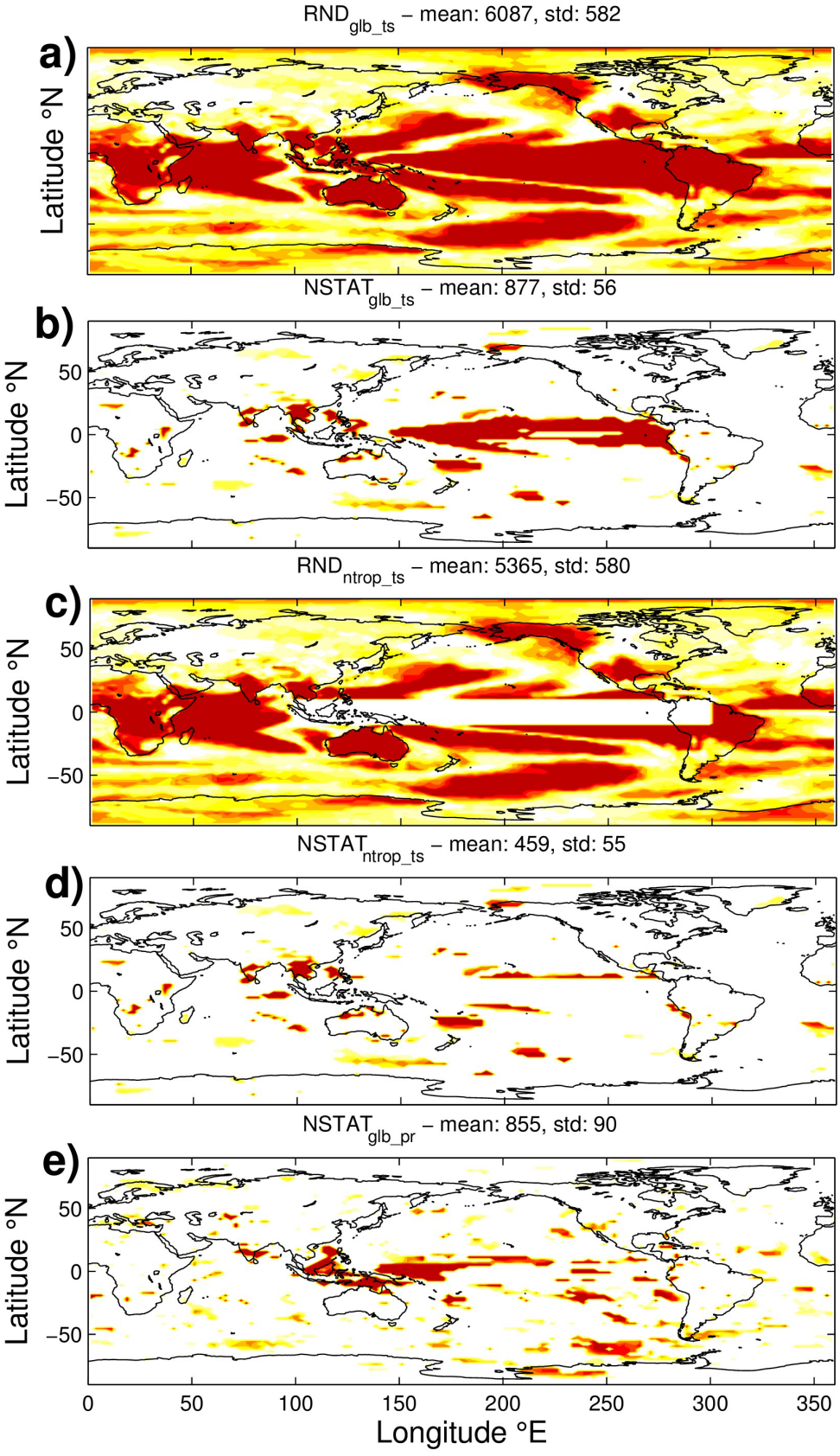

Figure S1. This is a density map of the locations of the possible pseudoproxies for each calibration window for different experiments. This map shows the variation in pseudoproxy selection criteria across calibration windows, and reflects some of the patterns in Figure 2a, as well as showing the physical effect of the experiment criteria. The numbers indicated in the titles refer the mean and standard deviation (std) across the different calibration window positions.

