Earth's Future

RESEARCH ARTICLE
10.1029/2021EF002481

Increasing Frequency of Anomalous Precipitation Events in Japan Detected by a Deep Learning Autoencoder

H. Murakami[1,2], T. L. Delworth[2,3], W. F. Cooke[2], S. B. Kapnick[2,4], and P.-C. Hsu[5]

[1]Cooperative Programs for the Advancement of Earth System Science, University Corporation for Atmospheric Research, Boulder, CO, USA, [2]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, NJ, USA, [3]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA, [4]Now at J. P. Morgan Chase & Co, New York, NY, USA, [5]Key Laboratory of Meteorological Disaster of Ministry of Education, Nanjing University of Information Science and Technology, Nanjing, China

Abstract The frequency of large-scale anomalous precipitation events associated with heavy precipitation has been increasing in Japan. However, it is unclear if the increase is due to anthropogenic warming or internal variability. Also, it is challenging to develop an objective methodology to identify anomalous events because of the large variety of anomalous precipitation cases. In this study, we applied a deep learning technique to objectively detect anomalous precipitation events in Japan for both observations and simulations using high-resolution climate models. The results show that the observed increases in anomalous heavy precipitation events in Western Japan during 1977–2015 were not made only by internal variability but the increases in anthropogenic forcing played an important role. Such events will continue to increase in frequency this century. The increases are attributable to the increasing frequency of tropical cyclones and enhanced frontal rainbands near Japan. These results highlight the mitigation challenge posed by the increasing occurrence of unprecedented precipitation events in the future.

Plain Language Summary The increasing frequency of heavy precipitation events in Japan arose public interest regarding its relevance to ongoing climate changes. However, little was known about how much of the observed increasing trend over the last 40 years can be statistically distinguishable from internal natural variability and attributable to climatic changes induced by human activity. Moreover, it was unclear if any rare precipitation events, other than intense precipitation events, have been increased in Japan. Here, we developed a new deep learning technique—autoencoder—that objectively identifies rare precipitation events. Observations and climate simulations suggest that the frequency of rare precipitation events has been increased in Japan over the last 40 years and attributable to human-induced climatic changes and will continue to increase in the future.

Key Points:

- A new deep learning technique—autoencoder—was developed to objectively identify anomalous precipitation events
- Frequency of anomalous precipitation events has been increased in Japan and will continue to increase in the future
- The increases are attributable to the anthropogenic forcing through the increased tropical cyclones and enhanced frontal rainbands

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

H. Murakami,
hir.murakami@gmail.com

Citation:

Murakami, H., Delworth, T. L., Cooke, W. F., Kapnick, S. B., & Hsu, P.-C. (2022). Increasing frequency of anomalous precipitation events in Japan detected by a deep learning autoencoder. *Earth's Future*, *10*, e2021EF002481. https://doi.org/10.1029/2021EF002481

Received 4 OCT 2021
Accepted 4 APR 2022

Author Contributions:

Conceptualization: H. Murakami
Data curation: H. Murakami, W. F. Cooke
Formal analysis: H. Murakami
Investigation: H. Murakami
Methodology: H. Murakami
Resources: H. Murakami
Software: H. Murakami
Validation: H. Murakami
Visualization: H. Murakami
Writing – original draft: H. Murakami
Writing – review & editing: H. Murakami, T. L. Delworth, S. B. Kapnick, P.-C. Hsu

1. Introduction

The effect of climate change on the frequency and intensity of anomalous hydroclimate events associated with heavy precipitation is of great interest and an important topic for society because of its substantial impact on hazards, water resources, agriculture, ecosystems, economies, insurance, and mitigation policy. The frequency of anomalous precipitation events around the world might have already increased and is anticipated to continue to increase in the future due to global warming (Fujibe, 2013; IPCC, 2013; Kharin & Zwiers, 2000; Min et al., 2011). One reason for this increase is thermodynamic: the maximum amount of water vapor in the air increases with temperature by approximately 6% K^{-1} at 298.15 K, as given by the Clausius–Clapeyron relation. However, some studies have reported a higher increase of precipitation with temperature than is expected from this relation, possibly associated with dynamic feedbacks such as circulation changes (Berg & Haerter, 2013; Emori & Brown, 2005) Therefore, the relationship between temperature and anomalous precipitation is not simple and varies regionally (IPCC, 2013; Shaw et al., 2011; Yu & Li, 2012).

In early July of 2020, a heavy precipitation event occurred in Japan, causing 37 casualties and damage to about 17,000 houses (Cabinet Office Japan, 2020). In early August of 2021, northern areas of Kyushu Island experienced an unprecedented heavy precipitation event that recorded more than 1,000 mm in total precipitation within 1 week which was about 50% of the mean annual total precipitation in the region. In addition to the 2020 and 2021 events, the scientific community is aware of more frequent heavy precipitation events over the last decade in Japan (e.g., July 2012, August 2014, June 2017, and July 2018), especially in Western Japan (see Figure 1 for

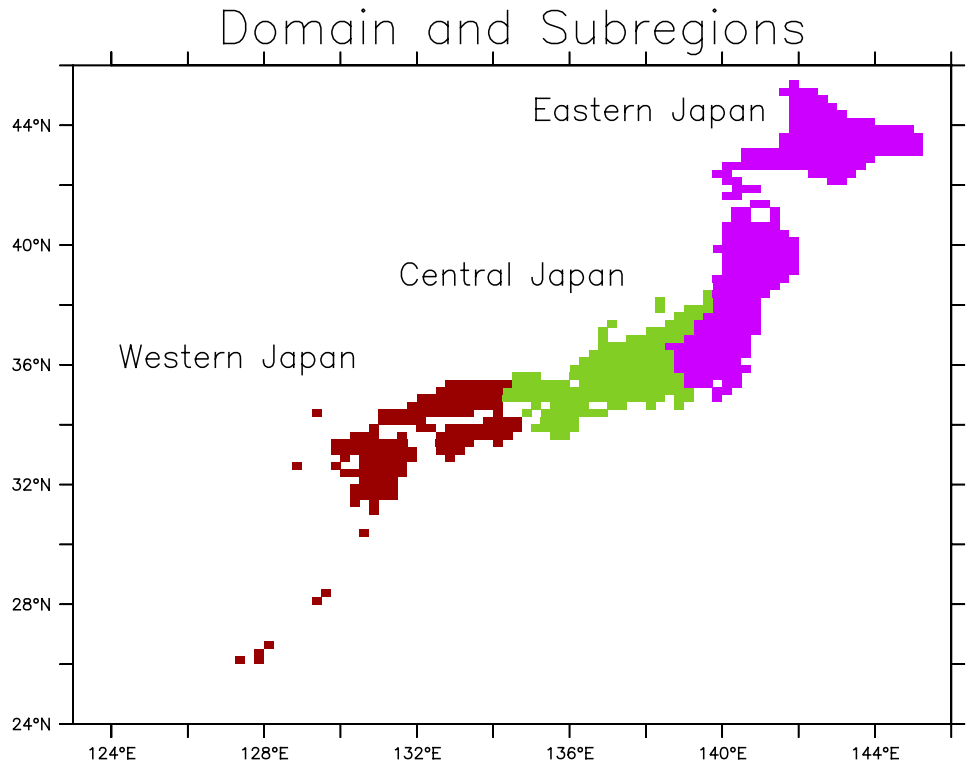

Figure 1. Subregion of Japan. There are 47 prefectures in Japan. These prefectures are grouped into three subregions as indicated by the colors and names. Also, shown is the domain used for the autoencoder.

regional names in Japan) at the end of the rainy season (i.e., between late July and early August; JMA, 2019; Kawano & Kawamura, 2020; Tsuji et al., 2020). Indeed, several studies have reported increasing trends in the observed frequency of heavy precipitation events in Japan (Fujibe, 2013; JMA, 2019; Ohba et al., 2015) and projected future increases by the end of this century (Hatsuzuka & Sato, 2019).

However, several issues remain to be addressed. First, there is limited literature attributing the observed trends in Japan to anthropogenic forcing or natural variability, although some studies have addressed the changes in the frequency of heavy precipitation in future projections (Hatsuzuka & Sato, 2019). Recent studies (Hatsuzuka et al., 2020; Kawase et al., 2019) reported the potential effect of anthropogenic warming on heavy precipitation on Kyushu Island over the past few decades. However, limitations in their experimental design made it difficult to separate the roles of internal variability and anthropogenic forcing in the observed and simulated trends of heavy precipitation. Moreover, it is unclear if other areas of Japan have also shown any significant trends in heavy precipitation, and if such trends are attributable to increases in anthropogenic forcing. Second, the horizontal resolutions of the models used in past studies to make future projections in heavy precipitation events were not fine enough to simulate intense tropical cyclones (TCs), which produce some of the heaviest precipitation events. In fact, a typical heavy precipitation event in Japan when there is a quasi-stationary front (QSF) is often accompanied by TCs near Japan, which enhance the frontal convective activity by transporting warm and moist air toward Japan (Utsumi et al., 2017). Thus, using a high-resolution climate model is vital for assessing the effect of climate change on anomalous precipitation in Japan. However, conducting large-ensemble climate simulations using high-resolution climate models is still a challenge owing to the limitations imposed by the current level of computational resources available. Lastly, it is not clear how to objectively identify anomalous precipitation events. Most previous studies defined an anomalous precipitation event in terms of intensity (Fujibe, 2013; Hatsuzuka et al., 2020; JMA, 2019; Kawase et al., 2019), such as $\geq$50 mm hr^{-1}. However, an anomalous event should be considered not only by its intensity but also by its spatial extent (IPCC, 2013). For example, even with a normal intensity of precipitation, we should consider it an anomalous event when the coverage area of the precipitation is unprecedented.

