

Supplementary Material:  
Future projections and uncertainty assessment of precipitation  
extremes in the Korean peninsula from the CMIP6 ensemble,  
with a statistical framework

Yonggwon Shin<sup>1</sup>,    Yire Shin<sup>1</sup>,    Juyoung Hong<sup>1</sup>,    Maeng-Ki Kim<sup>2</sup>,  
Young-Hwa Byun<sup>3</sup>,    Kyung-On Boo<sup>3</sup>,    Il-Ung Chung<sup>4</sup>,    Doo-Sun R. Park<sup>5</sup>,  
and    Jeong-Soo Park<sup>1,\*</sup>

*1: Department of Statistics, Chonnam National University, Gwangju 500-757, Korea.*

*2: Department of Atmospheric Science, Kongju National University, Gongju, Korea.*

*3: National Institute of Meteorological Sciences, Seogwipo, Korea.*

*4: Dept of Atmospheric and Environmental Sciences,  
Gangneung-Wonju National University, Korea.*

*5: Department of Earth Science Education, Kyungpook National University, Daegu, Korea.*

*\*: Corresponding author, E-mail: jspark@jnu.ac.kr, Tel: +82-62-530-3445*

January 11, 2021

## 1 Data and simulation models

Figure S1 shows examples of time series plots of the observations, APHRODITE data [1], and the bias-corrected data, for the AMP1. The APHRODITE values of the AMP1 are smaller than the observations when comparing those data at near stations. Because of this difference, we applied a bias correction technique to the APHRODITE data, based on the observations of nearest neighbor stations.

## 2 Multivariate bias correction

Some BC methods such as quantile mapping or delta change [2] make a perfect matching in the sense that the quantiles of the observations and the historical data are same. When the BC such as quantile mapping is used, most the model weights based on performance become

equal because of a perfect matching, and consequently, the prediction is the simple average of bias-corrected model outputs. This is approximately true for the MBC [4] employed in this study because the MBC is a multivariate generalization of quantile mapping. Thus the historical data is not bias corrected. No-bias-corrected historical data are utilized to calculate the performance weight of a model.

Chen et al.[3] found that the joint BC of precipitation and air temperature led to a much better performance than univariate BC, in terms of hydrological modelling for all their studying basins located in various climates except for the coldest Canadian basin. Cannon [4] proposes a multivariate generalization of quantile mapping (QM). It is an iterative method which conceptually lays between univariate bias correction (BC) methods and the empirical copula-based correction (EC-BC) [5]. For a univariate BC, the quantile delta mapping (QDM) [6] is used, which preserves trends of model data

It approximately preserves the multivariate dependence of the driving climate model. Here, an image processing technique—the N-dimensional probability density function transform (N-pdf)–designed to transfer color information from one image to another is adapted. In each iteration, univariate QM is first applied separately to each variable. Then a linear multivariate BC is applied by re-scaling the multivariate anomalies based on Cholesky decomposition of the covariance matrix. The algorithm ends when both the corrected marginals and the dependence structure are sufficiently close to their observed counter parts. A variant is based on ranks rather than on the actual values [2]. It provides a multivariate quantile delta mapping, referred to as MBCn (multivariate bias correction with N-pdf) algorithm. It consists, in each iteration, of a random orthogonal rotation of multivariate input data, a univariate quantile delta mapping on the rotated fields and the inverse rotation. This algorithm approximately preserves trends of model data. We used 'MBC' package [7] in R for computation. More details are found in Cannon [4].

### 3 Computing performance weights

To compute the performance of each model,  $T$ -year return levels are compared based on the GEVD fitting on the historical data and the observations. Let us denote  $r_T^i$  and  $r_T^0$  as  $T$ -year return level obtained from the historical data of  $i$ -th model and the observations, respectively. These values are normalized as follows to make it scale-free, for  $i = 0, 1, \dots, M$ :

$$\tilde{r}_T^i = \frac{r_T^i - med_i}{R_i}, \quad (1)$$

where

$$R_i = \begin{cases} max_i - med_i & \text{if } r_T^i \geq med_i, \\ med_i - min_i & \text{if } r_T^i < med_i, \end{cases} \quad (2)$$

and  $max_i$ ,  $min_i$ , and  $med_i$  are the maximum, the minimum, and the median of  $i$ -th model data. Other ways of standardizations are also possible.

The distance for performance measure is obtained by

$$D_i^2 = \sum_T (\tilde{r}_T^i - \tilde{r}_T^0)^2. \quad (3)$$

We set  $T = 2, 5, 10, 20, 30, 50$ , and  $100$ . Note that  $D_i$  does not depend on the shape parameter  $\sigma_D$ , and so obtained  $D_i$ s are fixed for the next computation.

## 4 Result

### 4.1 Relative change

The relative change of 20-year return level in the period P1 relative to the reference period P0 is defined as:

$$\delta R_{20}(P1) = \frac{R_{20}(P1) - R_{20}(P0)}{R_{20}(P0)} \times 100, \quad (4)$$

where  $R_{20}(P)$  is the 20-year return level in the period P.

