# Verifying the relationships among the variabilities of summer rainfall extremes over Japan in the d4PDF climate ensemble, Pacific sea surface temperature, and monsoon activity

Shao-Yi Lee[1], Sicheng He[1], Tetsuya Takemi[1]

[1]Disaster Prevention Research Institute, Kyoto University, Uji, Kyoto 611-0011, Japan

*Correspondence to*: Tetsuya Takemi (takemi@storm.dpri.kyoto-u.ac.jp)

**Abstract.**

Upper 99th percentile hourly, 90th percentile daily, and 90th percentile pentad rainfall was calculated over the four large Japanese islands for June–July every year in the 1952–2010 period, using 10 ensemble members of the 5 km resolution d4PDF (database for Policy Decision-making for Future climate changes) climate ensemble, and 126 rain-gauges. The HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) algorithm was used to cluster d4PDF grid-points. Spearman correlation was calculated between rainfall extremes of the d4PDF sets (clusters and unclassified points), and standardised scores of five modes from the rotated extended Principle Component Analysis of Pacific SST (Sea Surface Temperature) anomalies. These modes represent ENSO (El Niño-Southern Oscillation) growth, ENSO decay, warming trend, PDV (Pacific Decadal Variability), and non-canonical ENSO. Correlations from rain-gauges and d4PDF were at most weak ($|R| \leqslant 0.4$) and moderate ($|R| \leqslant 0.6$), respectively. The correlation patterns for the first two SST modes were somewhat similar between d4PDF and rain-gauges, but were biased spatially and magnitude-wise. Spatial patterns for higher modes were different between d4PDF and rain-gauges. Rainfall extremes were better related to regional monsoon indices; from rain-gauges, up to moderate correlation with the seasonal anomaly of monsoon jet zonal water vapour flux,. Pacific SST modes can modulate rainfall extremes through the monsoon, but were not major controlling factors in the occurrence of rainfall extremes. However, we hypothesize that multiple SST modes may stack to strengthen their modulating effect, and recommend further investigation with sensitivity simulations on case studies.

## 1 Introduction

Extreme rainfall events in recent summers have prompted the identification of regional conditions favorable to such developments (Harada et al. 2020, Yokoyama et al. 2020, Nayak and Takemi 2021, Naka and Takemi 2023). There is concern that climate change is increasing the strength of such events (Nayak and Takemi, 2019; Kamae et al., 2021; Mori et al., 2021), and if so, there is interest in what spatio-temporal scales of rainfall are impacted (Fujibe 2016, Unuma and Takemi 2021). The authors wish to evaluate the changes of monsoon rainfall extremes (if any) in the future warmer climate; for this purpose, a standard procedure of calculating climate projections begins by comparing and verifying historical or control climate simulation with observations. The possible dependence of such extremes on internal climate variability complicates the calculation of climate projections since climate variability itself may change in the future climate. Hence, in this study we evaluated the relationships between the variabilities of summer rainfall extremes over western Japan, monsoon activity, and Pacific sea surface temperature (SST).

Many hemispheric-scale modes of climate variability may interact with the regional- and global-scale signals over Japan; however, since Japan is an island located at the edge of the Pacific Ocean, it seems highly possible that the modes of Pacific SST have impacts on extreme rainfall over Japan. The dominant mode of Pacific variability is the El Niño-Southern Oscillation (ENSO). Although ENSO takes place in the tropical Pacific, it influences SSTs in the mid-latitudes and even globally in the form of the "Pacific-North America pattern" (Horel and Wallace 1981, Hoskins and Karoly 1981; see Alexander et al. 2002 for a literature review).

Over East Asia, the development of an El Niño during winter influences the following summer through the "Anomalous Philippine Sea Anti-cyclone", persistent anti-cyclonic wind anomalies over the western North Pacific to the east of the Philippines which result in a wetter monsoon (e.g. Wang et al. 2000). The persistence of the El Niño signal has been explained by the "Indo-western Pacific Ocean Capacitor", or teleconnections into and returning from the Indian Ocean (Kosaka et al. 2013, Xie et al. 2016), which excite a series of alternating high-and-low pressure, temperature and convection across the region termed as the "Pacific-Japan" (P-J) pattern (Nitta, 1987).

While the El Niño-associated anomalous anti-cyclone is one lobe of the of the P-J pattern, the P-J pattern is excited not only by ENSO. A positive Indian Ocean Dipole (IOD) occurring by itself without an El Niño would also have a teleconnection to the Pacific. An extremely warm Indian Ocean since 2019 likely contributed to the extreme rainfall event of 2020 (Takaya et al. 2020), in conjunction with other exacerbating factors like the Madden-Julian Oscillation (Zhang et al. 2021). The P-J pattern can also be excited by Rossby wave breaking (Takemura and Mukougawa 2022), which increases during Niña-like conditions in summer (Takemura et al. 2020). Such results suggest that it will be difficult to find clean and robust relationships between ENSO phase and rainfall over Japan.

An early study using data for a 30-year period (1951–1980) found no evident relationship between the Baiu (monsoon front) rainfall and the Southern Oscillation Index (SOI) of the season (Ninomiya and Mizuno, 1987). A study using another 30-year (1963–1992) period data found that the relationship changed from Baiu rainfall lagging the Nino3 index during 1963–1977, to leading Nino3 during 1978–1992 (Tanaka, 1997). A study of rainfall in Fukuoka (in Kyushu) for about a century (1890–2000) using categorised SOI found heavy rainfall to lag the "Strong La Niña" category (Kawamura et al. 2001, Jin et al. 2005). Yet another study for 54 years (1958–2011) found that weather patterns associated with heavy rainfall over Japan occurred more frequently when the Nino3.4 index of the season was high (Ohba et al. 2015). Such examples suggest that the relationship between ENSO and rainfall over Japan may depend on the period of data or even the location being studied. ENSO has a period of about 5–7 years, so even a 100-year period would contain only about 20 cycles. A 400-year study using paleoclimate proxies concluded that the strength and even sign of the relationship varied over time, and proposed that the Pacific Decadal Oscillation modulated the relationship (Sakashita et al. 2016). It appears that scientific literature is far from unanimous regarding the relationship between ENSO and rainfall variability over Japan.

Since we were unable to draw satisfactory conclusions from those past studies, we decided to investigate the relationships between Pacific climate variability and extreme rainfall variability in the monsoon season over Japan in observations first, before applying the same methodology to climate models. The methodology was based on two considerations:

Firstly, although the SOI or Niño indices indicate the ENSO-related state of the atmosphere and ocean respectively, ENSO phases are not instantaneous states but processes that develop and decay over the course of two years, even when SOI/Niño indices may have similar values before and after the ENSO peak. Clearer relationships may emerge between rainfall and an indicator measuring the direction of ENSO progress.

Secondly, the value of SOI/Niño indices may reflect not only the ENSO state, but also the states of other modes, such as the above-mentioned Pacific Decadal Oscillation. This study will directly decompose Pacific temperature anomalies through Principle Component Analysis and correlate between each mode and extreme rainfall, instead of using the SOI/Niño indices.

The structure of this paper is organized as follows: Section 2 describes the data and methods used. Section 3 describes the results from the Principle Component Analysis of Pacific SST, rainfall clustering, and correlation of SST modes with clustered rainfall extremes. In Section 4, we discuss how the SST modes may act on rainfall through monsoon activity. Finally, Section 5 summarises the findings of the study.

## 2 Data and Methods

### 2.1 Data and tools

#### 2.1.1 Rainfall data

The database for Policy Decision-making for Future climate changes (d4PDF) is a collection of ensemble climate simulations (Mizuta et al. 2017, Ishii and Mori 2020). A 100-member historical-warming (HPB) climate ensemble has been simulated by the Meteorological Research Institute (MRI) Atmospheric General Circulation Model (AGCM) version 3.2 (MRI-AGCM3.2; Mizuta et al. 2012), of which 12 members have been downscaled to hourly temporal and 5 km spatial resolutions (Kawase et al., 2023). "d4PDF" in this study will refer to the downscaled data; the AGCM will termed "d4PDF-AGCM". Ensemble mean rainfall extremes from 10 members of the 5 km d4PDF was compared against rain-gauge rainfall extremes from 126 meteorological stations in Japan. The analysis period was 1952–2010, or 59 years.

Rainfall during June and July was analysed. This is the season when the monsoon (Baiu) front passes over the study domain. An analysis of typhoon frequency was carried out for each month using the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al., 2018) for observations, and Webb et al. ( 2019) for d4PDF. The number of typhoons was found to increase substantially in August. Evaluating typhoon changes in the future climate is not trivial (e.g. Mori and Takemi, 2016 for a review), so the authors preferred to exclude typhoon-associated rainfall as much as possible, even though the typical calendar "summer" includes August. From a phenomenon perspective, August rainfall is climatologically quite distinct from June–July rainfall over Japan. The 99th upper percentile hourly, 90th upper percentile daily, and 90th upper percentile running-pentad rainfall values in June–July of every year were calculated. Percentiles were calculated inclusive of timesteps with no rainfall.

#### 2.1.2 Sea Surface Temperature (SST) and climate mode indices

The historical (HPB) d4PDF-AGCM ensemble was driven by randomly perturbed SST from the Centennial Observation-Based Estimates of SST version 2 (COBE-SST2; Hirahara et al., 2014). The SST data was of 1° spatial resolution and monthly temporal resolution. Data over the Pacific in the domain of 100°E–60°W, 20°S–60°N was used. Data on the Atlantic side was removed. Data over the Sea of Okhotsk (135°E–160°E, 45°N–60°N) was removed, in order to reduce the influence of sea ice. The same 59-year period as the rainfall period (1952–2010) was used for analysis.

The Southern Oscillation Index (SOI; Ropelewski and Jones, 1987) and Interdecadal Pacific Oscillation Tripole Index (TPI; Henley et al., 2015) were provided by the United States of America National Oceanic and Atmospheric Administration (NOAA). The non-standardised version of the SOI was used, but standardisation has no impact on correlation. The filtered COBE-SST version of the TPI was used, with units of degrees Celsius.

#### 2.1.3 Other meteorological variables

Meteorological variables in the region of (125–130°E, 20–50°N) from the Japanese 55-year Reanalysis (JRA55; Kobayashi et al., 2015) and d4PDF-AGCM were used to evaluate the monsoon activity. The datasets were obtained at 1.25° spatial resolution. 850 hPa daily mean temperature and specific humidity were used to calculate the monsoon front location. Zonal wind, meridional wind and specific humidity at pressure levels 1000, 925, 850, 700, 600 and 500 hPa were used to calculate water vapour flux. Monsoon activity in the simulations was evaluated using variables from d4PDF-AGCM using the same 10 ensemble members as the rainfall data. The period of 1958–2022 was used because JRA55 started from 1958.

### 2.1.4 Processing tools

Coastlines and administrative boundaries were obtained from the Database of Global Administrative Areas version 4.1 (GADM; https://gadm.org). Principle Component Analysis and part of the data processing was carried out using Max-Planck Institute for Meteorology's Climate Data Operators software package (CDO; Schulzweida, 2022). Equivalent potential temperature was calculated using the NCAR Command Language (NCL, 2019). The rest of the data processing was carried out in Python 3, using python libraries NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), pandas (McKinney, 2010) and xarray (Hoyer and Hamman, 2017). Clustering of rainfall and curve fitting was carried out using Scipy. Figures were prepared with the python libraries Matplotlib (Hunter, 2007) and Cartopy (Met Office, 2010–2015).

## 2.2 Method

### 2.2.1 Principle Component Analysis (PCA)

The SST field from COBE-SST was cropped to the domain of 100°E–60°W, 20°S–60°N. SST anomalies were calculated by subtracting the long-term monthly mean of the analysis period from the monthly SST, at each grid-point. The anomalies were areal-weighted with the cosine of latitude. PCA was performed on the concatenation of five seasonal mean anomalies, centered on the JJA season of the indexed year, i.e. the sample of 1952 would consist of D[-1]JF (December 1951, January 1952, February 1952), MAM (March 1952, April 1952, May 1952), JJA (March 1952, April 1952, May 1952), SON (September 1952, October 1952, November 1952), and DJ[+1]F[+1] (December 1952, January 1953, February 1953) seasonal mean anomalies. The anomalies were not standardised before PCA. PCA with time concatenation has been used in past studies to obtain the time evolution patterns of Pacific SST anomaly modes (Weare and Nasstrom 1982, Guan and Nigam 2008).

After PCA, areal-weighting was removed. Standardised scores (scores) and loadings were calculated from the Principle Components (PC) and Empirical Orthogonal Functions (EOFs), i.e. each PC (EOF) was divided (multiplied) by the square root of its corresponding eigenvalue to give the score (loading), so that the strength of the loading indicates its contribution to the total variance of temperature anomalies. A varimax rotation was performed on the 29 top modes and the rotated modes reordered by their explained variance. Twenty nine modes were used for rotation following Guan and Nigam (2008), who selected this number based on the number of modes being above "noise level".

COBE-SST (Ishii et al., 2005; Japan Meteorological Agency, 2006) was also compared with COBE-SST2, and produced similar climate modes except for the Trend-including mode (not shown). This mode was sensitive to the analysis time period, even for the same COBE-SST2 dataset, as will be described in Section 3.1 (Figs 1 and 2), where PCA was carried out for the 1921–2022 and 1952–2022 periods to examine the robustness of the resultant modes. The PCA results used for comparison with rainfall were based on PCA of the 1952–2010 period, the same period as the rainfall.

### 2.2.2 Rainfall clustering

Upper percentile rainfall in June–July of every year was calculated individual rain-gauges, as well as individual grid-points of d4PDF within a convex hull covering four large Japanese islands of Hokkaido (northern diamond-shaped island), Honshu (S-shaped large island), Shikoku (southern barbell-shaped island) and Kyushu (southwestern oval-shaped island). The last two days of May and the first two days of August were used in the calculation of running-pentads. The percentile calculation was inclusive of times with zero rainfall. For hourly rainfall, this was the 15th highest value of the June–July season. For daily and running-pentad rainfall, these were the sixth highest values. This produced time-series of 59 samples (years). For d4PDF, percentiles were calculated for each individual ensemble member, then the ensemble mean was taken.

