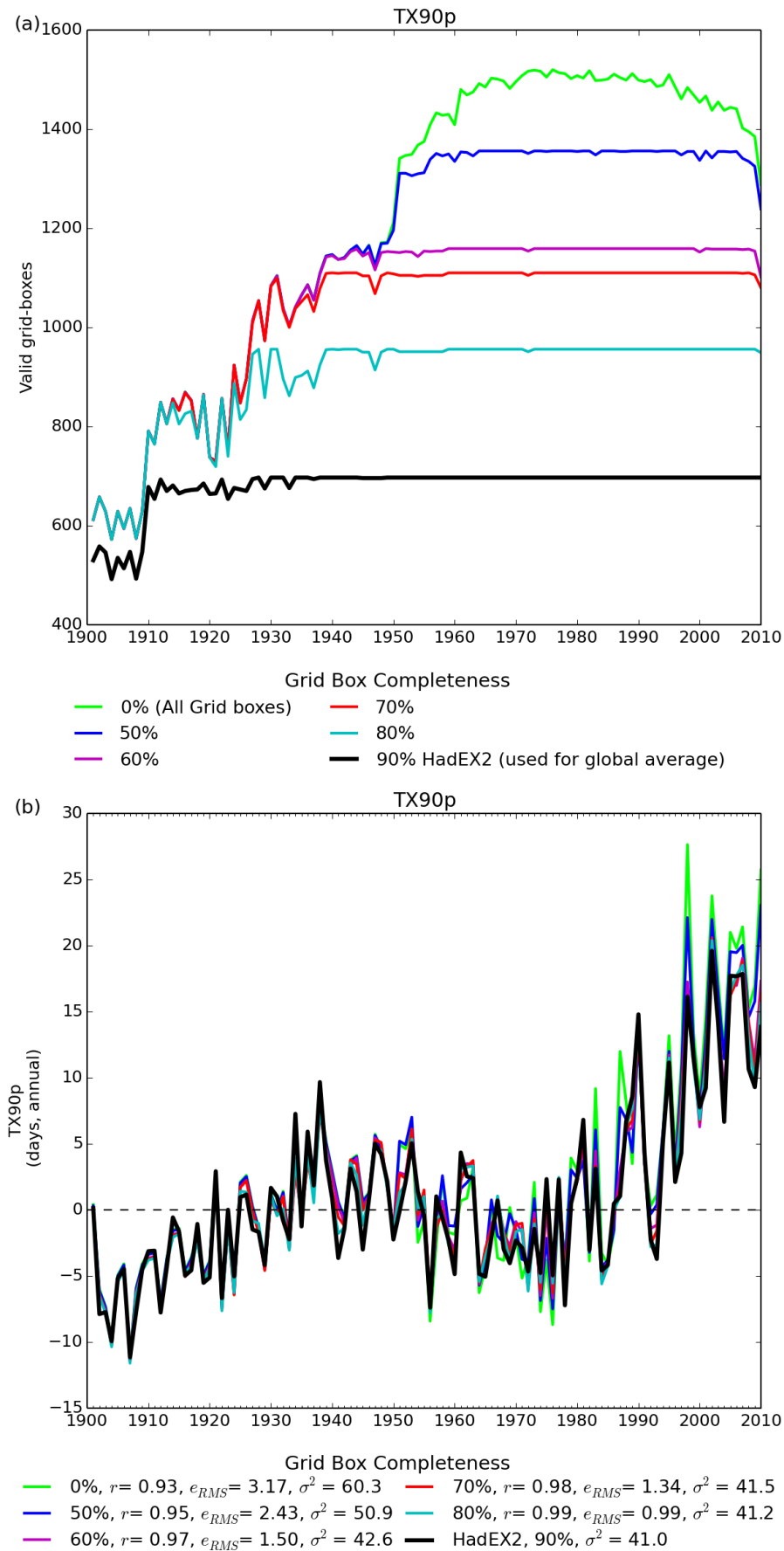


Figures for TX90p

November 10, 2014



2

Figure 1: **(a)** The coverage, showing the different curves for different completeness requirements, for TX90p. The total coverage is shown the curve in light green, whereas the coverage used when calculating global averages is in black, with a range of other completeness percentages shown in between. **(b)**: The time series for global average of TX90p. Also shown are the comparison of other completenesses to that used in HadEX2 (90 %) using the correlation coefficient, r , the root mean square error e_{ems} and the variance, σ^2 .

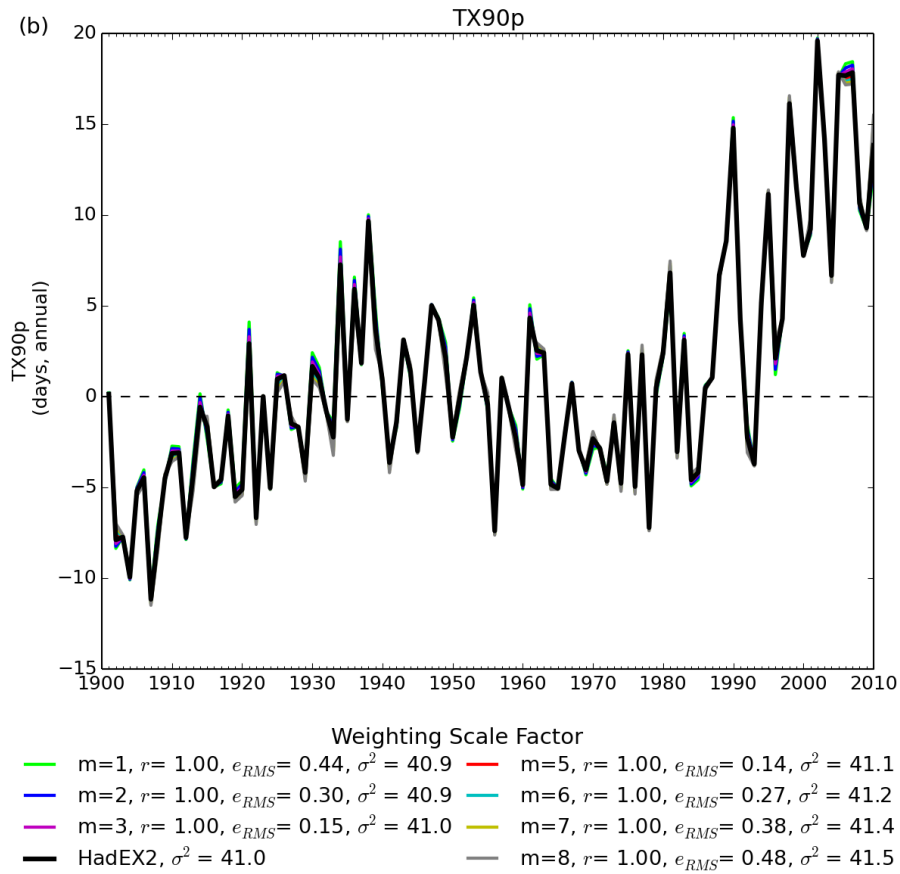
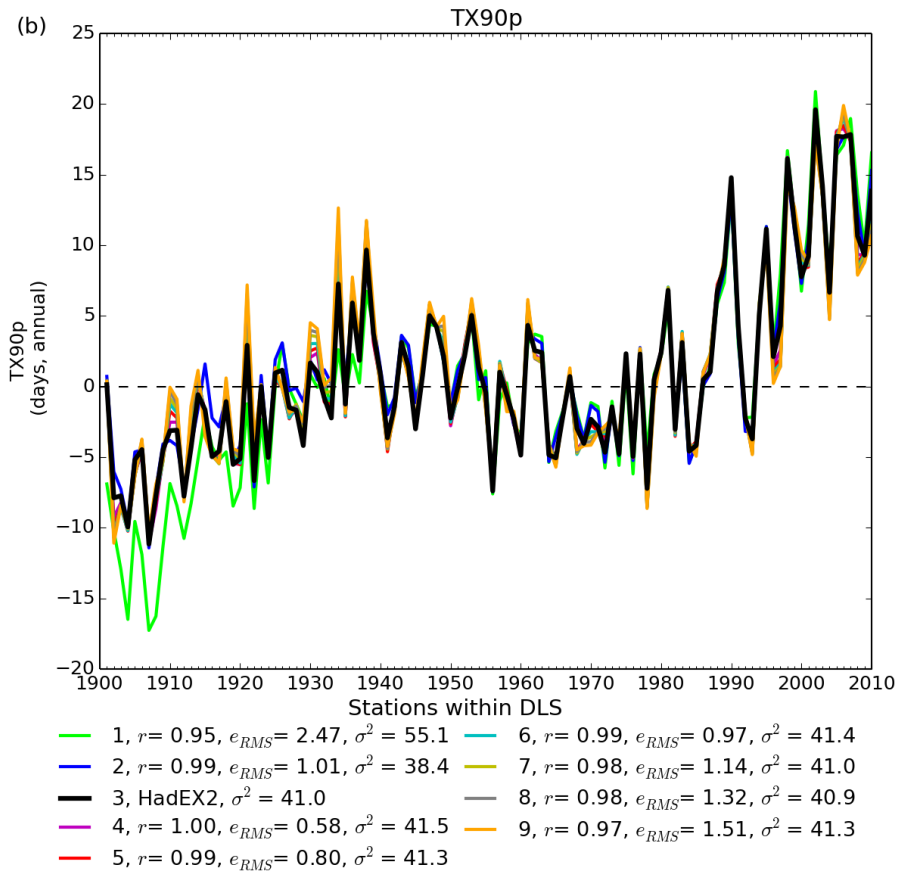
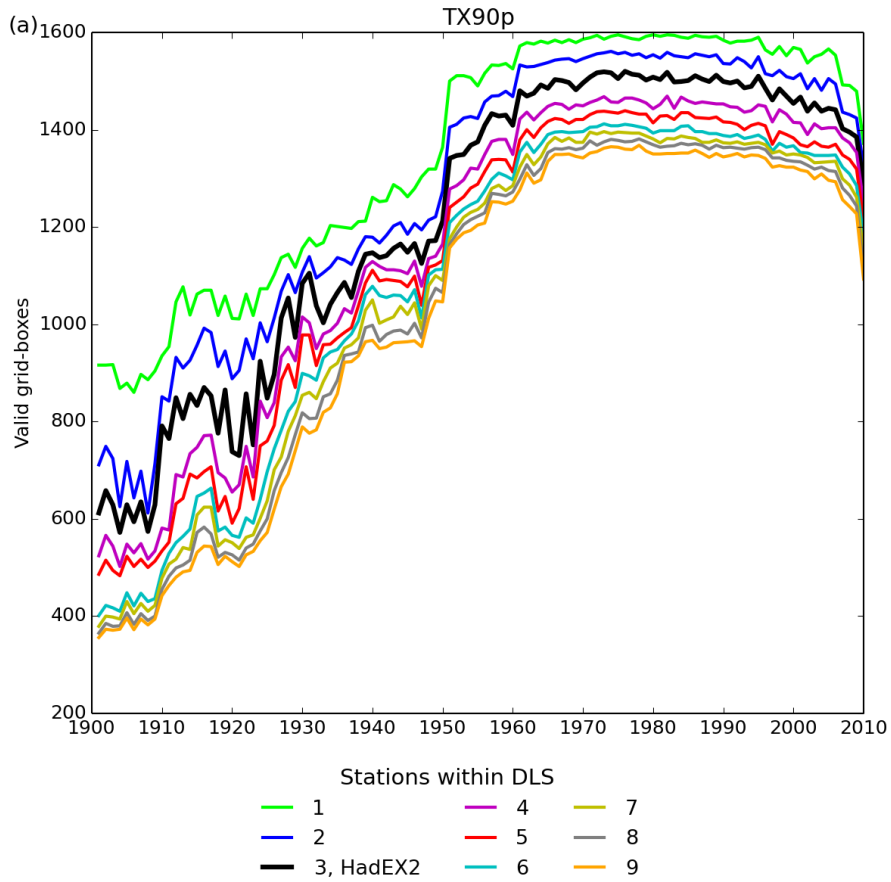


