Summer Temperature (JJA) - MPIGE (FULL)
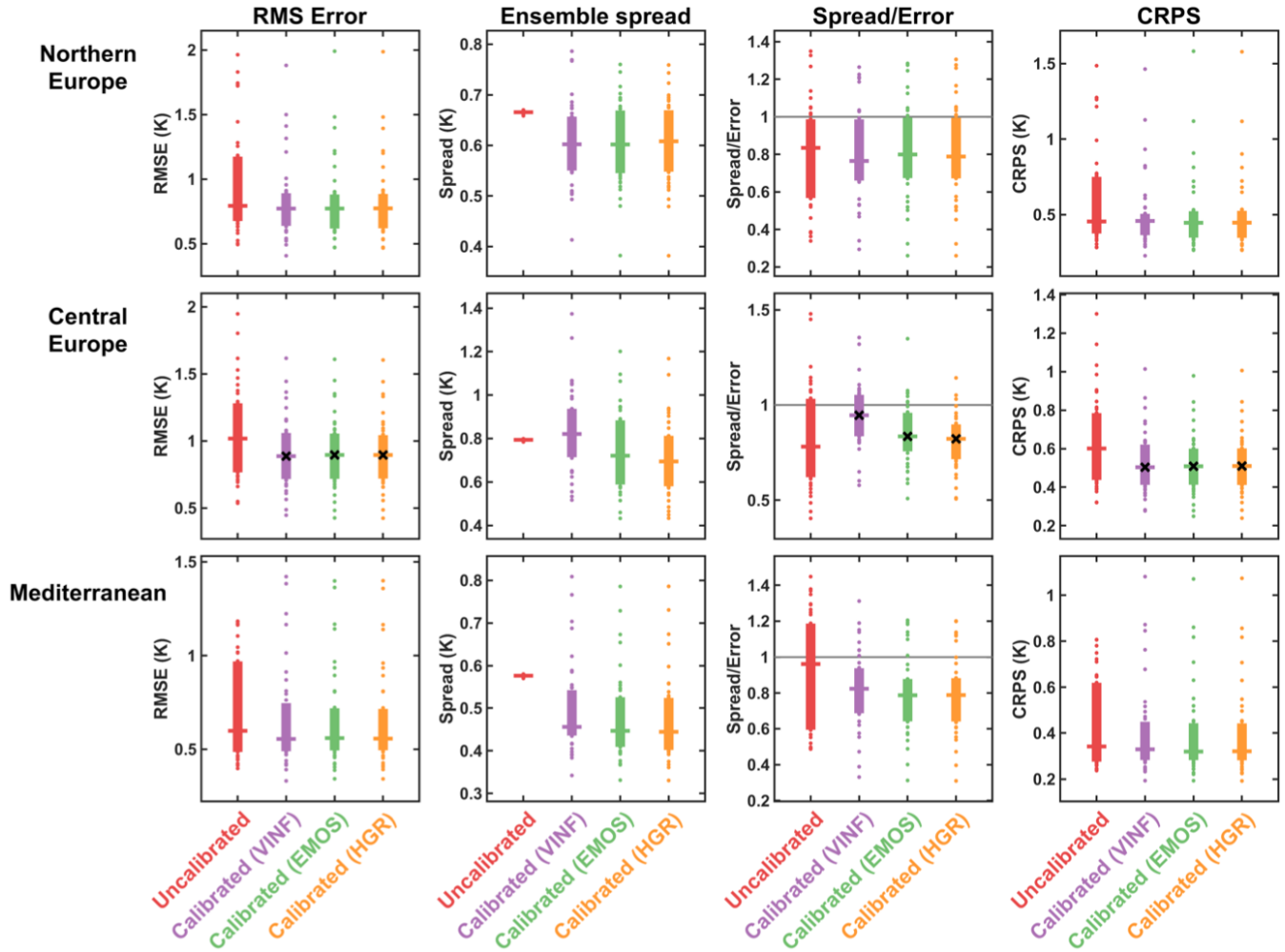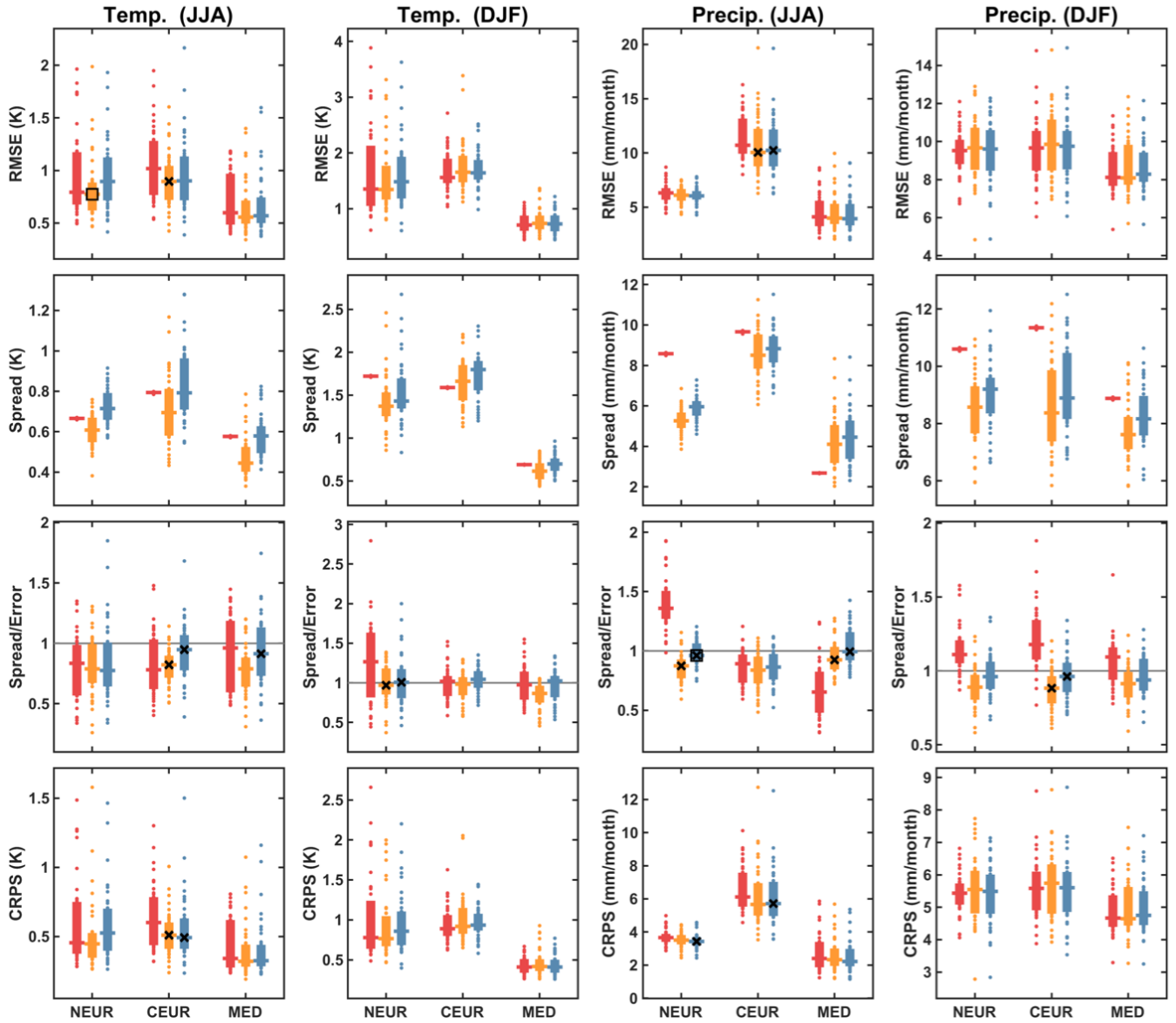

Figure S1: Comparison of calibration methods applied to the MPI-GE summer temperature projections calibrated to the CMIP5 models over the observational period (1920-2016) and verified using the 44 years in the out-of-sample period (1917-2060). The verification statistics for each of the individual CMIP5 models are shown in dots, the interquartile range of this distribution is shown by the solid bars and the median is indicated by the horizontal lines. For the calibrated RMS Error, spread/error and CRPS values, the black crosses indicate where the calibration represents a significant improvement over the uncalibrated (but bias-corrected) ensemble at the 90% significance level. The significance levels were calculated using the non-parametric Mann-Whitney U-test.

MPIGE - Uncalibrated, Calibrated (HGR) & Calibrated (HGR-decomp)

Figure S2: Overview of verification of the HGR and HGR-decomp calibration methods compared with the uncalibrated MPI-GE data in the European regions. Shown for all of the verification measures, for both summer and winter seasons and for temperature and precipitation. The verification statistics for each of the individual CMIP5 models are shown in dots, the interquartile range of this distribution is shown by the solid bars and the median is indicated by the horizontal lines. For the calibrated RMS Error, spread/error and CRPS values, the black crosses indicate where the calibration represents a significant improvement over the uncalibrated (but bias-corrected) ensemble at the 90% significance level. Black circles indicate where the calibration is significantly worse than the uncalibrated ensemble (at the 90% level). The black boxes show where one calibration method is found to be significantly better than the other calibration method for the same variable, season and region at the 90% significance level. The significance levels were calculated using the non-parametric Mann-Whitney U-test.
