Supplementary Information of Boreal Env. Res. Vol. 30: 93–109, 2025

Supplementary Information of

Less frequent but more intense summertime precipitation in Finland: results from a convection-permitting climate model

***Utriainen et al*.**

Correspondence to: Laura Utriainen (laura.utriainen@fmi.fi)

A HCLIM domain and its orography over Fennoscandia

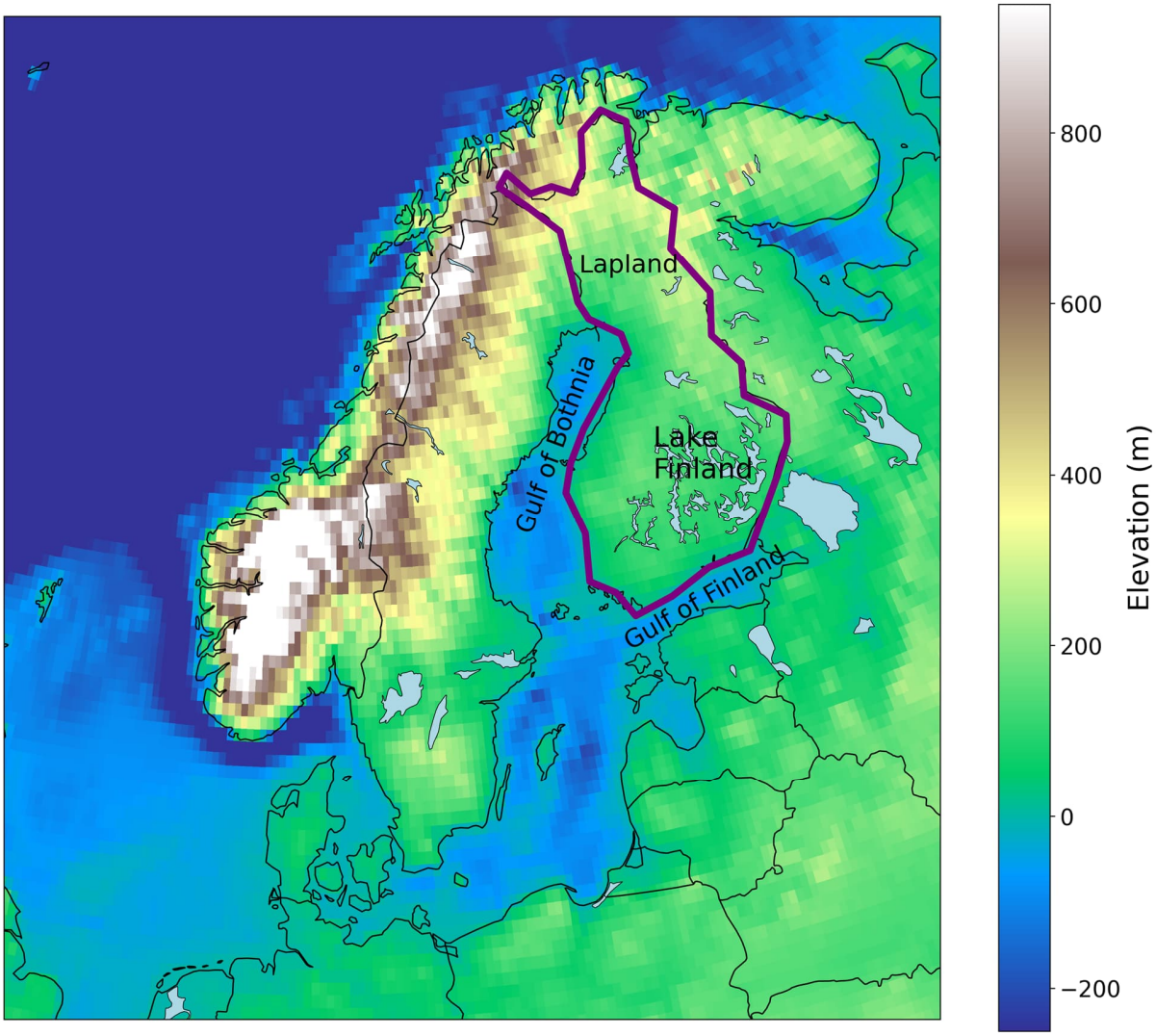

Figure A1. The simulation domain of the HCLIM38-AROME model at a 3 km resolution, showing the elevation of the study area in color. The domain outlined in purple corresponds to the land area of Finland analyzed in this study, with the geographical regions of Lapland, Lake Finland, Gulf of Bothnia, and Gulf of Finland highlighted.