To address the above issues, the present paper utilizes a deep learning technique—an autoencoder—to objectively identify anomalous events from daily precipitation data. Deep learning has begun to be applied more commonly

in Earth system science (Reichstein et al., 2019), with recent studies succeeding in applying deep learning architectures, such as an autoencoder, to objectively extract spatial features to identify and classify anomalous events (Liu et al., 2016; Racah et al., 2017). Such an approach could be superior to conventional methods in terms of improved accuracy of identification (W. Z. Zhao & Du, 2016) as well as independence from subjective human annotation or predefined arbitrary thresholds for anomalous events (Liu et al., 2016; Racah et al., 2017).

The developed autoencoder was applied to both observations and outputs from large-ensemble simulations using high-resolution climate models in which intense TCs and precipitation are well reproduced as observed. Large-ensemble simulations with and without anthropogenic forcing effect are also employed using the models to differentiate the potential influence of natural variability versus anthropogenic forcing on the frequency of anomalous events. The main objectives and scope of this study are as follows:

1. To develop a deep learning method to objectively identify anomalous precipitation events in Japan.
2. To identify if there are any significant trends in the frequency of anomalous precipitation events.
3. To elucidate if observed trends in the frequency of anomalous precipitation events in Japan are attributable to the increased level of anthropogenic forcing and the trends are beyond the influence of internal variability using model climate simulations.
4. To identify the physical causes of the trends in the frequency of anomalous precipitation events in Japan (e.g., changes in TC and meteorological front activities).
5. To project potential future changes in the frequency of anomalous precipitation events in Japan.

The remainder of this paper is organized as follows. Section 2 describes the methods including observed data, models, experimental design, deep learning autoencoder, and an automated method for tracking meteorological fronts. Section 3 shows the results. Section 4 provides brief conclusions and discussions.

2. Methods

2.1. Observed and Reanalysis Data

Daily mean high-resolution (0.05° × 0.05°) observed precipitation data over Japan were obtained from the Asian Precipitation—Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) project (version APHRO_JP_V1207 and V1207_R3; Kamiguchi et al., 2010) over the period 1950–2015. However, we utilized APHRODITE over the period 1977–2015 because the data before 1977 may include uncertainties (Kamiguchi et al., 2010). This gridded data were compiled from rain gauge observations.

The daily mean potential equivalent temperature from the Japanese 55-year Reanalysis (JRA-55; Kobayashi et al., 2015) over the period 1977–2015 was also used for an automated method for tracking meteorological fronts (to be described in Section 2.6 in more detail). The International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al., 2010), version 4, was used over the period 1977–2015 for the TC data.

2.2. Model and Large-Ensemble Experiments

The Geophysical Fluid Dynamics Laboratory (GFDL) Hi-resolution version of the Forecast-oriented Low Ocean Resolution model (HiFLOR, Murakami et al., 2015) and Seamless System for Prediction and Earth System Research (SPEAR, Delworth et al., 2020) were used for the large-ensemble climate simulations. HiFLOR incorporates a 25-km mesh atmosphere and land components, and 100-km mesh sea-ice and ocean components. SPEAR is the current operational seasonal and decadal prediction model developed at GFDL, consisting of the new AM4-LM4 atmosphere and land-surface model (M. Zhao et al., 2018a, 2018b), the MOM6 ocean model (https://github.com/NOAA-GFDL/MOM6), and the SIS2 sea-ice model (Adcroft et al., 2019). The horizontal resolution of the ocean component is identical to that of HiFLOR, whereas that of the atmosphere is a 50-km mesh.

We conducted two types of multidecadal simulations, AllForc (all forcing) and NatForc (natural forcing), using HiFLOR and SPEAR. A summary of the experiments is listed in Table 1. For the AllForc experiments, the historical time-varying anthropogenic forcing (e.g., greenhouse gases, ozone, and anthropogenic aerosols) and natural forcing (e.g., volcanic aerosols and solar constant) were prescribed for the period 1941–2004 for HiFLOR and 1921–2014 for SPEAR, and the anticipated future anthropogenic forcing was prescribed for the period

Table 1
Simulation Configuration

Type	Model	Period	External forcing	Ensemble member	Volcanic forcing
AllForc	HiFLOR	1941–2004	Time-varying historical natural and anthropogenic forcing	15	Yes
		2005–2050	Anthropogenic forcing under the RCP4.5 scenario		No
	SPEAR	1921–2014	Time-varying historical natural and anthropogenic forcing	30	Yes
		2015–2100	Anthropogenic forcing under the SSP5-8.5 scenario		No
NatForc	HiFLOR	1941–2020	Radiative forcing fixed at the 1941 level	15	Yes
	SPEAR	1921–2100	Radiative forcing fixed at the 1921 level	30	

Note. Two models (HiFLOR and SPEAR) were used for the large-ensemble simulations. The table summarizes the simulation types, models employed, periods of simulations, the external forcing, number of ensemble members, and whether or not there was volcanic forcing.

2005–2050 based on the Representative Concentration Pathway (RCP)4.5 scenario HiFLOR and 2015–2100 based on the Shared Socioeconomic Pathway 5-8.5 (SSP5-8.5; Kriegler et al., 2017; Riahi et al., 2017) for SPEAR. In the simulations, no volcanic event is assumed for HiFLOR and SPEAR after 2006 and 2015, respectively. The number of ensemble simulations was 15 for HiFLOR and 30 for SPEAR. Note that the simulations were initiated from the restart files derived from the 1,000-year preindustrial control experiments every 20 years from the year 101 for each ensemble member (e.g., year 101 for ensemble 1, year 121 for ensemble 2). Because ensemble members initiated with different states of initial conditions, they generate their own internal variability (i.e., butterfly effect) so that the simulated internal variability, such as Interdecadal Pacific Oscillation (Folland et al., 2002), was out of phase at a specific time among the ensemble members. Taking the mean of the ensemble members can filter out the internal variability; thus, the resultant mean field can be regarded as an estimated modeled response to the external forcing (Murakami et al., 2020).

The other simulation type comprised the NatForc experiments. The experimental settings for NatForc were identical to AllForc apart from the anthropogenic forcing such as greenhouse gases and anthropogenic aerosols were fixed at their levels in 1921 for SPEAR and 1941 for HiFLOR, but keeping the same time-varying natural forcing. The main difference between AllForc and NatForc was time-varying greenhouse gases, anthropogenic aerosols, and ozone. Note that the future scenarios are different between HiFLOR and SPEAR because we conducted the HiFLOR simulations when the SSP5-8.5 scenario was unavailable. Because the projected future level of anthropogenic forcing was different between AllForc SPEAR (SSP5-8.5) and AllForc HiFLOR (RCP4.5), we did not expect comparable magnitudes of trends in the future between the models. However, owing to the monotonic increases in anthropogenic forcing in both the scenarios, we did expect monotonic increases in the frequency of anomalous events in both future scenarios that could be largely different from NatForc where the fixed level of anthropogenic forcing is prescribed in the future. Although the AllForc and NatForc experiments in SPEAR were a complement to HiFLOR, SPEAR may still leave uncertainties in the trends of anomalous precipitation events because SPEAR cannot resolve intense TCs owing to its intermediate horizontal resolution (50-km mesh).

2.3. Precipitation Anomalies for the Input of Autoencoder

Our focus is on large-scale precipitation events with a relatively long duration than the short-term synoptic-scale weather events so we analyzed 5-day running-mean anomalies of daily mean precipitation during the summer season (1 May to 31 October) for both observations and model outputs. A 5-day running-mean filter is commonly used to remove the effect of short-term synoptic weather in many previous studies (e.g., Kikuchi et al., 2012; Neena et al., 2014; Tuan, 2019). The anomalies of daily mean precipitation were obtained by subtracting the moving 20-year climatological daily mean (derived from the preceding 20-year data) from the raw data. Then, 5-day running-mean anomalies were computed by averaging the anomalies using the previous 5 days to remove the short-term weather events (e.g., squall lines). Therefore, the target of our research is large-scale anomalous precipitation that lasts for a few days rather than short-term daily extreme precipitation events. We also preliminarily tested an alternative anomaly definition—that is, a deviation from the fixed 20-year average between 1981 and 2000—and the results led to qualitatively similar conclusions relative to the results with the 20-year moving averages (Figure S1 in Supporting Information S1).

For the inputs of the autoencoder, as will be described in Section 2.4, the original 5-day running-mean precipitation anomaly data of APHRODITE and HiFLOR were further interpolated into the 0.25° × 0.25° grid boxes over the domain near Japan (24°–46°N, 123°–146°E; 88 × 92 grids; Figure 1) that is exactly the whole domain of APHRO_JP_V1207. The interpolated anomaly data were then normalized by a min–max normalization $\frac{x-\min(x)}{\max(x)-\min(x)}$, where min($x$) is the minimum precipitation anomaly (i.e., −40 mm day^{-1}) and max(x) is the maximum precipitation anomaly (i.e., 140 mm day^{-1}) to obtain normalized values ranging nearly within 0–1.0. Because a specific number of grid sizes that is divisible by 12 is computationally efficient, zero values were further padded to the edges of each longitude and latitude axis after the normalization; namely, four grids on both sides of longitudinal edges and two grids on both sides of latitudinal edges were padded with zero values, resulting in 96 × 96 gridded data.

Because the horizontal resolution of SPEAR is a 50-km mesh, it is inadequate to interpolate the precipitation anomaly data by SPEAR into the 0.25° × 0.25° grid boxes so we slightly modified the autoencoder from HiFLOR (Tables S3 and S4 in Supporting Information S1). First, we interpolated the precipitation anomaly data into 0.5° × 0.5° grid boxes over the domain near Japan (24°–46°N, 123°–146°E; 44 × 38 grids). The interpolated anomaly data were then normalized in the same way as HiFLOR. Zero values were further padded to the edges of each longitude and latitude axis after the normalization; namely, two grids on both sides of longitudinal edges and one grid on both sides of latitudinal edges were padded with zero values, resulting in 48 × 40 gridded data.