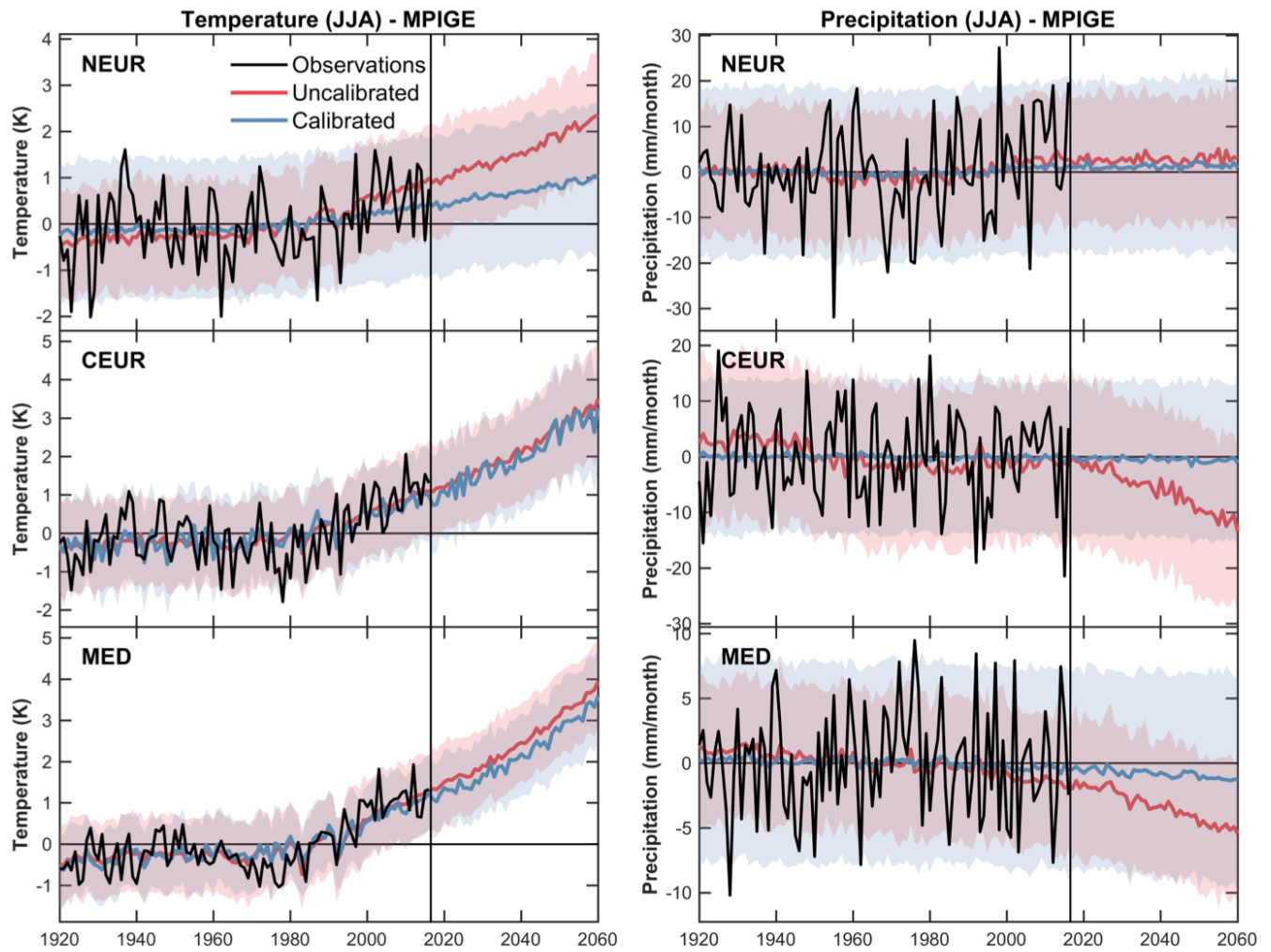
Figure S3: Uncalibrated and calibrated (HGR-decomp) MPI-GE projections, where here the calibrated projections have been calibrated against the observations over the period 1920-2016. The lines show the ensemble medians for the uncalibrated and calibrated ensembles. The shading shows the 90% range of the LENS ensemble. Based on the verification out-of-sample tests using the CMIP5 models the calibrated ensemble is expected to be more reliable than the uncalibrated ensemble, particularly for temperatures.