B *In-situ* gauge observations versus the HCLIM simulations

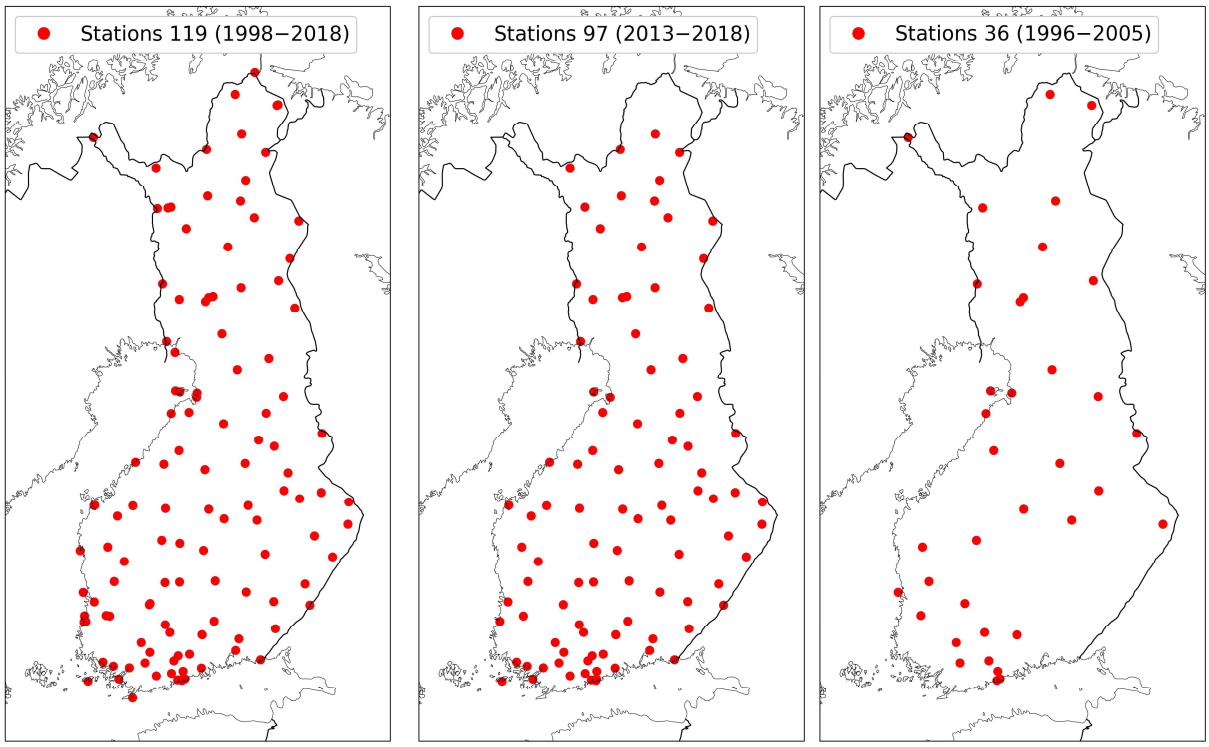

Figure B1. The locations of the weather stations for each is-situ gauge dataset presented in Table 2.

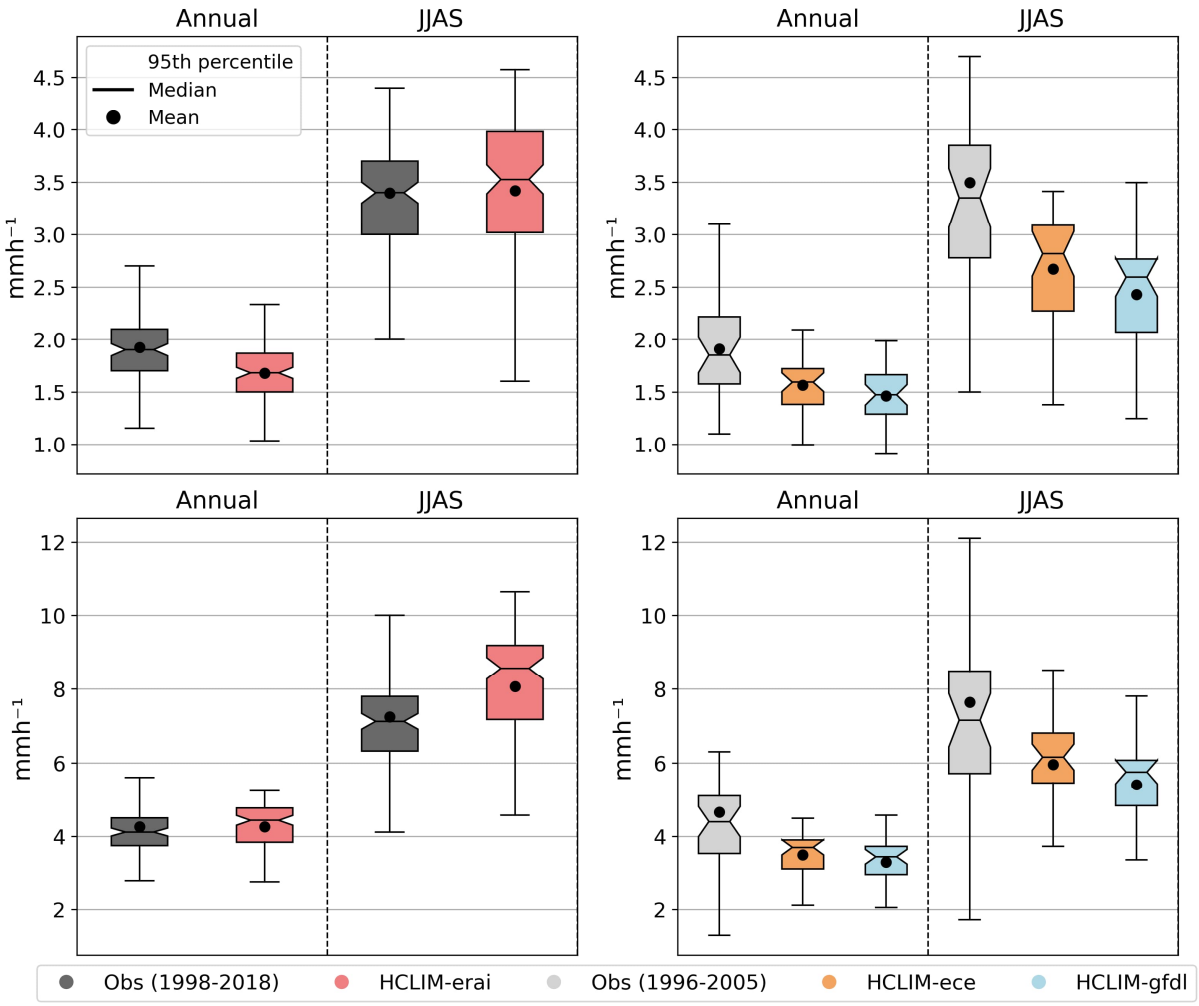

Figure B2. Distributions of the mean above the 95th percentile (top panels) and 99th percentile (bottom panels) of hourly precipitation intensity in the *in-situ* gauge observations and the HCLIM simulations. The left panels show comparisons between observations and HCLIM-erai, while the right panels show comparisons between observations and HCLIM-ece and HCLIM-gfdl. The results are presented separately for annual and JJAS (June--September) precipitation. HCLIM-erai, HCLIM-ece, and HCLIM-gfdl data are selected from grid points closest to the gauge observation stations. The extent of the boxplots represent the spatial variability in Finland. The median of the distribution is represented by the horizontal line whereas the mean is shown by the black dot. The upper and lower horizontal lines show the 75th and 25th percentile and the whiskers extend to data points within a range of 1.5 times the interquartile range (IQR) from the quartiles.