2.4. Developing an Autoencoder to Identify Anomalous Events

A flow chart describing the main processes of the autoencoder is provided in Figure 2a, and the general concept of the autoencoder for the detection of anomalous events is explained as follows.

An autoencoder is an unsupervised neural network (or deep learning) method that is commonly used for the detection of anomalies in various scientific fields (Alla & Adari, 2019; Chalapathy & Chawla, 2019). In the autoencoder procedure, the daily precipitation anomaly fields (i.e., size = 88 × 92 = 8,096, 25-km mesh), as described in Section 2.3, are converted to a low-dimensional latent representation (size = 64; i.e., encoding) that is a reduction to nearly 0.8% of the original data set (Figure 2a). The compressed low-dimensional data are therefore reconstructed to high-dimensional vectors with the same size as the original inputs (decoding). To build an autoencoder, a deep learning input precipitation data set attempts to reconstruct it, minimizes the mean square errors (MSEs), and determines millions of the parameters that configure the autoencoder. If an autoencoder is well trained by using large samples, it can reproduce the original input very well when a sample is frequently represented in the training data set (Figure 2b), resulting in a smaller MSE (Figure 2b) between before and after autoencoding. However, if a sample is very rare in the training data set, the autoencoder causes a substantial error when reconstructing the data, resulting in a large MSE (Figure 2c). Therefore, anomalous events can be detected by selecting the events with large MSEs.

A detailed and practical configure of the encoder for observations and HiFLOR is provided in Table S1 in Supporting Information S1. The encoder comprised 12 layers, including 3 2-D convolutional layers and 1 dense layer. The encoder inputs the normalized 96 × 96 precipitation anomaly data (see Section 2.1 for normalization), resulting in compression of the input data into 64 latent cells. For the decoder (Table S2 in Supporting Information S1), there were 14 layers, including 4 reversed 2-D convolution layers and 1 dense layer, resulting in the decompression of the latent data into 96 × 96 gridded data. In total, there were 2,712,385 parameters determined by deep learning for the autoencoder. Using TensorFlow, a mini-batch size of eight was set (i.e., batch size parameter), and 40 iterations were set for training (i.e., epoch parameter). We preliminarily assessed that our autoencoder does not show any overfitting. The patched edge data were then removed from the output data, resulting in the same dimensions as the original input data (88 × 92).

For SPEAR, there were nine layers in the encoder (Table S3 in Supporting Information S1), including two 2-D convolutional layers and one dense layer, resulting in the compression of the input data into 32 latent cells. For the decoder (Table S4 in Supporting Information S1), there were 11 layers, including 3 reversed 2-D convolutional layers and 1 dense layer, resulting in the decompression of the latent data into 48 × 40 gridded data. In total, there were 574,497 parameters determined through deep learning for the autoencoder for SPEAR.

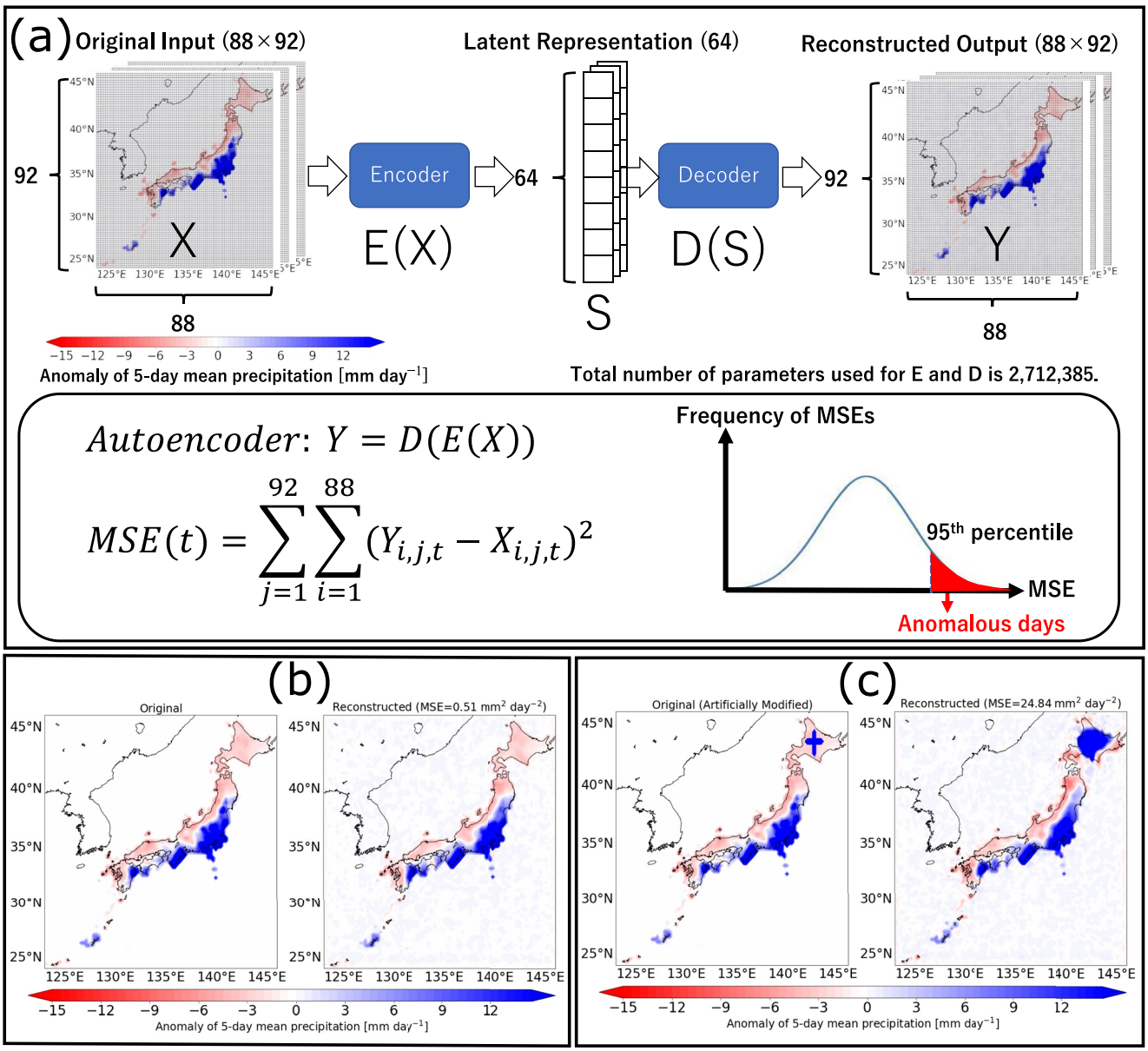

Figure 2. Flow chart and examples of using the autoencoder to identify abnormal precipitation events. (a) Flow diagram of the autoencoder. An autoencoder comprises an encoder (E) and a decoder (D). The encoder compresses the original input of precipitation with 88 × 92 grid cells (X) into a latent representation with 64 cells (S), whereas the decoder decompresses S into the original 88 × 92 grid cells (Y). An autoencoder is a series of compression and decompression [i.e., Y = D(E(X))]. The mean square error (MSE) is computed by comparing Y with X. Anomalous events are defined by their MSEs being greater than or equal to the 95th percentile. (b, c) Examples of autoencoder manipulation. In the case of (b), the autoencoder reconstructed (right) the original input (left) very well, with a small MSE of 0.51 $mm^2\ day^{-2}$, indicating that this kind of precipitation event is frequently represented in the training data set and the autoencoder can reproduce it. On the other hand, in the case of (c), artificial cross-shaped precipitation anomalies are embedded over Hokkaido Island (left). The resultant reconstruction (right) by the autoencoder shows a large MSE (24.84 $mm^2\ day^{-2}$) by misleading the reconstruction around the embedded region. This means that the embedded anomalies are very rare in the training data set, and the autoencoder did not reconstruct it well. Thus, the cases with large MSEs can be considered as anomalous events.

We trained autoencoders separately for observations, the AllForc ensemble of HiFLOR, and that of SPEAR, using the normalized precipitation anomalies between 1977 and 2015. The total number of sampling data for precipitation anomalies over the period 1977–2015 is 7,176 for observations, 107,640 (7,175 × 15 members) for HiFLOR, and 215,280 (7,175 × 30 members) for SPEAR. In the training process, two thirds of the precipitation anomalies were randomly selected and used for training, with the remaining data employed for the validation data set to construct an autoencoder. By using the constructed autoencoder, MSEs were then computed for all precipitation anomalies including training and validation over the period 1977–2015. We determined the 95th percentile

of MSEs calculated using the built-in function in Python numpy.quantile and considered the 95th percentile as the criterion for anomalous events (Figure 2a). If MSEs exceeded the 95th percentile, these events were considered as anomalous events. For the future period (i.e., 2016–2050, Table 1), MSEs are also computed by applying the same autoencoder. The same criterion of 95th percentile based on 1977–2015 is also applied for detecting anomalous events for the future period. Hereafter, we refer to anomalous events or anomalous precipitation events for the days when MSEs exceed the 95th percentile.

