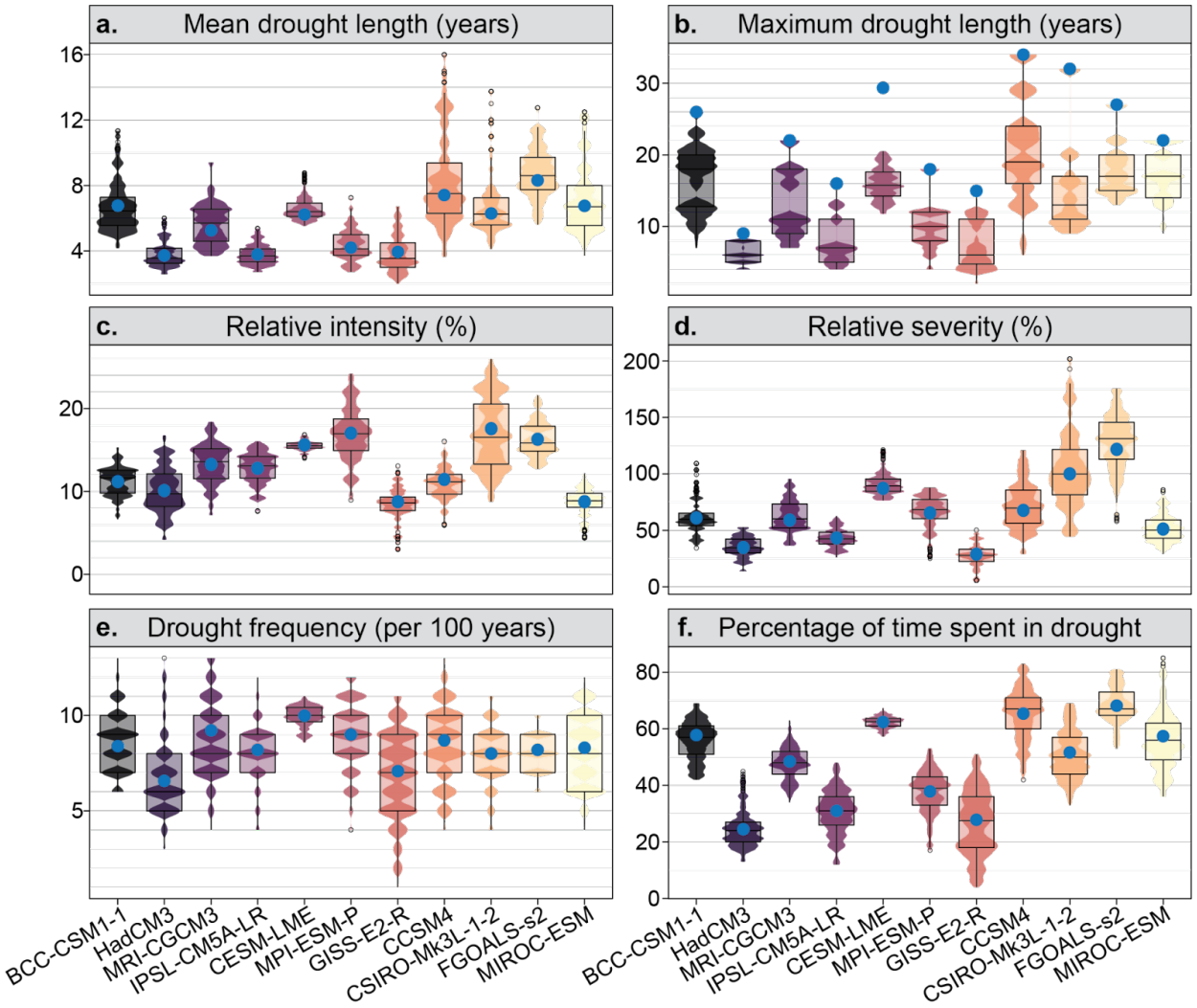
Supporting Fig 17: Frequency of multi-year droughts (per 100 years) in (a) observations (1900-2000) and (b-l) piControl simulations. Showing the CESM LME ensemble mean.



