Geophysical Research Letters

Supporting Information for

Observed emergence of the climate change signal: from the familiar to the unknown

E. Hawkins[1], D. Frame[2], L. Harrington[2,3], M. Joshi[4], A. King[5], M. Rojas[6], and R. Sutton[1]

[1] National Centre for Atmospheric Science, Dept. of Meteorology, University of Reading, UK.

[2] New Zealand Climate Change Research Institute, Victoria University of Wellington, New Zealand.

[3] Environmental Change Institute, University of Oxford, South Parks Road, Oxford, UK.

[4] Climatic Research Unit, School of Environmental Sciences, University of East Anglia, UK.

[5] School of Earth Sciences and ARC Centre of Excellence for Climate Extremes, University of Melbourne, Australia.

[6] Department of Geophysics and Centre for Climate and Resilience Research, CR2, University of Chile, Chile.

Contents of this file

S.1 Shifting distributions

The emergence of a signal can be visualised using shifting normal distributions (Fig. S1). *Frame et al.* (2017) described S/N>1 as a shift to an '*unfamiliar*' climate, S/N>2 as an '*unusual*' climate and S/N>3 as an '*unknown*' climate, in terms of an individual's lifetime. We add the term '*inconceivable*' for S/N>5, as the new mean climate would be experienced once every 3 million years in the old climate.

Two regional average examples are shown in Fig. S2, for tropical America and northern America, highlighting the differences in signal and noise characteristics. Even though northern America has a larger signal, the change is more apparent in tropical America.

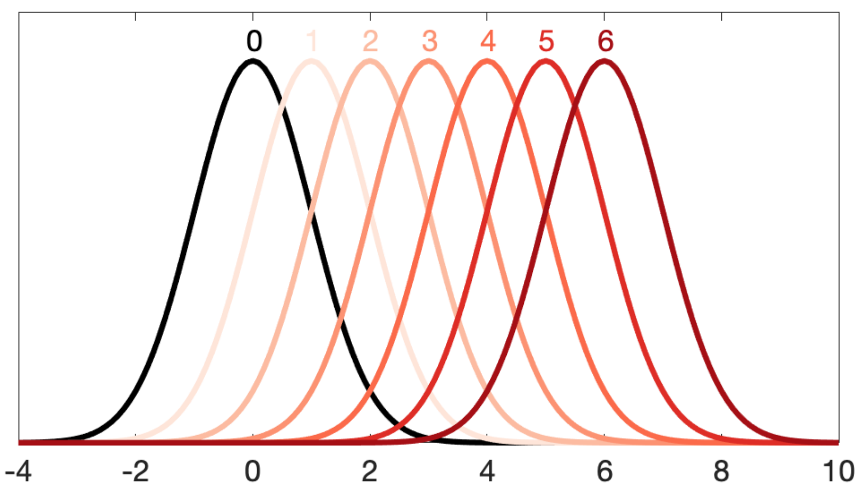

Figure S1: Shifting a normal distribution by 0 (black) to 6 (dark red) standard deviations.

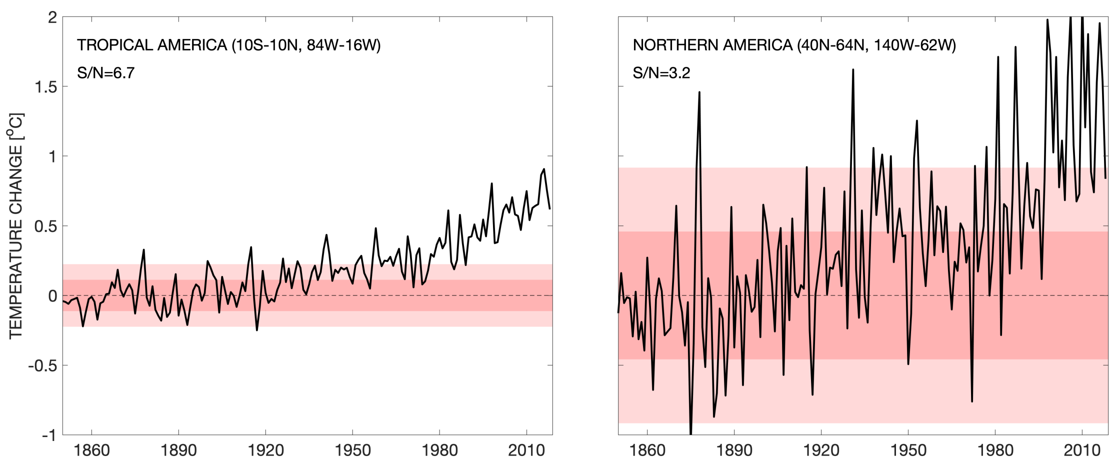

Figure S2: Two regional examples of how observed temperature changes have become apparent, using the Berkeley Earth land-only temperature dataset. The red shaded bands represent 1 and 2 standard deviations of the noise.