Note that an anomalous event might not always be linked to a heavy precipitation event; it could also be a drought event if an MSE is large during a dry period. Anomalous events are selected in terms of the rarity of events learned via the deep learning method. Therefore, an event with locally intense precipitation is not always selected as an anomalous event if similar events frequently occur. The autoencoder also considers the spatial distribution of precipitation. Even without intense precipitation, an event with a wide coverage area of precipitation is another consideration for selecting anomalous events (Figure S2 in Supporting Information S1). Also, note that our preliminary assessment revealed that applying either 90th or 99th percentile to the threshold led to the qualitatively same conclusion as in the case for the 95th percentile (Figure S3 in Supporting Information S1).

2.5. Tropical Cyclones Detection Method

Model-simulated TCs were obtained directly from 6-hourly outputs using the scheme documented in Murakami et al. (2015) and Harris et al. (2016). In short, the flood fill algorithm is applied to find closed contours of sea level pressure anomaly along with closed contours of 300–500-hPa temperature anomalies to identify the warm core. The storm detection must maintain above certain conditions as well as a specified wind speed criterion (17.5 m s^{-1} for HiFLOR and 15.75 m s^{-1} for SPEAR) for at least 36 consecutive hours. Following Murakami et al. (2020), TC positions were counted every 6 hr over each 5° × 5° grid box within the domain of Figure 1. The total count for each grid box was defined as the TC density. The TC density fields were smoothed using a nine-point moving average weighted by distance from the center of the grid box. Previous studies revealed the reasonable simulations by HiFLOR in terms of global TC distributions and intense TCs such as Category 4 and 5 hurricanes (Murakami et al., 2015); western North Pacific typhoons (Zhang et al., 2016); intensification rate of TCs (Bhatia et al., 2019); and spatial and temporal variation of hydroclimate (van der Wiel et al., 2016).

2.6. Automated Method for Tracking Meteorological Fronts

An automated method for tracking meteorological fronts is applied following a methodology documented in Schemm et al. (2015). This tracking method relies on the meridional gradient of daily equivalent potential temperature at the 850 hPa level interpolated on the 2.5° × 2.5° grid boxes. A thermal frontal parameter (TFP) is first defined as follows:

$$\mathrm{TFP} = -\nabla|\nabla\theta_e| \cdot \frac{\nabla\theta_e}{|\nabla\theta_e|}, \tag{1}$$

where θ_e stands for equivalent potential temperature at the 850 hPa level. We impose $|\nabla\theta_e| > 4\mathrm{K}\ (100\ \mathrm{km})^{-1}$ to exclude a priori regions of weak thermal gradients from the data. A front is located where TFP values are equal to zero. We fit a cubic spline through the connected grid points where TFP values are equal to zero to estimate the length of a front. We require all fronts to have a minimum length of 500 km. To separate QSFs from mobile fronts, the following parameter is considered:

$$v_f = v \cdot \frac{\nabla\mathrm{TFP}}{|\nabla\mathrm{TFP}|}, \tag{2}$$

where v is the horizontal wind at the 850 hPa level. This parameter allows us to distinguish cold ($v_f > +3$ m s^{-1}) from warm ($v_f < -3$ m s^{-1}) fronts. When $|v_f| \leq 3$ m s^{-1}, they are classified as QSFs. Figure S4 in Supporting Information S1 shows examples of fronts detected from the JRA-55 reanalysis data set compared with the actual weather maps, revealing that this automated method adequately tracks the recorded QSFs and mobile fronts. In addition, the 325-K isoline of equivalent potential temperature generally represents locations of QSFs as well as mobile fronts in the summer season near Japan. Figure S4 in Supporting Information S1 reveals that the automated method works well to detect meteorological fronts.

3. Results

3.1. Observed and Simulated Trends in Anomalous Events in Japan

The blue shaded areas in Figure 3a represent composites of the observed precipitation anomalies based on the anomalous events detected by the autoencoder derived from observed data. We preliminarily analyzed the detected anomalous events if they are accompanied by the occurrence of TCs, QSFs, mobile fronts, or their combinations using the observed TC record and the automated tracking method for meteorological fronts using the JRA55 reanalysis data set (Schemm et al., 2015).

It turned out that about 54% of the observed anomalous events occurred when TCs and QSFs simultaneously exist over the domain shown in Figure 1 and the other 45% of the events occurred when QSFs exist without TC occurrence (Figure 3b). Owing to the close linkage of TCs and QSFs with heavy rainfall, the locations of TCs

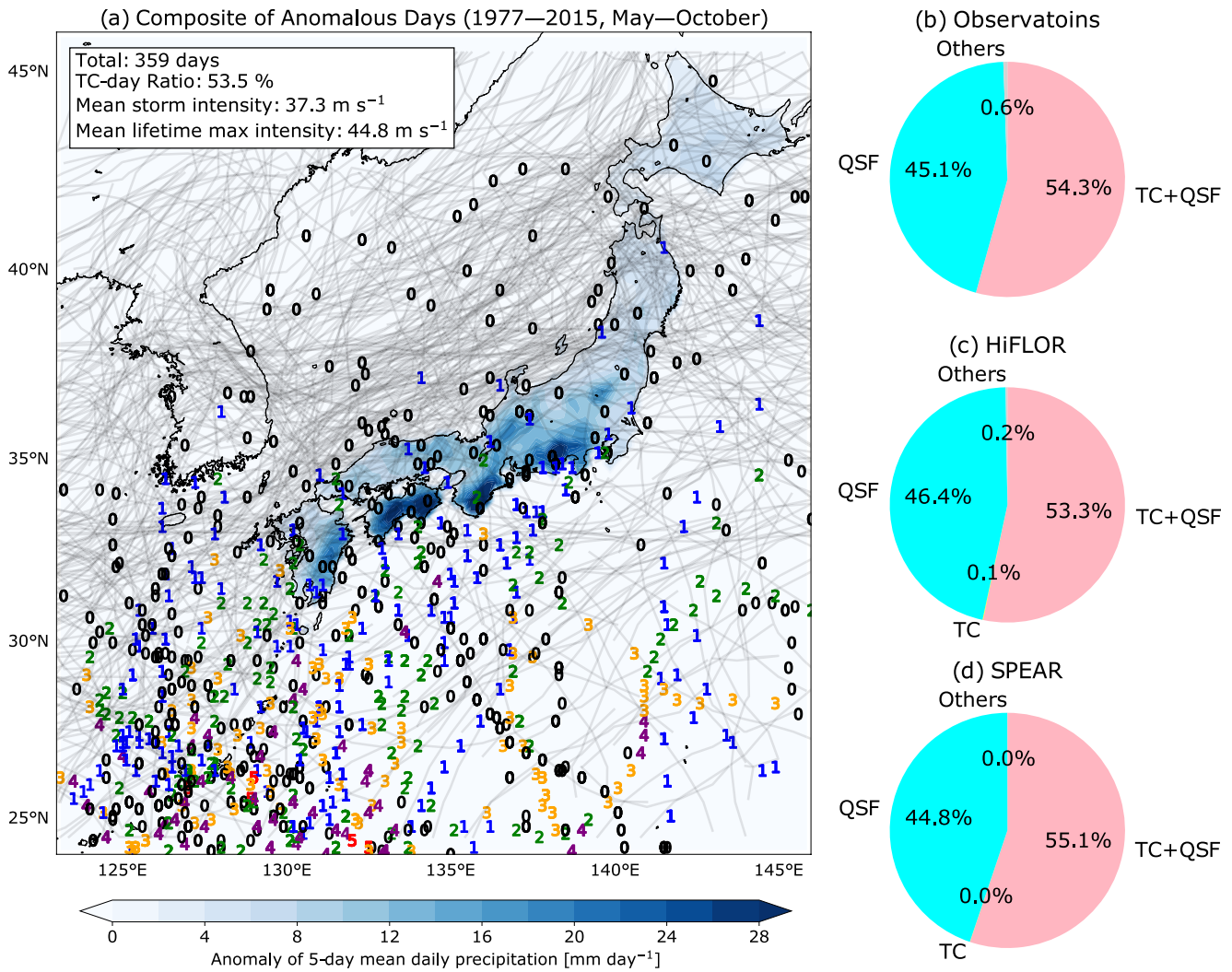

Figure 3. Composited spatial pattern of deep learning–detected anomalous precipitation events and breakdown of the events. (a) Composites of observed pentad-mean anomalies of precipitation (shading, mm day^{-1}), locations of quasi-stationary fronts (QSFs; black lines), and locations of tropical cyclones (TCs; colored numbers) for the anomalous precipitation events during the summer season (May–October) over the period 1977–2015 detected by the autoencoder. Color numbers represent the Saffir-Simpson TC intensity categories (0, tropical storms; 1–5, Categories 1–5, NOAA, 2021). The total frequency of anomalous events (359 days), the percentage of TC days relative to total days (53.5%), mean storm intensity (37.3 m s^{-1}), and mean lifetime maximum TC intensity (44.8 m s^{-1}) are listed in the top-left corner. (b) Breakdown of observed anomalous events detected by autoencoders over the period 1977–2015 using the observed TC record and automated meteorological fronts detection applying to the JRA55 reanalysis data set. Anomalous events were classified into four groups if the events are accompanied by the occurrence of TCs (TC), QSFs, their combinations (TC + QSF), and others (Others). Due to the small number, "Others" includes anomalous days associated with mobile fronts, non-TCs, and nonfrontal events. (c, d) As in (b), but for HiFLOR and SPEAR, respectively.

and quasi-stational fronts were superimposed if any TCs or QSFs were present near Japan (within the domain in Figure 3a) during the 5 days before the detected anomalous day. The total number of anomalous events for every year is defined as the frequency of anomalous events. The time series of the observed frequency in anomalous events is plotted in black in Figure 4a, showing a statistically significant positive trend over the period 1977–2015 at the 95% level using the Mann–Kendall test. The gray line represents the time series of the mean annual days of events per grid point with accumulated precipitation ≥100 mm in 5 days. The correlation coefficient between the black and gray lines over the period 1977–2015 is +0.56. Such moderate correlation indicates that the detected anomalous events are to a certain extent associated with extremely intense precipitation, but not always with intense precipitation events.