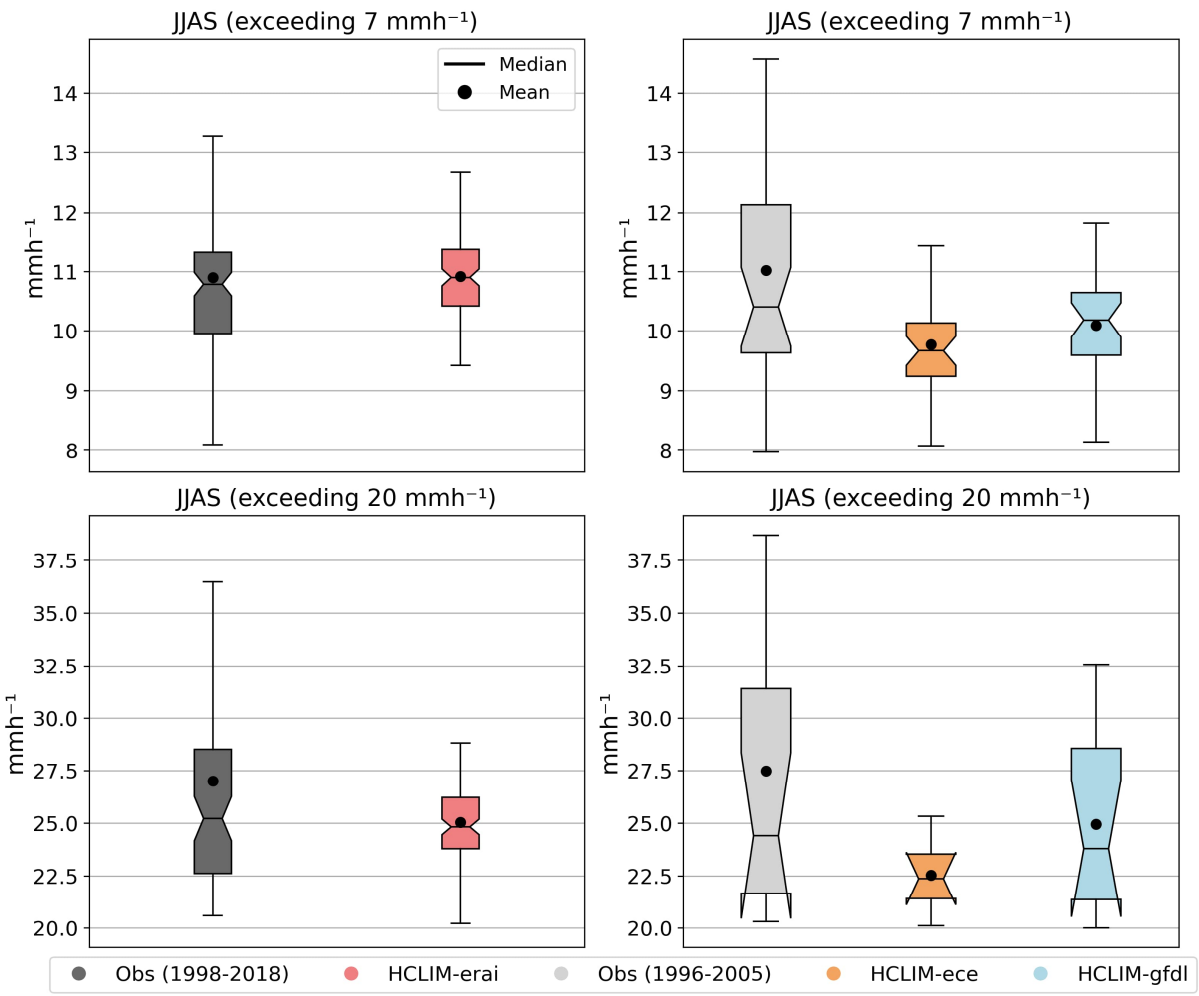

Figure B3. As in Figure B2, but for the distributions of mean hourly precipitation intensities exceeding 7 mm h^{-1} (top panels) and 20 mm h^{-1} (bottom panels) during JJAS (June--September) between *in-situ* gauge observations and model simulations. The left panels show comparison between observations and HCLIM-erai, while the right panels shows comparison between observations and HCLIM-ece and HCLIM-gfdl. HCLIM-erai, HCLIM-ece, and HCLIM-gfdl data are selected from grid points closest to the gauge observation stations. The extent of the boxplots represents the spatial variability within Finland.

C Description of radar data

We analyzed HCLIM-erai against radar-derived precipitation data for the warm season (JJAS) to examine the HCLIM's ability to replicate spatial precipitation patterns in Finland. We note that the comparison is possible only for a 6-year time period between 2013--2018, and due to this limited climatological representation, only a rough comparison of the mean wet hour precipitation intensity distribution between the radar observations and HCLIM-erai is possible.

The mean wet hour precipitation accumulation intensity were calculated from radar reflectivity composites created operationally from the Finnish national weather radar network measurements (Saltikoff et al. 2010). The reflectivity composites are constructed from the two lowest elevation angles measured by each radar by applying a weighted average rule based on the altitude of the radar measurement, range from radar and the quality of the measurements. The composite also includes a correction for the vertical profile of rain (VPR) that aims to correct for the radar beam overshooting the rain at long distances. The spatial resolution of the composites is 250 meters and the temporal resolution is 5 minutes.