Figure 2: The time series for global average TX90p showing the different curves for the different weighting values (colours are indicated below the plot). The figure also shows the correlation coefficient (r), variance (σ^2) and root-mean-square error (e_{RMS}) between each series and HadEX2. All choices have been normalised over the 1961-1990 period.



4

Figure 3: **(a)**: The number of non-missing grid boxes for TX90p showing the different curves for left numbers of stations within a DLS of the grid box centre. Note that the global average only takes grid boxes which have 90 per cent completeness or more, whereas all grid boxes are shown in the coverage series. **(b)**: The time series for global average TX90p.

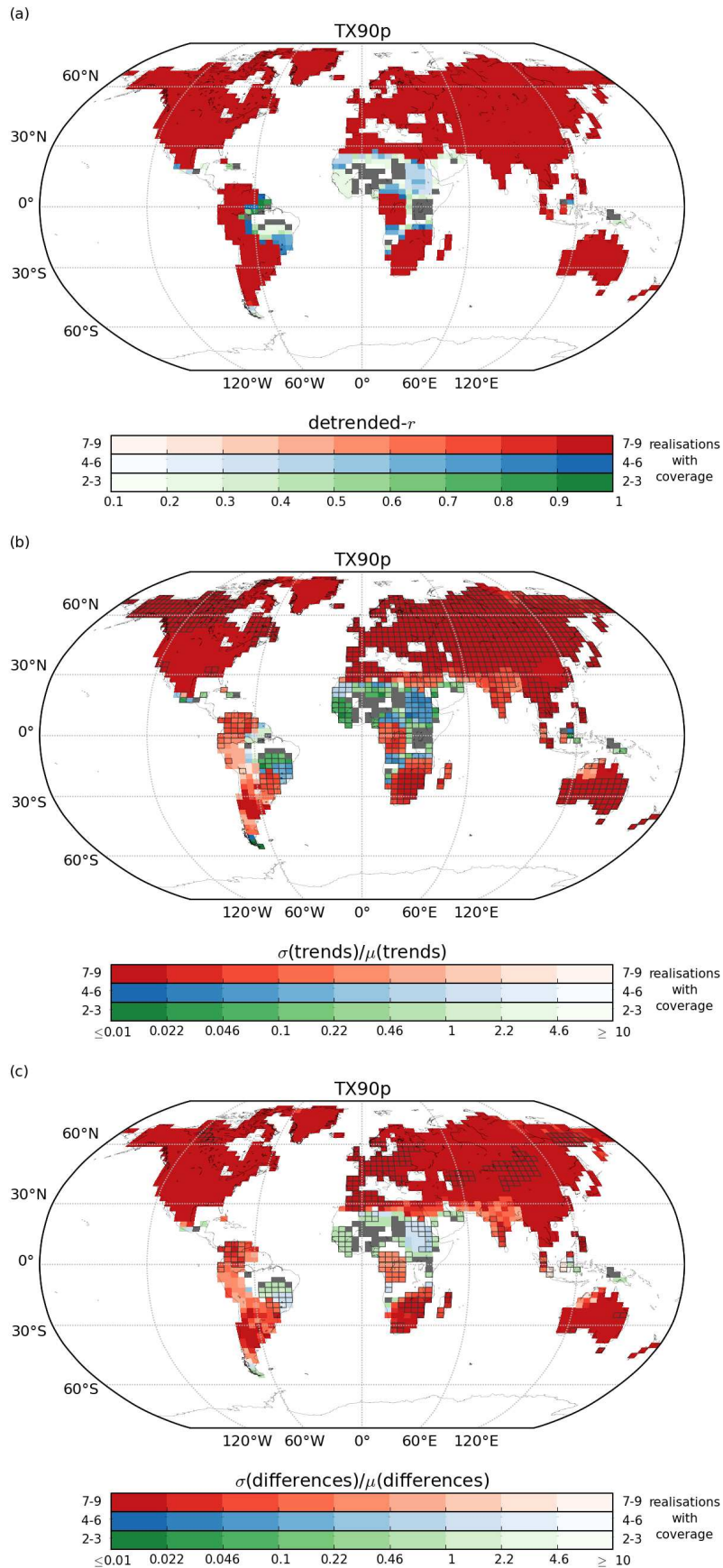
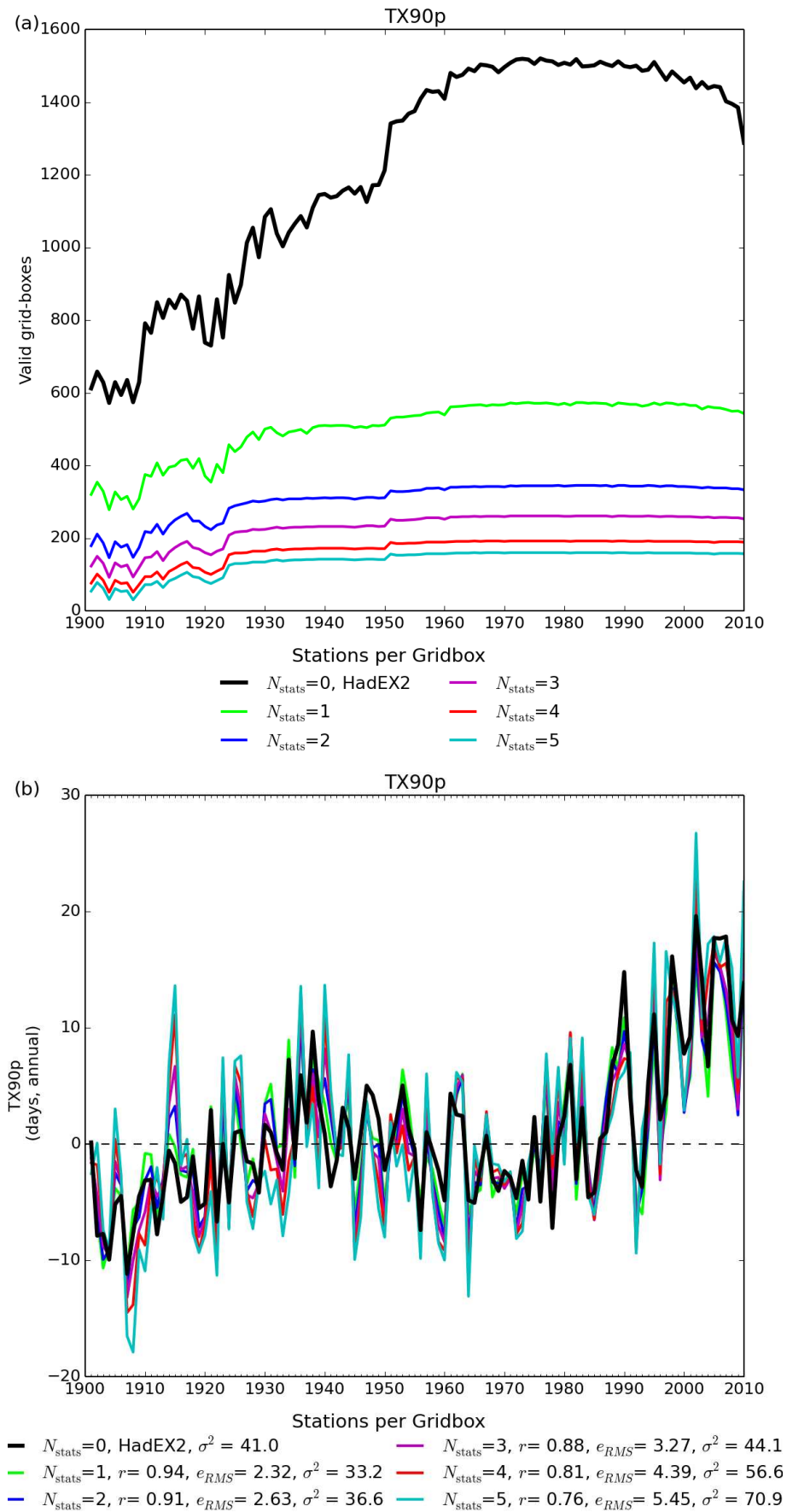


Figure 4: **(a)**: The mean detrended correlation coefficient, r for all the grid boxes; **(b)**: the standard deviation of the trends normalised by the mean trend (σ/μ) calculated for the period 1951 to 2010, and **(c)**: the standard deviations of the differences normalised by the mean differences for TX90p, where the differences have been calculated between the averages over the periods 1991-2010 and 1951-1970. Grey grid boxes are those where only one of the 9 possible stations per DLS choices results in a value. The grid boxes are green when 2-3, blue when 4-6 and red when 7-9 of the possible choices result in a value. In (b) and (c), boxes which have been outlined are those where there is high confidence in a non-zero trend in HadEX2. For details on the calculation method, see text.