The red line in Figure 4a represents the ensemble mean frequency of the anomalous events identified by the auto-encoder applied to the 15-member AllForc experiments using the GFDL-HiFLOR climate model. The simulated time series also shows a statistically significant positive trend, like the trend derived from observations. Note that averaging the ensemble members can filter out the effect of internal variability; thus, the ensemble mean is considered as the response to anthropogenic forcing (Murakami et al., 2020). Figure 4b reveals the observed and simulated linear trends in the frequency of anomalous events according to HiFLOR and SPEAR. The AllForc experiments by the two models show statistically significant trends in both the historical period (1977–2015) and the future period included (1977–2050). Although different levels of radiative forcing were used for AllForc in the future projections between the models, the significant positive trends in the frequency of anomalous events produced by AllForc are clearly different from those produced by NatForc, in which no significant trend is evident (Figure 5). These results suggest that the observed positive trend in the frequency of anomalous events over the period 1977–2015 cannot be explained only by the model generated internal variability, but the increases in anthropogenic forcing played an important role in the observed trend (i.e., attributable). Moreover, anomalous events will likely become more frequent toward the end of this century by anticipated increases in anthropogenic forcing.

To elucidate which regions of Japan show significant changes in anomalous events, the anomalous events are grouped into three different groups based on the location of maximum precipitation (Figure 6). When maximum precipitation is found in one of the three subregions defined in Figure 1, the event is assigned to the subregion. The frequencies of events in both observations and HiFLOR do not show significant trends for both Eastern and Central Japan (Figures 6a and 6b). In contrast, both observations and HiFLOR show a statistically significant positive trend in Western Japan (Figure 6c). Overall, the consistency between the observations and AllForc simulations by HiFLOR in terms of statistically significant trends in the same subregion highlights a good reproducibility of the observed trends in HiFLOR, and thereby increases in anthropogenic forcing might have played an important role in the observed trends in Western Japan. As for future projections, the AllForc experiments project an increasing frequency of extreme events regardless of the subregions (Figure 7), whereas the NatForc experi-

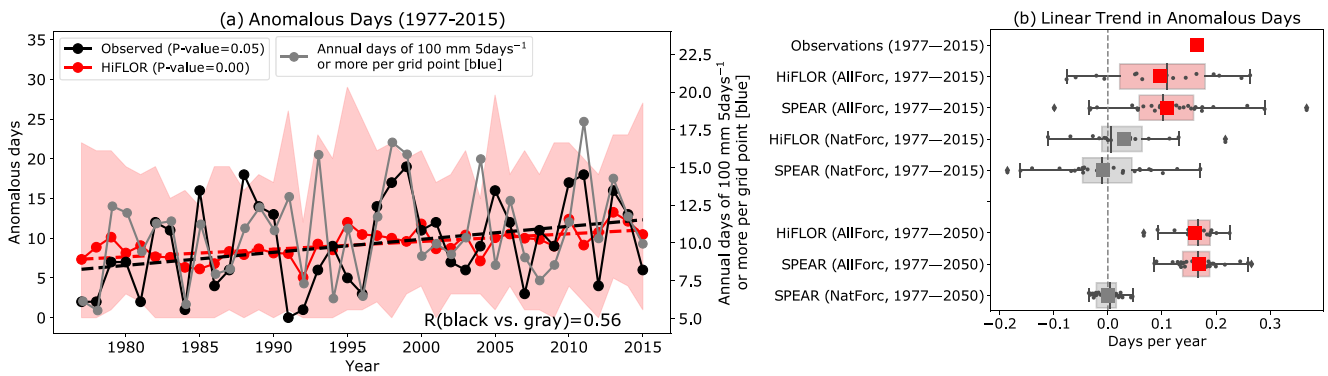

Figure 4. Time series and linear trends of deep learning–detected anomalous precipitation events. (a) Time series of annual frequency (days on the left *y*-axis) of anomalous precipitation events in the observations (black) and ensemble mean of the AllForc experiments in HiFLOR (red), along with the gray line representing the mean annual days of events (right *y*-axis) with precipitation ≥100 mm 5 days^{-1} per grid point. The Mann–Kendall significance test was applied, and the *p*-values for each trend are shown in the legends. Dashed lines were drawn when the linear trends were statistically significant at the 95% level. Red shading marks the range of minimum and maximum values of the ensemble members. (b) Linear trends in anomalous days for observations and models from 1977 to 2015 and 1977 to 2050. Red boxes or dots denote statistically significant positive trends for the ensemble means or observations at the 95% level by Mann–Kendall test. The whiskers represent the range of the 10% and 90% quantiles of ensemble members; the vertical lines show the median value; diamonds show outliers; dots show each ensemble member; and squares represent the ensemble mean or observed value. Units: days year^{-1}.

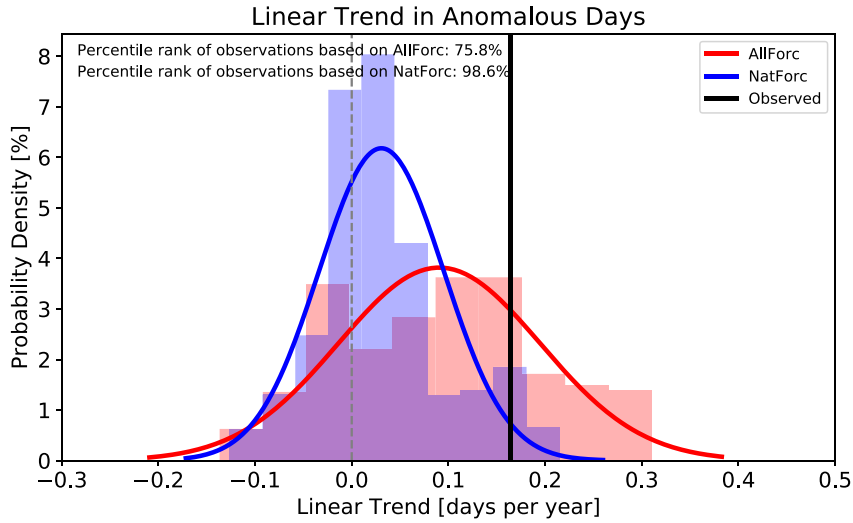

Figure 5. Histogram showing linear trends in annual anomalous days over the period 1977–2015 by the ensemble members of HiFLOR experiments and observations. The red (blue) line shows a fitted normal distribution applied to the 15 ensemble members for AllForc (NatForc) experiments. The black line denotes the observed trend over the same period. This figure shows that the observed trend is located below the 95th percentile of AllForc, but above the 95th percentile of NatForc, highlighting that the observed trend can be statistically distinguishable from internally generated noise and can be attributable to the increases in anthropogenic forcing.

ments show a steady frequency of anomalous events. These results highlight the dominant effect of anthropogenic forcing on increases in anomalous events everywhere in Japan in the future. A caveat is that the models do not always agree with observations quantitatively, specifically for the magnitude of trends and the fraction of TC days for each subregion (Figure 6c and Figure S5 in Supporting Information S1) probably due to the systematic biases in the mean state and/or TC frequency of occurrence as reported in the previous study (Murakami et al., 2015).

3.2. Effect of Tropical Cyclones and Quasi-Stationary Fronts on Trends in Anomalous Events

As mentioned earlier, most of the detected anomalous events are associated with TCs, QSFs, and their combinations. The simulated anomalous events by HiFLOR and SPEAR also showed a similar composition ratio as observed (Figures 3b–3d). We analyze the anomalous events when any TCs and QSFs exist at the same time over the analyzed domain in the 5 days before the extreme events (TC + QSF) as well as the events when only QSFs exist because the two types constitute the majority of the anomalous events. The observed frequency of TC + QSF events shows a statistically significant positive trend (Figures 8a and 8b). The historical simulations and future projections by AllForc HiFLOR and SPEAR also show statistically significant positive trends, despite slightly smaller slopes than observed, whereas there are no significant trends projected by the NatForc experiments. As for QSF events, both the observations and HiFLOR experiments do not show significant trends over the last few decades, although AllForc SPEAR shows statistically significant positive trends (Figure 8d). On the other hand, future projections show statistically significant positive trends in the AllForc experiments by both HiFLOR and SPEAR, whereas a significant positive trend is not observed by the NatForc experiment in SPEAR (Figure 8d). These results lead us to conclude that the frequency of anomalous TC + QSF events has increased over the last few decades and will continue increasing in the future due to increases in anthropogenic forcing.