S.2 Using model simulations to test the emergence methodology

We can test the robustness of the methodology to estimate the S/N using a large ensemble of model simulations. *Maher et al.* (2019) describe the 100-member ensemble of the MPI GCM, from which we use the simulated SAT for the historical period (1850-2005), extended to 2018 with the RCP4.5 scenario. First, we apply the same methodology used for the observations to each ensemble member individually. The ensemble mean S/N, which is expected to be smoother than the observed S/N due to averaging, is shown in Fig. S3a, and the spread in S/N across the ensemble is shown in Fig. S3c. The uncertainty in S/N is generally between 0.2-0.4 over land, which is typically far smaller than the mean S/N. The maritime continent, North Atlantic and Southern Ocean are regions with largest uncertainty in this GCM. The percentage uncertainty in S/N is less than 30% over most land areas (Fig. S3d). A simpler approach, which is not possible using observations, is to calculate the S/N by averaging the simulated temperature anomaly patterns in 2018, relative to the mean of 1850-1900, from all ensemble members, and dividing by the standard deviation of the 2018 anomalies (Fig. S3b). This pattern is virtually identical to Fig. S3a, highlighting that the regression approach produces S/N estimates that are robust. These results also demonstrate that the uncertainty in S/N due to simulated internal variability is relatively small.

Note that the patterns of simulated S/N in this ensemble are noticeably different from the observed patterns. One important example is in parts of west Africa where the MPI ensemble S/N is close to zero but is larger than 5 in the observations. India also has a low S/N in the ensemble, but significant values in the observations. This finding highlights the benefit of using the observations alone, as in the current study.

Fig. S4 shows the same maps for simulated precipitation change in the MPI ensemble. Again, the two methods produce similar patterns (Fig. S4a, b), with the ensemble method showing slightly larger values. The simulated uncertainty in S/N due to internal variability is typically 0.3-0.4 over land regions. The patterns are again different from that derived from the observations, especially in west Africa which is significantly wetter in the simulations but drier in the observations.