### 4.2 Return period and exceedance probability

We have experienced some technical flows in computing the waiting time or the return periods corresponding to a return value. For example, the resulting return period sometimes turns out to be greater than 500 years even though it is expected to correspond to 50 years. It may be due to the cumulation of truncation or rounding errors in computer, related to inverting the quantile function of the GEVD. A trouble caused by this flow does not vanish even applied to unequally weighted regional frequency analysis (RFA). In this study, we thus adopted the trimmed mean [8] in RFA in which unfairly very high estimates of return periods are deleted in computing the weighted average. The defects of return periods are described in Serinaldi [9].

The spatially averaged estimates of exceedance probability over the Korean peninsula are presented in Figure S6 and in Table S6.

### 4.3 Quantifying uncertainty

From the analysis of variance, Figure S7 shows the interaction plots between 21 CMIP6 and the latitude in which the latitude changes from  $33^\circ$  to  $43^\circ$ , for 20-year return levels (unit: mm).

## References

- [1] Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A (2012) APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull Amer Meteorol Soc* 93:1401–1415.
- [2] Maraun D, Widmann M (2018) *Statistical downscaling and bias correction for climate research*. Cambridge Univ Press.
- [3] Chen J, Li C, Brissette FP, Chen H, Wang M, Essou GRC (2018) Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling. *J Hydrol* 560:326–341.
- [4] Cannon AJ (2018) Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Clim Dyn* 50:31–49.
- [5] Vrac M, Friederichs P (2015) Multivariate-intervariable, spatial, and temporal-bias correction. *J Clim* 28:218–237.
- [6] Cannon AJ, Sobie SR, Murdock TQ (2015) Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *J Climate* 28:6938–6959.
- [7] Cannon AJ (2018) Package ‘MBC’ in R, version 0.10-4. <https://cran.r-project.org/web/packages/MBC/MBC.pdf>
- [8] Wilks D (2011) *Statistical Methods in the Atmospheric Sciences*, 3rd Ed., Academic Press, New York.
- [9] Serinaldi F (2015) Dismissing return periods! *Stoch Environ Res Risk Assess* 29:1179–1189. <https://doi.org/10.1007/s00477-014-0916-1>