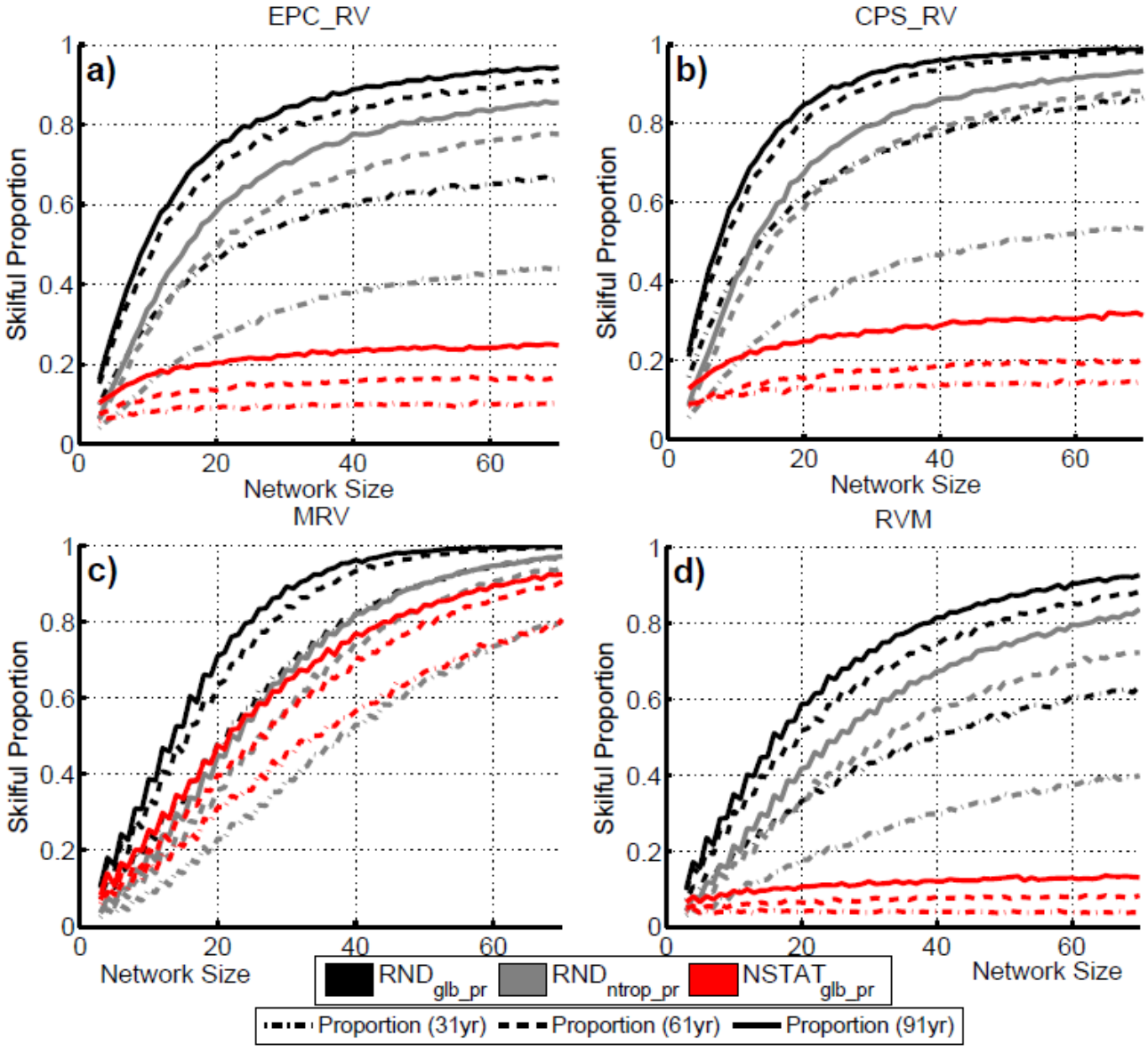

Figure S2. Reconstruction skill for different precipitation experiments (as indicated by colour) and reconstruction methods (each panel) using different calibration window lengths (see inset legend). The lines show the proportion of skilful reconstructions with skilful being defined as explaining greater than 50% of the explained variance of the Niño 3.4 running variance. This plot is the same format as Fig. 8, except that these are for the precipitation experiments.