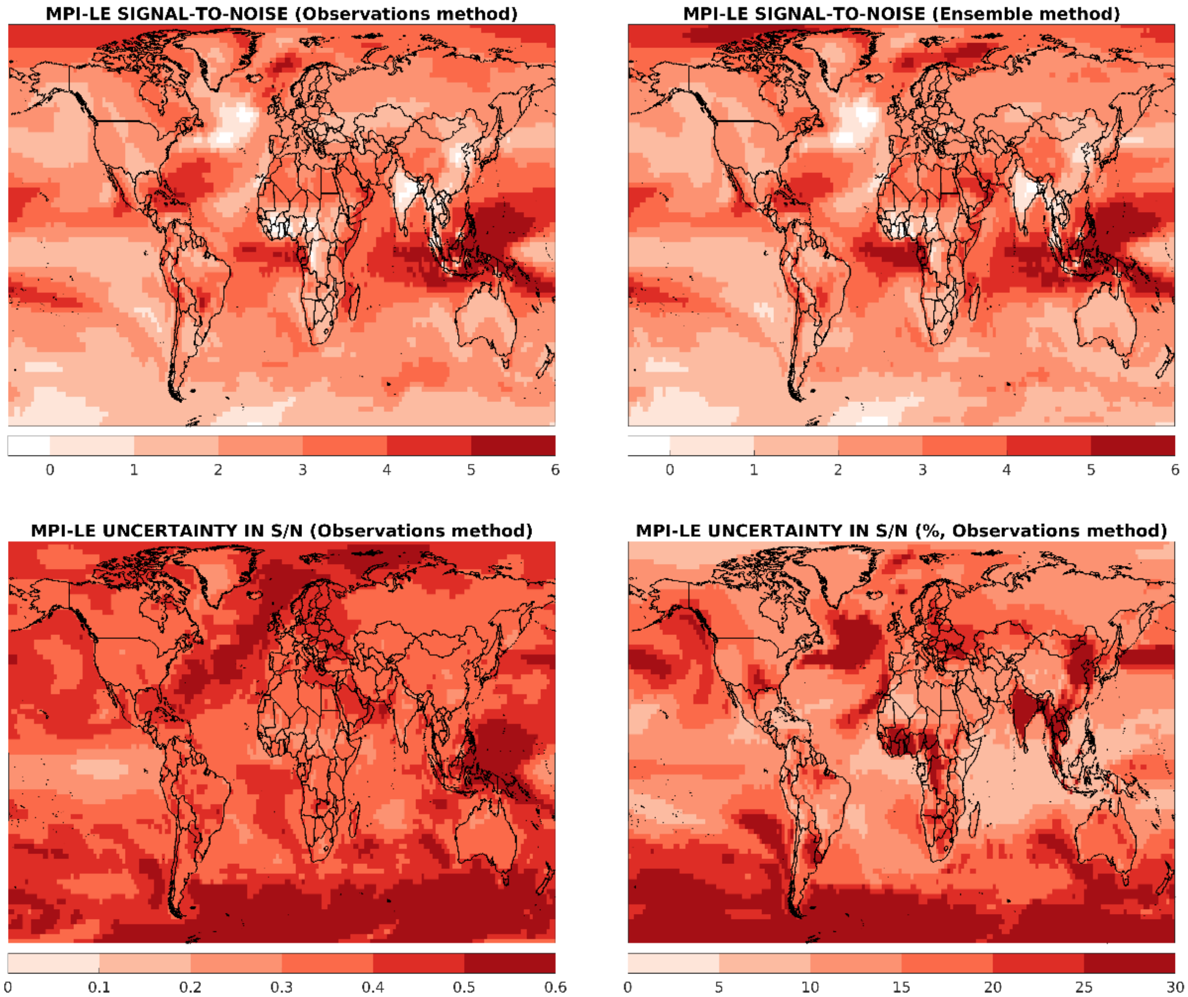

Figure S3: Testing the S/N methodology using the MPI Large Ensemble (*Maher et al.* 2019). (top left) S/N calculated as for the observations in each individual ensemble member, averaged across the 100-members. (top right) Mean simulated temperature in 2018 minus the average of 1850-1900 across all ensemble members, divided by the standard deviation of simulated temperature in 2018. (bottom left) Standard deviation in the S/N estimated using the observational method across the 100-members. (bottom right) The percentage uncertainty in S/N, i.e. bottom left panel divided by top left.