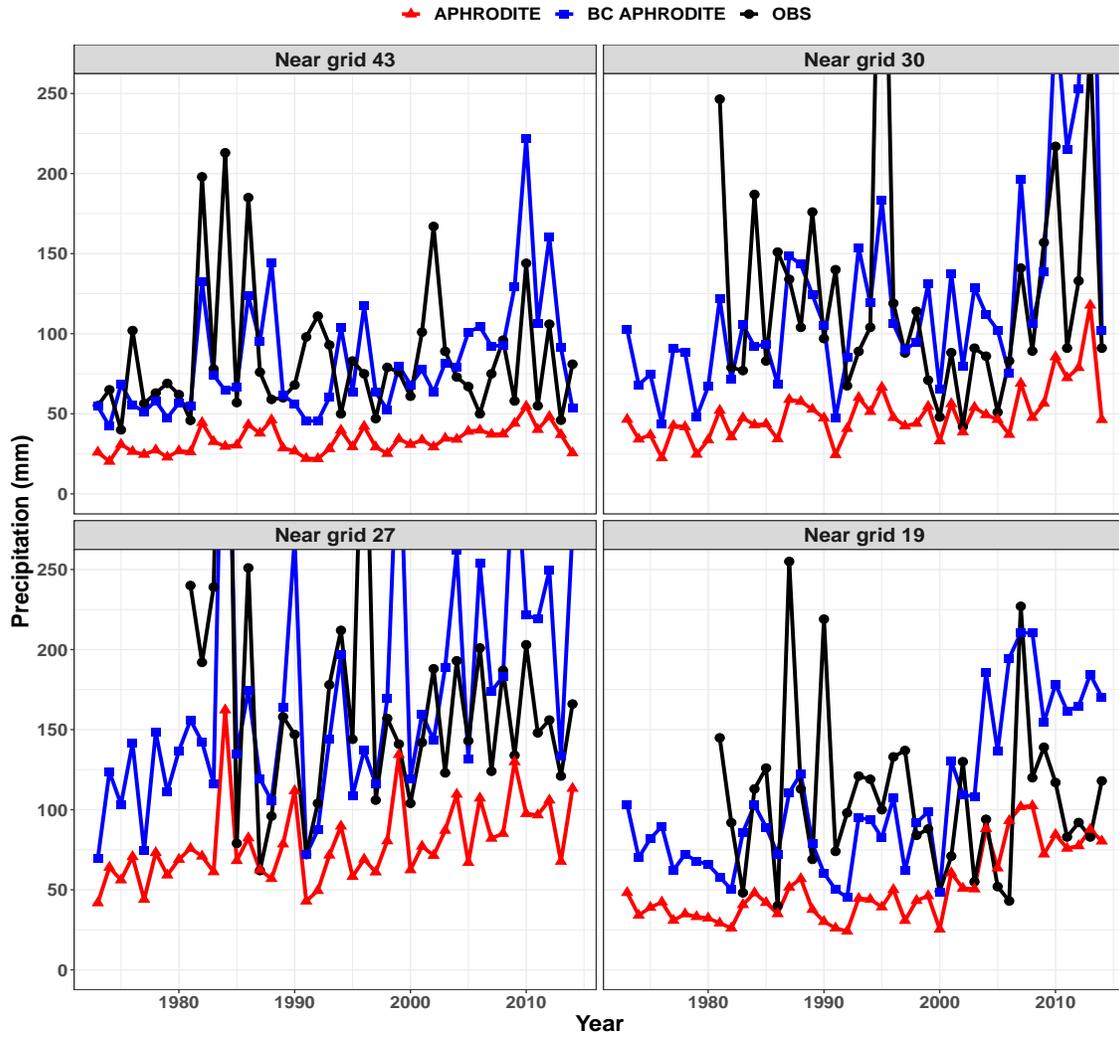


Figure S 1: Examples of time series plots of the observations (black line), APHRODITE data (red line), and the bias-corrected data (blue line) in the Korean peninsula.

Table S 1: The list of 21 CMIP6 (Coupled Model Intercomparison Project Phase 6) models analyzed in this study. The detailed information on each model are available at ESGF-node <https://esgf-node.llnl.gov/projects/cmip6/>.

Model Name	Institution	Country	Resolution (Lon × Lat Level)
MIROC6	JAMSTEC, AORI, NIES, R-CCS (MIROC)	Japan	256×128 L81(T85)
BCC-CSM2-MR	Beijing Clim Center	China	320×160 L46(T106)
CanESM5	Canadian Centre Clim Model & Analysis, Enviro & Clim Change (CCCma)	Canada	128×64 L49(T63)
MRI-ESM2.0	Meteoro Research Institute (MRI)	Japan	320×160 L80(TL159)
CESM2-WACCM	Nat Center for Atmos Res, Clim & Global Dynamics Lab (NCAR)	USA	288×192 L70
CESM2	Nat Center for Atmos Res, Clim & Global Dynamics Lab (NCAR)	USA	288×192 L32
KACE1.0-GLOMAP	National Inst of Meteo Sci/Meteo Admin, Clim Res Div (NIMS-KMA)	Korea	192×144 L85
UKESM1-0-N96ORCA1	MOHC & NERC, NIMS-KMA, NIWA	UK, Korea New Zealand	192×144 L85
MPI-ESM1.2-LR	Max Planck Inst for Meteo (MPI-M)	Germany	192×96 L47(T63)
MPI-ESM1.2-HR	Max Planck Inst for Meteo (MPI-M)	Germany	384×192 L95(T127)
INM-CM5-0	Inst for Numerical Math, Russian Acad of Sci (INM)	Russia	180×120 L73
INM-CM4-8	Inst for Numerical Math, Russian Acad of Sci (INM)	Russia	180×120 L21
IPSL-CM6A-LR	Institut Pierre Simon Laplace (IPSL)	France	144×143 L79
NorESM2-LM	NorESM Consortium of CICERO, MET-Norway, NERSC, NILU, UiB, UiO, UNI	Norway	144×96 L32
NorESM2-MM	NorESM Consortium of CICERO, MET-Norway, NERSC, NILU, UiB, UiO, UNI	Norway	288×192 L32
EC-Earth3-Veg	EC-Earth consortium, Swedish Meteo & Hydro Inst/SMHI, Sweden	EU	512×256 L91(TL255)
EC Earth 3.3	EC-Earth consortium, Swedish Meteo & Hydro Inst/SMHI, Sweden	EU	512×256 L91(TL255)
ACCESS-CM2	CSIRO, ARCCSS (Australian Res Council Centre of Excellence for Clim System Sci)	Australia	192×144 L85
ACCESS-ESM1-5	Commonwealth Scientific & Industrial Res Organ (CSIRO)	Australia	192×145 L38
GFDL-ESM4	National Oceanic & Atmos Admi, Geophy Fluid Dynamics Lab	USA	360×180 L49
FGOALS-g3	Chinese Academy of Sciences (CAS)	China	180×80 L26

Table S 2: The similarity distance metric  $S_{ij}$  between model  $i$  and model  $j$ . Small value indicates high dependency or high similarity between two models.