6

Figure 5: As for Fig. 3 but for the different numbers of stations within a grid box. Note that the global average only takes grid boxes which have 90 per cent completeness or more, whereas all grid boxes are shown in the coverage series.

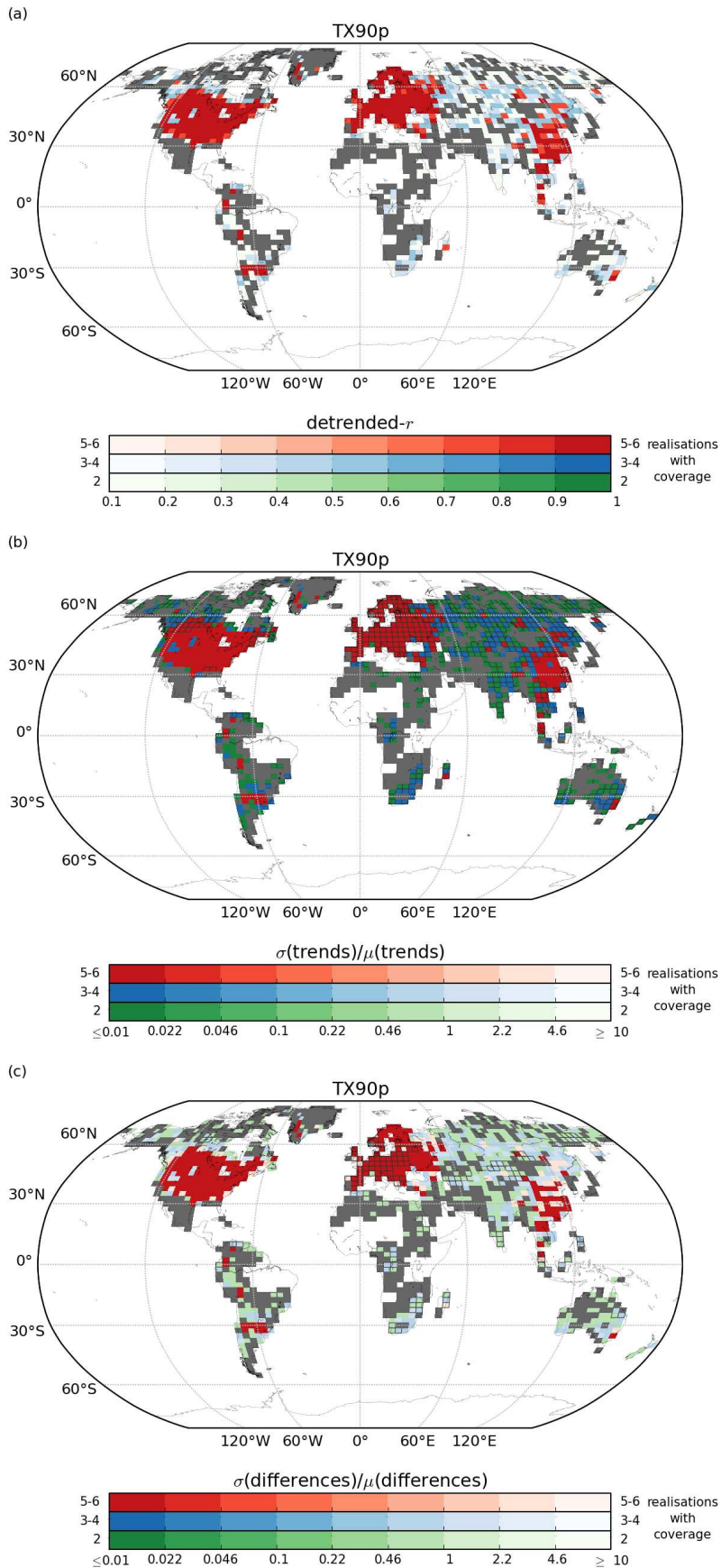
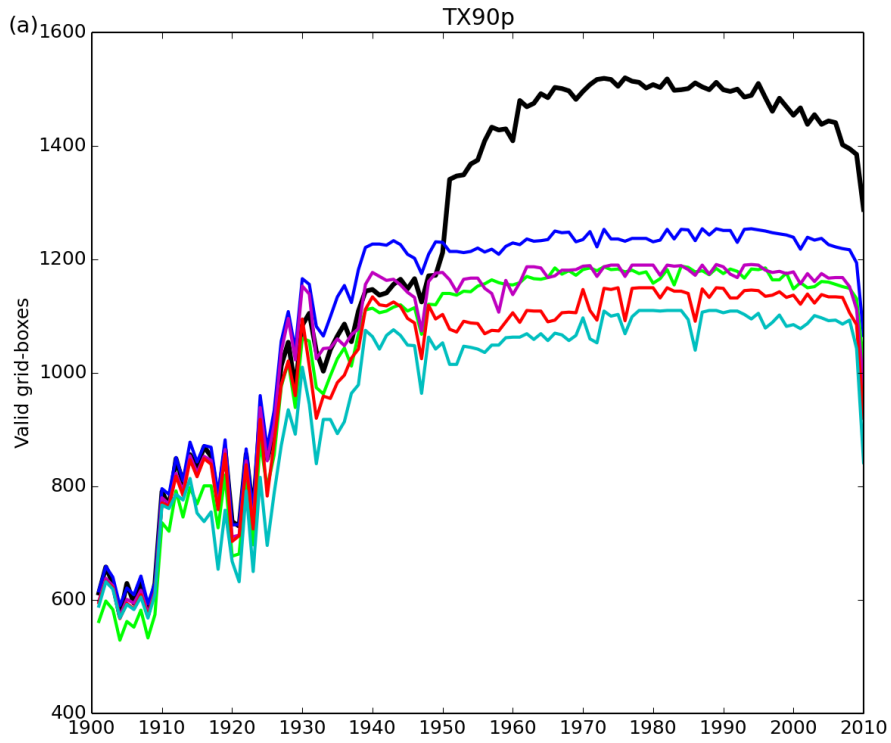
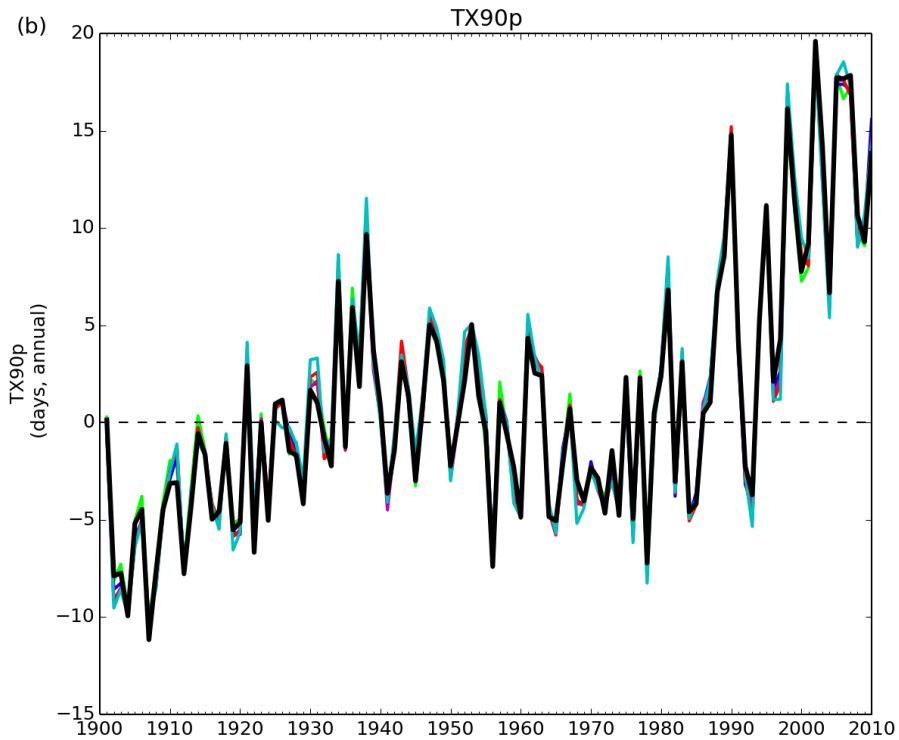


Figure 6: As for Fig. 4 but for TX90p restricted by the number of stations within a grid box. Grey grid boxes are those where only one of the 6 possible stations per grid box choices results in a value, green when 2, blue when 3-4 and red when 5-6 of the possible choices result in a value. Note that the HadEX2 method does not require any stations to be present within a grid box for a value to be calculated. In (b) and (c), boxes which have been outlined are ones where the trend was significant in the HadEX2 version. For details on the calculation method, see text.