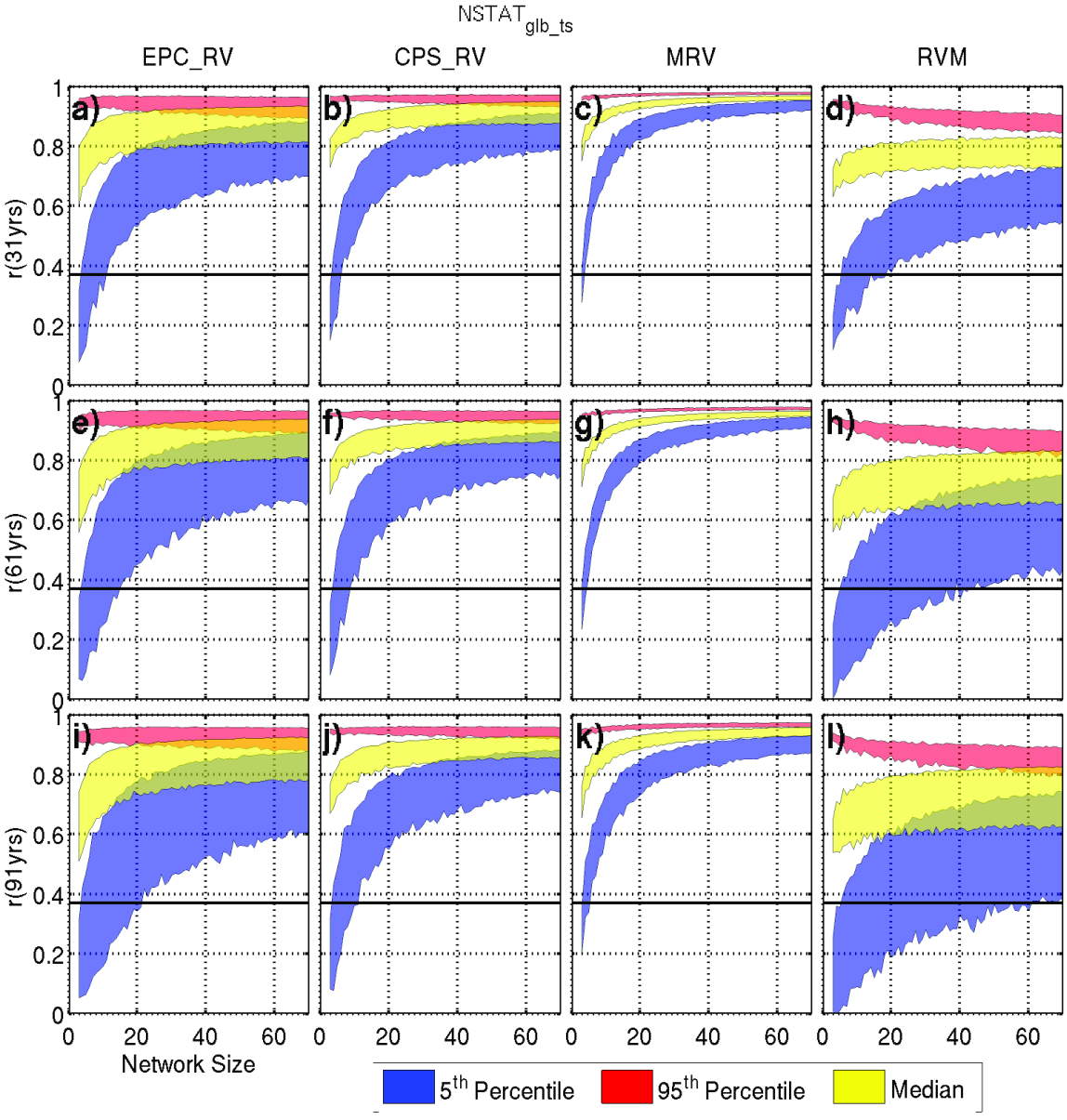

Figure S3. The 5th (blue), 50th (yellow) and 95th (red) percentiles of correlation coefficients calculated between the TS based $NST\ AT_{glb}$ experiment pseudo-reconstructions running variance and ENSO running variance (y-axis) plotted against the proxy network size (x-axis). The shaded range is the result of using ten different calibration windows within the 499 years of data, and is simply the range of these reconstructions' percentiles. Each column represents one reconstruction method (titled at the top of each column) while each row represents the length of the calibration windows (titled on the y-axis of each row). This figure is similar to Figure 4, but with the non-stationary pseudoproxy criteria.