	UKESM2	CanESM5	EC-Earth3-Veg	KACE-1-0-G	GFDL-ESM4	INM-CM5-0	MPI-ESM1-2-HR
UKESM	0.00	0.53	0.59	0.61	0.73	0.75	0.78
CanESM5	0.53	0.00	0.58	0.63	0.75	0.72	0.77
EC-Earth3-Veg	0.59	0.58	0.00	0.68	0.77	0.75	0.78
EC-Earth3	0.59	0.58	0.59	0.68	0.76	0.78	0.80
ACCESS-CM2	0.58	0.61	0.66	0.68	0.80	0.81	0.81
KACE-1-0-G	0.61	0.63	0.68	0.00	0.83	0.82	0.83
CESM2	0.66	0.66	0.70	0.74	0.79	0.81	0.83
IPSL-CM6A-LR	0.66	0.65	0.68	0.74	0.81	0.80	0.82
ACCESS-ESM1-5	0.65	0.67	0.70	0.74	0.78	0.82	0.83
CESM2-WACCM	0.65	0.66	0.70	0.75	0.79	0.84	0.83
NorESM2-MM	0.67	0.68	0.70	0.75	0.81	0.83	0.84
NorESM2-LM	0.68	0.66	0.70	0.75	0.81	0.83	0.83
MRI-ESM2-0	0.67	0.68	0.71	0.77	0.80	0.82	0.84
MIROC6	0.69	0.68	0.70	0.78	0.81	0.82	0.84
BCC-CSM2-MR	0.70	0.71	0.73	0.78	0.82	0.85	0.87
FGOALS-g3	0.72	0.70	0.73	0.79	0.84	0.84	0.86
INM-CM4-8	0.74	0.73	0.76	0.81	0.86	0.84	0.88
MPI-ESM1-2-LR	0.73	0.74	0.76	0.81	0.86	0.86	0.84
GFDL-ESM4	0.73	0.75	0.77	0.83	0.00	0.89	0.88
INM-CM5-0	0.75	0.72	0.75	0.82	0.89	0.00	0.88
MPI-ESM1-2-HR	0.78	0.77	0.78	0.83	0.88	0.88	0.00
SUM	13.38	13.39	13.95	14.96	16.18	16.36	16.63

Table S 3: Statistics of 20-year and 50-year return levels of the annual maximum daily precipitation (unit: mm) averaged over 46 grids in the Korean peninsula for the observations (OBS) and the future periods; P1 (2021-2050), P2 (2046-2075), and P3 (2071-2100) under the SSP2, SSP3, and SSP5 scenarios.

			SSP2-4.5			SSP3-7.0			SSP5-8.5		
Year	Statistic	OBS	P1	P2	P3	P1	P2	P3	P1	P2	P3
20-year	Mean	210	226	238	250	228	244	265	233	259	288
	Q1	162	188	196	207	190	203	217	184	207	239
	Median	229	239	251	261	241	260	279	250	276	303
	Q3	251	262	277	285	262	277	310	271	297	322
50-year	Mean	259	273	290	306	274	296	323	282	317	354
	Q1	200	235	245	258	230	250	268	228	254	302
	Median	274	283	300	309	287	308	340	294	336	369
	Q3	304	314	332	350	307	335	376	320	357	391

Table S 4: Relative change (unit: %) in 20-year and 50-year return levels of the annual maximum daily precipitation averaged over the Korean peninsula relative to 1973–2010.

		SSP2-4.5			SSP3-7.0			SSP5-8.5		
		P1	P2	P3	P1	P2	P3	P1	P2	P3
20-year	Mean	7.2	13.4	19.0	7.9	16.3	26.0	10.7	23.3	37.9
	Q1	4.7	9.5	14.2	7.6	12.0	22.1	6.2	17.2	26.8
	Median	7.8	13.0	18.4	8.7	18.3	25.9	12.4	25.7	41.7
	Q3	10.0	17.4	23.0	10.0	20.7	30.7	15.1	29.0	46.9
50-year	Mean	5.8	12.7	19.0	6.2	15.1	25.3	9.4	23.3	38.4
	Q1	2.4	8.4	13.1	5.4	11.2	21.2	3.9	16.2	24.0
	Median	6.2	12.1	18.9	7.6	15.7	26.3	12.0	25.7	44.0
	Q3	8.2	16.8	24.2	8.4	19.8	30.2	14.3	29.7	48.9

Table S 5: Statistics of 20-year and 50-year return periods (unit: year) of the annual maximum daily precipitation averaged over 46 grids in the Korean peninsula for the future periods P1 (2021-2050), P2 (2046-2075), and P3 (2071-2100) under the SSP2, SSP3, and SSP5 scenarios.

		SSP2-4.5			SSP3-7.0			SSP5-8.5		
		P1	P2	P2	P1	P2	P3	P1	P2	P3
20-year	Mean	17.2	12.3	10.5	15.0	11.7	8.7	14.4	9.8	6.7
	Q1	14.2	10.6	9.0	13.1	9.9	7.1	12.4	7.9	5.8
	Median	16.1	12.2	10.6	13.9	11.1	8.4	13.7	8.9	6.7
	Q3	18.4	13.6	11.5	15.7	13.4	10.0	15.4	11.0	7.4
50-year	Mean	39.3	33.7	26.9	40.3	29.7	20.5	36.0	24.6	15.6
	Q1	34.2	25.9	19.8	31.9	25.2	16.1	29.1	17.2	12.5
	Median	38.8	31.8	26.4	36.5	29.0	19.6	35.5	21.5	14.5
	Q3	44.0	40.7	30.4	47.3	33.1	23.6	40.2	29.0	18.4