Station year coverage

- HadEX2
- 1940-2000
- 1920-2000
- 1950-2000
- 1930-2000
- 1910-2000



Station year coverage

- HadEX2, $\sigma^2 = 41.0$
- 1950-2000, $r = 1.00$, $e_{RMS} = 0.54$, $\sigma^2 = 38.8$
- 1940-2000, $r = 0.99$, $e_{RMS} = 0.65$, $\sigma^2 = 40.8$
- 1930-2000, $r = 0.99$, $e_{RMS} = 0.70$, $\sigma^2 = 41.3$
- 1920-2000, $r = 0.99$, $e_{RMS} = 0.79$, $\sigma^2 = 41.7$
- 1910-2000, $r = 0.99$, $e_{RMS} = 1.06$, $\sigma^2 = 43.5$

Figure 7: As for Fig. 3 but for long-term stations selected using five time periods as detailed in the legend

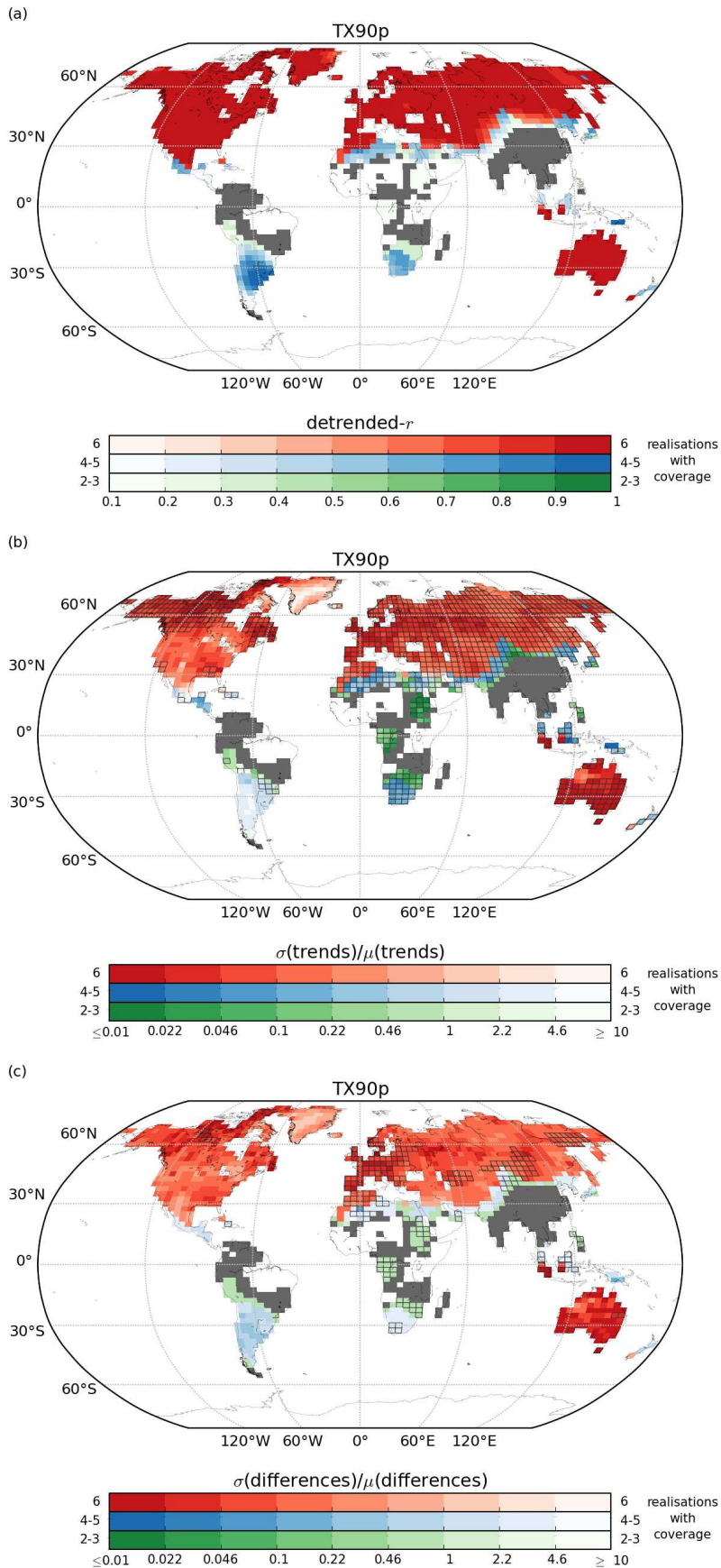


Figure 8: As for Fig. 4 but for long-term stations selected using five time periods, 1950-2000, 1940-2000, 1930-2000, 1920-2000 and 1910-2000. Grey grid boxes are those where only one of the 6 possible station reporting length choices results in a value, green when 2-3, blue when 4-5 and red when 6 of the possible choices result in a value. In (b) and (c), boxes which have been outlined are ones where the trend was significant in the HadEX2 version. For details on the calculation method, see text.

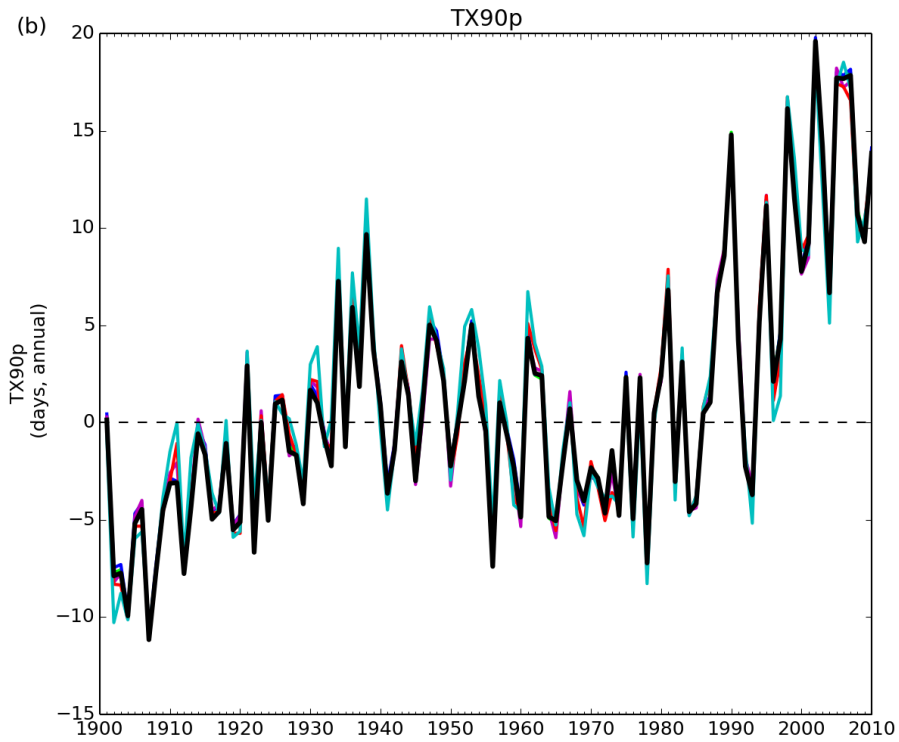
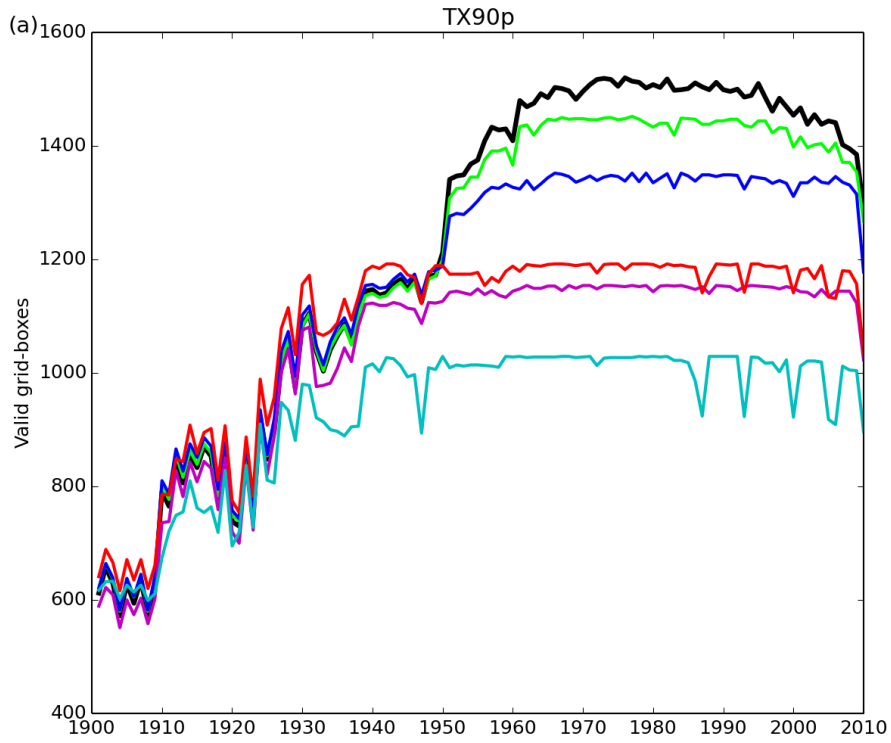


Figure 9: As for Fig. 3 but for long-term stations.