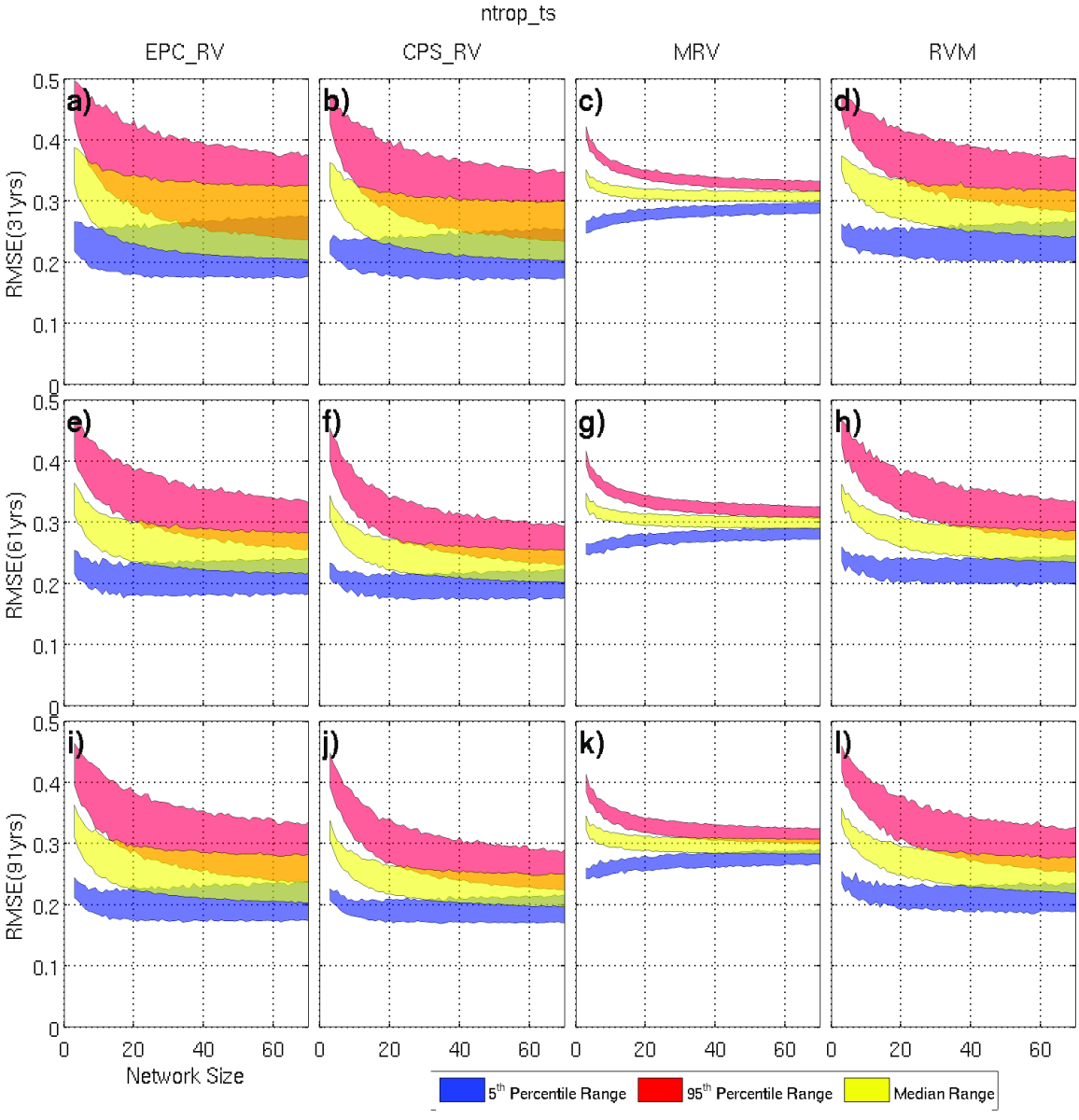

Figure S4. The 5th (blue), 50th (yellow) and 95th (red) percentiles of the root-mean-squared error statistic calculated between the TS based RND_{ntrop} pseudoproxy reconstructions running variance and ENSO running variance (y-axis) plotted against the proxy network size (x-axis). The shaded range is the result of using ten different calibration windows within the 499 years of data, and is simply the range of these reconstructions' percentiles. Each column represents one reconstruction method (titled at the top of each column) while each row represents the length of the calibration windows (titled on the y-axis of each row). It is similar to Figures 4-7, except that this shows RMSE instead of correlation for the RND_{glb} experiment.