The reflectivity composites were then transformed to precipitation intensity by applying the relation$Z = aR^b$, where the radar reflectivity Z is in linear units of millimetres to the sixth power per cubic meter, the rainfall rate R is in units of millimetres per hour, a = 223, and b = 1.53 when $R \leq 1.0$ mm h^{-1} and b = 1.43 when $R > 1.0$ mm h^{-1} (Leinonen et al. 2012). The hourly accumulations were calculated by summing the composites within each hour. The hourly precipitation accumulations were further corrected to remove bias compared to gauge observations. The gauge-correction was performed with a Kalman-filter based approach (Chumchean et al. 2006) where the rainfall composites are multiplied with a correction factor. The correction factor is estimated based on a logarithmic mean field bias between the radar-estimated rainfall and gauge measurements (Chumchean et al. 2006). The weather radar observations were further regridded from 250 m to 3 km to align with the model's resolution. The gauge observations were obtained from the FMI weather station network.

The data from 2013--2014 include measurements from eight radars, the data from 2015 from nine radars, and the data from 2016--2018 from ten radars. The quality of the radar measurements especially at the edges of the domain is impacted by the limited coverage of the Finnish radar network. The northernmost parts of Finland are beyond the radar coverage; areas beyond Lake Inari and also beyond Syväjärvi in Enontekiö have no data for

that reason. The limited radar coverage impacts also the measurements in eastern Finland. Specifically, in the analysis, the Kesälahti radar is included from 2015 onwards. The Nurmes radar was only installed in 2019 and as such is not included in the dataset. Therefore, in some areas in eastern Finland, the long distances from the closest radar lead to increasing measurement altitudes, which can cause precipitation to be missed or underestimated by the radar.

D Spatial distributions of HCLIM simulations

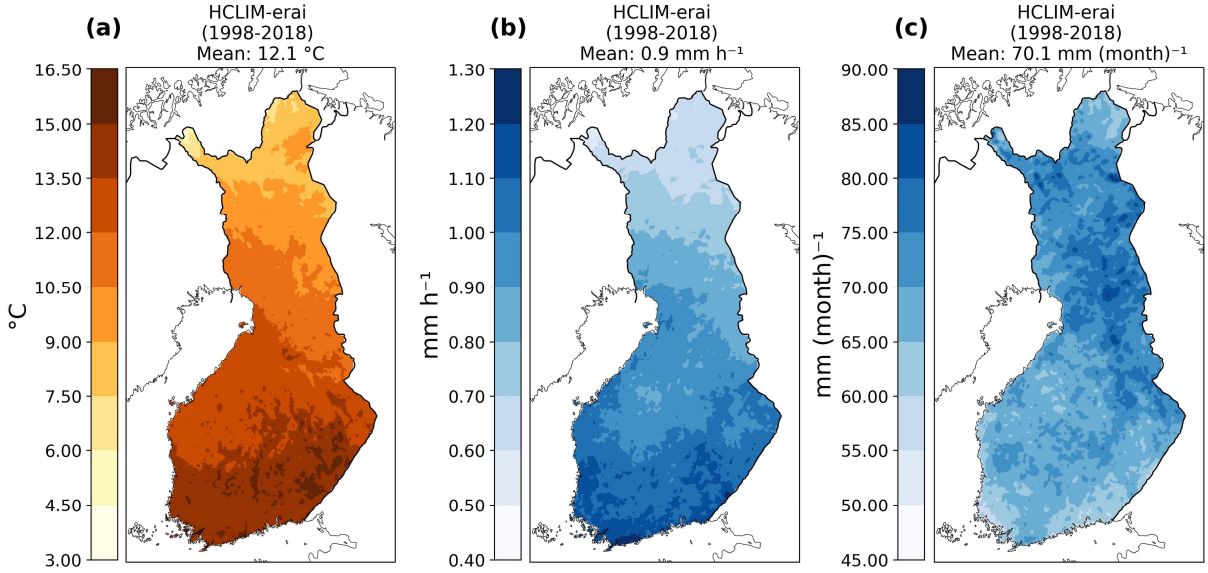

Figure D1. Spatial distributions of (a) mean temperature, (b) mean wet hour precipitation intensity, and (c) mean monthly precipitation intensity in JJAS in HCLIM-erai.