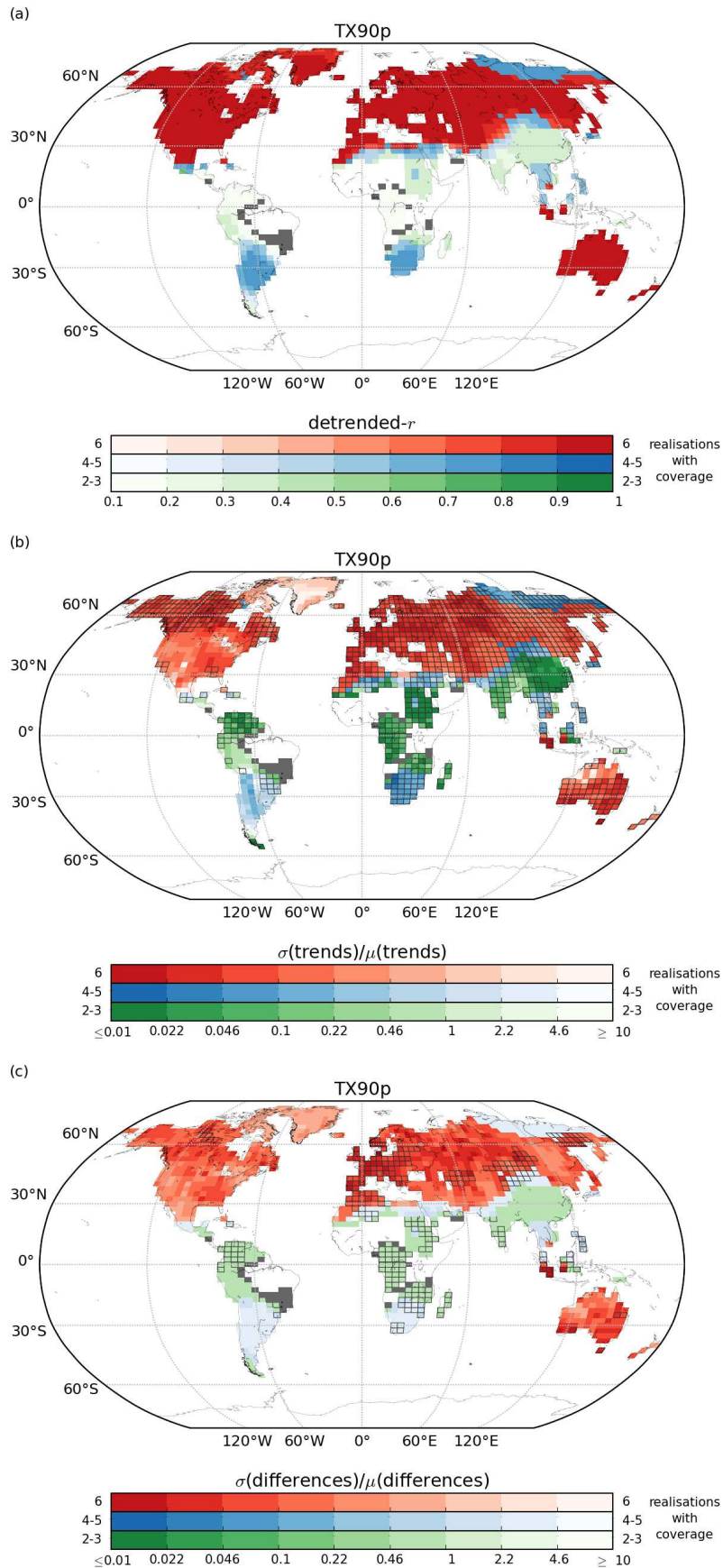
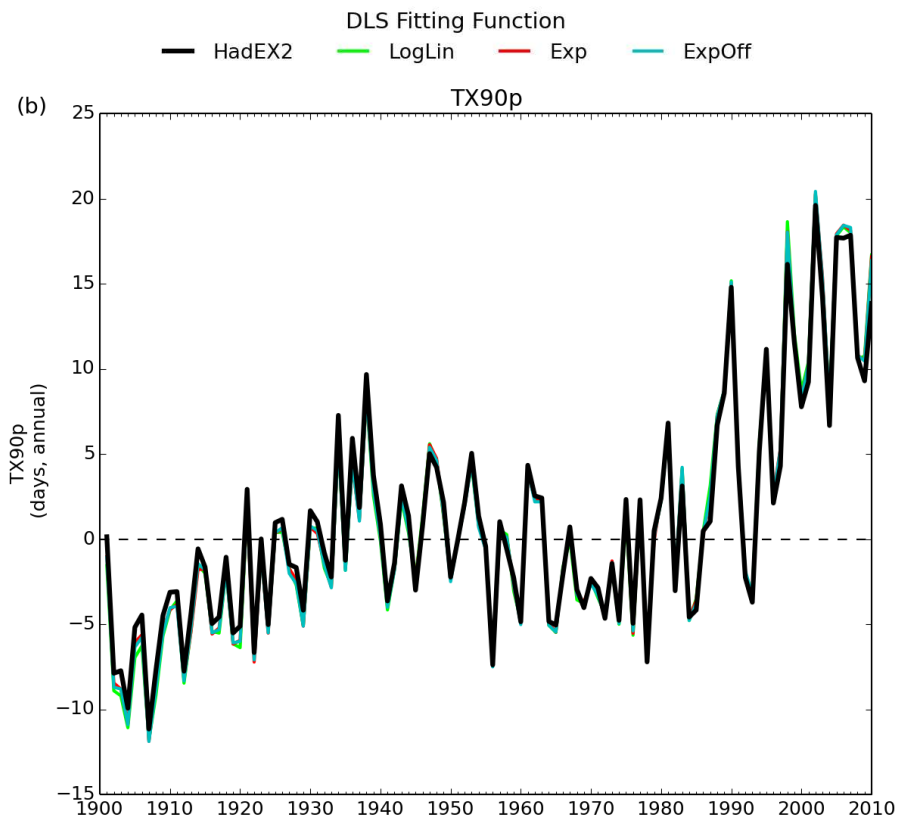
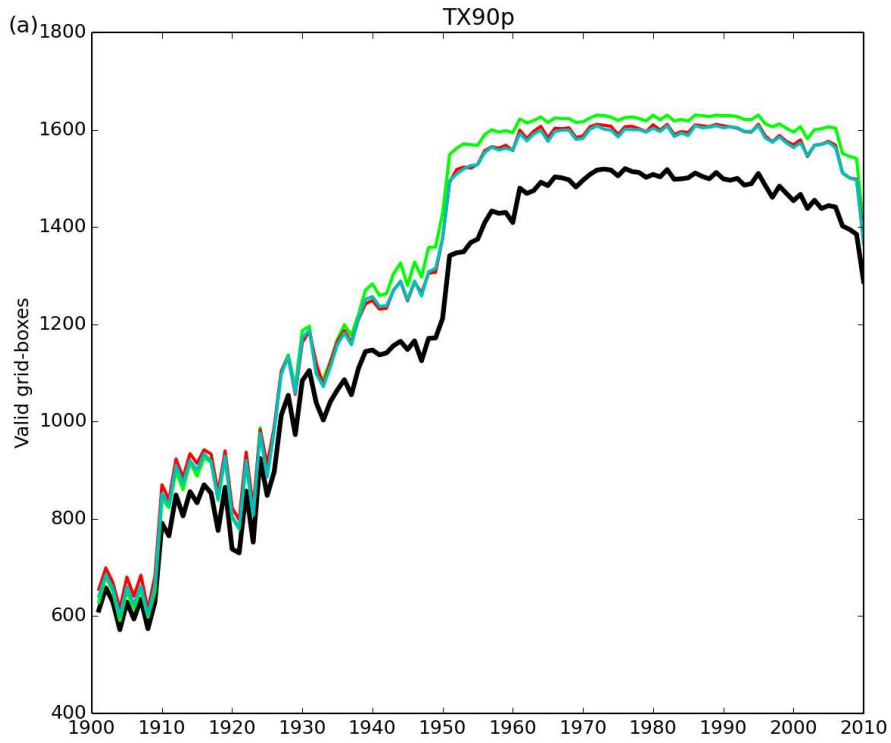


Figure 10: As for Fig. 4 but for long-term stations. Grey grid boxes are those where only one of the 6 possible station reporting length choices results in a value, green when 2-3, blue when 4-5 and red when 6 of the possible choices result in a value. In (b) and (c), boxes which have been outlined are ones where the trend was significant in the HadEX2 version. For details on the calculation method, see text.



DLS fitting function

— HadEX2, $\sigma^2 = 41.0$	— Exp, $r = 1.00$, $e_{RMS} = 0.73$, $\sigma^2 = 45.7$
— LogLin, $r = 0.99$, $e_{RMS} = 0.87$, $\sigma^2 = 46.5$	— ExpOff, $r = 1.00$, $e_{RMS} = 0.74$, $\sigma^2 = 46.0$

Figure 11: As for Fig. 3 but for choices in the DLS calculation method.