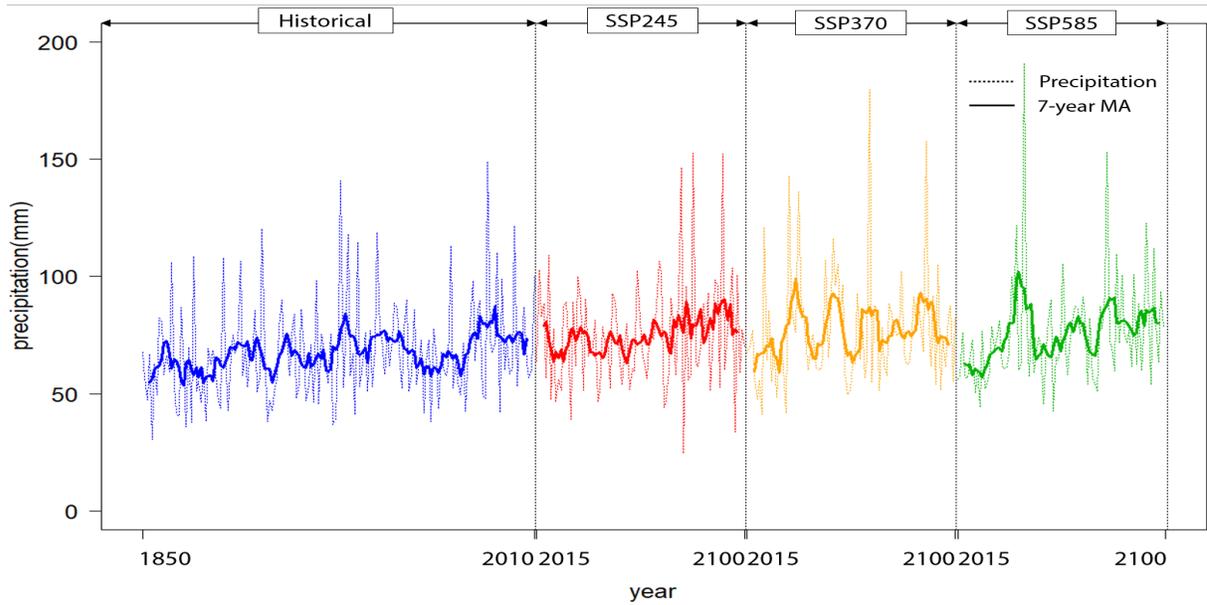


Figure S 2: Arrangement of data and 7-year moving averages composed of the historical data from 1850 to 2010 and the future data from 2015 to 2100 under SSP2, SSP3, and SSP5 scenarios for computing the Spearman correlation coefficient between models.

Table S 6: Spatially averaged the exceedance probability over the Korean peninsula for the annual maximum daily precipitation (AMP1) from 100mm to 500mm, obtained from the observations (OBS) and the CMIP6 models under the three scenarios for three future periods.

	AMP1	OBS	SSP2-4.5	SSP3-7.0	SSP5-8.5
Period 1	100 mm	0.550	0.715	0.705	0.741
	150 mm	0.227	0.227	0.249	0.274
	200 mm	0.091	0.069	0.076	0.077
	250 mm	0.033	0.024	0.023	0.029
	300 mm	0.013	0.011	0.010	0.013
	400 mm	0.002	0.003	0.002	0.004
	500 mm	0.001	0.001	0.001	0.002
Period 2	100 mm	0.550	0.741	0.771	0.799
	150 mm	0.227	0.256	0.317	0.335
	200 mm	0.091	0.090	0.098	0.125
	250 mm	0.033	0.032	0.035	0.040
	300 mm	0.013	0.013	0.017	0.020
	400 mm	0.002	0.003	0.004	0.006
	500 mm	0.001	0.001	0.002	0.002
Period 3	100 mm	0.550	0.756	0.837	0.848
	150 mm	0.227	0.272	0.403	0.402
	200 mm	0.091	0.095	0.142	0.169
	250 mm	0.033	0.037	0.059	0.067
	300 mm	0.013	0.018	0.024	0.035
	400 mm	0.002	0.005	0.005	0.009
	500 mm	0.001	0.002	0.002	0.003

Table S 7: The expected frequency of reoccurring years during 30 years for specific the annual maximum daily precipitation (AMP1) values from 100mm to 500mm in the Korean peninsula, obtained from the observations (OBS) and the CMIP6 models under the 3 scenarios for 3 future periods.

		SSP2-4.5			SSP3-7.0			SSP5-8.5		
AMP1	OBS	P1	P2	P3	P1	P2	P3	P1	P2	P3
100 mm	16.491	21.462	22.245	22.695	21.165	23.142	25.095	22.218	23.961	25.434
150 mm	6.810	6.807	7.695	8.172	7.470	9.504	12.093	8.235	10.062	12.060
200 mm	2.715	2.067	2.709	2.862	2.271	2.940	4.266	2.313	3.759	5.085
250 mm	0.978	0.720	0.945	1.107	0.699	1.044	1.779	0.873	1.206	2.004
300 mm	0.384	0.324	0.396	0.543	0.291	0.513	0.723	0.393	0.597	1.038
400 mm	0.063	0.075	0.102	0.147	0.063	0.117	0.150	0.132	0.180	0.282
500 mm	0.015	0.027	0.036	0.051	0.027	0.045	0.054	0.057	0.060	0.099