The spatial points were clustered using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) function, Scipy's implementation of the hierarchical density-based clustering algorithm developed by Campello et al. (2013). In this algorithm, a hierarchical tree of all the samples is first built by searching around the vicinity of each point at increasing radius based on a distance metric, merging them into larger and larger clusters, until a single cluster that includes all points is formed. The most persistent clusters along the depth of the tree that fulfill the user-provided minimum cluster size ($c_{min}$) are finally selected. We defined the distance metric between two points as $D = 1 - R$ if the Spearman correlation $R$ was statistically significant at $\alpha$=0.05 and the two points were neighboring points, $D$=2.0 otherwise. In the case of rain-gauges, neighboring points were determined through Delaunay triangulation. In the case of d4PDF gridded data, neighboring points were those within a circle of radius 0.15° centered on the grid-point. The choice of 0.15° was based on the estimated size of the smallest rain-gauge cluster produced by HDBSCAN. For each rain-gauge cluster, the upper bound of cluster size was estimated as maximum of the distances between pairs of stations in the cluster. The smallest cluster was the one with the smallest upper bound of cluster size, or approximately 0.15°.

The tree-construction algorithm was set to brute-force ("brute"), i.e. the tree was explicitly constructed without the use of approximations that would speed up calculation. In the case of rain-gauges, neighboring points were determined through Delaunay triangulation, so the algorithm was set to return at least one triangle (three samples, or $c_{min}$=3). Increasing $c_{min}$ resulted in more unclassified rain-gauges when their groupings were not large enough, which due to the sparse distribution of rain-gauges. For this reason, $c_{min}$=3 was used for rain-gauges to include all possible triangles. For d4PDF, the initial setting was $c_{min}$=3 and the algorithm was iterated with larger $c_{min}$, until the threshold was reached when the algorithm returned a singular cluster consisting of almost all grid-points in the domain. The previous $c_{min}$ was then selected. The smallest parameter amongst the three time resolutions was used as the final parameter applied to all three time resolutions. The manner in which clusters grew to fulfill the requirement of increasing $c_{min}$ depended on the minimum sample size, which was set to 1, i.e. even one sample that was significantly correlated with any member of the cluster would be added to the cluster, allowing the cluster to keep expanding until no more such samples could be found.

A time series of rainfall was calculated for each cluster or unclassified point ("set"). At each timestep, the maximum rainfall in each set was selected. For example, in a cluster of three points, this would be the point with the highest rainfall and the other two points would be discarded. This was based on the concept that a set included locations which experienced rainfall from the same event, and the maximum rainfall of that event was sampled by an observer that could "see" the entire location rather than only one specific point. The number of samples every year would also be only along the time dimension and remain the same between different sets for the calculation of percentile rainfall. Upper percentile rainfall for the set was then calculated every year from the set's rainfall time series. Spearman correlation was calculated between the upper percentile rainfall of each set, and the score of each mode from the rotated PCA.

### 2.2.3 Selection of analysis regions

The spatial patterns that result from clustering can be compared between rain-gauges and d4PDF, but since rain-gauges are sparsely distributed point, the rainfall sampled by the points may not be representative of the regional spatial pattern. As an additional check, d4PDF sets corresponding to rain-gauge locations were selected for each rain-gauge. This was defined as sets that contain any grid-points within a circle of radius 0.10° centered on the rain-gauge. A circle of radius 0.05° which approximated the 5 km spatial resolution of d4PDF sometimes resulted in only a single grid-point (a set of one unclassified grid-point) being selected. Such a spatial coverage could be unrepresentative if there were even minor spatial biases in d4PDF, so the radius was doubled.

Since there were over a hundred rain-gauges, we would like to group them into regions that persisted across the three time resolutions. Rain-gauges belonging to the same set of clusters produced by HDBSCAN at three time resolutions were grouped together ("3Class"). To illustrate this, if (A, B, C, D, E, F, G) was an hourly cluster, (A, B, C) and (E, F, G) daily clusters, and (A, B, D) and (C, E, F, G) were pentad clusters, the 3Class regions would be (A, B) and (E, F, G). A group must contain at least two rain-gauges. Each gauge and time-step was treated as a rainfall sample in the year, e.g. a region containing two rain-gauges each with 61 samples of daily rainfall per year would result in a total of 122 daily rainfall samples per year. This was because many rain-gauges in the regions were sparsely located and distant from one another. The time series of annual rainfall extremes at two rain-gauges in a group could be significantly correlated, but typical rainfall levels could differ between the rain-gauges. If the maximum rainfall between rain-gauges was selected at each time-step, one rain-gauge or a subset of rain-gauges in the region could excessively dominate the time samples. In contrast, d4PDF sets consisted of geographically nearby grid-points at high spatial resolution, so the typical rainfall levels were expected to be similar between grid-points in a set. The individual d4PDF sets that corresponded to the rain-gauge locations in a 3Class region were treated as separate rainfall samples. To illustrate this, if a 3Class region contained rain-gauges (A, B), with d4PDF sets (a1, a2, a3) near A's location and (b1, b2) near B's location, then the corresponding d4PDF 3Class region would contain sets (a1, a2, a3, b1, b2) providing 5×61=305 daily samples per year.

**Table 1. The 20 regions in Japan selected for analysis, and the rain-gauges in regions, as shown in Fig. 3d.**

| Region | Rain-gauge locations |
|---|---|
| 1 | Kitami-ekiko, Asahikawa, Rumoi, Wakkanai, Haboro, Obu |
| 2 | Kutchan, Suttsu, Otaru, Iwamizawa, Sapporo |
| 3 | Urakawa, Tomakomai |
| 4 | Mutsu, Hakodate, Esashi |
| 5 | Morioka, Akita, Sakata |
| 6 | Sendai, Ishinomaki, Fukushima |
| 7 | Katsuura, Mito, Choshi, Tateno |
| 8 | Tokyo, Yokohama, Kumagaya |
| 9 | Lake Kawaguchi, Kofu, Chichibu |
| 10 | Fushiki, Toyama, Kanazawa, Takada |
| 11 | Tsuruga, Takayama, Fukui |
| 12 | Kyoto, Nagoya, Gifu, Hikone |
| 13 | Ueno, Osaka, Nara, Himeji, Tsu, Kobe |
| 14 | Wakayama, Tokushima, Sumoto |
| 15 | Kure, Okayama, Hiroshima, Fukuyama |
| 16 | Sakai, Matsue, Yonago |
| 17 | Shimonoseki, Hagi |
| 18 | Iizuka, Sasebo, Saga, Hirado, Hita |
| 19 | Kumamoto, Nagasaki, Mount Unzen |
| 20 | Miyazaki, Makurazaki, Aburatsu, Miyakonojo, Kagoshima |

### 2.2.4 Evaluation of correlation coefficients (CCs)

In this study, Spearman correlation was used, which measured only monotonicity between two variables without need for linearity. The strength of the relationship was defined as "perfect" if $|CC| = 1.0$, "very strong" if $1.0 > |CC| \geq 0.8$, "strong" if $0.8 > |CC| \geq 0.6$, "moderate" if $0.6 > |CC| \geq 0.4$, "weak" if $0.4 > |CC| \geq 0.2$, and "no relation" otherwise. Statistical significance was tested at $\alpha$=0.05, such that the thresholds for statistical significance were approximately at "weak" level ($|CC| > 0.22$) for the 1952–2010 period.

When multiple significance tests are repeatedly performed, a number of false positives may occur. The number of positives must exceed a threshold for there to be statistical significance for the entire field, i.e. field significance. Rain-gauge CCs were tested for field significance following the methodology of Livezey and Chen (1983). The method requires samples to be mutually independent, so sets produced by clustering were used as samples.

### 2.2.5 Monsoon activity indices

The structure of the monsoon front during a heavy rainfall episode has been described by Matsumoto et al, (1971) as being characterised more so by a low-level jet than temperature gradient, with heavy rainfall brought about by mesoscale systems associated with 1000 km wavelength disturbances. We defined indices that described the location and stability of the monsoon front, as well as the location, stability and strength of the low-level jet.

The daily location of the monsoon front was determined based on the methodology of Li et al. (2018). The 850 hPa equivalent potential temperature $\theta_e$ was calculated from daily mean temperature and specific humidity, based on equation 39 of Bolton (1981). At each discrete longitude value in the range 125–130 °E, the latitudinal rate of change of $\theta_e$, $(\partial \theta_e / \partial \varphi)_\lambda$, was then calculated. The monsoon belt was defined as present in the latitudes where $(\partial \theta_e / \partial \varphi)_\lambda \leq -0.04$ K km$^{-1}$, and the location of the front was defined as the mean latitude if more than one latitude point was found. If only one latitude point was found, then the front was defined as not present. After this, the presence of the front was checked at the two points directly west and east of each longitude point. If the front was absent at both sides, then the front was defined as not present at the central point. The zonal mean of the front latitude was defined to be FLat.

The meridional and zonal components of the lower tropospheric water vapour flux (henceforth, just "WVflux") at each grid-point were calculated for each day by vertically integrating the product of daily mean specific humidity $q$ and vector wind $(u,v)$ through six pressure levels (1000, 925, 850, 700, 600, and 500 hPa). In a discrete form, the zonal WVflux is $\sum_i (q_i u_i + q_{i+1} u_{i+1})\ (p_{i-1} - p_i) / 2g$, for the $i$=1...5 pressure levels. The magnitude of the WVflux was calculated from its zonal and meridional components. At each discrete longitude value in the range 125–130 °E, the discrete latitude location with the maximum WVflux was located, then its zonal and meridional components were extracted at that location. The zonal means of the WVflux jet latitude and vector components were defined to be JLat and (QU, QV), respectively.

For each of the 61 days in June–July, the means of Flat, JLat, and (QU,QV) were taken over all the years of data, to obtain a 61-day time series. For d4PDF, this was done individually for each of the 10 ensemble members, to obtain 10×61=610 days. The time series QU and QV were fitted to sine functions $A sin[\pi(t\text{-}T)/L]+B$, with A, B, L and T as fit parameters. The time series JLat and FLat were fitted to logistic functions $A[1 + e^{-k(t-T)}]^{-1}+B$, with A, B, k and T as fit parameters. For d4PDF, the 610 days of all ensemble members were fitted together. The fits were defined to be the climatological seasonal cycles $(QU)_0$, $(QV)_0$, $(JLat)_0$, and $(FLat)_0$. For any day of any year, the deviation from the climatological seasonal cycle was calculated to obtain anomalies $(QU)_a$, $(QV)_a$, $(JLat)_a$, and $(FLat)_a$. For example, $(QU)_a = QU - (QU)_0$. For each year, the seasonal mean anomalies $\mu_{QU}$, $\mu_{QV}$, $\mu_{JLat}$, $\mu_{FLat}$, and seasonal variances $\sigma^2_{Q(U,V)}$, $\sigma^2_{JLat}$, $\sigma^2_{FLat}$ were calculated from the anomalies to obtain seven annual time indices. For d4PDF, the indices were calculated individually for each member of the ensemble, then the ensemble mean was taken. For example, $\sigma^2_{Q(U,V)} = [\ \sum_i (\sigma^2_{QU} + \sigma^2_{QV})_i\ ]/N$ for $i$=1...N ensemble members (N=10).

## 3 Results

### 3.1 Principal Component Analysis (PCA) of Pacific SST anomalies

PCA was carried out on three time periods, 1921–2022 (102 years), 1952–2022 (71 years), and 1952–2010 (59 years). The top five modes appeared distinct from the remaining modes in terms of explained variance. For the 1921–2022 period, the top five modes accounted for 24%, 16%, 8%, 6%, and 5% of the variance. For the 1952–2022 period, they accounted for 28%,16%, 7%, 6%, and 4% of the variance. For the 1952–2010 period, they accounted for 26%, 16%, 9%, 5%, 7% of the variance, with the order of the fourth and fifth modes swapped to match the order from the other two periods. The remaining modes each explained 3% or less of the total variance. The top five modes were similar to those from Guan and Nigam (2008), which had identified the first five modes of their rotated eigenvectors as ENSO+ (growth), ENSO- (decay), Trend, and Pan-Pacific Decadal Variability. and non-canonical ENSO. In this study, these modes will be termed canonical ENSO growth (ENSO+), canonical ENSO decay (ENSO-), Trend-containing (Trend+), Pacific Decadal Variability (PDV), and non-canonical ENSO (ENSO-NC). Figures 1 and 2 show the scores and loadings, respectively.

The scores for ENSO+, ENSO-, PDV, and ENSO-NC were almost identical where they overlapped between the three time periods. This was particularly so for the four ENSO and PDV modes, where the relationships were almost perfect (CC>0.95) between time periods (Figs 1a, 1b, 1c, 1e). The loadings were similar between time periods as well (not shown). Hence, these four modes were robust irrespective of the time period used for PCA. To ascertain the identity of the modes, their scores were correlated against SOI and TPI for the 1952–2010 period. The negative of SOI was used such that positive correlation with -SOI indicated better correspondence to canonical El Niño. The correlation between ENSO+ score and -SOI was strong (CC=0.78; Fig. 1a), and the loading clearly showed a canonical El Niño developing over Eastern Tropical Pacific (Figs 2a-e). The correlation between ENSO- score and -SOI leading by one year was also strong (CC=0.70; Fig. 1b), and the loading clearly showed a decaying canonical El Niño (Figs 2f-j). However, the ENSO-NC had no relationship with -SOI, and only weak anti-correlation with -SOI leading by one year (CC=-0.25; Fig. 1e). The score was strongly positive in 1983 and 1998, years containing the tail end of strong prolonged El Niño events. In addition, the loading showed the Niño-like warm SST anomalies over the Eastern Tropical Pacific from March to August (Fig 2z-D). This indicated the ENSO-NC mode describes the tail end of an atypically strong El Niño that extends into summer months. The correlation between PDV score and TPI was moderate (CC=0.49; Fig. 1d), but this was the strongest CC between TPI and the SST modes (ENSO+ 0.29, ENSO- -0.27, Trend+ -0.27, ENSO-NC -0.18).