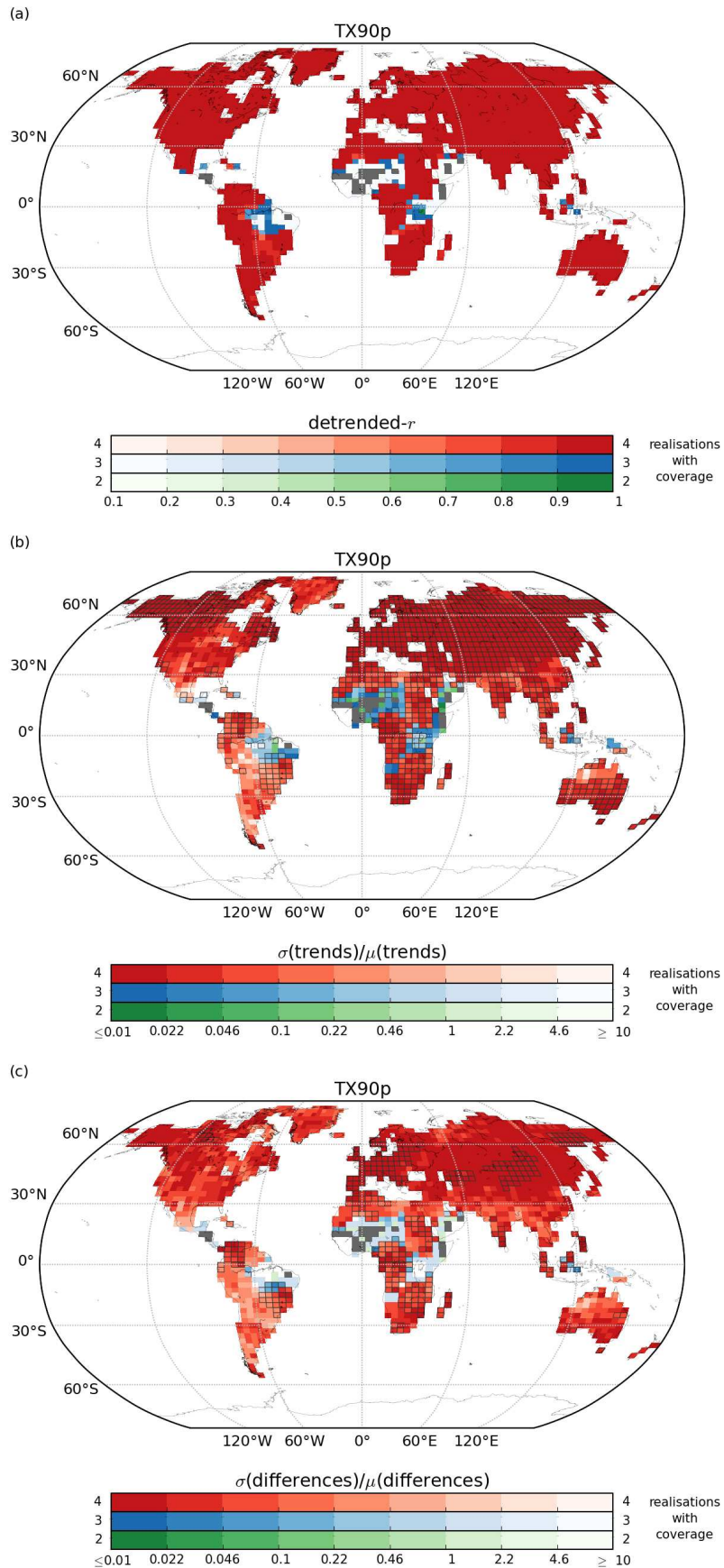


Figure 12: As for Fig. 4 but for choices of the DLS fitting method. Grey grid boxes are those where only one of the four possible DLS calculation choices results in a value, green when 2, blue when 3 and red when all 4 of the possible choices result in a value. In (b) and (c), boxes which have been outlined are ones where the trend was significant in the HadEX2 version. For details on the calculation method, see text.

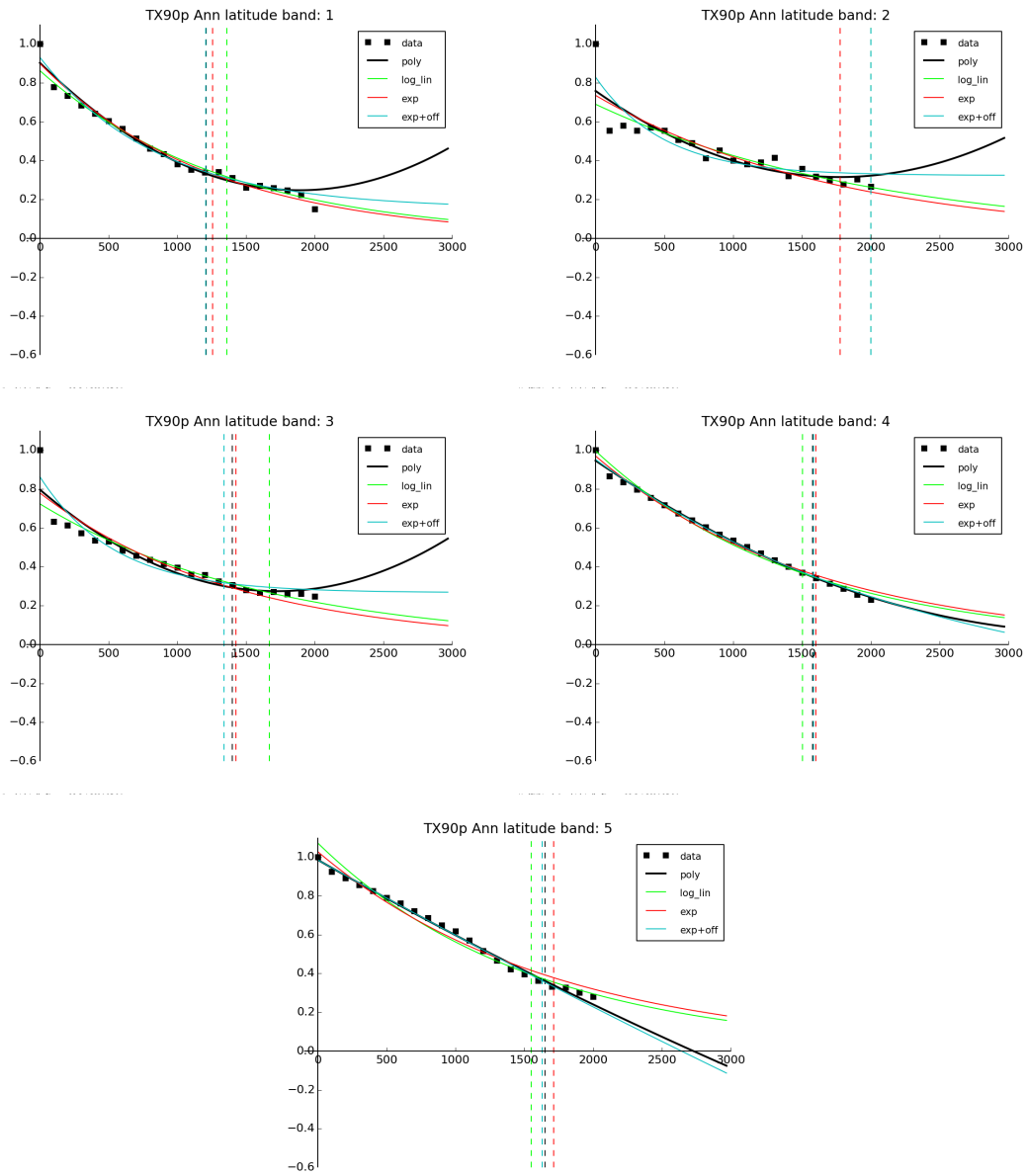


Figure 13: The results of the DLS fitting for each of the five latitude bands: 1 = 90S to 30S; 2 = 30S to 0; 3 = 0 to 30N; 4 = 30N to 60N and 5 = 60N to 90N.

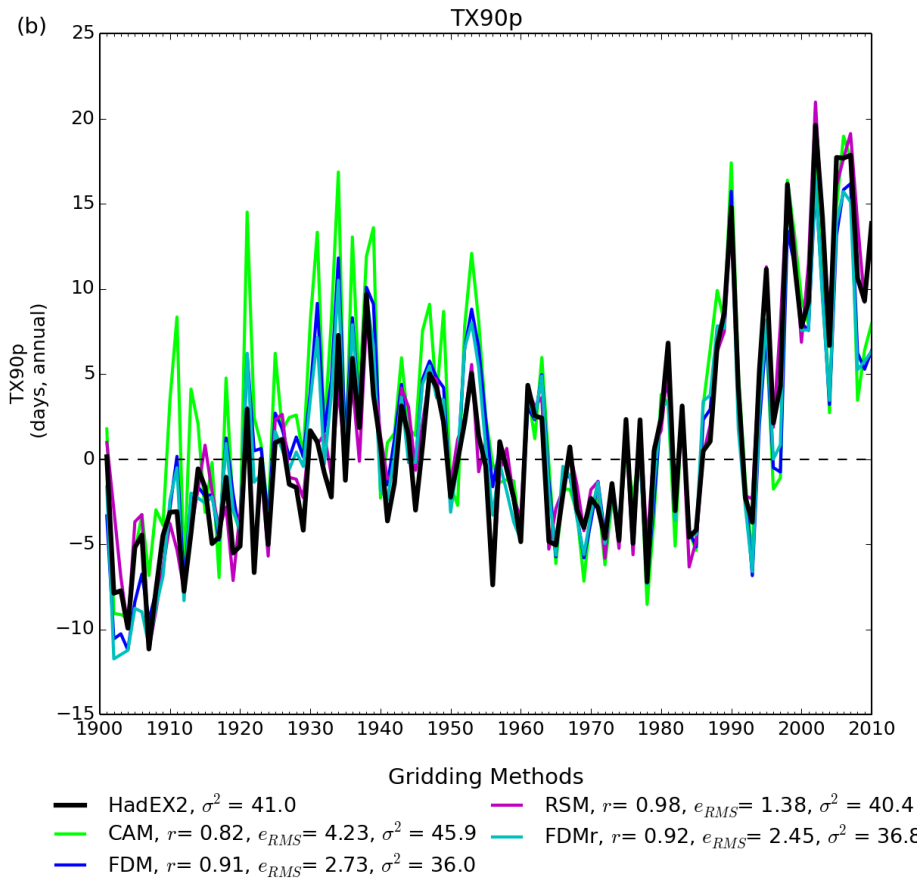
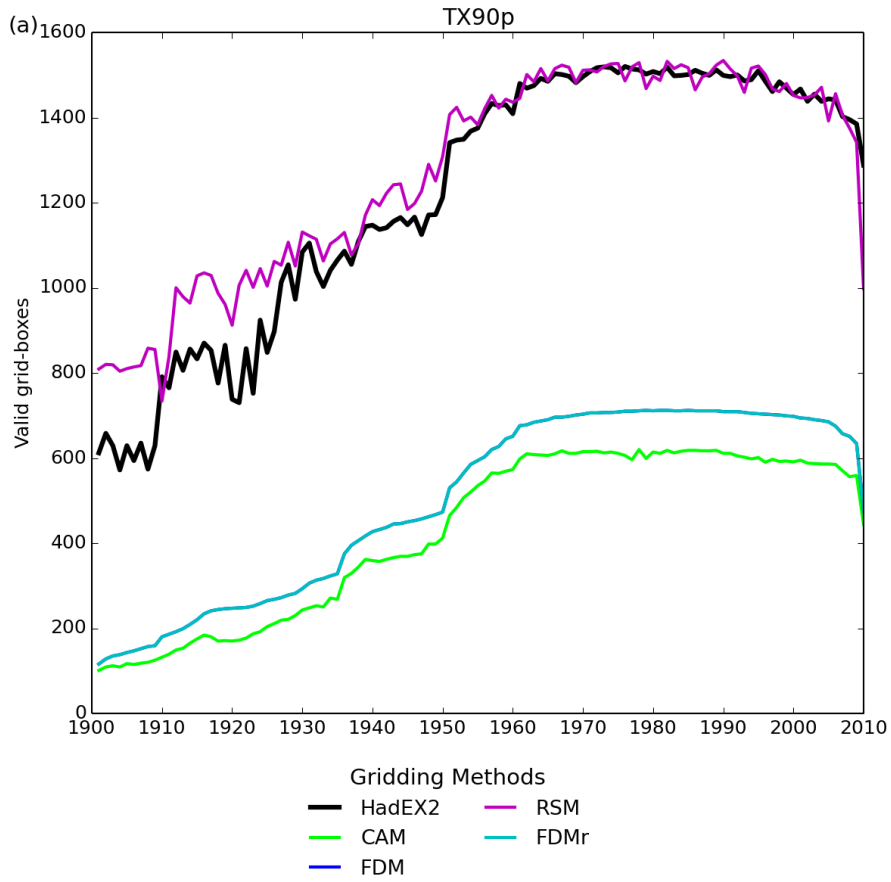