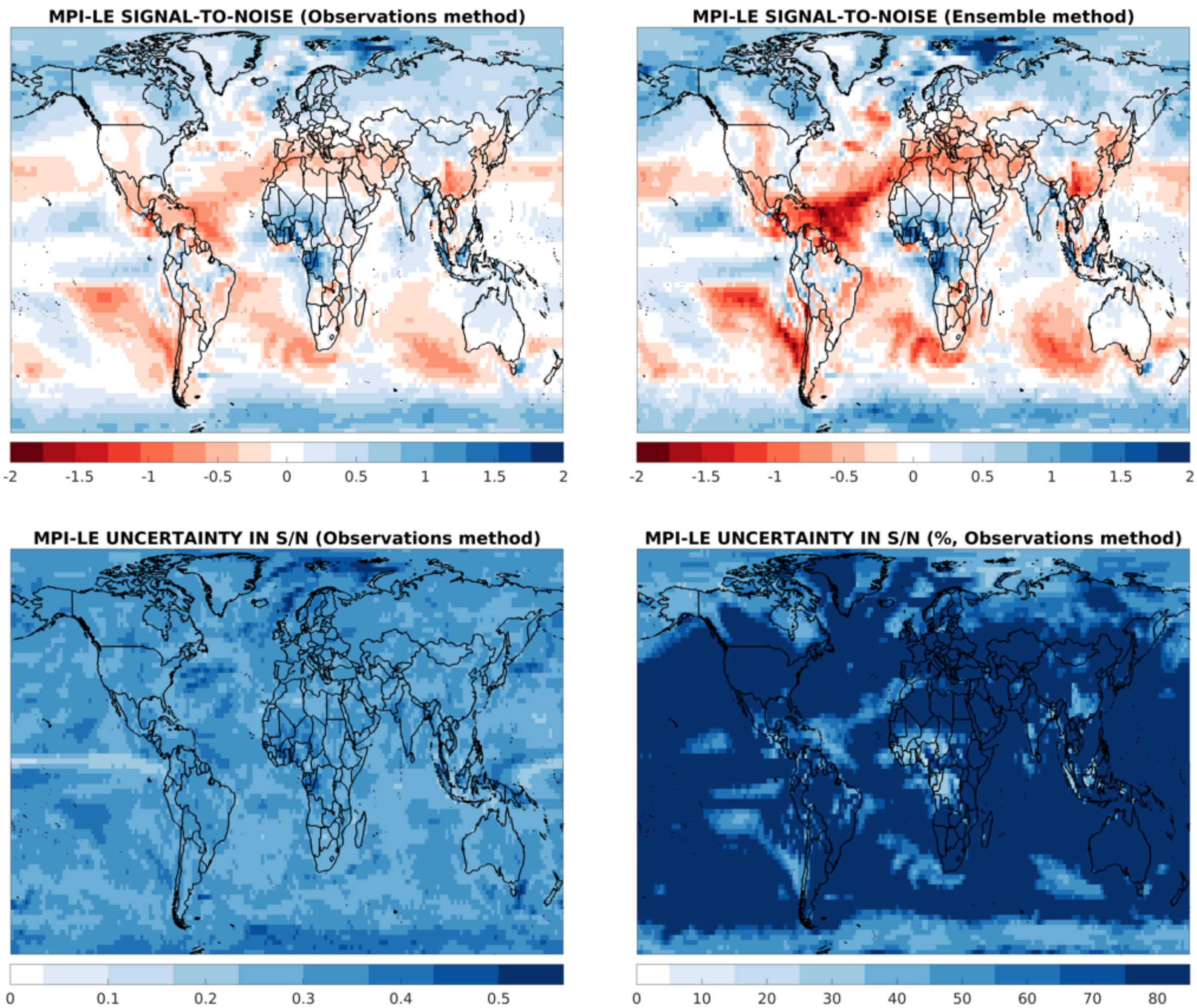

Figure S4: as Fig. S3 for precipitation**.**

S.3 Additional metrics

Figure S5 shows the fraction of land area which has a S/N for temperature exceeding the value indicated, using the Berkeley Earth dataset. For the annual mean, around 15% of the land area has a S/N larger than 5, and 40% shows a S/N larger than 2 for the warmest climatological month of the year. The warmest months tend to show larger S/N values than the coldest months.

Figure S6 repeats the S/N temperature analysis using other datasets: HadCRUT4 (*Morice et al.* 2012), *Cowtan & Way* (2014, hereafter CW14) infilled version of HadCRUT4, GISTEMP (*Lenssen et al.* 2019) and NOAA GlobTemp (*Zhang et al.* 2019). For this sensitivity test we have used the same smoothed GMST from Berkeley Earth in all cases. These datasets generally produce similar patterns to that from Berkeley Earth (Fig. 2c), but with varying amplitudes. NOAA GlobTemp has larger S/N values in the tropics than the other datasets and Berkeley Earth has larger S/N for the south-east USA. There are other notable differences for west Africa and parts of south America, mainly due to different estimates for the signal, rather than the noise (not shown). There is consistent agreement that the tropical Atlantic and Indian Oceans exhibit the highest S/N for the ocean areas, and that there has been very little warming overall in the central North Atlantic.

Figure S7 shows the S/N patterns for precipitation in different seasons, highlighting that the west Africa signals are present in all seasons except DJF, and the south-west Australia drying signal is mainly present in JJA. The wetter northern latitude signal is mainly present in DJF and MAM.

Figure S8 shows the S/N patterns for UK mean precipitation in different seasons. There are tendencies towards wetter seasons, except for JJA where the S/N is rarely significant. Note that the observed signal in southern UK is for drier summers but it has not yet emerged.

Figure S9 shows the UK mean RX1day time-series with maps for two example years.