We address why Western Japan specifically has experienced an increasing frequency of anomalous events over the last few decades. One of the reasons is the increasing frequency of TC occurrence (i.e., TC density) near Japan. Figures 9a and 9b show the spatial pattern of linear trends for TC frequency of occurrence for all storms (i.e., maximum wind speed greater than 17.5 m s^{-1}) between 1977 and 2015. The observed trend shows a wave-like spatial pattern with a positive trend in southwest Japan, although the positive trend does not pass the significance test. The simulated trend derived from the ensemble mean of the AllForc HiFLOR experiments over the same period moderately reproduces the increasing trends in southwest Japan (Figure 9b), albeit the magnitude

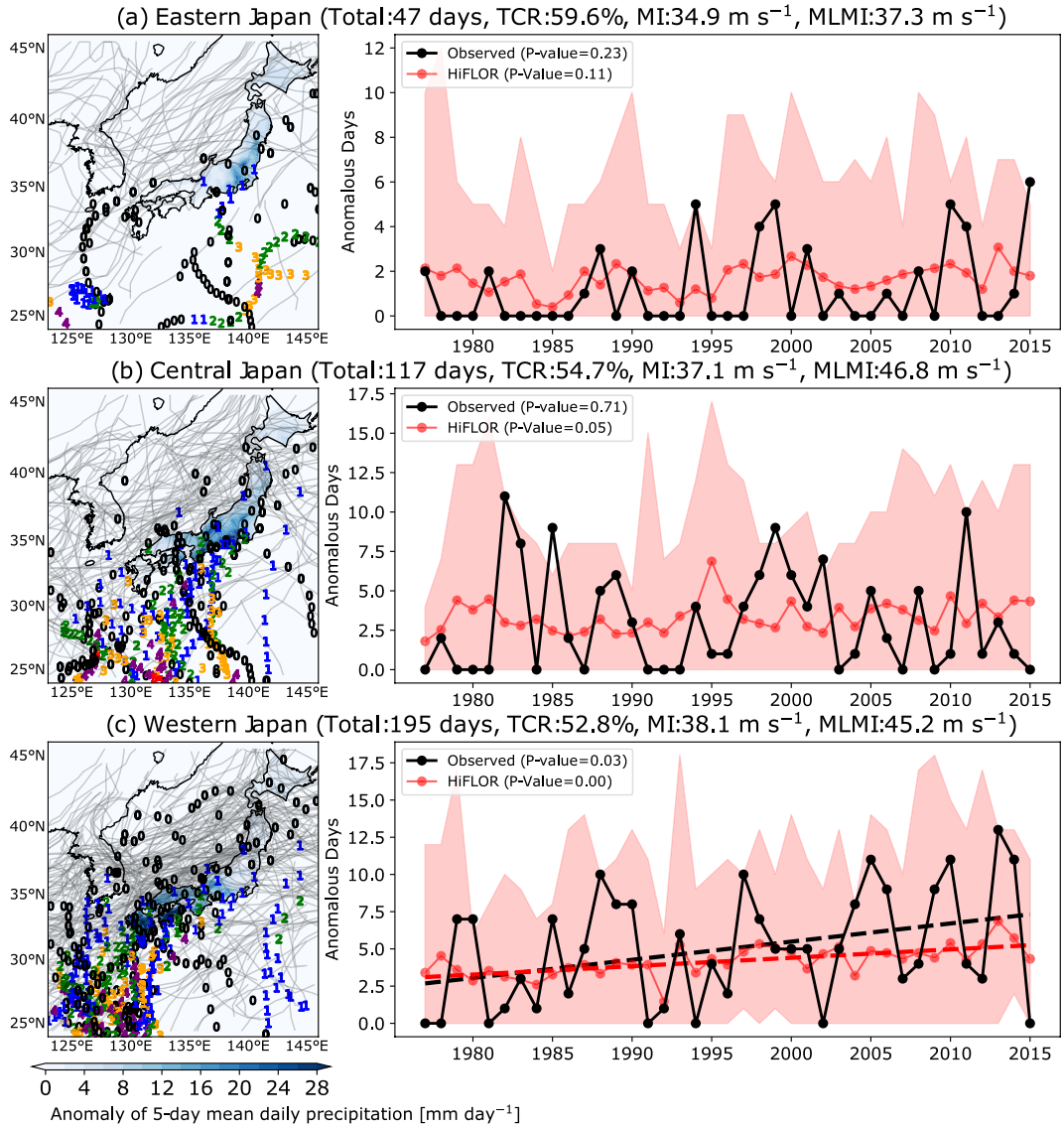

Figure 6. Three patterns for anomalous precipitation events. The observed anomalous events are assigned to one of the three subregions based on the location of maximum precipitation. The left and right panels are the same as Figures 3a and 4a, respectively, but for the results of each subregion. Numbers in each subtitle denote the total frequency of anomalous events (Total, days), the percentage of TC days relative to total days (TCR, %), mean maximum intensity of TCs (MI, m s^{-1}), and mean lifetime maximum intensity of TCs (MLMI, m s^{-1}) for these TC days during the anomalous precipitation events. The same assignment was applied to the anomalous events derived from the historical simulations by HiFLOR, resulting in similar trends to observations (red lines, Figure S5 in Supporting Information S1).

of the simulated trends and spatial pattern are somewhat inconsistent with observations (note that the color scale is different in Figures 9a and 9b). In terms of the changes in the frequency of heavy precipitation events, a change in intense TCs should be an important factor. HiFLOR is a valuable high-resolution global coupled climate model capable of simulating intense storms, such as major hurricanes (Murakami et al., 2015, 2016; Zhang et al., 2016), and is therefore used to quantify changes in intense storms. Figures 9c and 9d are the same as Figures 9a and 9b, except for intense storms with maximum wind speed greater than 49 m s^{-1} that are equivalent to major hurricanes according to Saffir-Simpson's intensity scale (NOAA, 2021). AllForc HiFLOR reproduces the spatial distribution of the observed trends, showing overall increases in intense storms over the whole of the ocean basin, although NatForc HiFLOR does not show any significant trend (figure not shown). Therefore, these consistent spatial patterns between observations and AllForc HiFLOR suggest that increases in anthropogenic

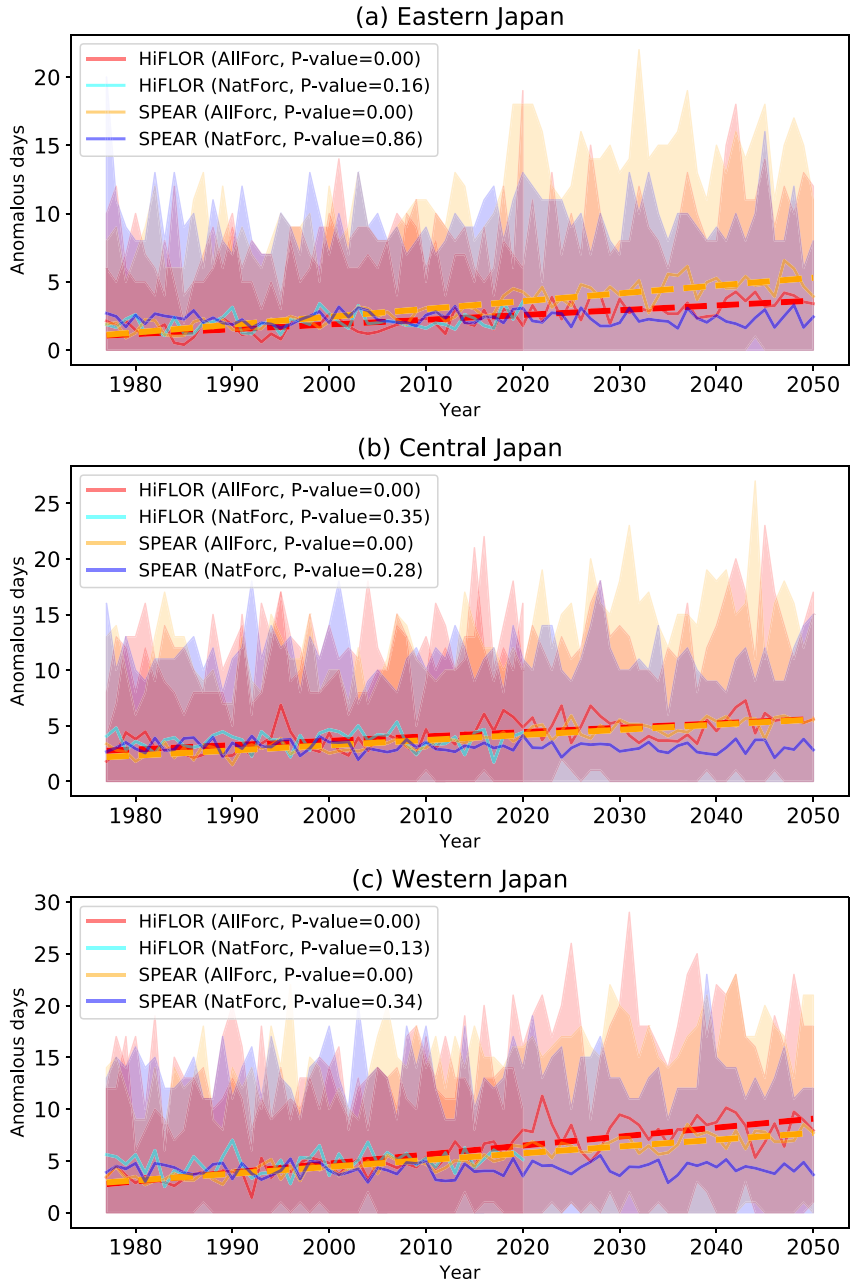

Figure 7. Time series of the frequency of extreme events for each subregion for the period including the future (1977–2050). As in Figure 4a, but for the AllForc future projections by HiFLOR (red) and SPEAR (orange), and NatForc projections by SPEAR (blue) for individual subregions (a–c).

forcing played an important role in the observed increase in TC frequency, specifically intense TCs in the western quadrant of the Pacific, and in turn, this may lead to significant increases in anomalous events in Western Japan. Moreover, Figures 9e and 9f show the linear trends between 1977 and 2050 projected by AllForc HiFLOR. Both figures show a similar spatial pattern of the trends between 1977 and 2015, highlighting continuing increases in the frequency of occurrence of TCs and intense TCs over southwest Japan in the future. This would lead to more frequent anomalous events in this region in the future.