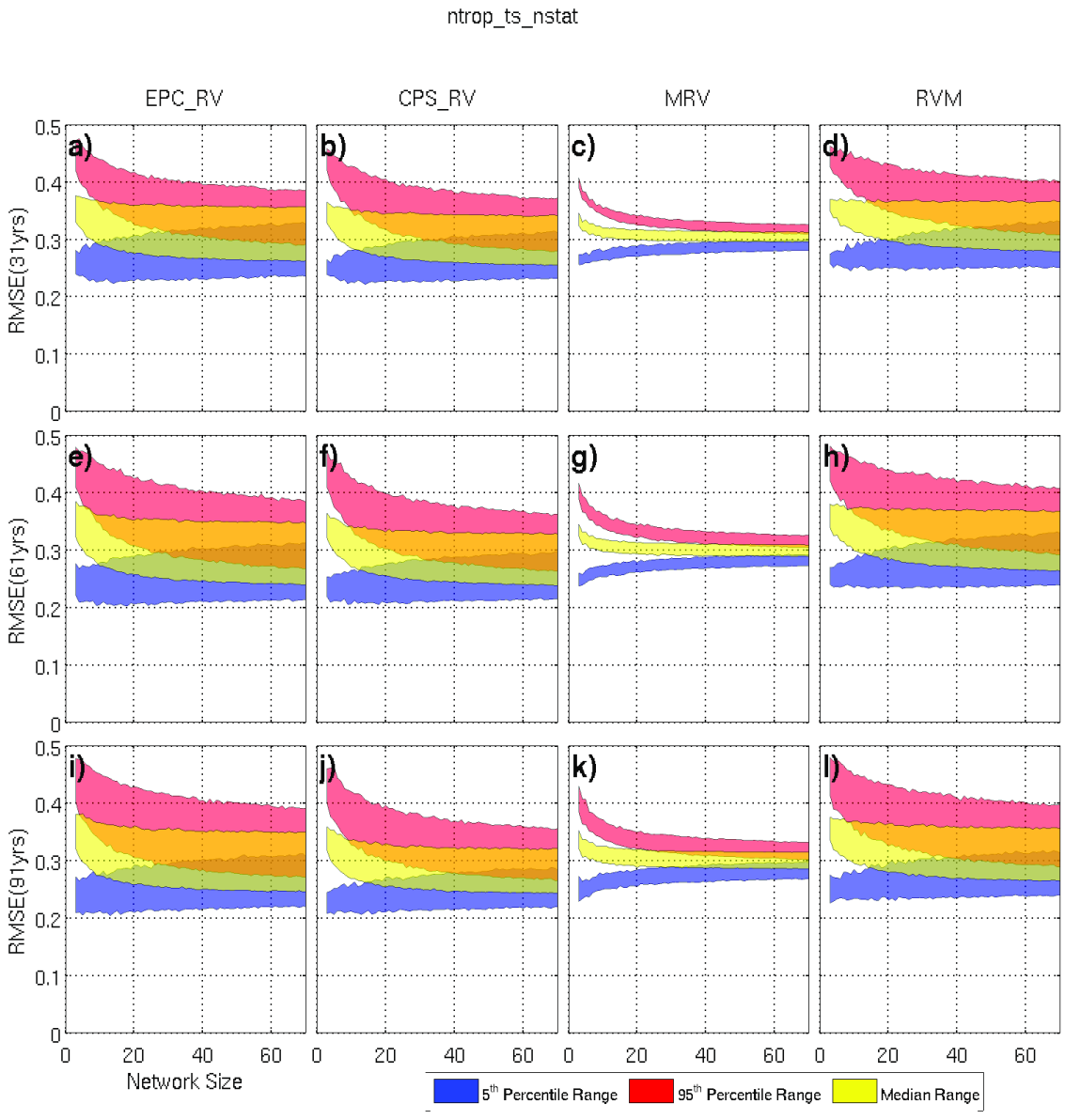

Figure S5. The 5th (blue), 50th (yellow) and 95th (red) percentiles of the root-mean-squared error statistic calculated between the TS based $NSTAT_{ntrop}$ pseudoproxy reconstructions running variance and ENSO running variance (y-axis) plotted against the proxy network size (x-axis). The shaded range is the result of using ten different calibration windows within the 499 years of data, and is simply the range of these reconstructions' percentiles. Each column represents one reconstruction method (titled at the top of each column) while each row represents the length of the calibration windows (titled on the y-axis of each row). It is similar to Figures 4-7, except that this shows RMSE instead of correlation for the $NSTAT_{ntrop}$ experiment.