The score and loading of the Trend+ mode was sensitive to the analysis period. As shown in Fig. 1c, the scores for the 1952–2010 and 1952–2022 periods were strongly correlated (CC=0.95), but both were different from the score for the 1921–2022 period (CC=0.62 with the score for the 1952–2010 period). Although the scores of Trend+ overall increased in all three analysis periods, different interannual and decadal variability could be seen, which may contribute to the lower CCs. The spatial loading of Trend+ was not homogenous but showed seasonal and spatial variability, with cool anomalies over the tropical Pacific resembling the La Niña and warm anomalies over the northwest Pacific (Figs 2k-o). Although the Niña-like pattern was consistent with some past studies, other studies have also reported a Niño-like pattern, with the direction of the pattern depending on the choice of data used (e.g. Vecchi et al. 2008, Lee et al. 2022). The loading from the 1921–2022 period showed stronger warm anomalies over Northwest Pacific, but weaker “horseshoe-shape” cool anomalies over East Pacific (Figs 2p-t). Hence, the spatial patterns in Fig. 1k-o may be residues from interannual and decadal variability. The actual global warming pattern associated with only the warming trend may be much more homogenous. For this reason, the mode was termed “Trend-containing (other climate variability)” rather than just “Trend”.

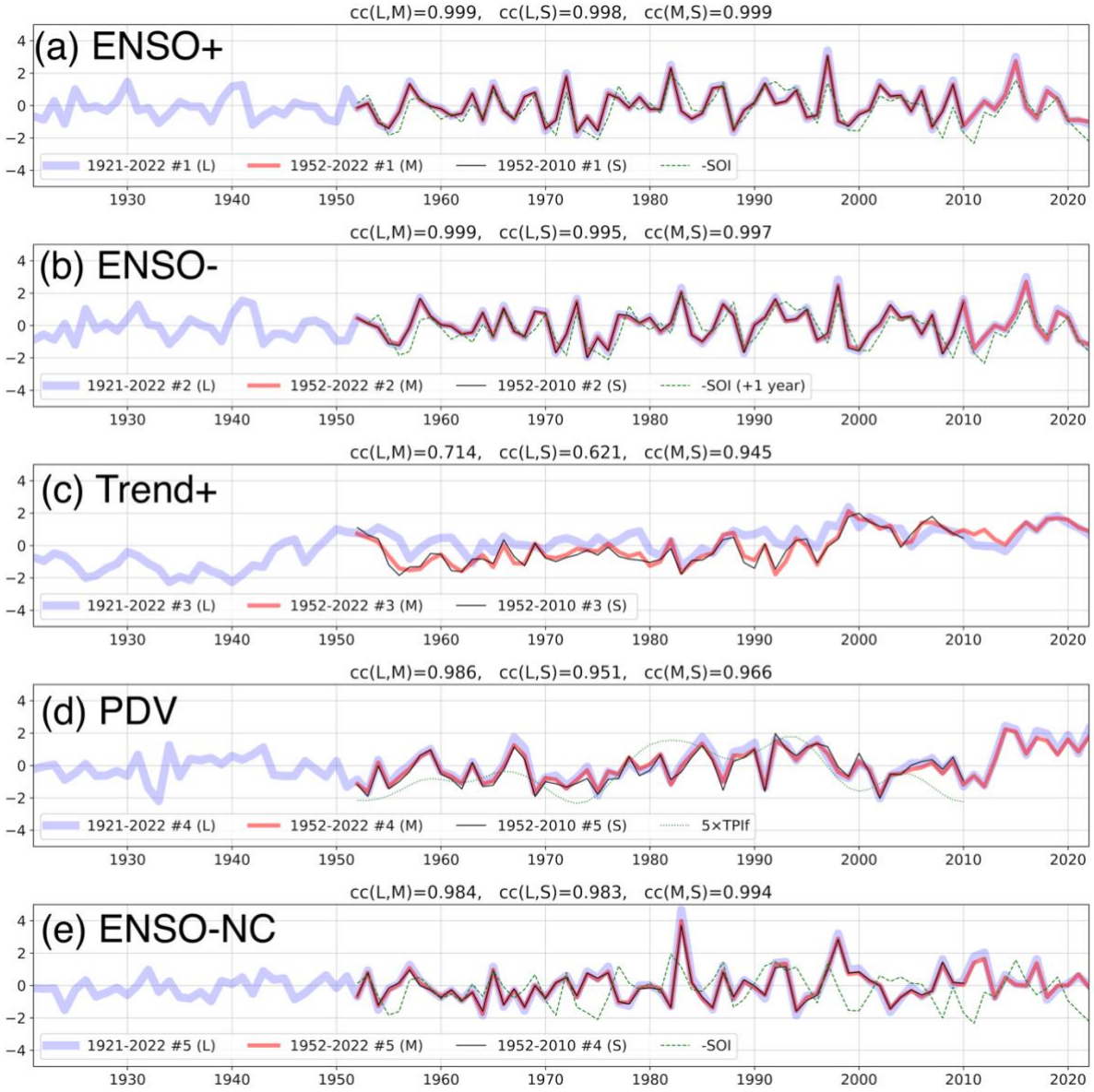



**Figure 1. Scores of the first five modes from the rotated Principal Component Analysis of Pacific Sea Surface Temperature anomalies for the 1921–2022 (blue), 1952–2022 (red), and 1952–2010 (black) periods. (a) Canonical ENSO-growth mode "ENSO+";**
330 **(b) Canonical ENSO-decay mode "ENSO-"; (c) Trend-containing mode "Trend+"; (d) Pacific Decadal Variability mode "PDV"; (e) non-canonical ENSO "ENSO-NC". Green dashed line shows negative of the Southern Oscillation Index "-SOI". Green dotted line shows Interdecadal Pacific Oscillation Tripole Index "TPI" (in °C scaled by a factor of 5). The correlation coefficients between the scores from 1921–2022 "L", 1952–2022 "M", and 1952–2010 "S" for overlapping periods are shown on the top of the panels.**

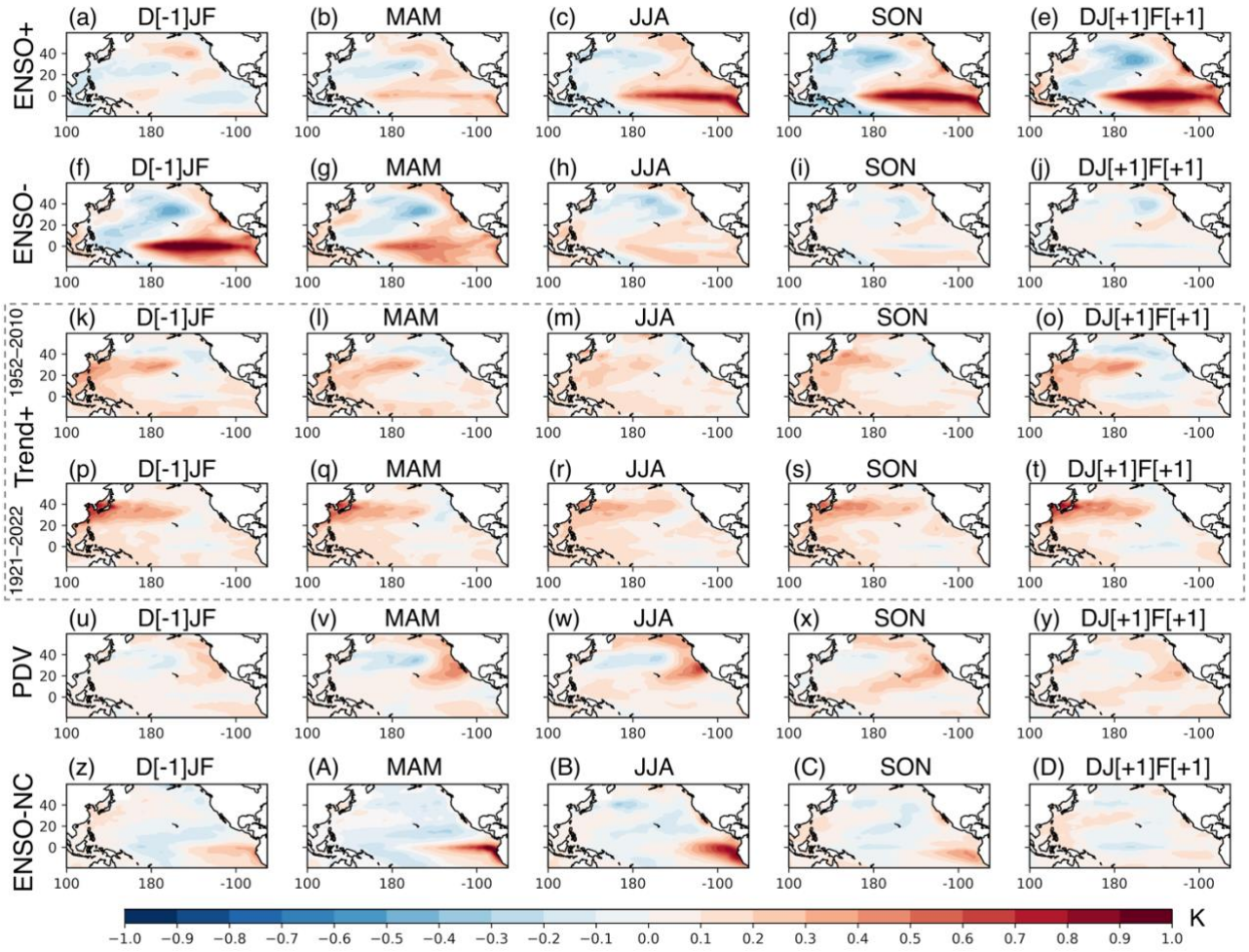



**Figure 2. Loadings of the first five modes from the rotated Principal Component Analysis of Pacific Sea Surface Temperature anomalies for the 1952–2010 period, and only the Trend+ mode for the 1921–2022 period. Rows show the modes. Row 1 (a-e) ENSO+; row 2 (f-j) ENSO-; row 3 (k-o) Trend+; row 4 (p-t) Trend+ for 1921–2022; row 5 (u-y) PDV; row 6 (z-D) ENSO-NC. Columns show anomalies of five seasons centered on June-July-August "JJA".**

### 3.2 Rainfall clustering

The results of clustering on three time resolutions of rain-gauge rainfall are shown in Figs 3a–c. The three results were similar despite some minor differences. There were three clusters in Hokkaido island from north to south (clusters 5–7 in Fig. 3a; clusters 2–4 in Fig. 3b; clusters 1–3 in Fig. 3c). There were two clusters over the Tohoku region (clusters 1 and 4 in Fig. 3a; clusters 1 and 6 in Fig. 3b; clusters 4 and 5 in Fig. 3c). The southern-most Hokkaido cluster sometimes extended into Tohoku. The rain-gauges in near the Kanto region either formed one large cluster or a few smaller clusters (clusters 2 and 11 in Fig. 3a; clusters 5, 9–11 in Fig. 4b; cluster 7 in Fig. 3c). This was also the case for the Kinki region (cluster 8 in Fig. 3a; cluster 12 in Fig. 3b; clusters 12, 14 and 15 in Fig. 3c). There were one or two clusters along the Sea of Japan side in the Chubu region (clusters 3 and 9 in Fig. 3a; cluster 8 in Fig. 3b; clusters 8 and 9 in Fig. 3c). Rain-gauges in the Chugoku region formed one single hourly cluster (cluster 13 in Fig. 3a), but otherwise were arranged into two clusters, one along the Sea of Japan coast and one along the inner sea coast (clusters 16 and 15 in Fig. 3b; clusters 13 and 10 in Fig. 3c). Rain-gauges at the southwestern end of Chugoku and the northern end of Kyushu island formed one cluster (cluster 15 in Fig. 3a; cluster 17 in Fig. 3b; cluster 17 in Fig. 3c). The remaining rain-gauges in Kyushu formed two to three clusters along the west coast (clusters 16, 14 and 10 in Fig. 3a; clusters 18, 14 and 13 in Fig. 3b; clusters 16 and 11 in Fig. 3c). Most rain-gauges along the Pacific coast of Shikoku island and Kinki were unclassified. Figure 3d shows twenty regions were defined based on rain-gauges in the same clusters at three time resolutions.

During the iterative clustering process on d4PDF, the number of clusters decreased as $c_{min}$ was increased, until all points merged into one or a few extremely large clusters. This typically happened suddenly when a $c_{min}$ threshold was crossed, because the number of neighboring points with significant CCs became less than $c_{min}$. At this point, the points would be clustered at $D$=2.0, returning one large cluster which contained all the points. For example, for running-pentad extremes 1004 clusters were returned at $c_{min}$=9, but only four clusters were returned at $c_{min}$=10, one of which contained almost all the grid-points. The same parameter of $c_{min}$=9 was used to cluster hourly and daily rainfall extremes, although these two time resolutions had different and higher $c_{min}$ thresholds. This was done to ensure that the clustering algorithm was standardised across time resolutions and to increase comparability between time resolutions. The caveat is that the hourly and daily rainfall extremes were not as optimally clustered.

Figure 4 shows the results of the above-described $c_{min}$=9 clustering on three time resolutions of d4PDF rainfall. The spatially dense data revealed regional patterns that changed from hourly to running-pentad time resolution. For hourly extremes (Fig. 4a), large clusters with similar (purple) labels were found along Kyushu and the Pacific coast of western Japan. Clusters inland and along the Sea of Japan coast were relatively smaller, and had labels more similar to clusters in Tohoku and Hokkaido. For daily extremes (Fig. 4b), clusters along the Pacific coast clusters became smaller, and the coast-aligned linear pattern started to break apart. In particular, cluster labels over Kyushu were split into western (purple) and eastern (red–orange) groups. A large cluster formed over the northeastern coast of Hokkaido. For running-pentad extremes (Fig. 4c), the separation between western (purple) and eastern (orange–yellow) groups over Kyushu became more apparent, with the western group extending along the Sea of Japan coast into Chugoku. These clusters had closer labels with regions along the Sea of Japan coast. The clusters region near Tokyo in Kanto formed a similarly labelled group (red-purple). Thus, a new coast-aligned linear patterned formed.