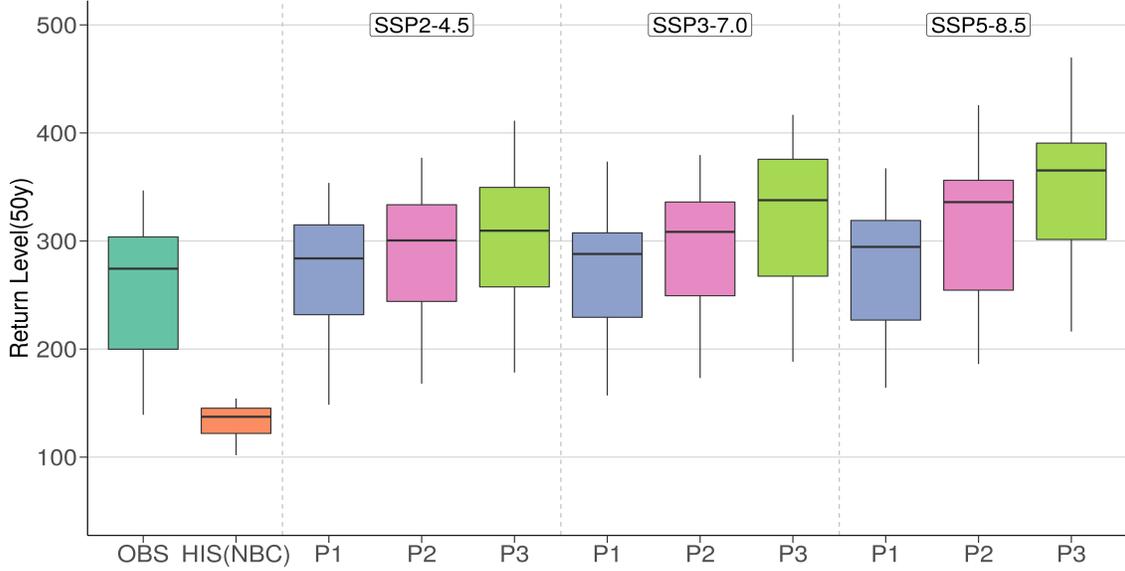


Figure S 3: Schematic box-plots of 50-year return levels of the annual maximum daily precipitation (unit: mm) averaged over 46 grids in the Korean peninsula for the future periods P1 (2021-2050), P2 (2046-2075), and p3 (2071-2100) under the SSP2, SSP3, and SSP5 scenarios. OBS and HIST(NBC) stand for the observations and the historical data without bias correction.

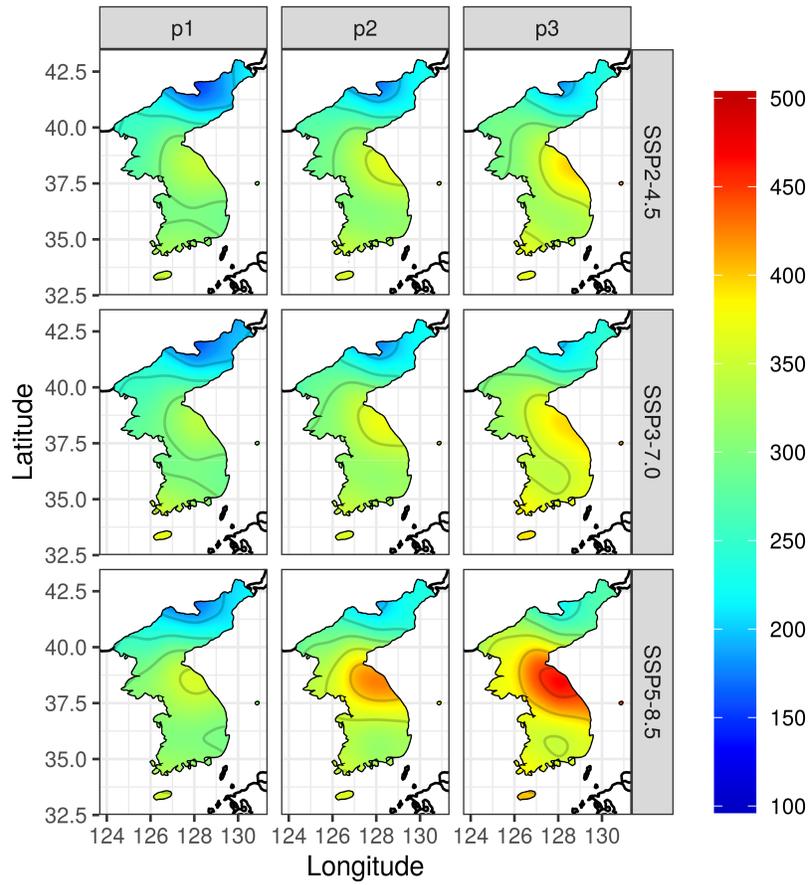


Figure S 4: Isopluvial maps of 50-year return levels of the annual maximum daily precipitation for 46 grids over the Korean peninsula for the future periods P1 (2021-2050), P2 (2046-2075), and P3 (2071-2100) under the SSP2, SSP3, and SSP5 scenarios.