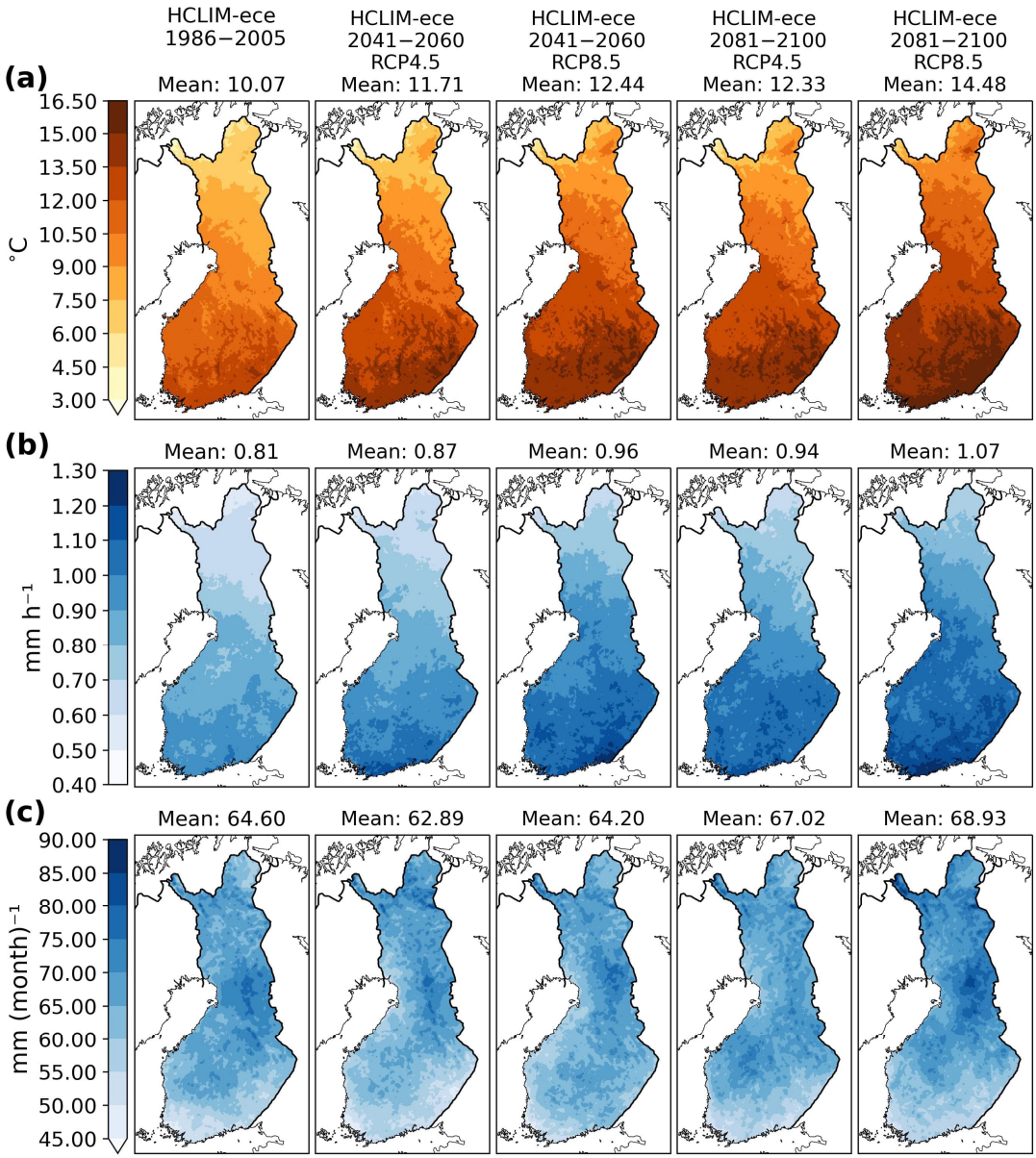

Figure D2. HCLIM-ece's spatial distributions in the historical, mid-century and late century time periods in JJAS and RCP4.5 and RCP8.5 scenarios for (a) mean temperature, (b) mean wet hour precipitation intensity, and (c) mean monthly precipitation intensity.

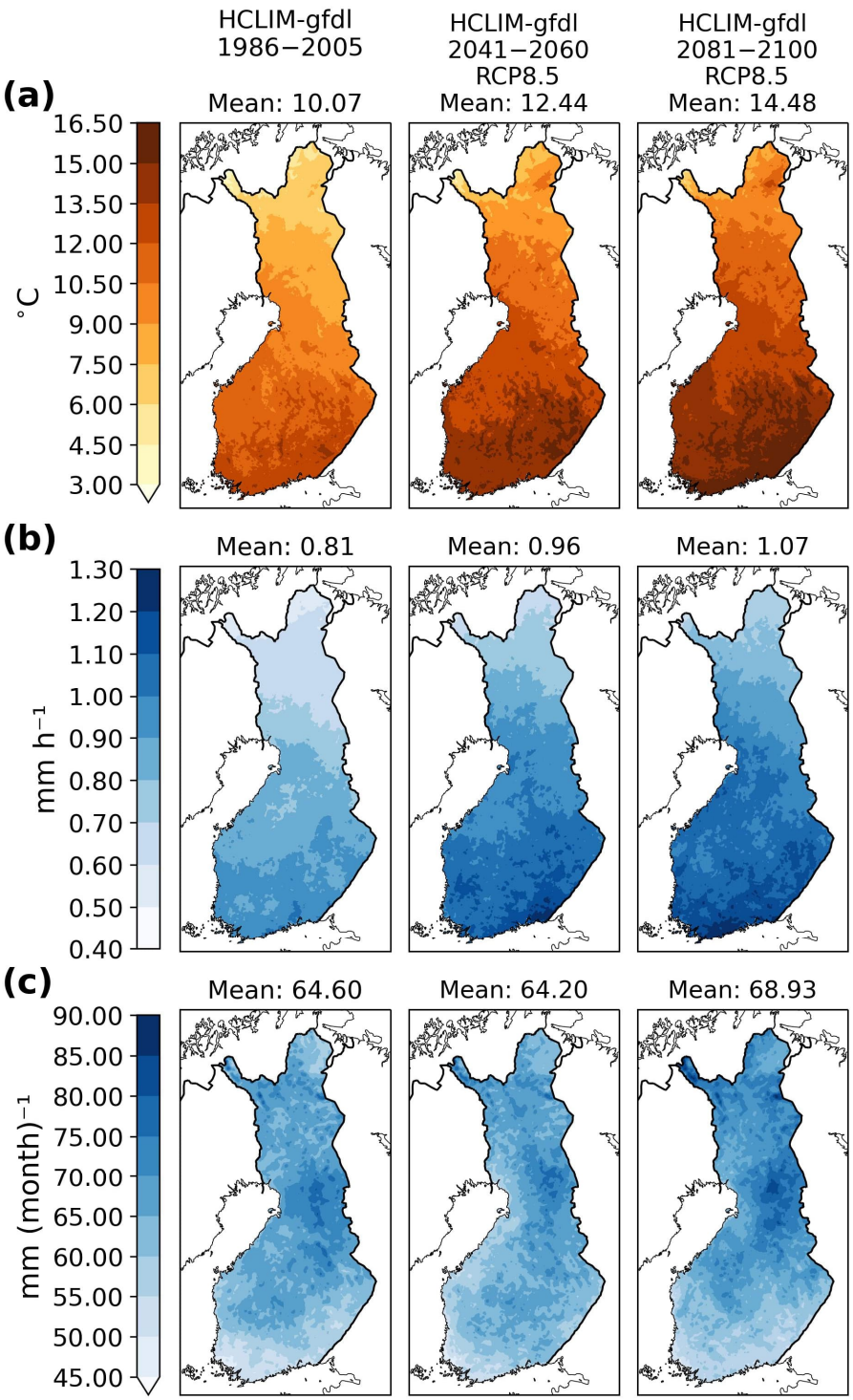

Figure D3. HCLIM-gfdl's spatial distributions in the historical, mid-century, and late-century time periods in JJAS and RCP8.5 scenario for (a) mean temperature, (b) mean wet hour precipitation intensity, and (c) mean monthly precipitation intensity.