Another reason for the increasing trend in anomalous events in Western Japan could be related to changes in precipitation associated with QSFs. The rainy or the rainband associated with QSFs over Japan from June to July

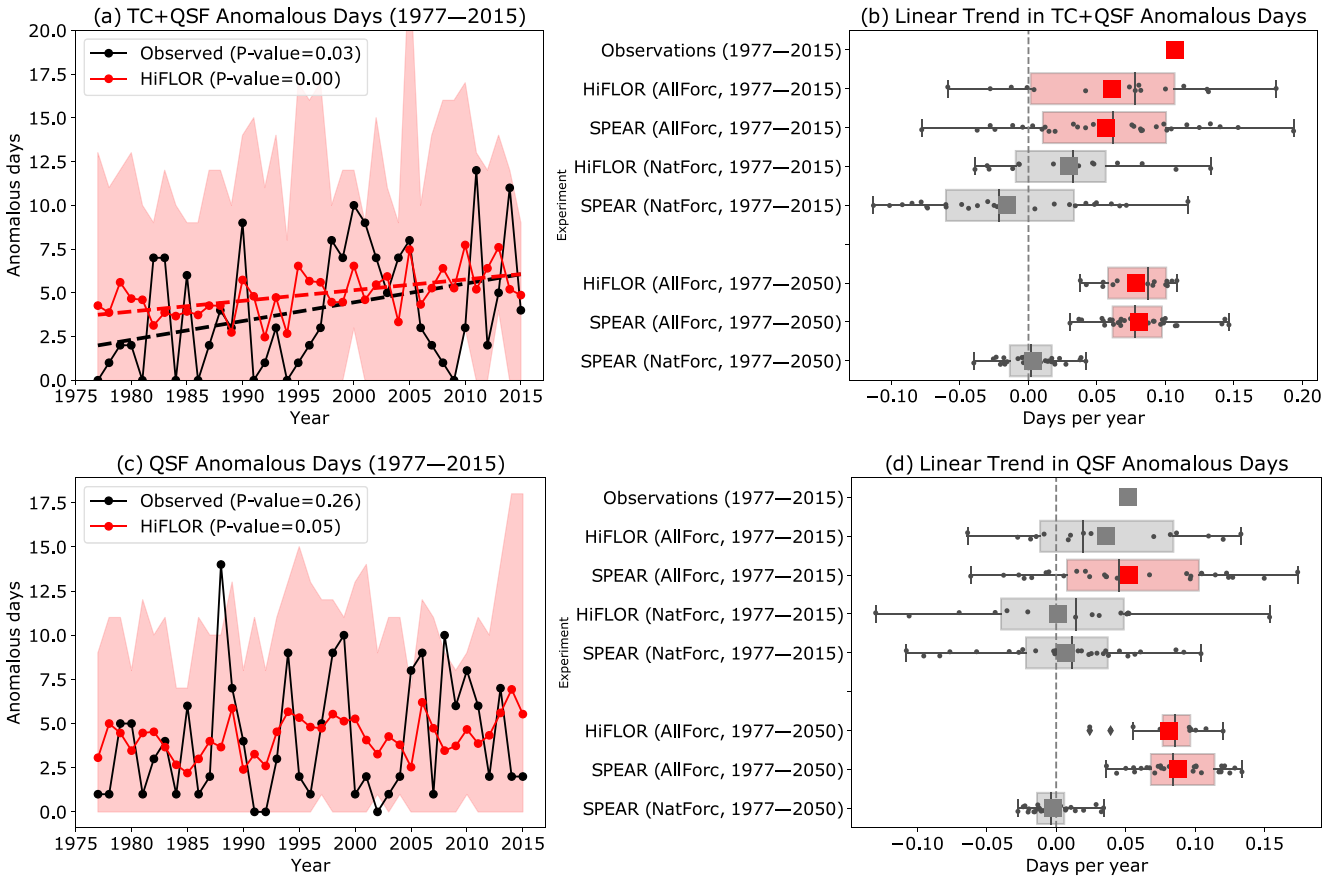

Figure 8. Trends in observed and simulated frequency of TC + QSF and QSF anomalous precipitation events. Anomalous events were grouped into events with TCs and quasi-stationary fronts (TC + QSF) or these with quasi-stationary front (QSF) depending on whether TCs and QSF simultaneously appeared or only QSF appeared over the domain of Figure 1 in the 5 days before the anomalous event. (a, b) As in Figures 4a and 4b, but for the TC + QSF events. (c, d) As in (a) and (b), but for the QSF events.

is called the Baiu. Figure 10 shows the observed and simulated seasonal evolution of precipitation for the longitudinal sector including Western Japan. Observations show a climatological northward migration of the Baiu rainband from June and July (shading in Figure 10a). There are significant positive trends in observed precipitation from June to early July around 28°–36°N that including Western Japan (black contours in Figure 10a). Although the same figure by AllForc HiFLOR somewhat shows inconsistencies in terms of the statistical significance relative to observations, AllForc HiFLOR qualitatively agrees with the observed trends, showing increasing precipitation in the same latitudes associated with a more active Baiu rainband (Figure 10b). The significant trends simulated by AllForc HiFLOR are clearer for the future projection (Figure 10c). Previous studies on future projections are also consistent with this finding, showing projected increases in the mean precipitation around Western Japan from June to the middle of August that represents a delay in the termination of the Baiu season (Kusunoki et al., 2011).

4. Conclusions and Discussions

We have developed a new deep learning method—autoencoder—to objectively detect anomalous precipitation events from observations and model outputs. The detected anomalous events are mostly associated with extremely intense precipitation events that are concurrent with the existence of TCs and QSFs near Japan. There is a significant positive trend in the frequency of anomalous events in observations over the period 1977–2015, and the trend is significant specifically in Western Japan. We found that our chosen climate models, HiFLOR and SPEAR, faithfully reproduced the observed increasing trends in anomalous precipitation events in Japan,

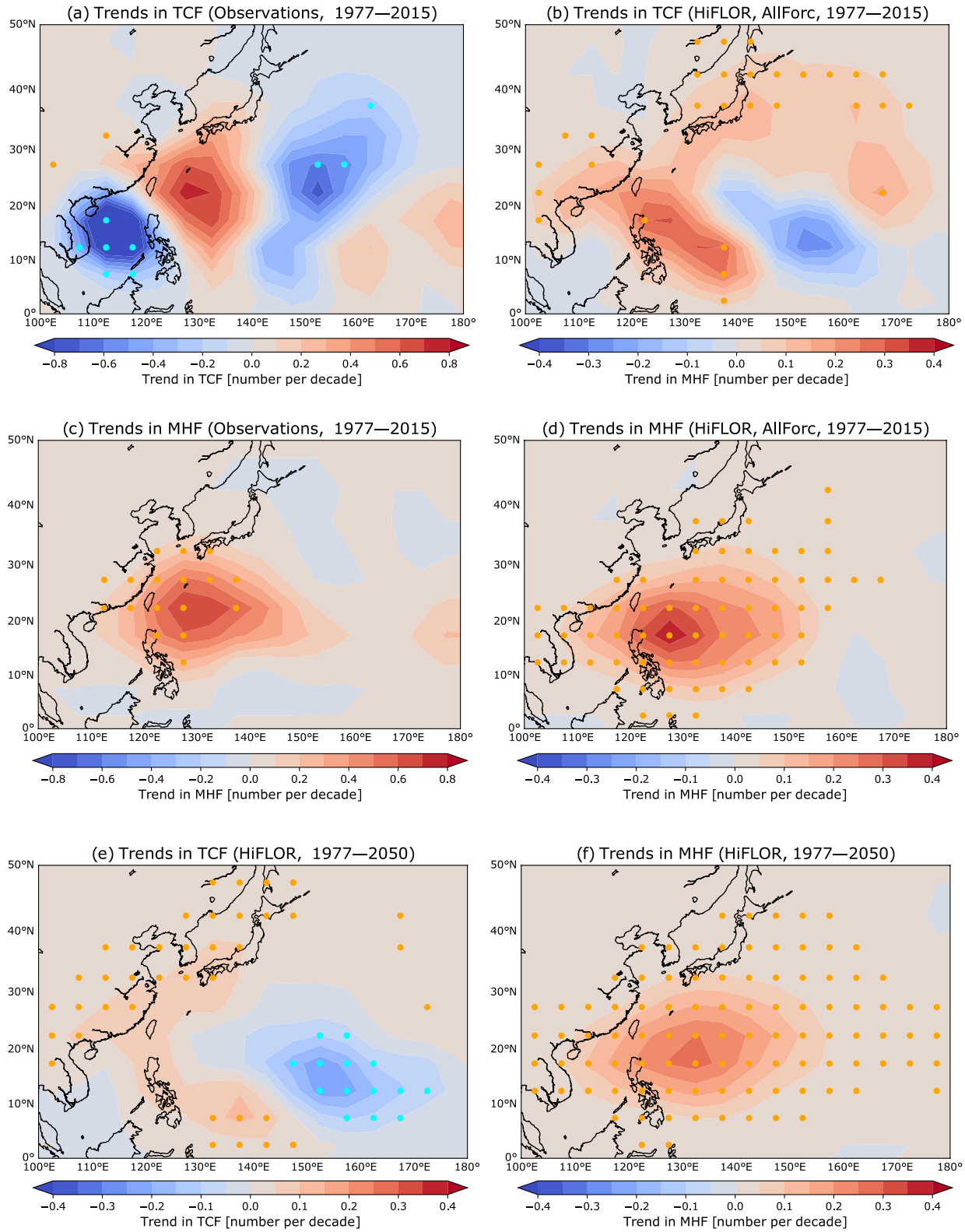

Figure 9. Linear trends in observed and simulated TC density. TC density is defined as the total count of 6-hourly TC tracks in each 5° × 5° grid box. (a) Observed linear trends in TC density (≥17.5 m s^{-1}) over the period 1977–2015. (b) As in (a), but for trends simulated by AllForc experiments by HiFLOR. (c, d) As in (a) and (b), but for intense TCs (≥49 m s^{-1}). (e, f) As in (b) and (d), but for the trends from 1977 to 2050. Red (blue) dots indicate the positive (negative) trends that are statistically significant at the 95% level according to the Mann–Kendall test. Note that the color scale is different between observations and HiFLOR.

