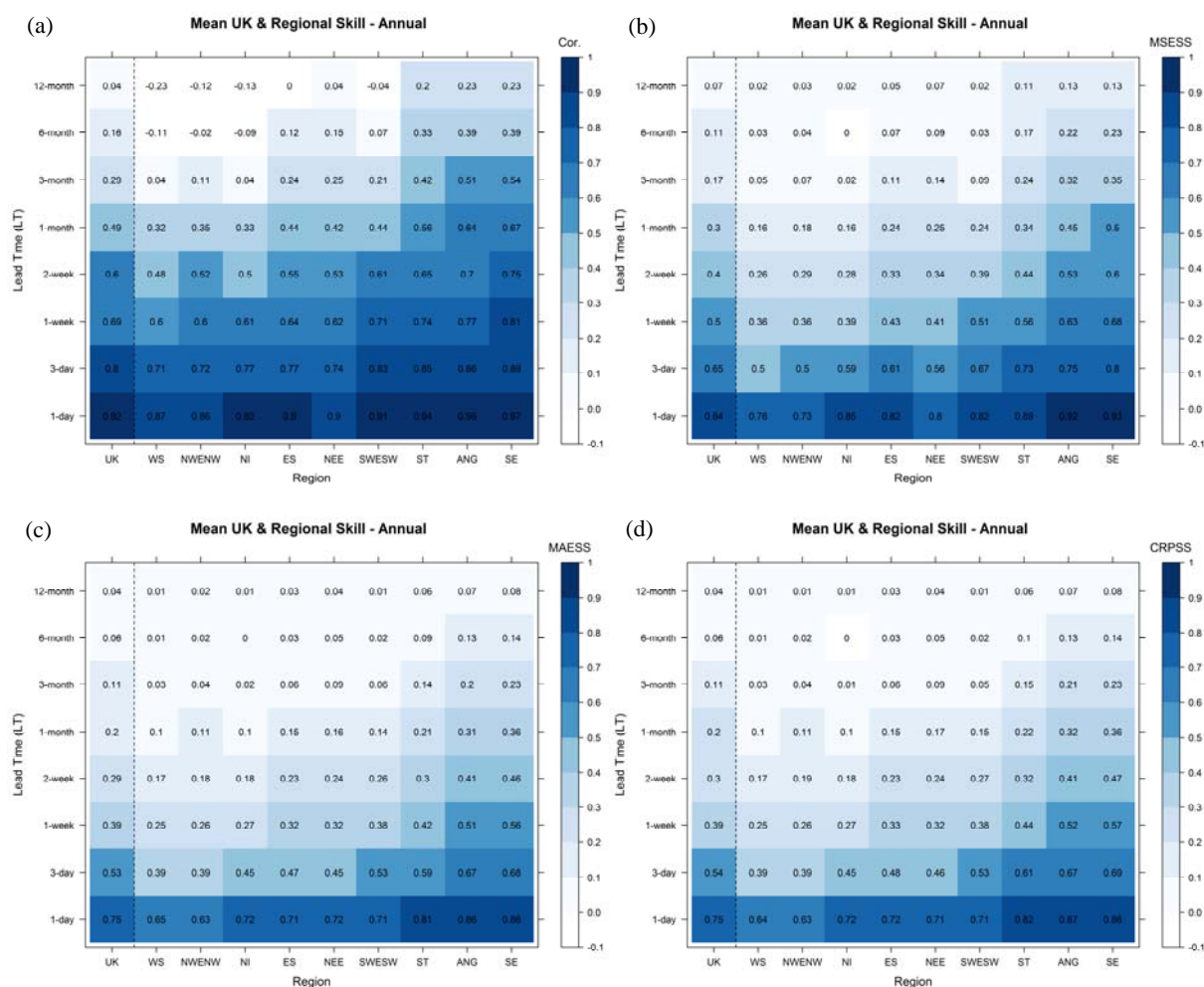**Supplementary Figure S1:** Spatial distribution of GR4J model performance for 314 catchments over the calibration (Cal CP [WY1983-2014], top row), and two evaluation periods (Eval P1 [WY1983-1998], middle row and Eval P2 [WY1999-2014], bottom row) for the modified Kling-Gupta Efficiency applied to root squared transformed flows (KGE$_{mod}$[sqrt]) and Nash-Sutcliffe Efficiency (NSE) model performance metrics. UK-wide Summary statistics are given in the bottom left for the median and 5th and 95th percentiles.

Supplementary Figure S2: Mean ESP skill across all 12 forecast initialisation months for the UK and for each of the nine hydroclimate regions ordered from least to most skilful (horizontal axis) at eight sample lead times (vertical axis). Skill is given by the a.) Pearson correlation coefficient (Cor.), b.) Mean Squared Error Skill Score (MSESS), c.) Mean Absolute Error Skill Score (MAESS), and d.) Continuous Ranked Probability Skill Score (CRPSS). Darker (lighter) shades showing higher (lower) skill; mean skill score values are shown within each cell.