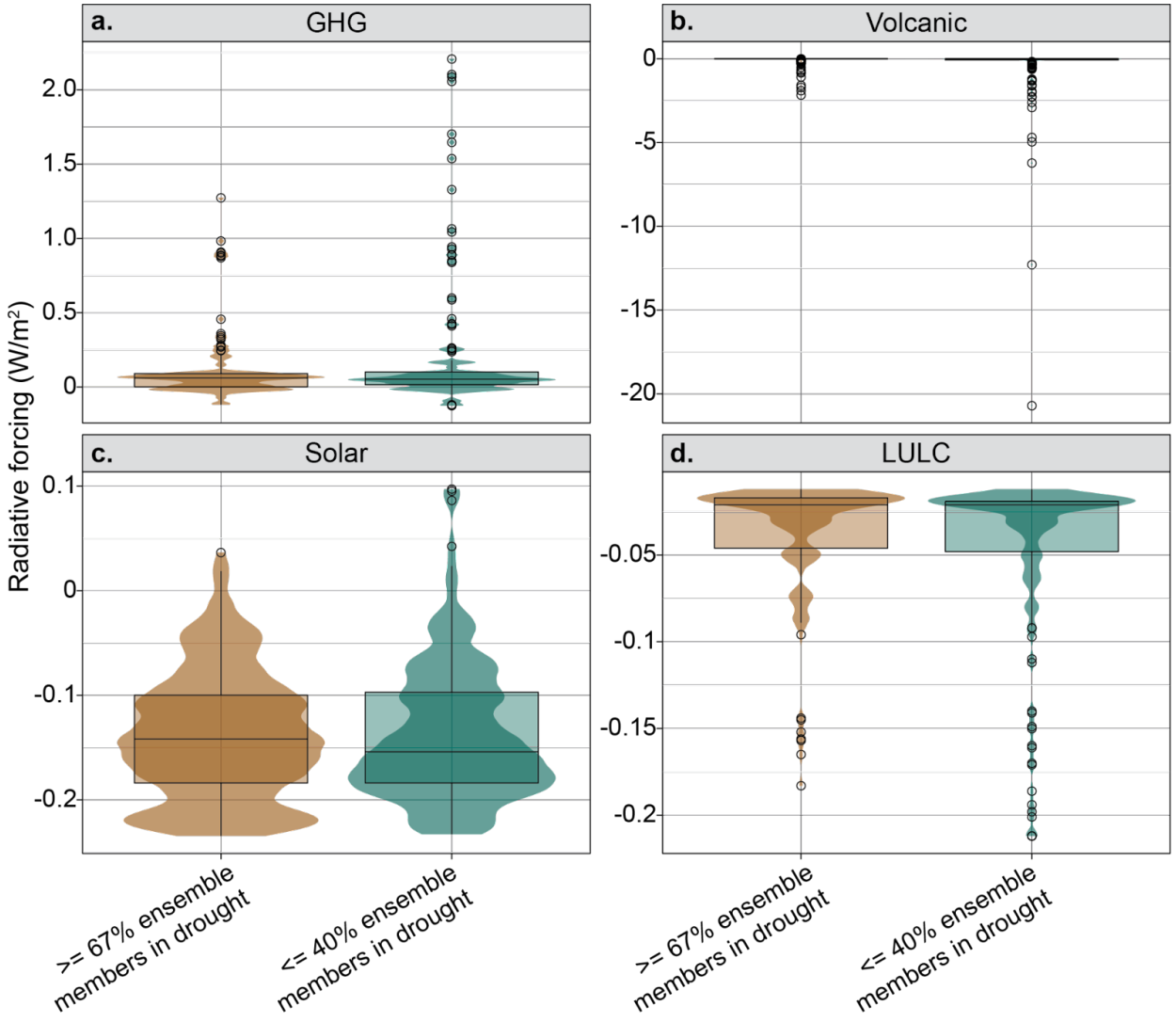
Supporting Fig 18: Total years spent in drought (per 100 years) in (a) observations (1900-2000) and (b-l) piControl simulations. Showing the CESM LME ensemble mean.



Supporting Fig. 19. Summed percent bias for PMIP3/CMIP5 models and the CESM LME, simulating drought metrics (CV, MAP, mean length, maximum length, intensity, severity, frequency, and proportion of time in drought) in the MDB. Percent bias values for each model were calculated relative to the observed value in the twentieth century (1900-2000). Total bias was calculated using the absolute value for each metric. Showing the ensemble mean for the CESM LME. Models are ordered from lowest to highest total bias along the x-axis; this ordering corresponds to the ordering of models in Fig. 6.



85 **Supporting Fig 20**. Comparison of MDB drought metrics calculated on the LM simulations from each model (850-1849) versus shorter intervals subsampled from the full LM simulations. Violin-and-boxplots ('voxplots') show the distribution of values for each drought metric, for 500 randomly-sampled 101-year segments of the 1000-year-long LM simulations, for each model. Large blue spots show the mean value for the *full* pre-industrial last millennium, for each model. Showing CESM LME ensemble mean. Models are ordered according to their summed 90 percent bias across all drought metrics in the twentieth century (as per Fig. 8).



Supporting Fig 21. Magnitudes of external radiative forcings during periods where most ensemble members in the full CESM LME (n=30) are in drought (brown voxplots) or not in drought (green voxplots). In the brown voxplots, at least 20 of the 30 ensemble members are in drought. In the green voxplots, 12 or fewer of the 30 ensemble members are in drought. This represents the upper and lower 10% of possible percentages of ensemble members in drought in any one year (i.e., colours in main text Fig. 10b). GHG = well-mixed greenhouse gases. LULC = land use/land cover changes.