Rain-gauge and d4PDF reflected similar pattern shifts. Firstly, Rain-gauge hourly clusters along the Pacific coast had similar labels (clusters 10–12 in Fig. 3a). Secondly, two larger clusters formed over Hokkaido for running-pentad extremes (clusters 1 and 2 in Fig. 3c). Thirdly, one large cluster formed over Kanto for running-pentad extremes (cluster 7 in Fig. 3c). We could not ascertain whether an east-west split developed over Kyushu, since there were few rain-gauges in the eastern side.

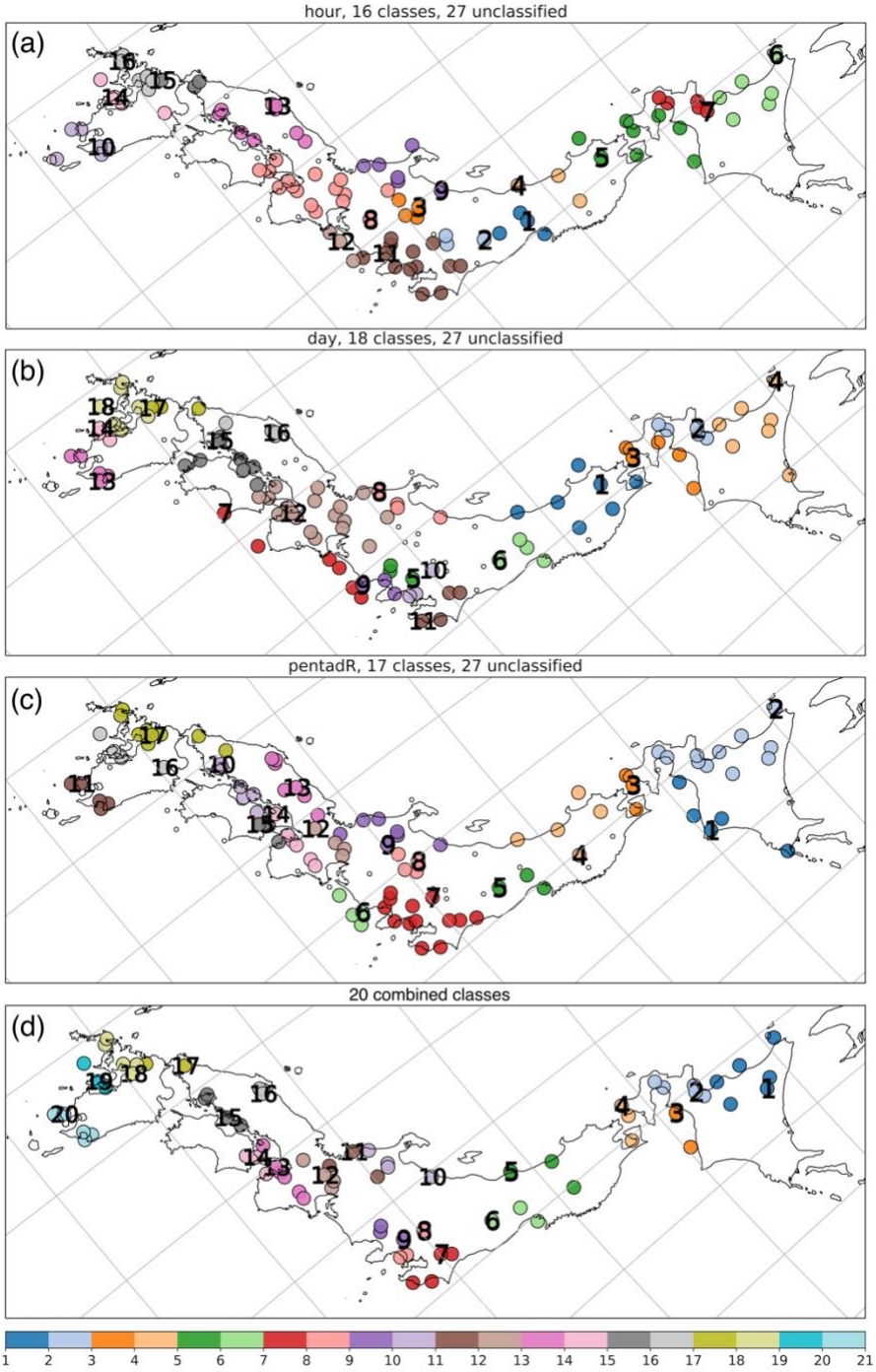



**Figure 3. Rain-gauge clusters. Labels of clusters based on the time series of seasonal (a) 99th percentile hourly, (a) 90th percentile daily, and (a) 90th percentile running-pentad rainfall. Colours only reflect the cluster numbering and do not reflect any units or magnitudes. Small white markers show unclassified rain-gauges. (d) The 3Class analysis regions described in Section 2.2.3. Longitude-latitude gridlines are shown as grey lines; the map is a rotated projection.**

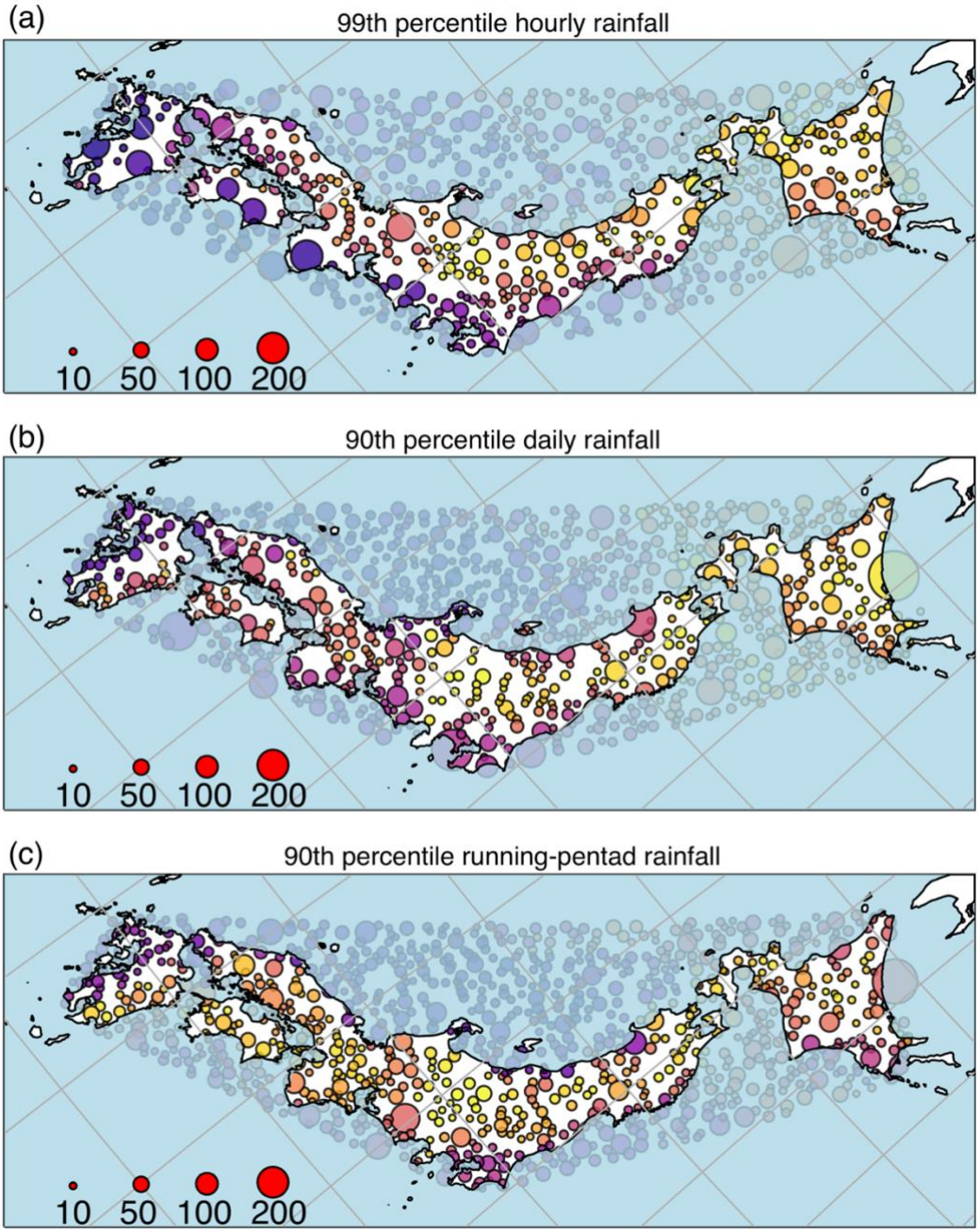



**Figure 4. d4PDF clusters. Labels and sizes of clusters based on the time series of seasonal (a) 99th percentile hourly, (a) 90th percentile daily, and (a) 90th percentile running-pentad rainfall. Colours only reflect the cluster numbering and do not reflect any units or magnitudes. Circle marker sizes reflect cluster sizes (legend). A translucent blue mask was used to cover data within the convex hull used for clustering, but falling inside the ocean. Longitude-latitude gridlines are shown as grey lines; the map is a rotated projection.**

### 3.3 Correlation Coefficients (CCs)

Figures 5–7 show the CCs between rainfall extremes and scores of the Pacific SST modes, at three time resolutions. The left columns show CCs for individual rain-gauges, which were at most weak. The only exception is moderate correlation between hourly extremes and Trend+ at one rain-gauge at the southern tip of Kyushu (Fig. 5c). With regards to the first two ENSO modes, rain-gauges showed no domain-wide field significance at any of the three time resolutions, but smaller regions contained rain-gauges with significant CCs.

With ENSO+, d4PDF showed similar correlation signs to rain-gauges over northeastern, eastern and southwestern Japan, but with spatial biases. Anti-correlated rain-gauges were found at western Hokkaido for hourly and running-pentad extremes (Figs 5a and 7a; Table 2 regions 1–2), but at eastern Hokkaido for daily extremes (Fig. 6a). Anti-correlated rain-gauges were scattered over central Japan, Kanto and southern Kyushu (Figs 5a, 6a, 7a; Table 2 region 20). Correlated rain-gauges were found along the Sea of Japan coast (Table 2 region 5), and at northwestern Kyushu (Table 2 regions 17–19). Similar to the rain-gauges, d4PDF also showed anti-correlation over Hokkaido for all three time resolutions, although not at the same locations as rain-gauges (Figs 5f, 6f, 7f; Table 2 region 1). Similar to the rain-gauges, there were also some anti-correlated sets near Kanto, particularly for hourly extremes, although not at the same locations as rain-gauges. Similar to the rain-gauges, there were correlated sets at Kyushu, although on the eastern side. Unlike the rain-gauges, d4PDF showed widespread correlation over western Japan (Table 2 regions 12–16).

With ENSO-, d4PDF was excessively correlated across most of Japan, except at the eastern side of topography with some anti-correlated sets. Anti-correlated rain-gauges were found at northwestern Kyushu, for all three time resolutions (Figs 5b, 6b, 7b; Table 2 region 19). Anti-correlated rain-gauges were also scattered along the Pacific coast and near the Japanese Alps central of the domain, particularly for daily extremes (Fig. 6b). Similar to the rain-gauges, d4PDF also showed anti-correlation along the Pacific coast for daily and pentad extremes, although only on the east side of topography (Figs 6g, 7g). Similar to the rain-gauges, there were some anti-correlation sets near the Japanese Alps for all three time resolutions. Unlike the rain-gauges, d4PDF showed widespread correlation across Japan (Table 2 regions 1–5, 7–13, 15–20). Reversed the rain-gauges, d4PDF showed correlation at northwestern Kyushu for daily and running-pentad extremes (Figs 6g, 7g).

Considering clustered results for hourly extremes, there was field significance over the study domain for correlation with Trend+ and ENSO-NC, i.e. at least 5 sets with significant CCs. For daily extremes, there was field significance for anti-correlation with PDV. For running-pentad extremes, no mode showed field significance, although there were a number of individual rain-gauges correlated with Trend+ and anti-correlated with PDV. As discussed in Section 2.2.4, the field significance test had the disadvantage of treating large clusters and single rain-gauges equally, so field significance tests were also carried out individual rain-gauges under the assumption that they were mutually independent. Since this is a false assumption, such a test would be too permissive, but the results may be useful as a comparison. For hourly extremes, there was field significance over the domain for correlation with Trend+, i.e. at least 11 rain-gauges with significant CCs. For daily extremes, there was field significance for anti-correlation with PDV. For running-pentad extremes, no mode showed field significance. Hence, field significance results were consistent between the two methods, except for ENSO-NC.

The right columns of Figs 5–7 show CCs for sets in d4PDF. CCs were weak over most of Japan, but could be moderate in some locations, particularly in Kanto with ENSO-NC (Figs 5j and 7j). For the higher modes of Trend+, PDV, and ENSO-NC, the patterns of CCs in d4PDF were similar between time resolutions, but did not seem to resemble rain-gauge patterns. With Trend+, rain-gauges showed correlation outside of Hokkaido (Fig. 5c; Table 2 regions 8–9 Kanto, 19–20 southwestern Kyushu), but d4PDF showed correlation in Hokkaido (Fig. 5h; Table 2 regions 1–4). With PDV, rain-gauges showed anti-correlation across Japan (Fig. 6d; Table 2 regions 1, 2, 11, 14, 19), but d4PDF showed mostly correlation near the Pacific coast (Fig. 6i; Table 2 regions 3, 8, 12, 20). There was only one anti-correlated region at the Sea of Japan coast (Table 2 region 16). With ENSO-NC, rain-gauges showed correlation across Japan (Fig. 5e; Table 2 regions 3, 4, 6 ,8, 19), but d4PDF showed correlation only at northwestern Kyushu (Table 2 region 17–19). Northeastern and eastern Japan were anti-correlated (Table 2 regions 2, 4, 6–11), particularly over Kanto (Fig. 5j; Table 2 region 8).