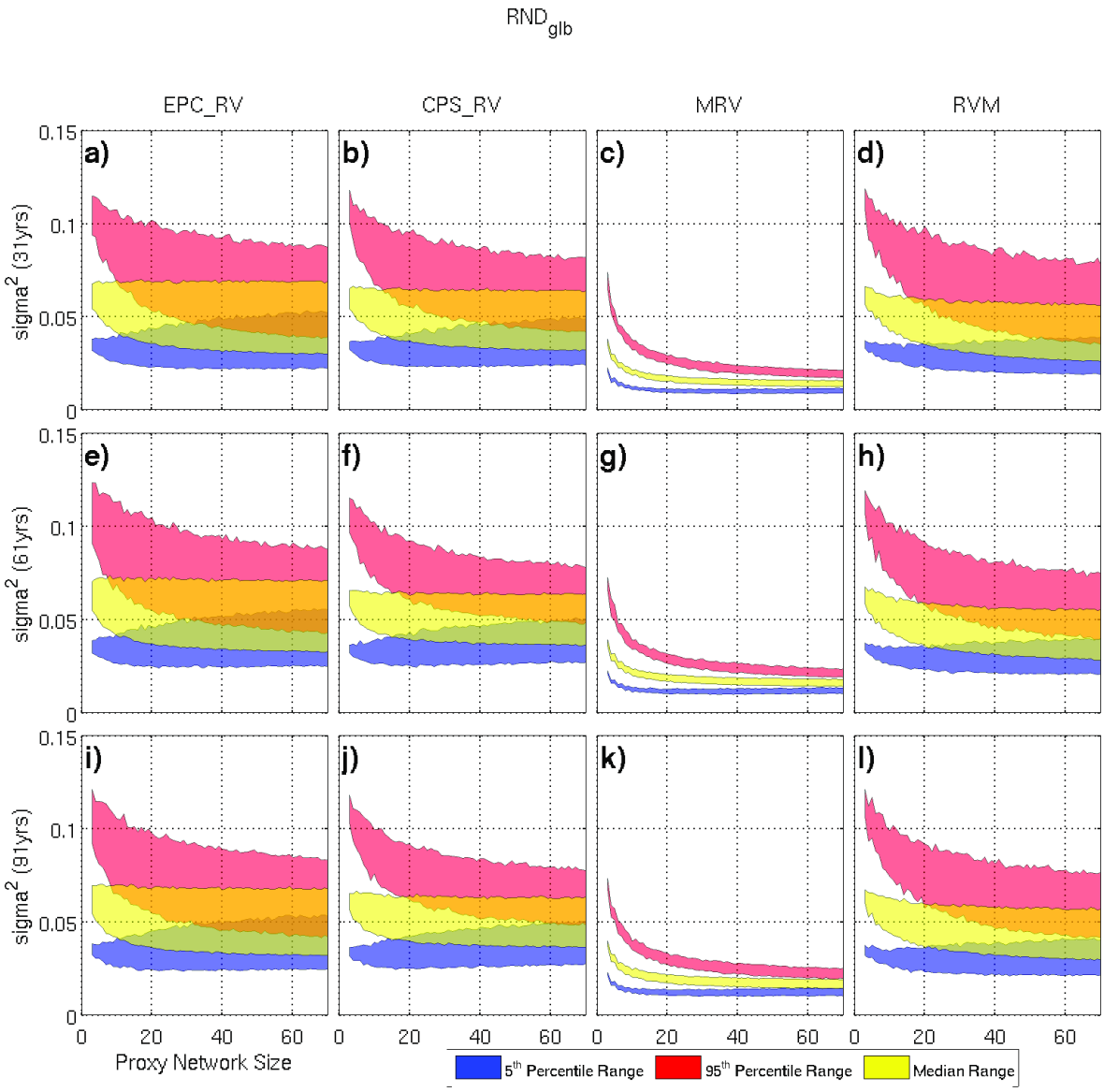

Figure S6. The 5th (blue), 50th (yellow) and 95th (red) percentiles of the variance of the TS based RND_{glb} pseudoproxy reconstructions running variance (y-axis) plotted against the proxy network size (x-axis). The shaded range is the result of using ten different calibration windows within the 499 years of data, and is simply the range of these reconstructions' percentiles. Each column represents one reconstruction method (titled at the top of each column) while each row represents the length of the calibration windows (titled on the y-axis of each row). It is similar to Figures 4-7, except that this shows variance instead of correlation for the RND_{glb} experiment.