E Gridded daily precipitation versus HCLIM

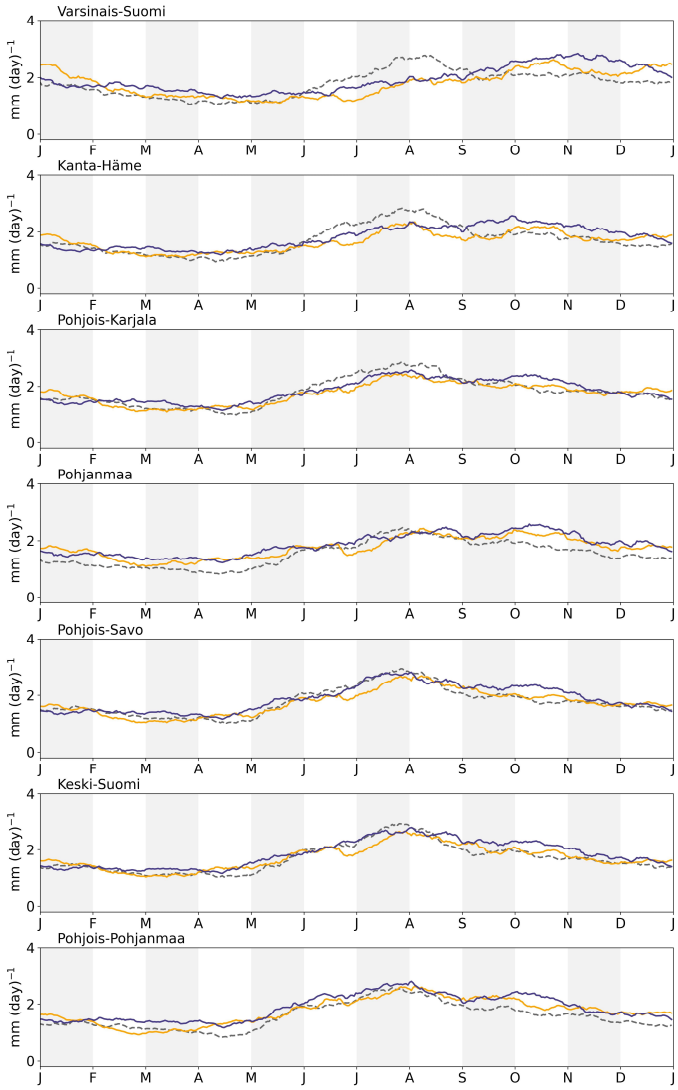

Figure E1. The annual cycle of the 30 day running mean of daily precipitation (mm day^{-1}) calculated from the gridded daily precipitation observations (grey dashed line), HCLIM-ece (orange line) and HCLIM-gfdl (blue line) in 1986--2005 in the regions of Varsinais-Suomi, Kanta-Häme, Pohjois-Karjala, Pohjanmaa, Pohjois-Savo, Keski-Suomi, and Pohjois-Pohjanmaa.

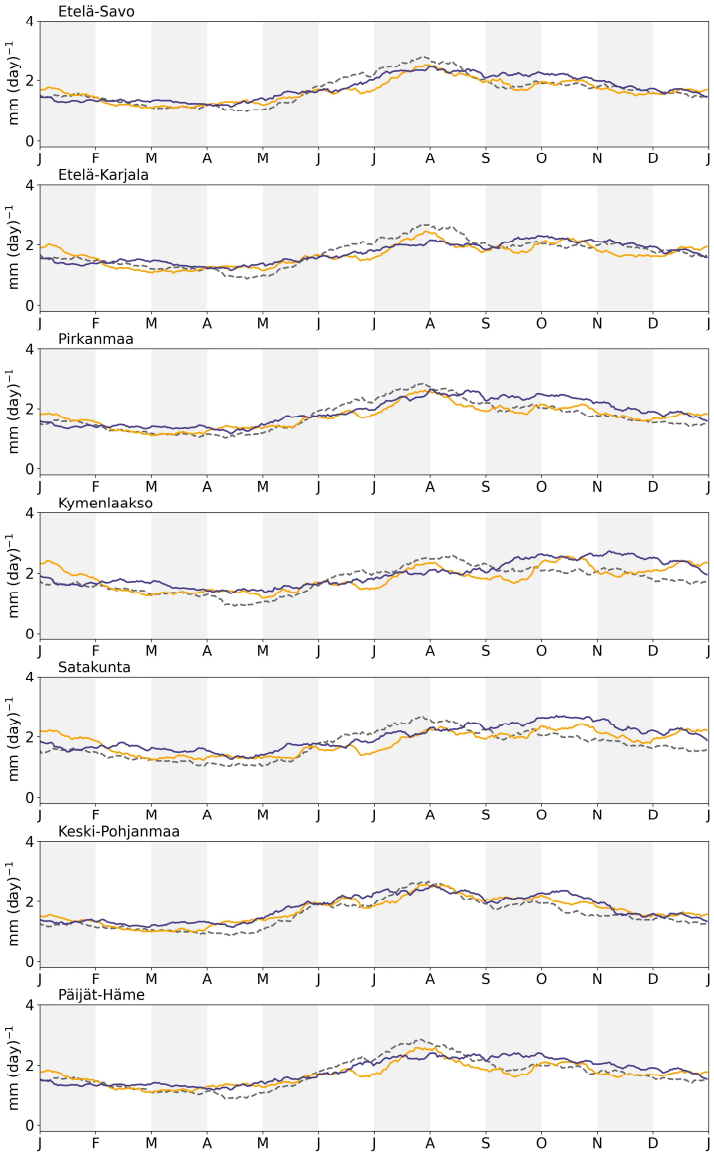

Figure E2. The annual cycle of the 30 day running mean of daily precipitation (PrD; mm day^{-1}) calculated from the gridded daily precipitation observations (grey dashed line), HCLIM-ece (orange line) and HCLIM-gfdl (blue line) in 1986--2005 in the regions of Etelä-Savo, Etelä-Karjala, Pirkanmaa, Kymenlaakso, Satakunta, Keski-Pohjanmaa, and Päijät-Häme.