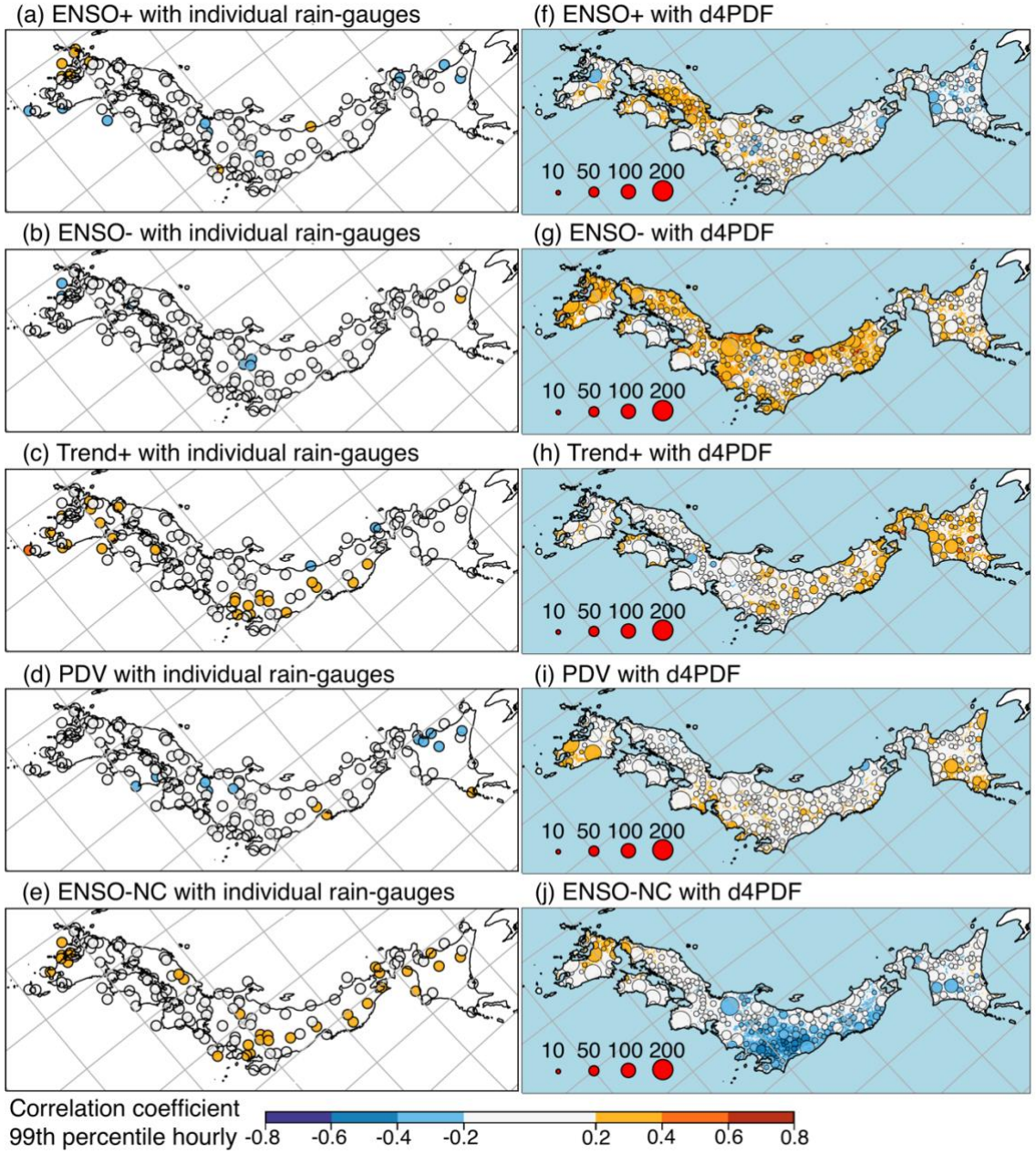



**Figure 5. Spearman CCs for the 1952–2010 period between the scores of five Pacific SST modes (rows) and seasonal 99th upper percentile hourly rainfall from rain-gauges (left column) or d4PDF (right column). June–July rainfall extremes of individual rain-gauges with (a) ENSO+, (b) ENSO-, (c) Trend+, (d) PDV, (e) ENSO-NC. June–July rainfall extremes of d4PDF sets with (f) ENSO+, (g) ENSO-, (h) Trend+, (i) PDV, (j) ENSO-NC. In the right column, circle marker sizes reflect cluster sizes, while single points are shown as coloured dots. A blue mask was used to cover d4PDF data inside the ocean. Longitude-latitude gridlines are shown as grey lines; the map is a rotated projection.**

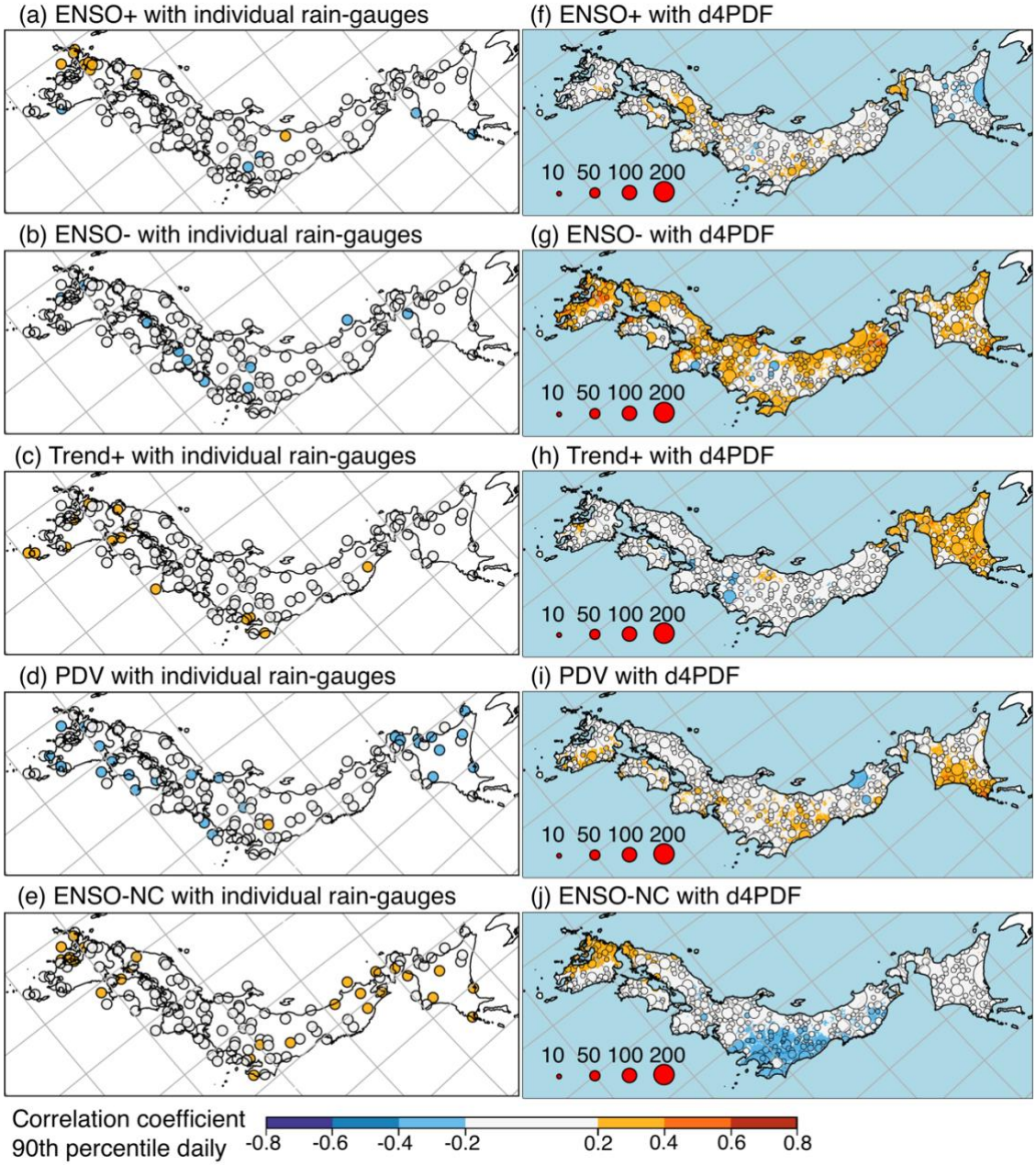
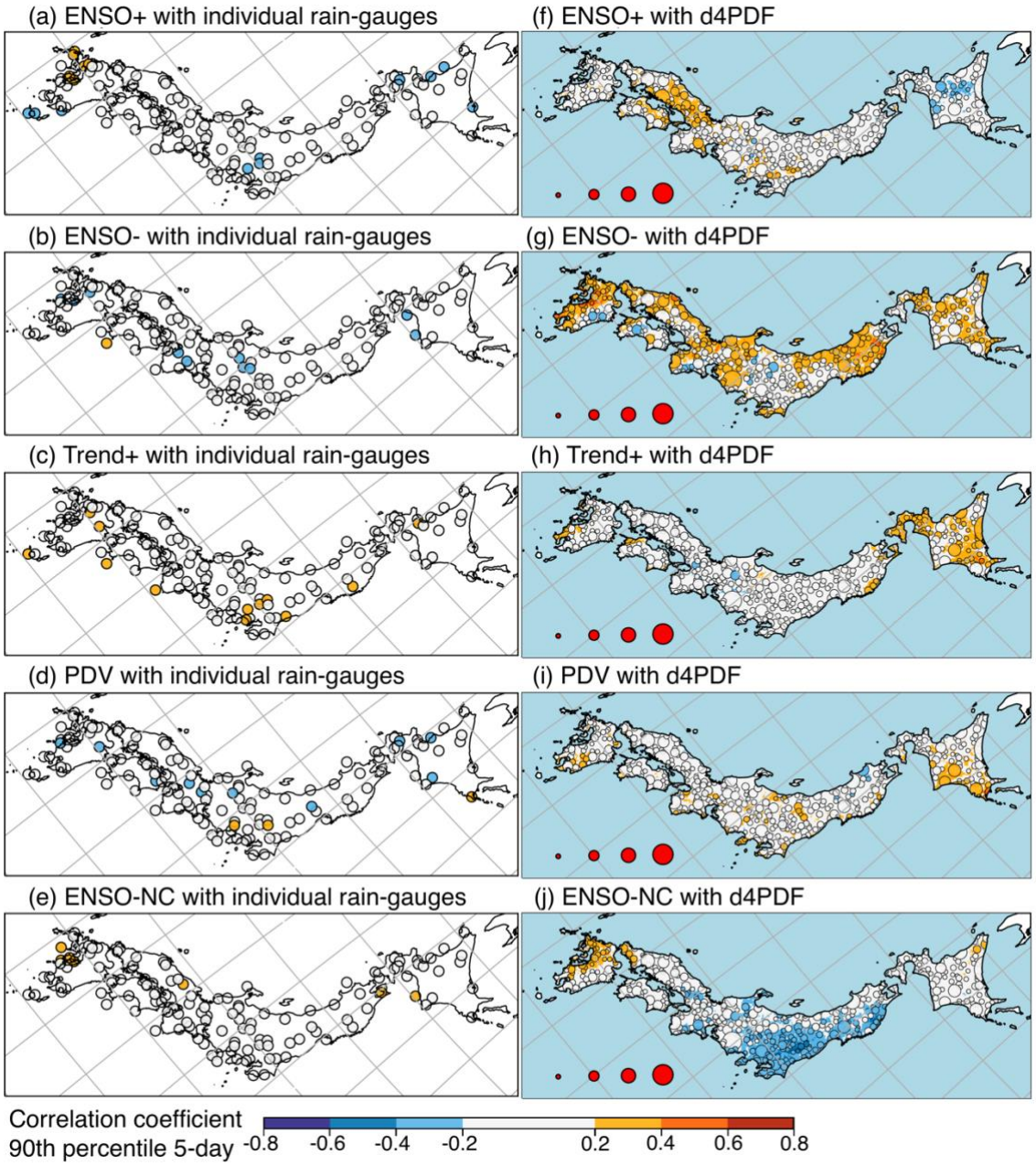



**Figure 6. Like as Fig. 5, except for seasonal 90th upper percentile daily rainfall.**




460 **Figure 7. Like Fig. 5, except for seasonal 90th upper percentile running-pentad rainfall.**

**Table 2. Spearman CCs between June–July rainfall extremes in the 20 analysis regions (rows) and the scores of the first five Pacific SST modes (columns). Under the "time" column, "H" indicates 99th upper percentile hourly rainfall, "D" indicates 90th upper percentile daily rainfall, and "P" indicates 90th upper percentile running-pentad rainfall. The CC strengths are shaded using the same colors as Figs 5–7.**