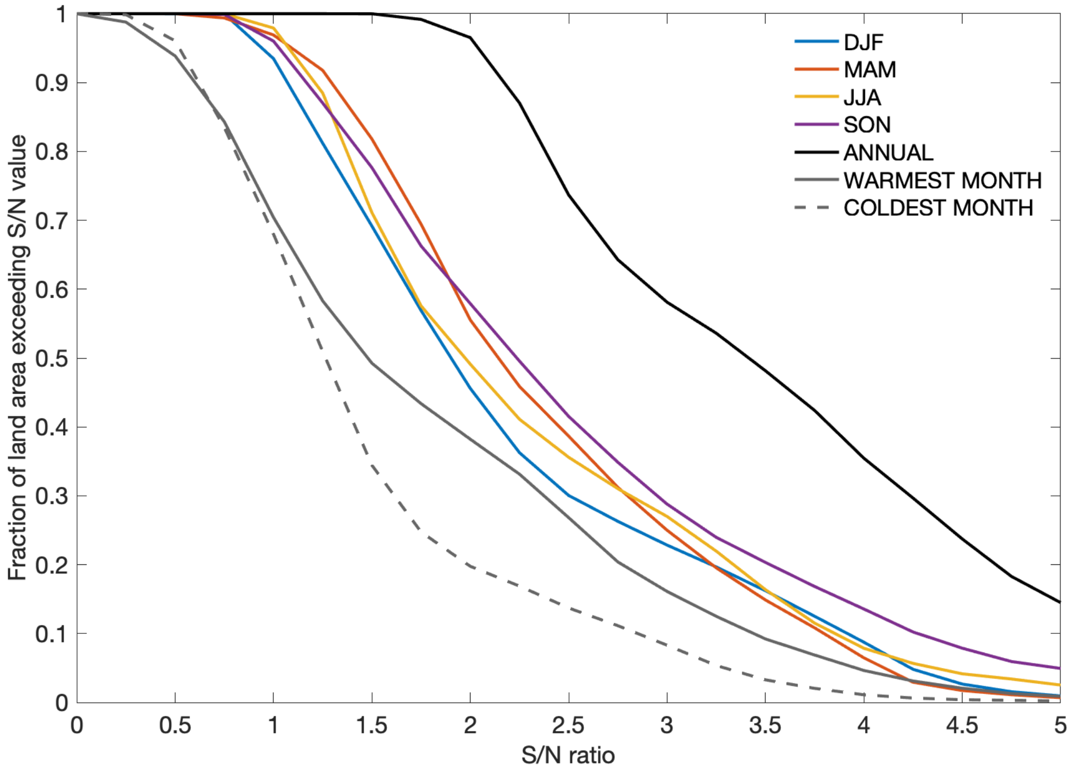

Figure S5: The fraction of land area with an observed temperature S/N larger than the ratio shown, for different seasons, the annual average, and warmest and coldest months (using the Berkeley Earth dataset).