Figure 14: As for Fig. 3 but for the four different gridding methods.

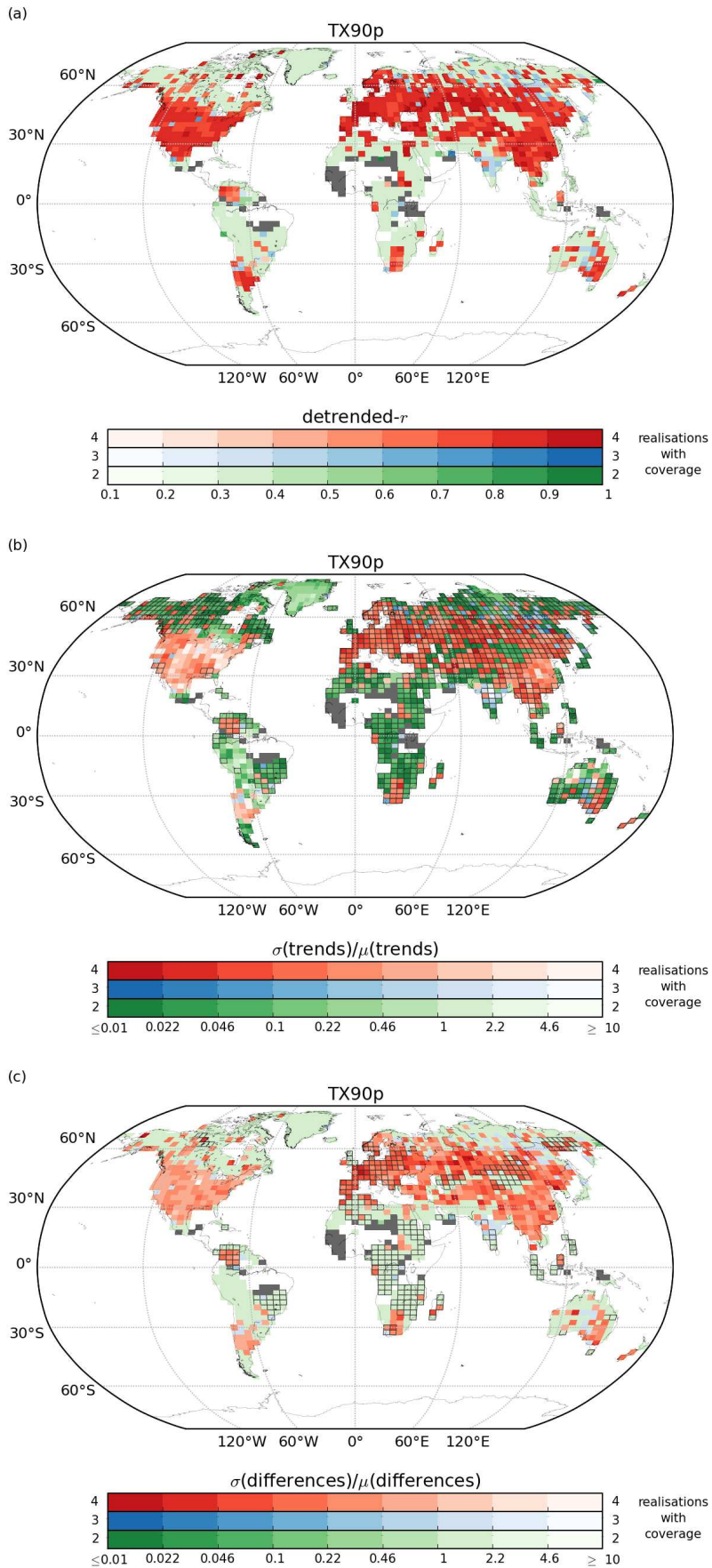


Figure 15: As for Fig. 4 but for the four different gridding methods. Grey grid boxes are those where only one of the four possible gridding methods results in a value, green when 2, blue when 3 and red when all 4 of the possible choices result in a value. In (b) and (c), boxes which have been outlined are ones where the trend was significant in the HadEX2 version. For details on the calculation method, see text.

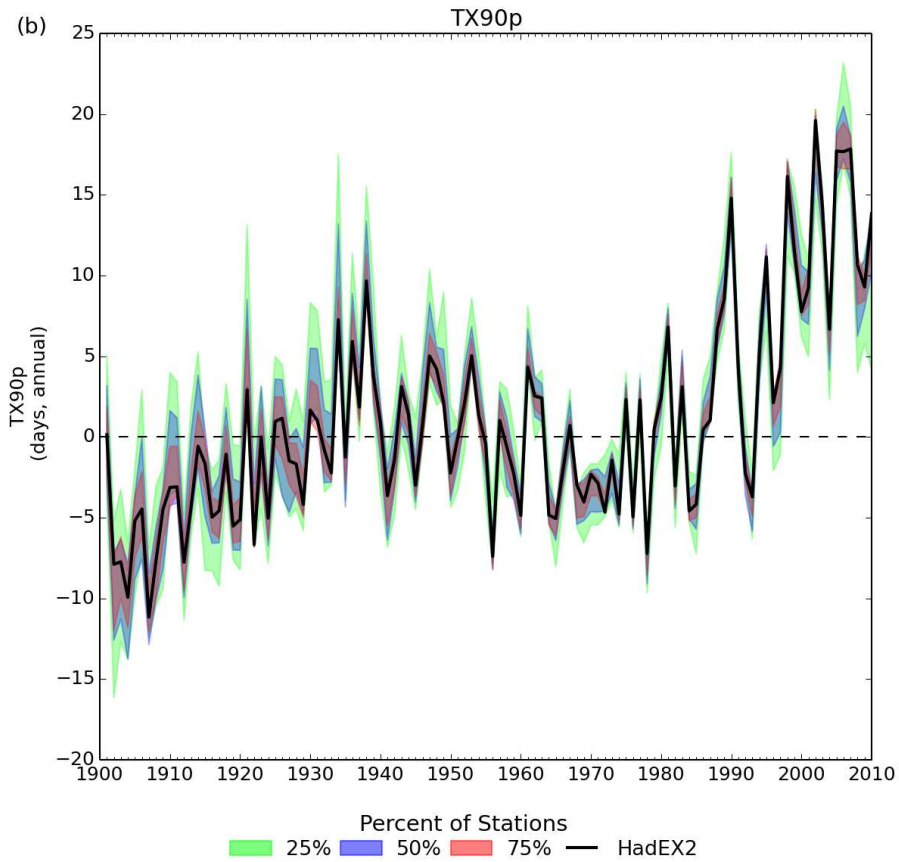
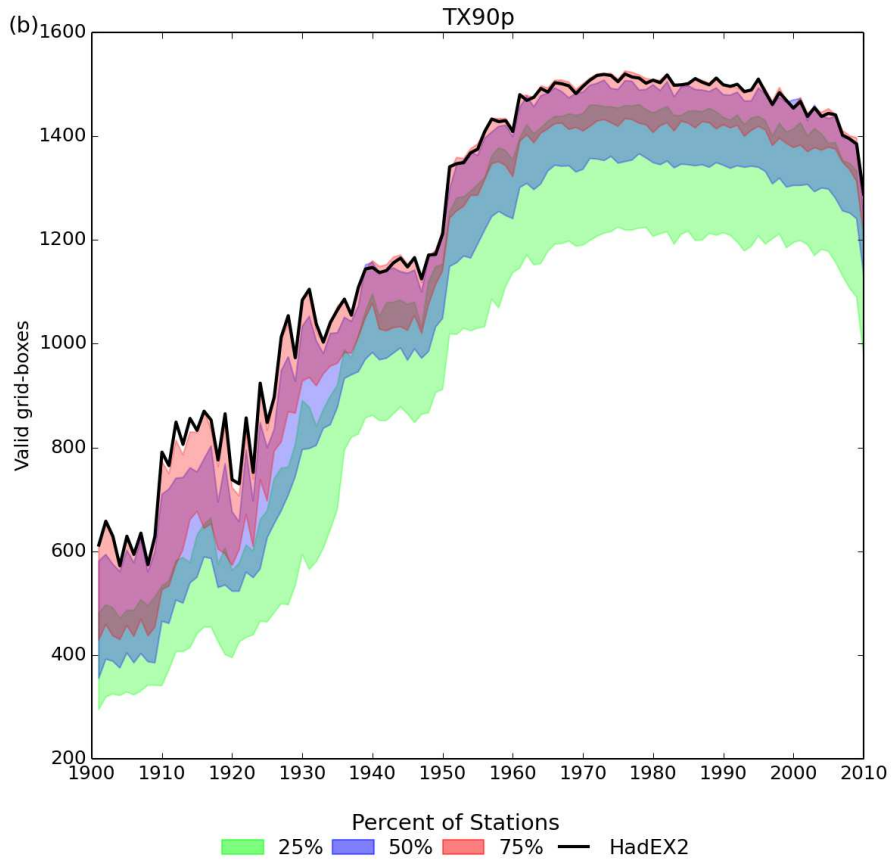


Figure 16: **(a)**: The number of non-missing grid boxes and **(b)**: the time series for the global average of TX90p for the sub-sampling runs. Each colour shows the maximum range of the time series for each of the three sets of runs, 25% of stations in green, 50% in blue and 75% in purple respectively. Note that the global average only takes grid boxes which have 90 per cent completeness or more, whereas all grid boxes are shown in the coverage series.