| | | ENSO+ | | ENSO- | | Trend+ | | PDV | | ENSO-NC | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| region | time | gauge | d4PDF | gauge | d4PDF | gauge | d4PDF | gauge | d4PDF | gauge | d4PDF |
| 1 | H | -0.16 | -0.21 | 0.04 | 0.10 | 0.01 | 0.15 | -0.34 | 0.06 | 0.14 | -0.13 |
| | D | -0.12 | -0.09 | -0.04 | 0.25 | -0.05 | 0.27 | -0.26 | 0.07 | 0.17 | 0.08 |
| | P | -0.24 | -0.16 | -0.04 | 0.26 | 0.07 | 0.22 | -0.25 | 0.01 | 0.15 | 0.04 |
| 2 | H | -0.31 | 0.15 | -0.04 | 0.19 | 0.03 | 0.28 | -0.25 | 0.04 | 0.06 | -0.22 |
| | D | -0.16 | 0.04 | -0.02 | 0.28 | 0.02 | 0.32 | -0.33 | 0.17 | 0.05 | 0.00 |
| | P | -0.25 | -0.04 | -0.06 | 0.20 | 0.13 | 0.35 | -0.18 | 0.03 | 0.01 | -0.04 |
| 3 | H | -0.17 | -0.02 | -0.20 | 0.20 | -0.01 | 0.13 | -0.13 | 0.09 | 0.16 | -0.13 |
| | D | -0.09 | -0.07 | -0.18 | 0.06 | -0.11 | 0.21 | -0.11 | 0.29 | 0.21 | 0.01 |
| | P | -0.06 | -0.05 | -0.36 | 0.12 | -0.04 | 0.15 | -0.09 | 0.16 | 0.26 | 0.00 |
| 4 | H | -0.04 | 0.01 | -0.11 | 0.21 | -0.11 | 0.24 | -0.15 | 0.04 | 0.25 | -0.20 |
| | D | 0.00 | 0.14 | -0.16 | 0.19 | 0.01 | 0.20 | -0.14 | 0.00 | 0.32 | 0.08 |
| | P | -0.13 | 0.09 | -0.06 | 0.09 | 0.00 | 0.24 | -0.10 | 0.01 | 0.23 | 0.03 |
| 5 | H | 0.22 | 0.01 | -0.01 | 0.30 | -0.15 | 0.13 | -0.01 | -0.03 | 0.13 | -0.13 |
| | D | 0.16 | 0.11 | 0.02 | 0.38 | -0.04 | 0.13 | -0.09 | 0.01 | 0.15 | -0.04 |
| | P | 0.12 | 0.09 | 0.02 | 0.36 | -0.07 | 0.09 | -0.15 | -0.10 | 0.16 | -0.05 |
| 6 | H | 0.06 | 0.06 | -0.02 | 0.10 | 0.15 | -0.01 | 0.25 | 0.12 | 0.22 | -0.23 |
| | D | 0.07 | 0.05 | 0.11 | 0.18 | 0.14 | -0.07 | 0.02 | 0.16 | 0.01 | -0.21 |
| | P | 0.00 | 0.01 | 0.11 | 0.14 | 0.06 | 0.01 | 0.13 | 0.11 | 0.13 | -0.34 |
| 7 | H | 0.06 | 0.07 | 0.03 | 0.25 | 0.14 | 0.04 | 0.07 | 0.13 | 0.19 | -0.27 |
| | D | 0.11 | 0.19 | 0.01 | 0.24 | 0.10 | 0.05 | -0.09 | 0.13 | 0.19 | -0.36 |
| | P | -0.04 | 0.14 | 0.05 | 0.20 | 0.19 | 0.03 | 0.08 | 0.16 | 0.17 | -0.31 |
| 8 | H | -0.08 | 0.02 | -0.07 | 0.21 | 0.15 | 0.07 | 0.16 | 0.24 | 0.27 | -0.40 |
| | D | -0.09 | 0.14 | 0.01 | 0.04 | 0.27 | 0.02 | 0.10 | 0.22 | 0.16 | -0.33 |
| | P | -0.10 | 0.13 | -0.14 | 0.00 | 0.23 | 0.08 | 0.14 | 0.16 | 0.15 | -0.35 |
| 9 | H | -0.03 | 0.12 | -0.06 | 0.20 | 0.20 | 0.15 | 0.09 | 0.10 | 0.19 | -0.32 |
| | D | -0.19 | 0.15 | -0.03 | 0.07 | 0.07 | 0.02 | 0.06 | 0.13 | 0.04 | -0.35 |
| | P | -0.17 | 0.17 | -0.07 | -0.02 | 0.18 | 0.05 | 0.15 | 0.08 | 0.08 | -0.33 |
| 10 | H | 0.03 | -0.03 | -0.11 | 0.29 | -0.09 | 0.15 | -0.05 | 0.12 | 0.15 | -0.21 |
| | D | 0.04 | 0.08 | -0.16 | 0.24 | -0.04 | 0.12 | -0.08 | 0.12 | 0.12 | -0.15 |
| | P | 0.03 | 0.08 | -0.16 | 0.24 | -0.08 | 0.07 | -0.04 | 0.11 | 0.15 | -0.27 |
| 11 | H | -0.03 | 0.08 | -0.09 | 0.33 | 0.01 | -0.03 | -0.18 | 0.05 | 0.06 | -0.21 |
| | D | -0.10 | 0.08 | 0.02 | 0.34 | -0.15 | -0.02 | -0.26 | 0.02 | 0.05 | -0.06 |
| | P | -0.03 | 0.17 | -0.03 | 0.29 | -0.09 | -0.08 | -0.20 | 0.05 | 0.06 | -0.24 |
| 12 | H | -0.05 | 0.28 | -0.12 | 0.16 | -0.06 | -0.04 | -0.14 | 0.22 | 0.14 | -0.04 |
| | D | -0.03 | 0.15 | -0.04 | 0.21 | 0.00 | -0.07 | -0.12 | 0.18 | 0.09 | -0.07 |
| | P | -0.09 | 0.21 | -0.05 | 0.17 | 0.04 | -0.08 | -0.16 | 0.13 | 0.08 | -0.08 |
| 13 | H | -0.07 | 0.17 | -0.05 | 0.23 | 0.00 | 0.00 | 0.00 | 0.10 | 0.01 | -0.09 |
| | D | -0.14 | 0.16 | -0.19 | 0.23 | -0.01 | -0.09 | -0.11 | 0.18 | 0.07 | -0.04 |
| | P | -0.08 | 0.23 | -0.11 | 0.14 | -0.01 | -0.04 | -0.03 | 0.17 | -0.01 | -0.07 |
| 14 | H | -0.12 | 0.23 | -0.11 | 0.13 | 0.11 | 0.12 | -0.18 | -0.01 | 0.00 | 0.02 |
| | D | -0.18 | 0.14 | -0.04 | 0.18 | 0.12 | -0.01 | -0.26 | 0.18 | 0.05 | -0.01 |
| | P | -0.03 | 0.27 | 0.03 | 0.10 | 0.01 | 0.05 | -0.15 | 0.03 | -0.01 | -0.02 |
| 15 | H | 0.13 | 0.21 | -0.14 | 0.22 | 0.00 | 0.19 | -0.14 | 0.01 | 0.06 | 0.12 |
| | D | 0.11 | 0.17 | -0.08 | 0.21 | 0.13 | -0.02 | -0.12 | 0.10 | 0.11 | 0.19 |
| | P | 0.10 | 0.22 | -0.06 | 0.29 | 0.00 | 0.09 | -0.08 | 0.08 | 0.02 | 0.06 |
| 16 | H | 0.11 | 0.11 | -0.09 | 0.44 | -0.01 | 0.10 | -0.05 | -0.20 | 0.08 | 0.11 |
| | D | 0.13 | 0.03 | 0.01 | 0.32 | -0.02 | 0.00 | -0.06 | 0.01 | 0.11 | 0.06 |
| | P | 0.09 | 0.23 | 0.02 | 0.32 | 0.04 | 0.01 | -0.06 | -0.09 | 0.06 | 0.00 |
| 17 | H | 0.14 | 0.10 | -0.09 | 0.34 | 0.12 | 0.10 | -0.10 | 0.02 | 0.04 | 0.24 |
| | D | 0.20 | -0.07 | -0.11 | 0.37 | 0.05 | 0.03 | -0.04 | 0.04 | 0.13 | 0.31 |
| | P | 0.10 | -0.03 | -0.08 | 0.34 | 0.07 | 0.07 | -0.02 | 0.06 | 0.05 | 0.26 |
| 18 | H | 0.20 | 0.08 | -0.13 | 0.34 | 0.17 | 0.14 | -0.10 | 0.15 | 0.15 | 0.20 |
| | D | 0.26 | -0.05 | -0.12 | 0.34 | 0.15 | 0.08 | -0.14 | 0.04 | 0.17 | 0.28 |
| | P | 0.23 | -0.03 | -0.05 | 0.35 | 0.16 | 0.13 | -0.04 | 0.05 | 0.11 | 0.26 |
| 19 | H | 0.25 | 0.08 | -0.22 | 0.41 | 0.07 | 0.17 | -0.13 | 0.04 | 0.28 | 0.20 |
| | D | 0.17 | 0.04 | -0.18 | 0.42 | 0.22 | 0.16 | -0.23 | 0.03 | 0.33 | 0.27 |
| | P | 0.22 | 0.05 | -0.17 | 0.37 | 0.02 | 0.09 | -0.14 | 0.11 | 0.21 | 0.26 |
| 20 | H | -0.19 | 0.01 | -0.11 | 0.38 | 0.26 | 0.17 | -0.10 | 0.21 | 0.18 | 0.10 |
| | D | -0.13 | -0.02 | 0.03 | 0.38 | 0.17 | 0.16 | -0.19 | 0.22 | 0.06 | 0.03 |
| | P | -0.19 | 0.05 | -0.06 | 0.37 | 0.19 | 0.17 | -0.14 | 0.23 | 0.07 | 0.13 |

## 4 Discussion

The lack of strong CCs between rainfall extremes and the Pacific SST modes indicates that the latter are not major controlling factors for extreme rainfall over Japan. These results are sensible in that extreme rainfall typically occurs in mesoscale systems which may be associated with certain synoptic scale conditions, but are much less likely to be directly influenced by hemispheric or planetary scale climate variability. In this section, rainfall extremes were correlated against indices reflecting the regional state of the monsoon near Japan, to check if there were stronger relationships with regional conditions. Rain-gauge rainfall extremes were correlated with monsoon indices calculated from JRA55, and these relationships were considered as "observations". d4PDF rainfall extremes were correlated with monsoon indices calculated from d4PDF-AGCM. The results were similar between the three time resolutions, so only the results for daily extremes are described below (Figs 8, 9).

Based on observations, CCs with the seasonal anomaly of the monsoon front latitude ($\mu_{FLat}$; Fig. 8a), seasonal variance of the monsoon front latitude ($\sigma^2_{FLat}$; Fig 8b), seasonal anomaly of the monsoon jet latitude ($\mu_{JLat}$; Fig. 8c), and seasonal anomaly of the monsoon jet meridional strength ($\mu_{QV}$; Fig. 9b) were no stronger than CCs with the SST modes. The spatial patterns of correlation were similar between d4PDF and rain-gauges for the $\mu_{FLat}$, $\mu_{JLat}$, and $\mu_{QV}$, showing correlation along the Sea of Japan coast and anti-correlation along the Pacific coast (Figs 8e, 8g, 9e). However, d4PDF CCs were much stronger than observations, showing up to strong CCs in some locations.

Compared to SST modes, rainfall extremes were observed to have better relationships with the seasonal variance of monsoon jet latitude ($\sigma^2_{JLat}$;Fig. 8d) and the seasonal variance of monsoon jet strength ($\sigma^2_{Q(U,V)}$; Fig. 9c), both of which were indicators of monsoon jet instability. Spatially widespread weak anti-correlation with $\sigma^2_{JLat}$ was observed along the Pacific coast, for more rain-gauges than with any SST mode. With $\sigma^2_{Q(U,V)}$, two moderately anti-correlated rain-gauges were seen at southern Kyushu. This meant that larger variance or greater instability of the monsoon jet did not correlate with stronger rainfall extremes in the observations. In contrast to observations, d4PDF showed up to strong anti-correlation with $\sigma^2_{JLat}$ along the opposite Sea of Japan coast, but correlation over the Pacific coast east of topography and at eastern Japan (Fig. 8h). d4PDF also showed mainly correlation with $\sigma^2_{Q(U,V)}$, particularly at eastern Kyushu (Fig. 9f).

The strongest observed relationships were with the seasonal anomaly of monsoon jet zonal strength ($\mu_{QU}$; Fig. 9a). Widespread correlation of up to moderate strength was observed along the Pacific coast, particularly at southwestern Kyushu. In contrast to observations, d4PDF showed up to strong correlation along the opposite Sea of Japan coast, but anti-correlation along the Pacific coast east of topography.

The spatially widespread and up to strong CCs seen in d4PDF indicated that model extreme rainfall had much better relationships with regional monsoon indices than with Pacific SST modes. This was consistent with observations which also showed stronger relationships between extreme rainfall and the indices, and supported the hypothesis that Pacific SST modes act indirectly through the regional monsoon. Next, correlations between the monsoon indices and scores of the Pacific SST modes were calculated. CCs from JRA55 and d4PDF were at most weak, with only one exception from d4PDF.

The exception of moderate CC from d4PDF was 0.46 between ENSO- and $\mu_{QU}$, in contrast to the JRA55 significant value of 0.30. For d4PDF, ENSO- also had significant CCs with $\sigma^2_{JLat}$ and $\mu_{QV}$ at -0.32 and 0.32, respectively. In contrast, observations showed no relationships with $\sigma^2_{JLat}$ and $\mu_{QV}$. The d4PDF spatial patterns of $\sigma^2_{JLat}$ and $\mu_{QV}$ were respectively reversed to (Fig. 8h) and similar to (Fig. 9e) the pattern of $\mu_{QU}$ (Fig. 9d). The stacking of these three patterns could account for the d4PDF spatial patterns between ENSO- and rainfall extremes (Fig. 6g).

From JRA55, PDV was correlated with $\sigma^2_{JLat}$ at CC=0.20, and significantly anti-correlated with $\mu_{QV}$ at CC=-0.24. The observed spatial patterns of $\sigma^2_{JLat}$ and $\mu_{QV}$ were respectively anti-correlated (Fig. 8d) and correlated (Fig. 9b) with rainfall extremes along the Pacific coast. The stacking of two reversed patterns could account for observed spatial pattern between PDV and rainfall extremes (Fig. 6d).

From d4PDF, ENSO-NC had significant CCs with $\sigma^2_{JLat}$ and $\mu_{QU}$ at -0.36 and 0.31, respectively. In contrast, JRA55 showed no relationship with $\sigma^2_{JLat}$ and $\mu_{QU}$. The d4PDF spatial pattern of $\sigma^2_{JLat}$ was anti-correlation along Sea of Japan coast and correlation at Kanto (Fig. 8h), while the spatial pattern of $\mu_{QU}$ was reversed (Fig. 9e). The stacking of two reversed patterns could account for the d4PDF spatial pattern between ENSO-NC and rainfall extremes (Fig. 6e).

For Trend+, JRA55 and d4PDF identically had CC=0.21 with $\mu_{QV}$. However, neither of the spatial patterns (Figs 9b, 9e) could not account for difference between observed and d4PDF, in terms of pattern between Trend+ and rainfall extremes.