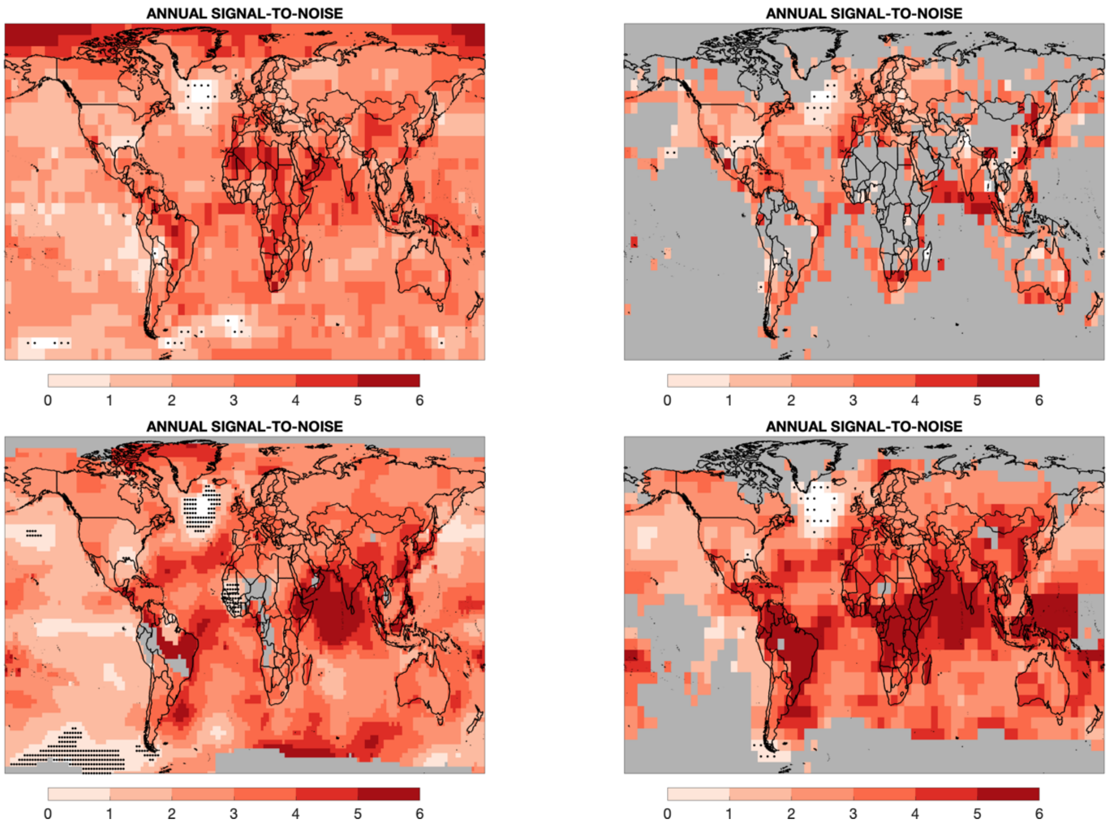

Figure S6: Observed S/N for temperature using the CW14 dataset (top left), HadCRUT4 (top right), GISTEMP (bottom left) and NOAA GlobTemp (bottom right). Stippled cells indicate that the regression coefficient is not statistically significant. Grey regions are where there is less than 100 years of data in that location for that dataset.

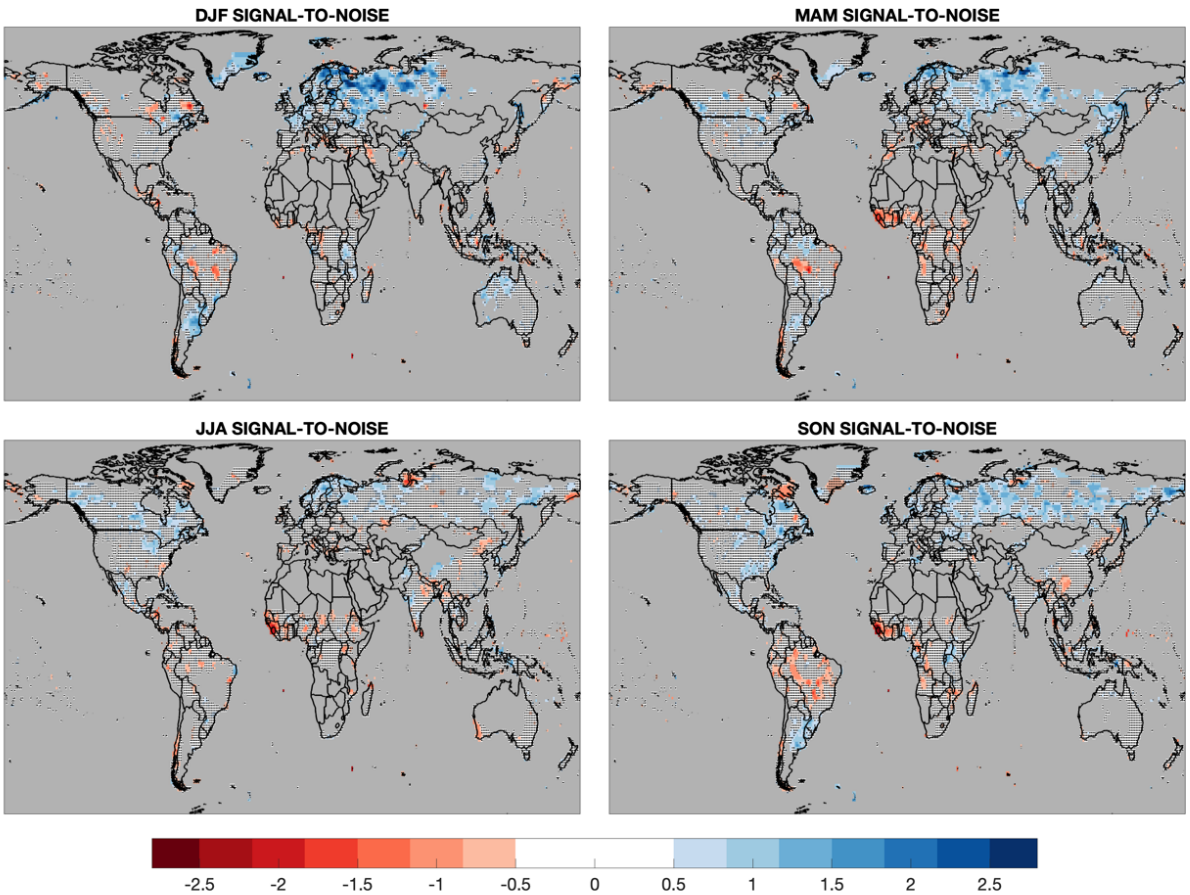

Figure S7: Signal-to-noise for precipitation in different seasons. Grey regions are either unobserved (oceans), have a seasonal precipitation of less than 62.5mm or annual precipitation less than 250mm. Stippled regions denote areas where the regression parameter is not statistically significant.