F Statistical tests

The boxplots in Figure 1a--b represent the mean distribution of the in-situ gauge observations calculated yearly. The spread reflects inter-station variability. Consequently, autocorrelation is primarily related to spatial correlation between stations, rather than temporal autocorrelation. Stations located near one another may exhibit similar values due to shared weather systems. If spatial autocorrelation exists, nearby stations may not be independent. Moran's I was used to evaluate spatial autocorrelation, and a 20 km threshold distance for clustering in block bootstrapping was selected based on p-values. Stations within 20 km were grouped into the same cluster, while stations farther apart were treated as independent. Bootstrap samples were generated by resampling these spatial blocks, and the Two-Sample Kolmogorov-Smirnov (KS) test was applied to assess the maximum difference between their cumulative distribution functions at a 95 % confidence level. The 95 % confidence interval (red and blue dashed lines in Figure F1 (left)) indicates the range of KS statistics plausible under the assumption that the two distributions are similar. If the observed KS statistic falls within this range, it suggests no significant differences between the model and observations. A p-value greater than 0.05 further supports this, indicating that the null hypothesis (similarity between the model and observations) cannot be rejected. Results in Table F1 show that the model generally represents the observational data well, with no statistically significant differences for most simulations. However, a significant difference was observed for HCLIM-gfdl during the JJAS season.

The scatter plot illustrates the relationship between mean wet hour precipitation values from HCLIM-ERAI grid points and weather radar observations (Fig. F2). A statistically significant positive correlation (r=0.48) is evident, as indicated by the red line of best fit. This suggests a moderate alignment between the two datasets, emphasizing the consistency of modeled and radar observation patterns.

Table F2 summarizes the results of the Mann-Whitney U test, assessing differences in the frequency of wet hours across HCLIM simulations under various climate scenarios. Significant differences ($p < 0.05$) are observed between future projections under RCP8.5 and historical period (HCLIM-ece and HCLIM-gfdl), while no significant differences are noted for RCP4.5 (HCLIM-ece).

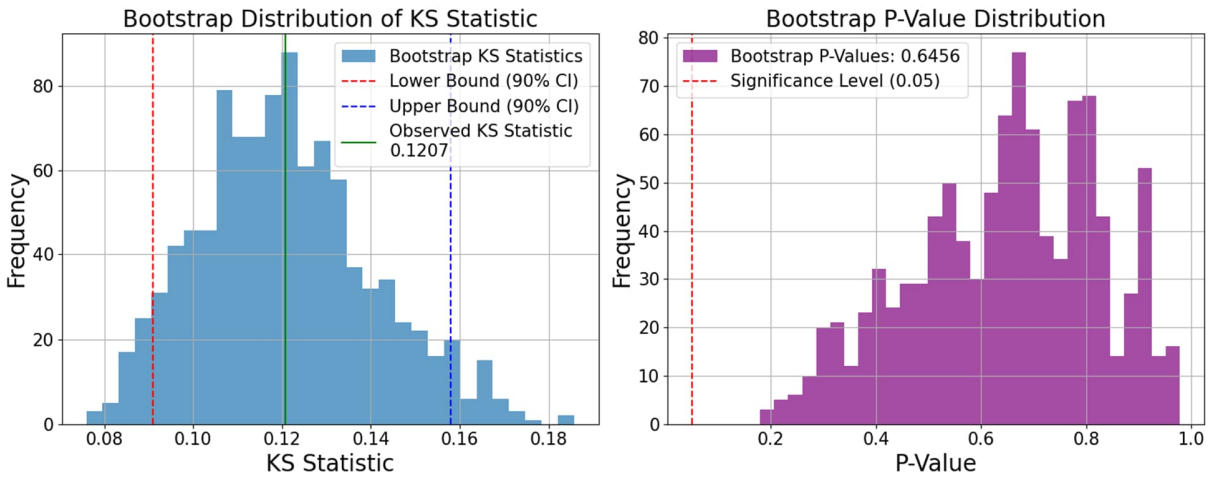

Figure F1. Distribution of the Kolmogorov-Smirnov (KS) test statistics (left) and corresponding p-values (right) for the comparison between HCLIM-erai and observations in the annual analysis. The histogram on the left represents the bootstrap distribution of KS statistics, with the observed KS statistics indicated by the vertical dashed line. The histogram on the right shows the bootstrap distribution of p-values, with the red dashed line marking the significance threshold of 0.05. The remaining results are summarized in Table F1.

Table F1. Results of the Two-Sample Kolmogorov-Smirnov (KS) test comparing the cumulative distribution functions of HCLIM simulations and observations for different time periods and seasons (see Figure 1a--b). Table shows KS statistic, mean p-value, and conclusion regarding statistical significance at the 90 % confidence level. A p-value greater than 0.05 indicates no significant difference, while a p-value less than or equal to 0.05 suggests a statistically significant difference.