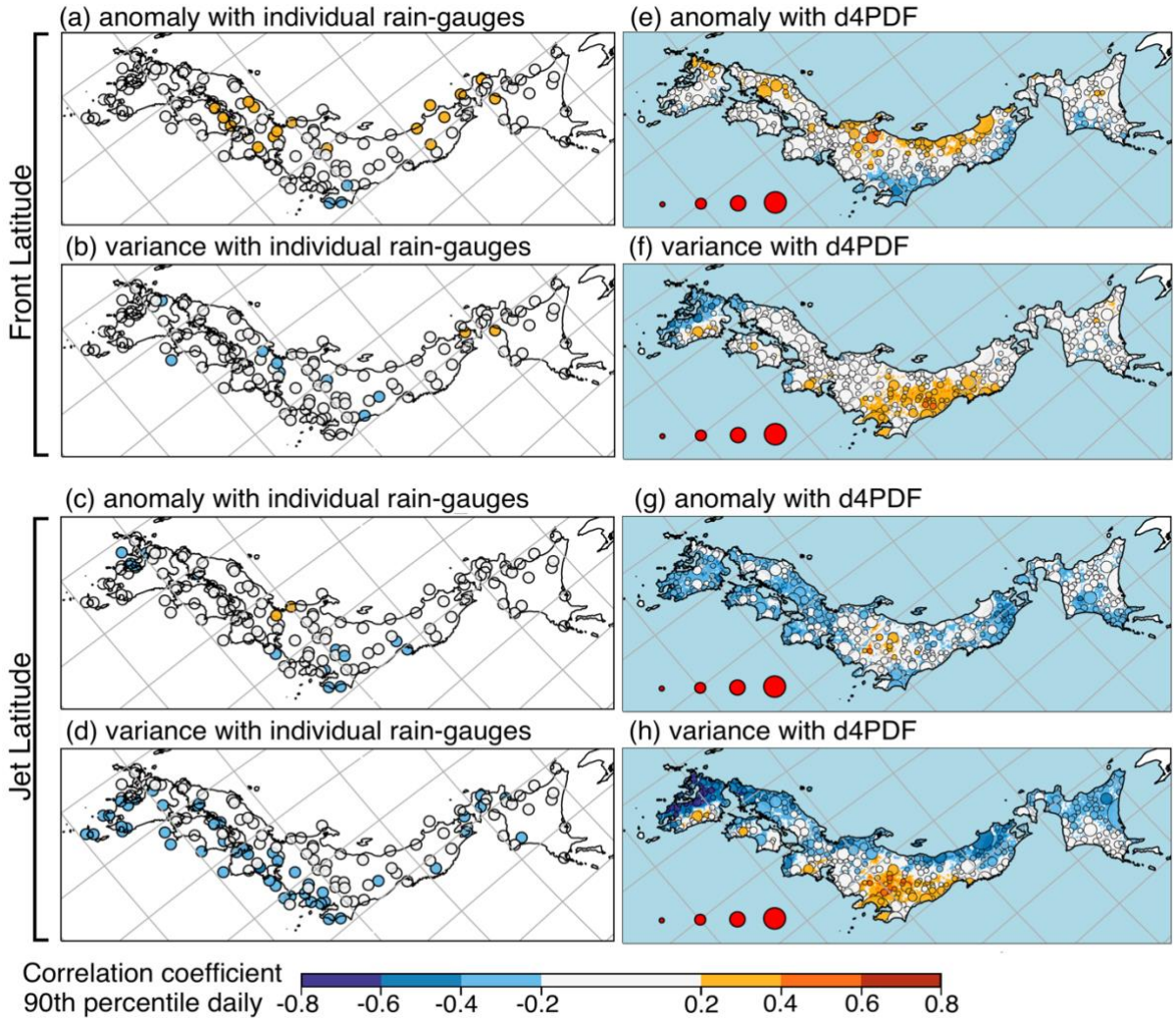



**Figure 8. Spearman CCs for the 1958–2010 period between monsoon indices related to position (rows) and June–July 90th upper percentile daily rainfall from rain-gauges (left column) or d4PDF (right column). Rainfall extremes of individual rain-gauges with JRA55 indices; (a) seasonal anomaly of monsoon front latitude $\mu_{FLat}$, (b) seasonal variance of monsoon front latitude $\sigma^2_{FLat}$, (c) seasonal anomaly of monsoon jet latitude $\mu_{JLat}$, (d) seasonal variance of monsoon jet latitude $\sigma^2_{JLat}$. Rainfall extremes of d4PDF sets with d4PDF indices; (e) $\mu_{FLat}$, (f) $\sigma^2_{FLat}$, (g) $\mu_{JLat}$, (h) $\sigma^2_{JLat}$. In the right column, circle marker sizes reflect cluster sizes, while single points are shown as coloured dots. A blue mask was used to cover d4PDF data inside the ocean. Longitude-latitude gridlines are shown as grey lines; the map is a rotated projection.**

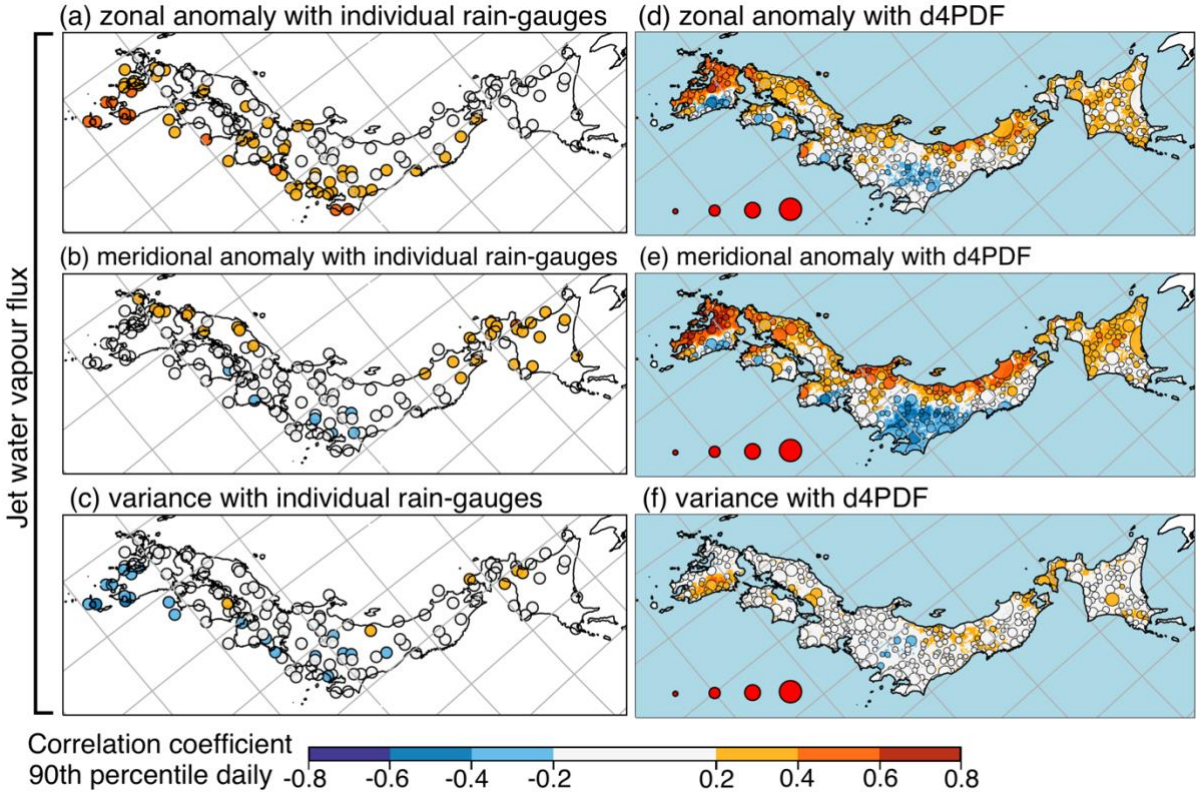



**Figure 9. Like Fig. 8, but with monsoon indices related to the strength of the monsoon jet. Rainfall extremes of individual rain-gauges with JRA55 indices; (a) seasonal anomaly of the monsoon jet zonal water vapour flux $\mu_{QU}$, (b) seasonal anomaly of the monsoon jet meridional water vapour flux $\mu_{QV}$, (c) seasonal variance of the monsoon jet water vapour flux $\sigma^2_{Q(U,V)}$. Rainfall extremes of d4PDF sets with d4PDF indices; (d) $\mu_{QU}$, (e) $\mu_{QV}$, (f) $\sigma^2_{Q(U,V)}$.**

## 5 Conclusion

This study investigated the relationships between the extreme rainfall over Japan and the hemispheric-scale climate variability of SST anomalies over the Pacific basin. The June–July 99th upper percentile hourly rainfall, June–July 90th upper percentile daily rainfall, and June–July 90th upper percentile running-pentad rainfall in the 59-year period of 1952–2010) was compared between d4PDF and rain-gauges (observations). d4PDF and observations were clustered using the HDBSCAN algorithm based on the Spearman correlation of rainfall extremes. Clusters were similar between d4PDF and observations; aligned along coasts and long axis of the islands, and migrating from Pacific coast to Sea of Japan coast from hourly to running-pentad timescales. Using each d4PDF "set" (cluster or unclassified point) as a spatial sample, Spearman correlation was calculated between the rainfall extremes and the scores of five Pacific SST modes. The modes were obtained by the varimax rotation of extended PCA on Pacific SST anomalies from COBE-SST2, and reflected canonical ENSO growth, canonical ENSO decay, warming trend, PDV, and a non-canonical ENSO that persisted into summer.

Observed correlation coefficients (CCs) were at most weak ($|CC| < 0.4$) over most of Japan. d4PDF CCs were similarly weak, although there were a few regions with moderate CCs ($0.4 \leq |CC| < 0.6$). These regions were correlation with ENSO- around Matsue (Table 2 region 16) and Nagasaki (Table region 19), and anti-correlation with ENSO-NC around Tokyo (Table 2 region 8). For the first two ENSO modes, domain-wide field significance was not observed, but smaller regions contained groups of rain-gauges with weak CCs. The correlation direction in d4PDF mostly matched observations near to these smaller regions. However, d4PDF also showed relationships at regions where observations showed none. This was particularly severe for ENSO-, where d4PDF showed a strong bias towards correlation. Domain-wide field significance was observed for the three higher SST modes, but d4PDF did not reproduce the correlation spatial patterns.

Rainfall extremes were correlated with regional monsoon indices calculated from JRA55 and d4PDF-AGCM. Up to moderate CCs were found for JRA55 with $\mu_{QU}$ (the seasonal anomaly of monsoon jet zonal water vapour flux), at southern Kyushu and a few locations along the Pacific coast (Fig. 9a). From JRA55, $\mu_{QU}$ was weakly anti-correlated and correlated with ENSO+ and ENSO-, respectively. From d4PDF, $\mu_{QU}$ was moderately correlated with ENSO-. These relationships are consistent with Naoi et al. (2020) who found atmospheric rivers to be more frequent in years when El Niño transited into La Niña; An El Niño transitioning into a La Niña may be interpreted as combining the reversed effects of the ENSO+ mode and the effects of the ENSO- mode, and associated with increased $\mu_{QU}$.

Rainfall extremes were observed to be weakly anti-correlated with PDV over many regions, particularly at daily time resolution (Fig. 6d). While the relationships with individual modes were weak, they may stack since the SST anomalies of the PDV are quite ENSO-like. The relationships found in this study are consistent with Kawamura et al. (2001) and Jin et al. (2005), which found strong rainfall over Fukuoka after strong La Niña events; A La Niña transitioning into an El Niño may be interpreted as combining the effects of the ENSO+ mode (Fig. 6a) and the reversed effects of the ENSO- mode (reversed Fig. 6a). Should this event also occur during the Niña-like phase of the PDV (reversed Fig. 6d), the three modes would stack in their strengthening influence on extreme rainfall such as over northwestern Kyushu.

The Pacific SST modes then seemed to modulate rainfall extremes over Japan through the water vapour flux of the regional monsoon. Since SST anomalies rarely occur as pure modes but as a combination of modes, in a particular year a few climate modes may stack to create a stronger modulating effect on rainfall. Such cases would be exacerbated by the influence of the Trend+ mode in the future warmer climate, specifically domain-wide field significant increase of hourly extremes (Fig. 5c). However, this is still a hypothesis based on this study's statistical analysis. The statistical results should be investigated through sensitivity tests using physics-based numerical weather prediction models. Recently, large ensemble climate prediction datasets such as d4PDF have been widely used in order to extract probabilities of occurrence of extreme weather such as tropical cyclones and heavy rainfalls. Attributing of such extreme events to global warming has been one of the foci in climate and impact assessment studies. In identifying the relationship of extreme events with specific climate modes, special care in evaluating the representation and performance of climate prediction datasets is desirable. This study demonstrates one of the approaches for such evaluations.

## Code availability

All software used to process the data are publicly available from their respective project webpages.

## Data availability

COBE-SST may downloaded from https://ds.data.jma.go.jp/tcc/tcc/products/elnino/cobesst/cobe-sst.html. The d4PDF data may be downloaded from https://diasjp.net/en/service/d4pdf-data-download. Rainfall from meteorological stations data may be downloaded from the JMA website at https://www.data.jma.go.jp/gmd/risk/obsdl/index.php. SOI and TPI climate indices may be downloaded from the NOAA website https://psl.noaa.gov/data/climateindices/list.

## Author contribution

Conceptualisation and Writing: Lee and Takemi. Investigation: Lee. Supervision and Project administration: Takemi.

## Competing interests

The authors declare that they have no conflict of interest.

## Acknowledgements

This work was supported by the MEXT-Program for the advanced studies of climate change projection (SENTAN) Grant Number JPMXD0722678534 and was also supported by the Environment Research and Technology Development Fund Number 2-2303 of the Environmental Restoration and Conservation Agency.

## References

Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N. C., and Scott, J. D.: The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *J. Clim.*, **15(16)**, 2205-2231, https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2, 2002.

Bolton, D.: The computation of equivalent potential temperature. *Mon. Wea. Rev.*, **108(7)**, 1046-1053, https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2, 1980.

Campello, R. J. G. B., Moulavi, D., Sander, J.: Density-Based Clustering Based on Hierarchical Density Estimates, in: Advances in Knowledge Discovery and Data Mining. PAKDD 2013. Lecture Notes in Computer Science, vol 7819, edited by Pei, J., Tseng, V.S., Cao, L., Motoda, H., Xu, G., Springer, Berlin, Heidelberg, Germany, 160–172. https://doi.org/10.1007/978-3-642-37456-2_14, 2013.