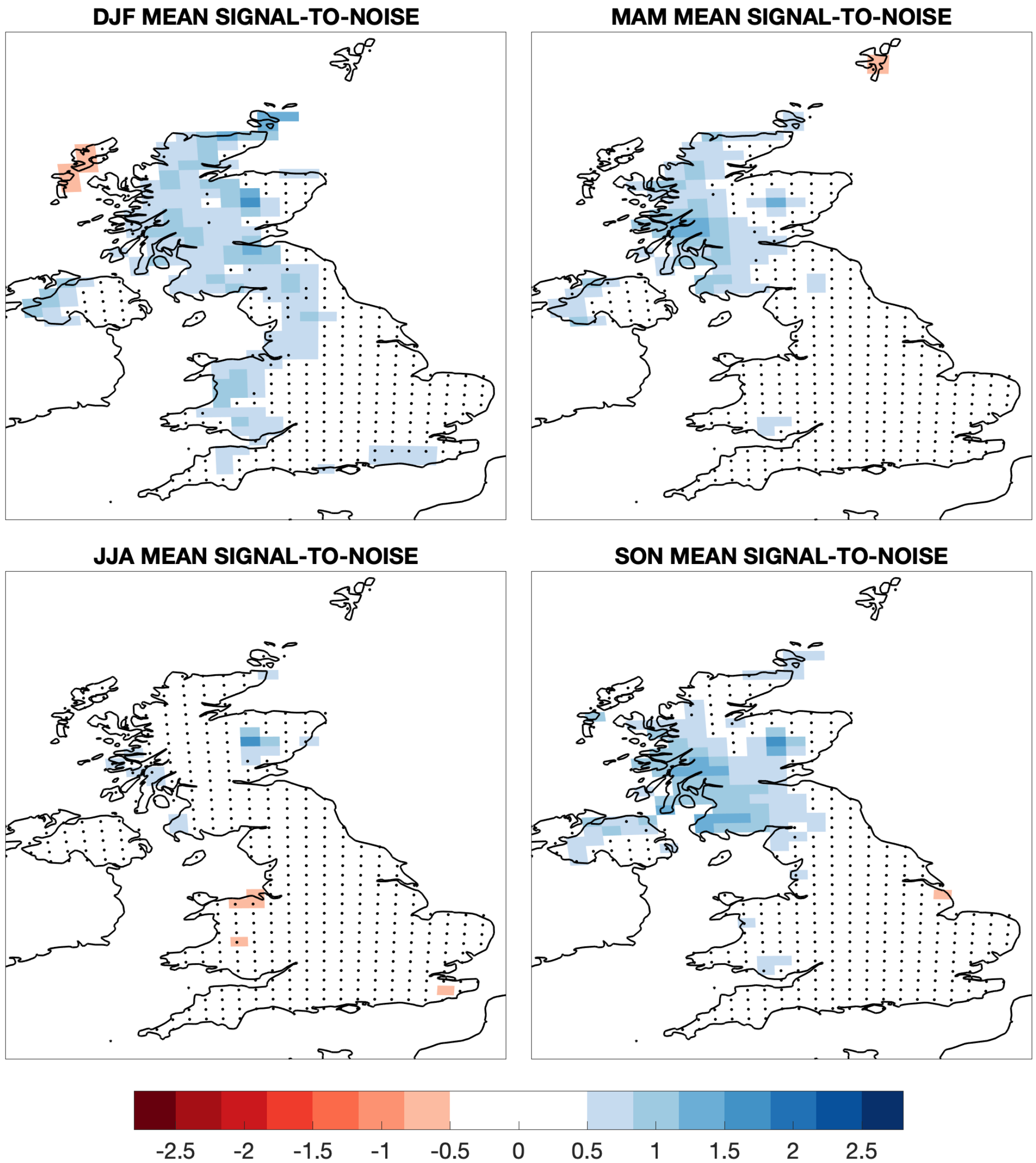

Figure S8: Signal-to-noise for UK mean precipitation in different seasons. Stippled regions denote areas where the regression parameter is not statistically significant.