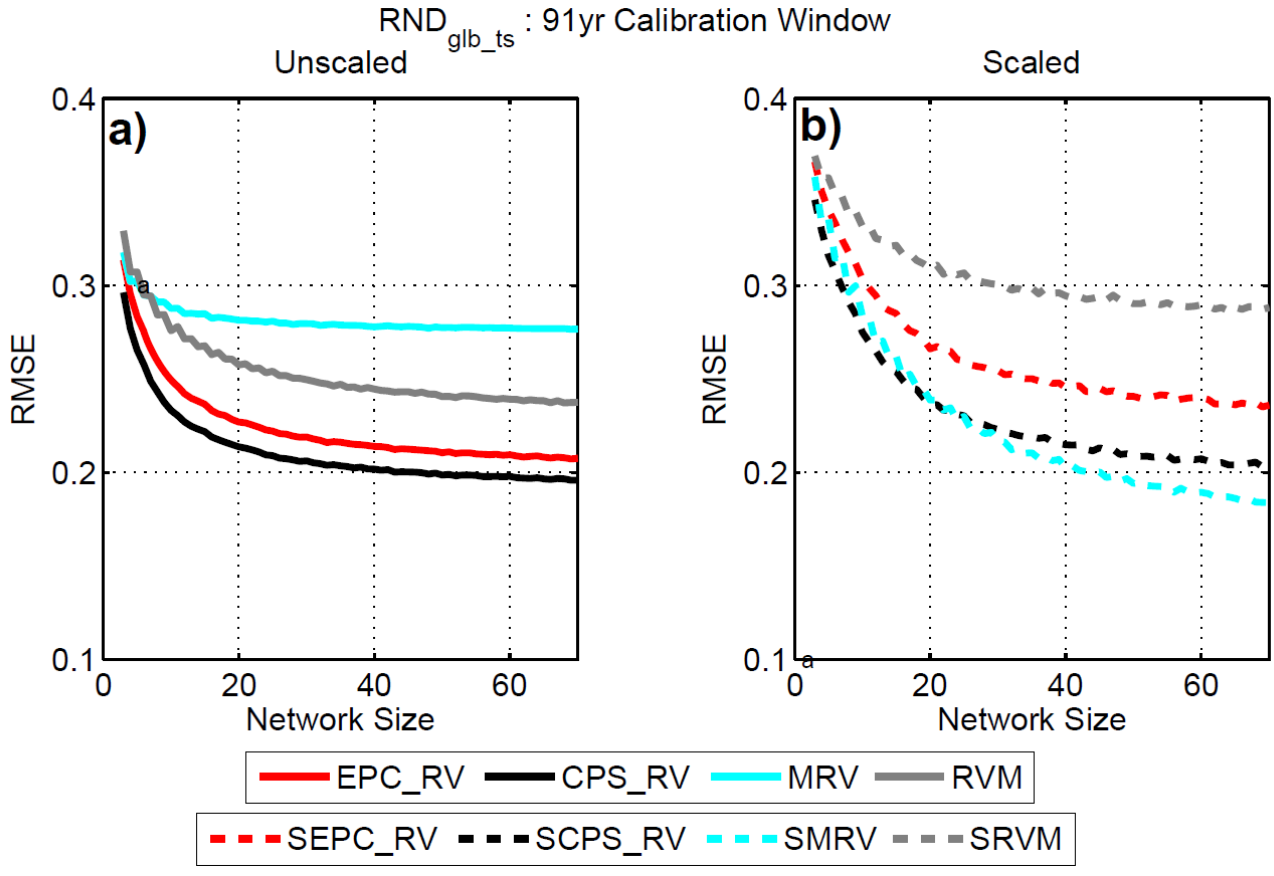

Figure S7. The root-mean-squared error statistic (RMSE) calculated between the TS based RND_{glb_ts} pseudoproxy reconstructions running variance and ENSO running variance (y-axis) plotted against the proxy network size (x-axis). The lines are the median RMSE of the ten thousand reconstructions for each reconstruction method, for the 91 year long calibration window. In panel a), the solid lines are the normal unscaled reconstruction methods, while in panel b), the dashed lines are the scaled variants indicated by the legend.