Met Office: Cartopy: a cartographic python library with a Matplotlib interface, https://scitools.org.uk/cartopy, 2010 – 2015.

Fujibe, F.: Annual variation of extreme precipitation intensity in Japan: Assessment of the validity of Clausius-Clapeyron scaling in seasonal change. *SOLA*, **12**, 106-10, https://doi.org/10.2151/sola.2016-024, 2016.

Guan, B. and Nigam, S.: Pacific sea surface temperatures in the twentieth century: An evolution-centric analysis of variability and trend. *J. Clim.*, **21(12)**, 2790-809, https://doi.org/10.1175/2007JCLI2076.1, 2008.

Harada, Y., Endo, H., and Takemura, K.: Characteristics of large-scale atmospheric fields during heavy rainfall events in western Japan: Comparison with an extreme event in early July 2018. *J. Meteorol. Soc. Japan. Ser. II*, **98(6)**, 1207-29, https://doi.org/10.2151/jmsj.2020-062, 2020.

Harris, C. R., and Coauthors: Array programming with NumPy. *Nature*, **585**, 357–62, https://doi.org/10.1038/s41586-020-2649-2.

Henley, B. J., and Coauthors: A Tripole Index for the Interdecadal Pacific Oscillation, *Clim. Dyn.*, **45(11-12)**, 3077-3090, http://dx.doi.org/10.1007/s00382-015-2525-1, 2015. Available at https://psl.noaa.gov/data//data/timeseries/IPOTPI.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Clim.*, **27**, 57–75, https://doi.org/10.1175/JCLI -D-12-00837.1, 2020.

Horel, J. D., and Wallace, J. M.: Planetary-scale atmospheric phenomena associated with the interannual variability of sea surface temperature in the equatorial Pacific. *Mon. Wea. Rev.*, **109**, 813–829, https://doi.org/10.1175/1520-0493(1981)109<0813:PSAPAW>2.0.CO;2, 1981.

Hoskins, B. J., and Karoly, D. J.: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179–1196, https://doi.org/10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2, 1981.

Hoyer, S. and Hamman, J.: xarray: N-D labeled Arrays and Datasets in Python. *Journal of Open Research Software*, **5(1)**, 10, https://doi.org/10.5334/jors.148, 2017.

Hunter, J. D.: Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, **9(3)**, 90–5, https://doi.org/10.1109/MCSE.2007.55, 2007.

Ishii, M., Shouji, A., Sugimoto, S., and Matsumoto, T.: Objective Analyses of Sea-Surface Temperature and Marine Meteorological Variables for the 20th Century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865-879, https://doi.org/10.1002/joc.1169 , 2005.

Ishii M., and Mori, N.: d4PDF: large-ensemble and high-resolution climate simulations for global warming risk assessment. *Prog. Earth Planet Sci.*, **7(1)**, 1-22, https://doi.org/10.1186/s40645-020-00367-7, 2020.

Japan Meteorological Agency: Characteristics of Global Sea Surface Temperature Analysis Data (COBE-SST) for Climate Use. Monthly Report on Climate System Separated Volume, 12, 116pp, 2006.

Jin, Y. H., Kawamura, A., Jinno, K., and Berndtsson, R.: Quantitative relationship between SOI and observed precipitation in southern Korea and Japan by nonparametric approaches. *Journal of Hydrology*, **301(1-4)**, 54-65, https://doi.org/10.1016/j.jhydrol.2004.06.026, 2005.

Kamae, Y., Imada, Y., Kawase, H., and Mei, W.: Atmospheric rivers bring more frequent and intense extreme rainfall events over East Asia under global warming. *Geophys. Res. Lett.*, **48**, e2021GL096030, https://doi.org/10.1029/2021GL096030, 2021.

Kawamura, A., Eguchi, S., and Jinno, K.: Correlation between Southern Oscillation and monthly precipitation in Fukuoka. *Journal of Japan Society of Civil Engineers*, **2001(691)**, 153-8. https://doi.org/10.2208/jscej.2001.691_153, 2001.

Kawase, H., and CoAuthors: Identifying robust changes of extreme precipitation in Japan from large ensemble 5-km-grid regional experiments for 4K warming scenario. *J. Geophys. Res.-Atmos*, **128(18)**, https://doi.org/10.1029/2023JD038513, 2023.

Kobayashi, S., and CoAuthors: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteorol. Soc. Japan. Ser. II*, **93(1)**, 5-48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Kosaka, Y., Xie, S. P., Lau, N. C., and Vecchi, G. A.: Origin of seasonal predictability for summer climate over the Northwestern Pacific. *Proc. Natl. Acad. Sci. U.S.A.*, **110(19)**, 7574-9, https://doi.org/10.1073/pnas.1215582110, 2013.

Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., and Schreck, C. J.: International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4. NOAA National Centers for Environmental Information, https://doi.org/10.25921/82ty-9e16, 2018.

Lee, S., L'Heureux, M., Wittenberg, A. T., Seager, R., O'Gorman, P. A., and Johnson, N. C.: On the future zonal contrasts of equatorial Pacific climate: Perspectives from Observations, Simulations, and Theories. *npj Clim. Atmos. Sci.*, **5**, 82, https://doi.org/10.1038/s41612-022-00301-2, 2022.

Li, Y., Deng, Y., Yang, S., and Zhang, H.: Multi-scale temporospatial variability of the East Asian Meiyu-Baiu fronts: characterization with a suite of new objective indices. *Clim. Dyn.*, **51**, 1659–1670, https://doi.org/10.1007/s00382-017-3975-4, 2018.

Livezey, R. E., and W. Y. Chen: Statistical field significance and its determination by Monte Carlo techniques, *Mon. Wea. Rev.*, **111**, 46–59, 1983.

Matsumoto, S., Ninomiya, K., and Yoshizumi, S.: Characteristic features of "Baiu" front associated with heavy rainfall. *J. Meteorol. Soc. Japan. Ser. II*, **49(4)**, 267-281, https://doi.org/10.2151/jmsj1965.49.4_267, 1971.

McKinney W.: Data structures for statistical computing in python, in Proceedings of the 9th Python in Science Conference. p. 51–6. https://doi.org/10.25080/Majora-92bf1922-00a, 2010.

Mizuta, R., and Coauthors: Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bull. Amer. Meteor. Soc.*, **98(7)**,1383-98, https://doi.org/10.1175/BAMS-D-16-0099.1, 2017.

Mizuta, R., and Coauthors: Climate simulations using MRI-AGCM3.2 with 20-km grid. *J. Meteorol. Soc. Japan. Ser II*, **90**, 233–258, https://doi.org/10.2151/jmsj.2012- A12, 2012.

Mori, N., and Takemi, T.: Impact assessment of coastal hazards due to future changes of tropical cyclones in the North Pacific Ocean. *Weather Clim. Extremes*, **11**, 53-69, https://doi.org/10.1016/j.wace.2015.09.002, 2016.

Mori, N., and Coauthors: Recent nationwide climate change impact assessments of natural hazards in Japan and East Asia. *Weather Clim. Extremes*, **32**, 100309, https://doi.org/10.1016/j.wace.2021.100309, 2021.

Naka, N., and Takemi, T.: Characteristics of the environmental conditions for the occurrence of recent extreme rainfall events in northern Kyushu, Japan. *Sci. Online Lett. Atmos.*, **19A**, 9-16, https://doi:10.2151/sola.19A-002, 2023.

Naoi, M., Kamae, Y., Ueda, H., and Mei, W.: Impacts of seasonal transitions of ENSO on atmospheric river activity over East Asia. *J. Meteor. Soc. Japan*, **98**, 655–668, https://doi.org/10.2151/jmsj.2020-027, 2020.

Nayak, S., and Takemi, T.: Dependence of extreme precipitable water events on temperature. *Atmósfera*, **32(2)**, 159-65, https://doi.org/10.20937/atm.2019.32.02.06, 2019.

Nayak, S., and Takemi, T.: Atmospheric driving mechanisms of extreme precipitation events in July of 2017 and 2018 in western Japan. *Dynamics of Atmospheres and Oceans*, **93**, 101186, https://doi.org/10.1016/j.dynatmoce.2020.101186, 2021.

The NCAR Command Language (Version 6.6.2) [Software]. Boulder, Colorado: UCAR/NCAR/CISL/TDD. https://doi.org/10.5065/D6WD3XH5, 2019.

Ninomiya, K., and Mizuno, H.: Variations of Baiu precipitation over Japan in 1951-1980 and large-scale characteristics of wet and dry Baiu. *J. Meteorol. Soc. Japan. Ser. II*, 1987, **65(1)**, 115-27, https://doi.org/10.2151/jmsj1965.65.1_115, 1987.

Nitta, T.: Convective activities in the tropical western Pacific and their impact on the Northern Hemisphere summer circulation. *J. Meteorol. Soc. Japan. Ser. II*, **65(3)**, 373-90, https://doi.org/10.2151/jmsj1965.65.3_373, 1987.

Ohba, M., Kadokura, S., Yoshida, Y., Nohara, D., and Toyoda, Y.: Anomalous weather patterns in relation to heavy precipitation events in Japan during the Baiu season. *J. Hydrometeorol*, **16(2)**, 688-701, https://doi.org/10.1175/JHM-D-14-0124.1, 2015.

Ropelewski, C. F., and Jones, P. D.: An extension of the Tahiti-Darwin Southern Oscillation Index, *Mon. Wea. Rev.*, **115**, 2161-2165, https://doi.org/10.1175/1520-0493(1987)115<2161:AEOTTS>2.0.CO;2,1987.

Sakashita, W., Yokoyama, Y., Miyahara, H., Yamaguchi, Y. T., Aze, T., Obrochta, S. P., and Nakatsuka, T.: Relationship between early summer precipitation in Japan and the El Niño-Southern and Pacific Decadal Oscillations over the past 400 years. *Quaternary International*, **397**, 300-6, https://doi.org/10.1016/j.quaint.2015.05.054, 2016.

Sasaki, H., Murata, A., Hanafusa, M., Oh'izumi, M., and Kurihara, K.: Reproducibility of present climate in a non-hydrostatic regional climate model nested within an atmosphere general circulation model. *SOLA*, **7**, 173–176, https://doi.org/10.2151/sola.2011-044, 2011.

Schulzweida, U.: CDO User Guide (2.1.0). Zenodo [code], https://doi.org/10.5281/zenodo.7112925, 2022.

Takaya, Y., Ishikawa, I., Kobayashi, C., Endo, H., and Ose, T.: Enhanced Meiyu-Baiu rainfall in early summer 2020: Aftermath of the 2019 super IOD event. *Geophys. Res. Lett.*, **47**, e2020GL090671, https://doi.org/10.1029/2020GL090671, 2020.

Takemura, K., Mukougawa, H., and Maeda, S.: large-scale atmospheric circulation related to frequent rossby wave breaking near Japan in boreal summer. *J. Clim.*, **33(15)**, 6731-44, https://doi.org/10.1175/JCLI-D-19-0958.1, 2020.

Takemura, K., Mukougawa, H.: A new perspective of Pacific–Japan pattern: Estimated percentage of the cases triggered by Rossby wave breaking. *J. Meteorol. Soc. Japan. Ser. II*, **100(1)**, 115-39, https://doi.org/10.2151/jmsj.2022-006, 2022.

Tanaka, M.: Interannual and interdecadal variations of the western North Pacific monsoon and Baiu rainfall and their relationship to the ENSO cycles. *J. Meteorol. Soc. Japan. Ser. II*, **75(6)**, 1109-23, https://doi.org/10.2151/jmsj1965.75.6_1109, 1997.

Unuma, T., and Takemi, T.: Rainfall characteristics and their environmental conditions during the heavy rainfall events over Japan in July of 2017 and 2018. *J. Meteorol. Soc. Japan. Ser. II*, **99(1)**, 165-80, https://doi.org/10.2151/jmsj.2021-009, 2021.

Vecchi, G. A., Clement, A., and Soden, B. J.: Examining the tropical Pacific's response to global warming. *Eos, Transactions American Geophysical Union*, **89(9)**, 81-3, https://doi.org/10.1029/2008EO090002, 2008.

Virtanen, P., and Coauthors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, **17**, 261–72, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wang, B., Wu, R., and Fu, X.: Pacific–East Asian teleconnection: how does ENSO affect East Asian climate?. *J. Clim.*, **13(9)**, 1517-36, https://doi.org/10.1175/1520-0442(2000)013<1517:PEATHD>2.0.CO;2, 2000.

Webb, A., Shimura, T., and Mori, N.: Global tropical cyclone track detection and analysis of the d4PDF mega-ensemble projection. *Journal of Japan Society of Civil Engineers, Ser. B2 (Coastal Engineering)*, 75(2), I_1207-12, https://doi.org/10.2208/kaigan.75.I_1207, 2019.

Weare, B. C., and Nasstrom, J. S.: Examples of extended empirical orthogonal function analyses. *Mon. Wea. Rev.*, **110(6)**, 481-5, https://doi.org/10.1175/1520-0493(1982)110<0481:EOEEOF>2.0.CO;2, 1982.

Xie, S. P., Kosaka, Y., Du, Y., Hu, K., Chowdary, J. S., and Huang, G.: Indo-western Pacific Ocean capacitor and coherent climate anomalies in post-ENSO summer: A review. *Adv. Atmos. Sci.*, **33(4)**, 411-32, https://doi.org/10.1007/s00376-015-5192-6, 2016.

Yokoyama, C., Tsuji, H., and Takayabu, Y. N.: The effects of an upper-tropospheric trough on the heavy rainfall event in July 2018 over Japan. *J. Meteorol. Soc. Japan. Ser. II*, **98(1)**, 235-55, https://doi.org/10.2151/jmsj.2020-013, 2020.

Zhang, W., Huang, Z., Jiang, F., Stuecker, M. F., Chen, G., and Jin, F.-F.: Exceptionally persistent Madden-Julian Oscillation activity contributes to the extreme 2020 East Asian summer monsoon rainfall. *Geophys. Res. Lett.*, **48**, e2020GL091588, https://doi.org/10.1029/2020GL091588, 2021.