Period	1. sample	2. sample	KS statistics	Mean p-value	Conclusion
1998-2018	HCLIM-erai annual	Obs annual	0.12	0.65	Not significantly different
1998-2018	HCLIM-erai JJAS	Obs JJAS	0.19	0.20	Not significantly different
1996-2005	HCLIM-ece annual	Obs annual	0.22	0.64	Not significantly different
1996-2005	HCLIM-ece JJAS	Obs JJAS	0.41	0.07	Not significantly different
1996-2005	HCLIM-gfdl annual	Obs annual	0.28	0.37	Not significantly different
1996-2005	HCLIM-gfdl JJAS	Obs JJAS	0.53	0.01	Significantly different

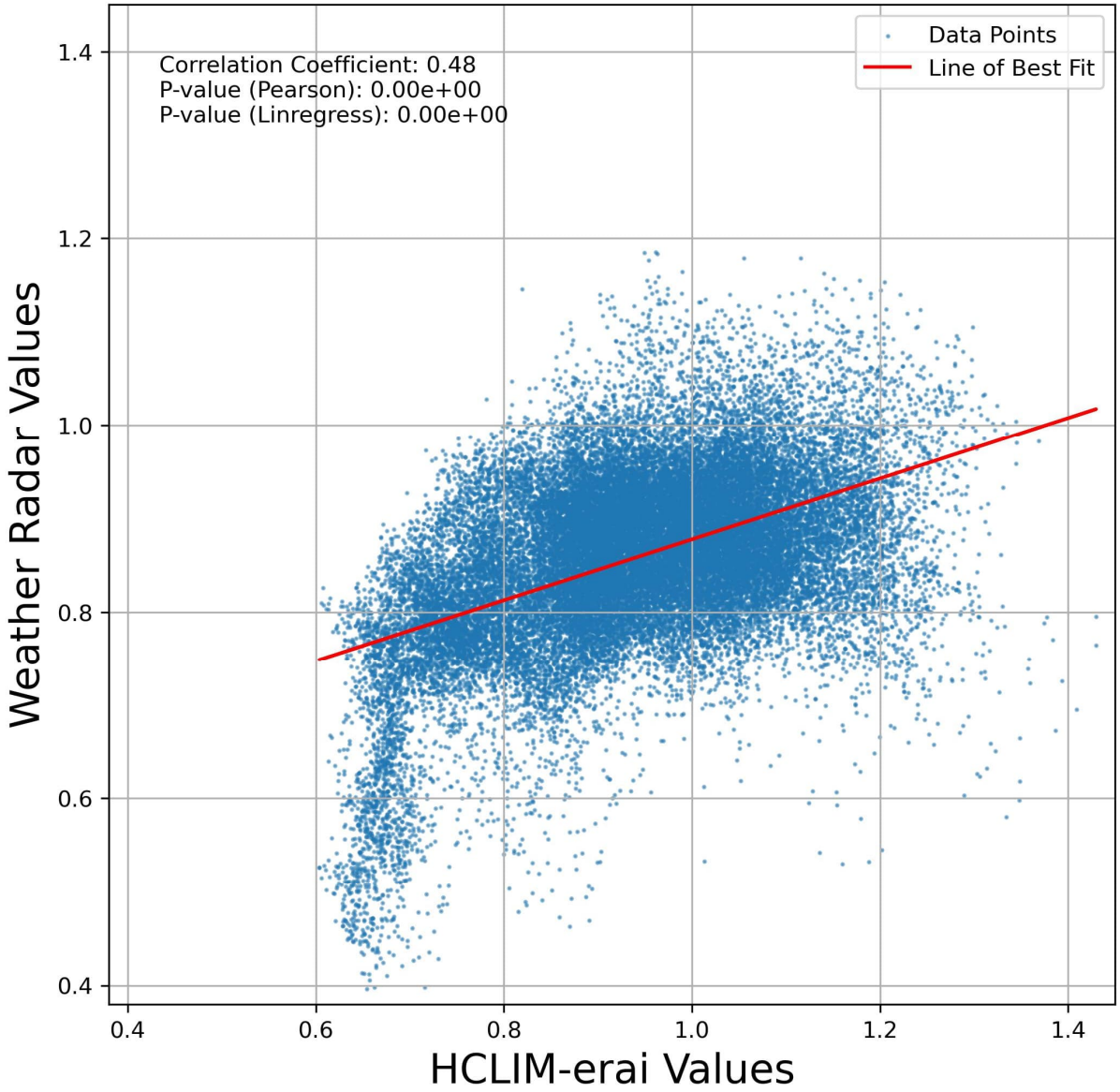

Figure F2. Scatter plot showing the relationship between mean wet hour precipitation values of HCLIM-erai grid points and weather radar observations. Each blue dot represents a data point, while the red line indicates the line of best fit. The correlation coefficient (0.48) and corresponding p-values for Pearson correlation and linear regression highlight a statistically significant positive correlation between the two datasets.

Table F2. Results of the Mann-Whitney U test evaluating the statistical significance of differences in the number of wet hours in HCLIM simulations under different climate scenarios and periods (see Figure 4). The test assesses the null hypothesis that the underlying distributions of the two samples are identical. A p-value less than 0.05 denotes a significant difference, as highlighted in 'Conclusion' column. Results demonstrate significant differences between future climate projections with scenario RCP8.5 and historical HCLIM-ece and HCLIM-gfdl simulations, while RCP4.5 shows no significant differences compared to the historical simulation.

1. sample	2. sample	U-statistics	P-value	Conclusion
HCLIM-ece 1986-2005	HCLIM-ece 2041-2060 (RCP4.5)	260.5	0.3	Not significantly different
HCLIM-ece 1986-2005	HCLIM-ece 2041-2060 (RCP8.5)	298.5	5.0e-02	Significantly different
HCLIM-ece 1986-2005	HCLIM-ece 2081-2100 (RCP4.5)	284.5	0.1	Not significantly different
HCLIM-ece 1986-2005	HCLIM-ece 2081-2100 (RCP8.5)	307.5	2.9e-02	Significantly different
HCLIM-gfdl 1986-2005	HCLIM-gfdl 2041-2060 (RCP8.5)	326.5	7.9e-03	Significantly different
HCLIM-gfdl 1986-2005	HCLIM-gfdl 2081-2100 (RCP8.5)	381.5	5.4e-05	Significantly different