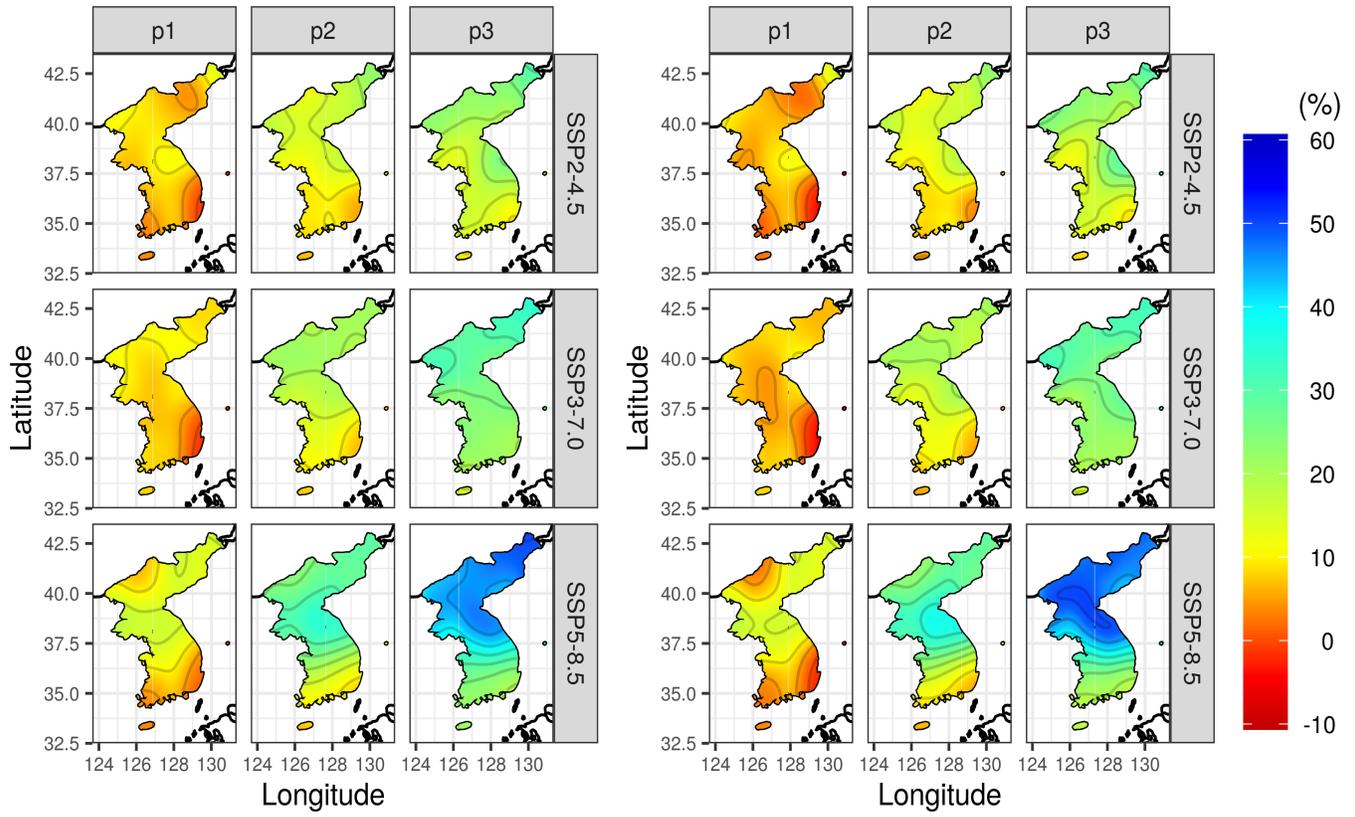


Figure S 5: Isopluvial maps of for the relative changes (unit: %) of 20-year and 50 return levels relative to 1973–2010 for the annual maximum daily precipitation for 46 grids over the Korean peninsula for the future periods P1 (2021-2050), P2 (2046-2075), and p3 (2071-2100) under the SSP2, SSP3, and SSP5 scenarios.