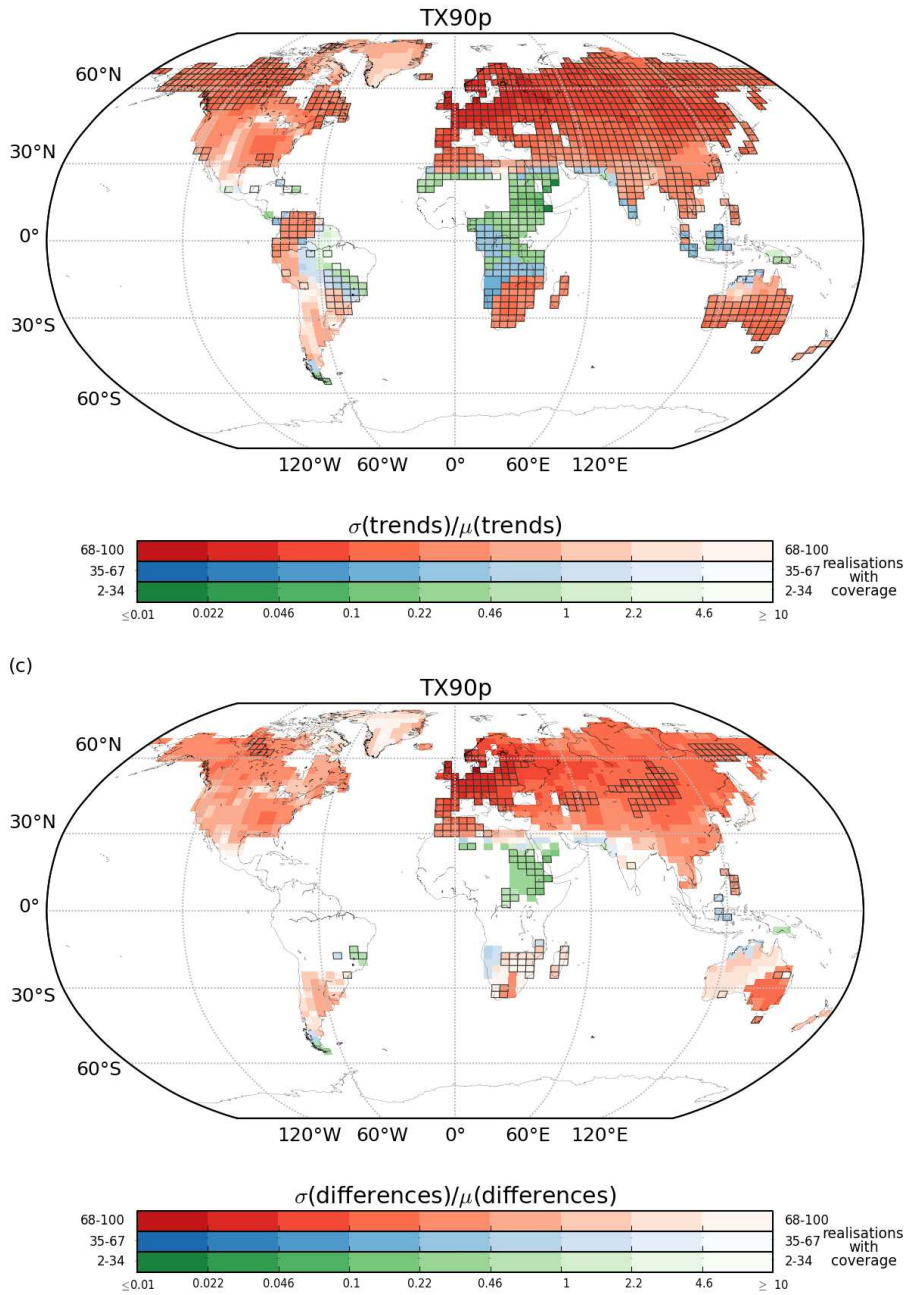


Figure 17: As for Fig. 4b and c but for the 25 per cent runs of the sub-sampling analysis. Boxes which have been outlined are ones where the trend was significant in the HadEX2 version. For details on the calculation method, see text.

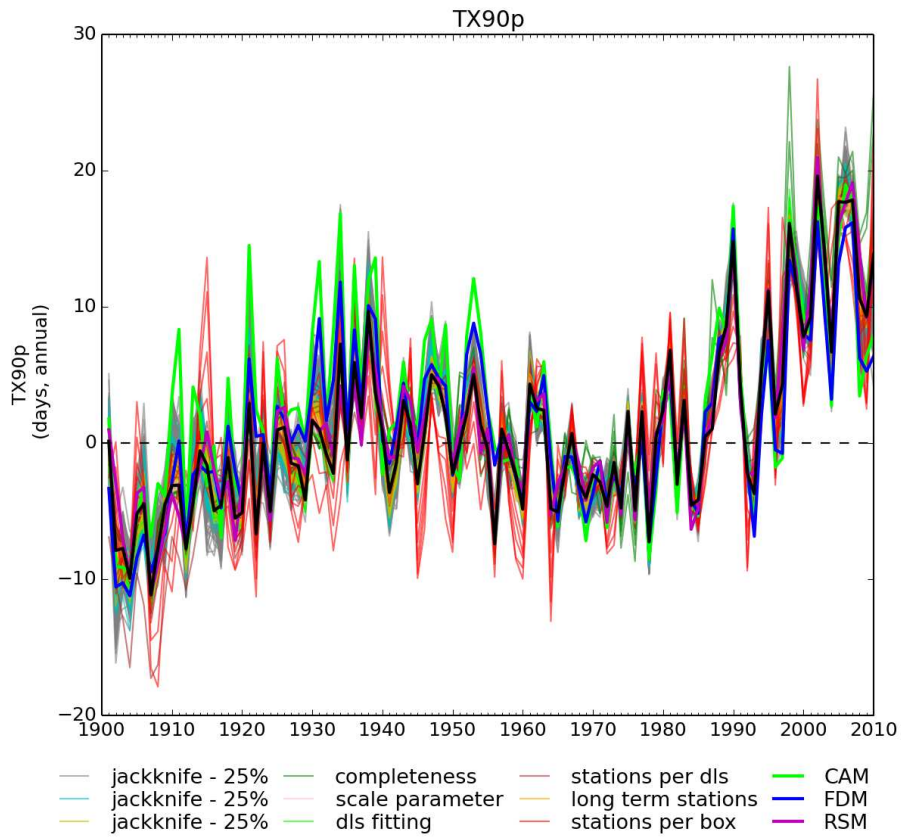


Figure 18: The timeseries curves for all of the different methodological choices assessed during this work, with HadEX2 shown by the thick black line.

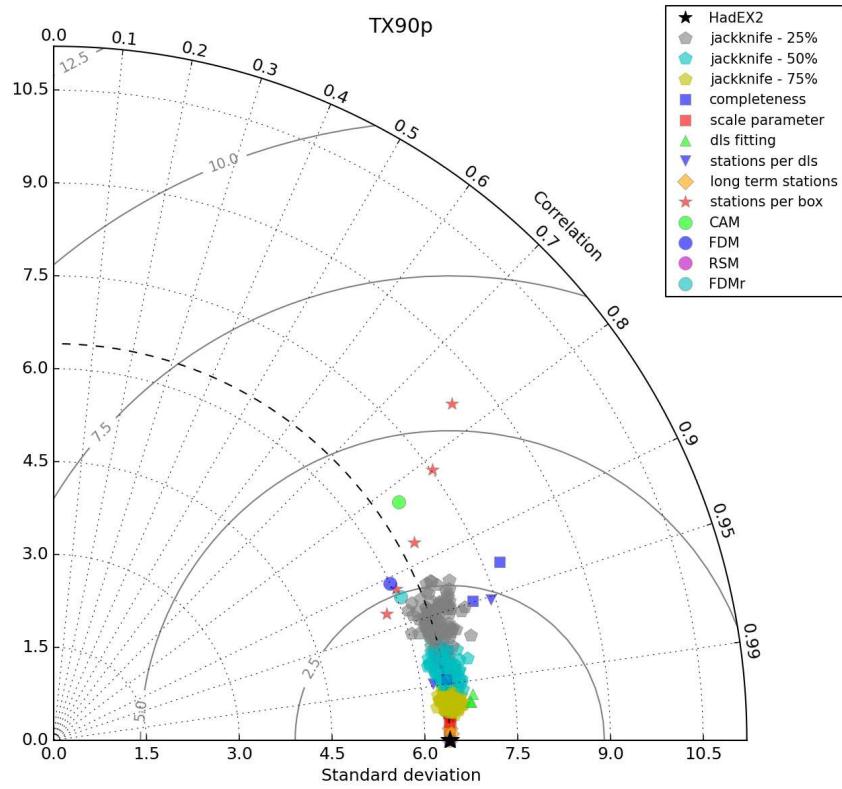


Figure 19: The Taylor Diagram for TX90p calculated using the global average timeseries for each choice. Each of the different methodological choices are shown using a different symbol and colour as indicated in the legend.