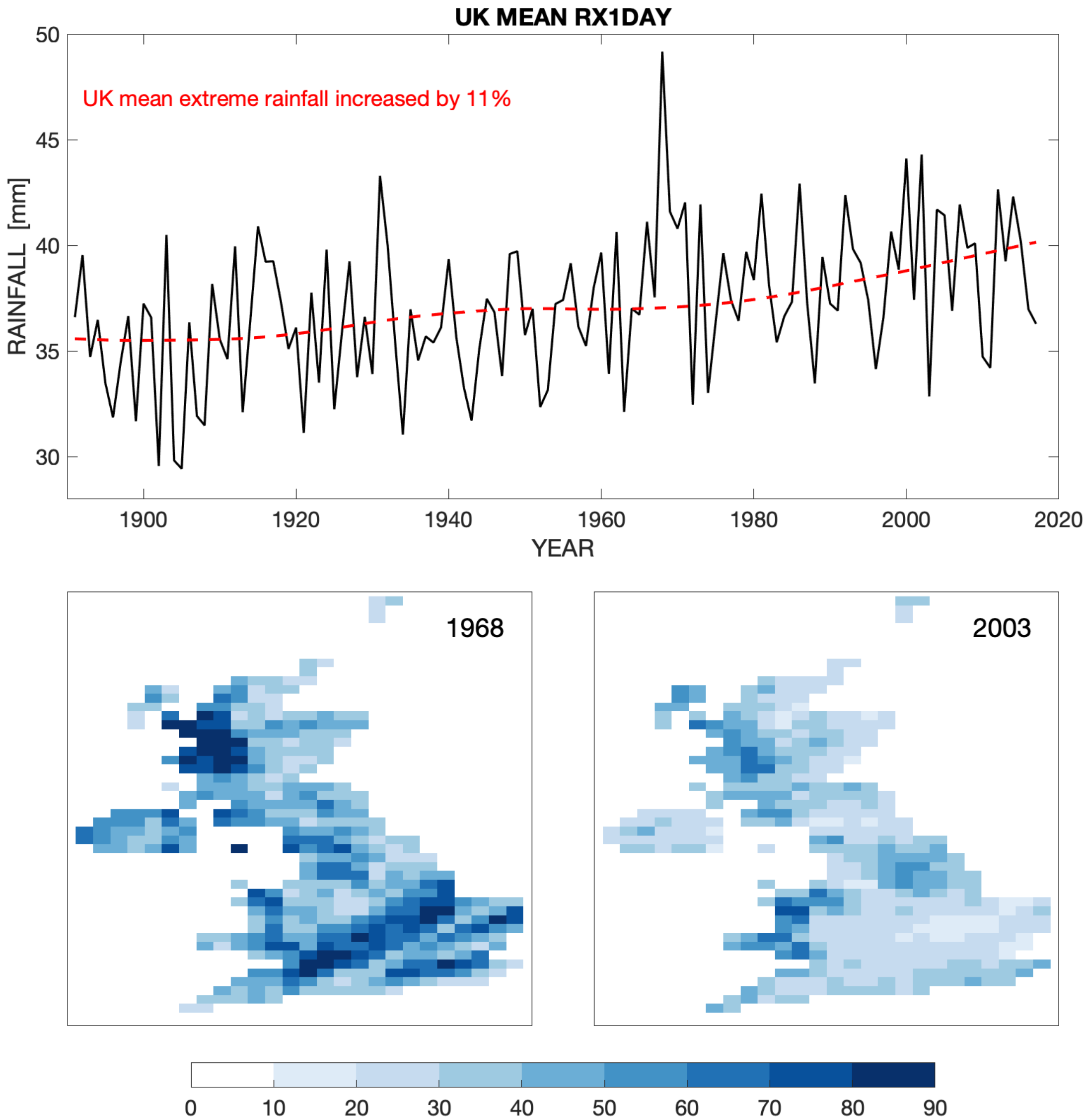

Figure S9: UK extreme rainfall (RX1day, mm): average across the UK (1891-2017, black line) with regression on smoothed GMST (red dashed line), and maps for two example years (1968 and 2003). 1968 shows the effect of three significant storm events, in contrast to 2003 which mainly shows larger rainfall over higher orographic features.