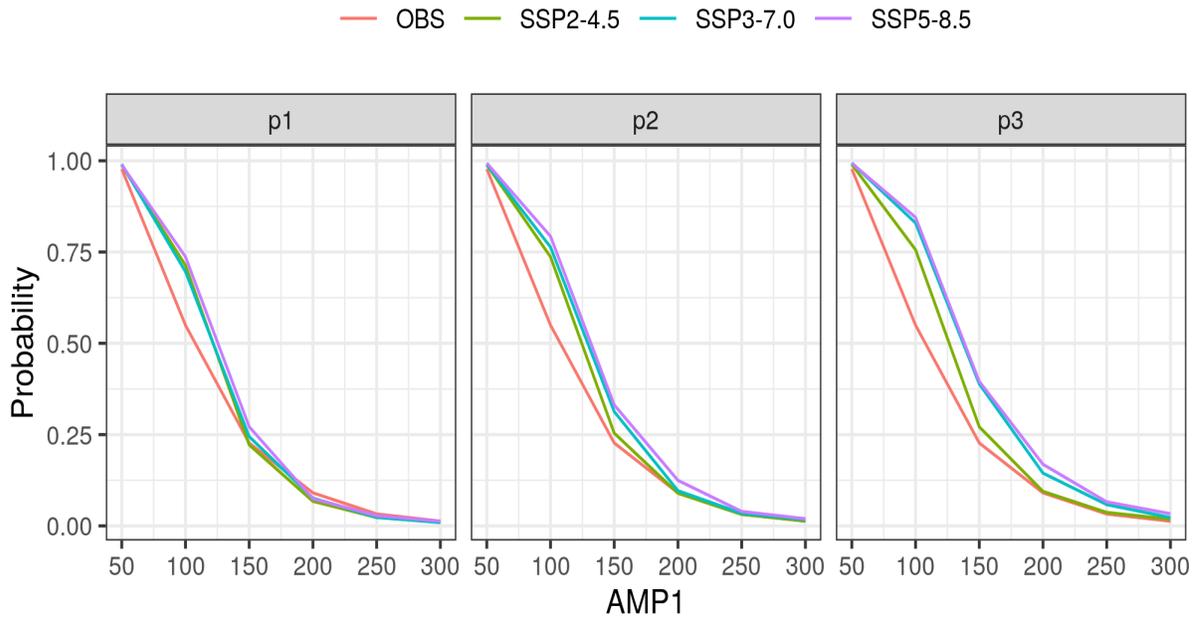


Figure S 6: The exceedance probability plots for the annual maximum daily precipitation (AMP1) from 50mm to 300mm in the Korean peninsula, obtained from the observations (OBS) and the CMIP6 models under the three scenarios for three future periods.

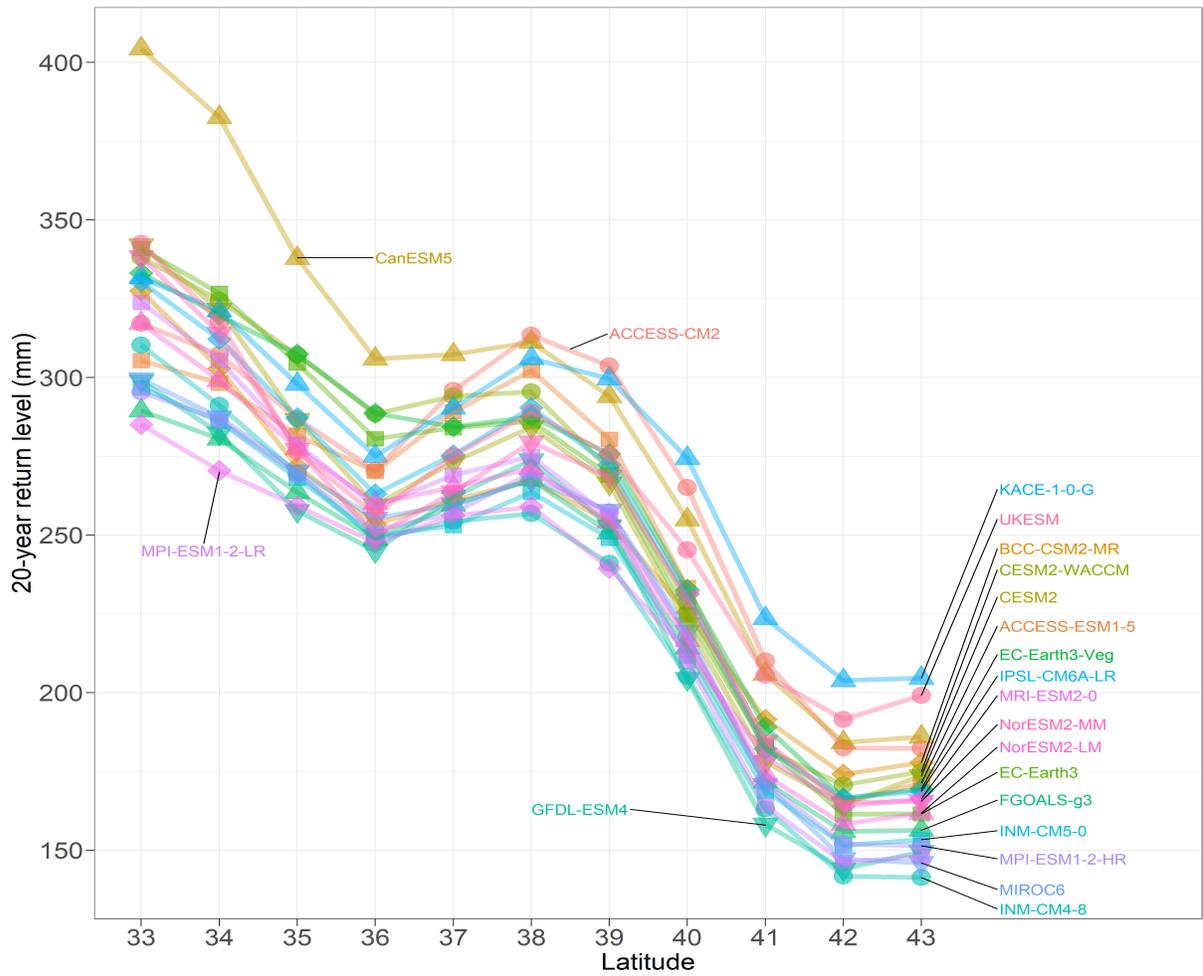


Figure S 7: Interaction plots between 21 CMIP6 models and the latitude in which the latitude changes from 33° to 43°, for 20-year return levels (unit: mm) computed over the Korean peninsula.