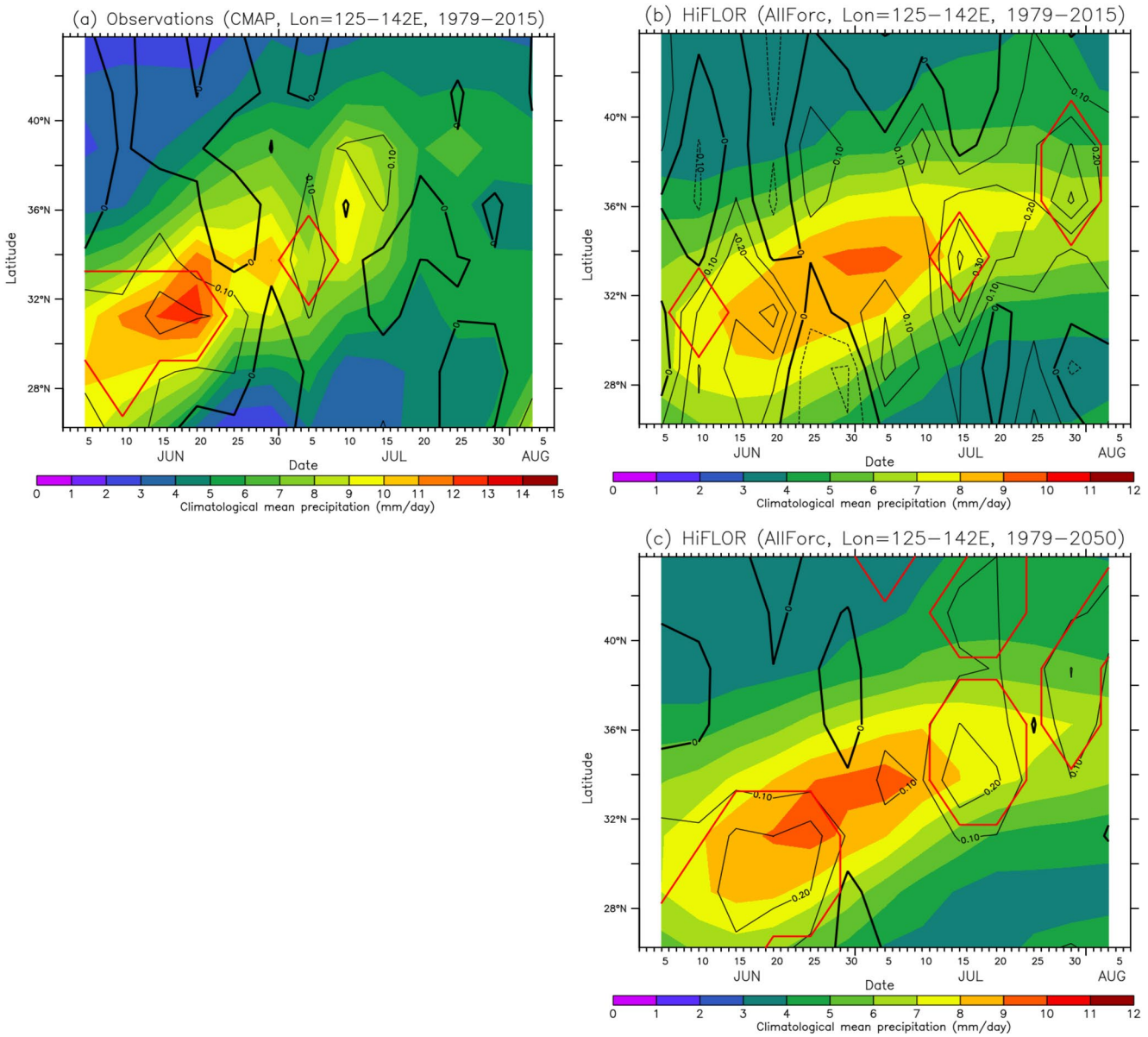

Figure 10. Time–latitude cross section of the precipitation averaged for 125°–142°E. The period is from pentad 32 (5–11 June) to 44 (4–9 August). (a) Observations by CPC Merged Analysis of Precipitation (CMAP, 2.5°, 1979–2015). Shadings show climatological mean values (mm day^{-1}), whereas black contours show linear trends over the period 1977–2015 (mm day^{-1} per year). Red contours show statistically significant trends at the 95% level according to the Mann–Kendall test. (b) As in (a), but for the ensemble mean of AllForc HiFLOR experiments over the same period. (c) As in (b), but for the future period of 1977–2050. CMAP is available online at https://psl.noaa.gov/data/gridded/data.cmap.html.

specifically Western Japan. In contrast, the idealized experiments by HiFLOR and SPEAR in which anthropogenic forcing was not increased over the period 1977–2015 did not show any significant trends in the frequency of anomalous events, suggesting a substantial influence of anthropogenic forcing on the increase in anomalous precipitation events beyond the influence of internal variability. These increases have been caused by the increased frequency of occurrence of intense TCs along with the enhanced activity of QSFs near Japan. The climate models also projected continuing increases in the frequency of occurrence of intense TCs near Japan and enhancement in QSF activities in the future due to the anticipated anthropogenic climate changes. They will in turn lead to increases in anomalous precipitation events in Japan in the future.

We found that deep learning is a useful technique to objectively screen anomalous events from observed and climate modeling data sets. However, a caveat worth noting is that anomaly detection using an autoencoder is a work in progress, and far from a fully solved area of machine learning (Chalapathy & Chawla, 2019). Many other methods can be used for anomaly detection, depending on the nature of the data (Chalapathy & Chawla, 2019). For instance, convolution networks, such as the autoencoder developed in this study, are preferred for image data sets, while "long short-term memory" (LSTM)–based models tend to produce good results for sequential data. Although our preliminary results suggest that consistent results can be obtained using a spatiotemporal LSTM (Chong & Tay, 2017; Figures S6 and S7 in Supporting Information S1), it is necessary to verify the differences between the methods. As indicated in Section 1, an autoencoder can objectively identify anomalous events. On the other hand, some anomalous cases identified by an autoencoder are sometimes difficult to perceive with our own eyes. Therefore, further development of deep learning, its verification, and comparisons with different methods are important future considerations for identifying anomalous events in the field of climate science.

Note that we interpolated the original precipitation data onto the different grid spacing between APHRODITE and HiFLOR ($0.25° \times 0.25°$) and SPEAR ($0.5° \times 0.5°$) to separately develop autoencoders. This special treatment for SPEAR was considered because the horizontal resolution of SPEAR is coarser than 0.25°. To investigate the dependence of the results on grid spacing, we repeat the same analysis using APHRODITE except that precipitation anomalies were interpolated onto $0.5° \times 0.5°$. We found that the 0.5° resolution showed consistent results with the 0.25° resolution in terms of the spatial structure of anomalous events and trends in the frequency of anomalous events (Figures S8 and S9 in Supporting Information S1). We speculate that taking a 5-day running mean of the precipitation anomaly can smooth out fine-scale precipitation structure well so that the impact of the resolution is minimal at least in this study.

Anthropogenic forcing includes various forcings such as greenhouse gases, anthropogenic aerosols, and ozone. In the future, it would be desirable to estimate the impacts of individual forcings on anomalous events by carrying out additional large-ensemble simulations forced with time-varying single forcings only. Further enhancements in supercomputing would enable us to conduct such expensive simulations to further understand the physical mechanisms of climate changes in anomalous events.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The Asian Precipitation—Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) project (version APHRO_JP_V1207 and V1207_R3) is available at http://aphrodite.st.hirosaki-u.ac.jp/download/. The Japanese 55-year Reanalysis (JRA-55) is available at http://search.diasjp.net/en/dataset/JRA55. The International Best Track Archive for Climate Stewardship (IBTrACS), version 4, is available at https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access. The large-ensemble simulations by SPEAR are available at https://www.gfdl.noaa.gov/spear_large_ensembles/. The model source code for HiFLOR is available at https://www.gfdl.noaa.gov/cm2-5-and-flor/. The data sets, codes, and autoencoder generated during the current study are available at https://doi.org/10.7910/DVN/TWZR75. These uploaded files are freely available.

Acknowledgments
The authors thank Dr V. Balaji and Dr Kai-Chih Tseng for their suggestions and comments. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration or the U.S. Department of Commerce. P.-C. Hsu was supported by the National Natural Science Foundation of China (42088101).

References

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., et al. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*, 3167–3211. https://doi.org/10.1029/2019MS001726

Alla, S., & Adari, S. K. (2019). *Beginning anomaly detection using Python-based deep learning: With Keras and PyTorch*, (p. 146). New York: Apress. https://doi.org/10.1007/978-1-4842-5177-5

Berg, P., & Haerter, J. O. (2013). Unexpected increase in precipitation intensity with temperature. A result of mixing of precipitation types? *Atmospheric Research*, *119*, 56–61. https://doi.org/10.1016/j.atmosres.2011.05.012

Bhatia, K. T., Vecchi, G. A., Knutson, T., Murakami, H., Kossin, J., Dixon, K., & Whitlock, C. (2019). Recent increases in tropical cyclone intensification rates. *Nature Communications*, *10*, 635. https://doi.org/10.1038/s41467-019-08471-z

Cabinet Office Japan. (2020). *Online available in Japanese*. Retrieved from http://www.bousai.go.jp/updates/r2_07ooame/pdf/r20703_ooame_25.pdf

Chalapathy, R., & Chawla, S. (2019). *Deep learning for anomaly detection: A survey*. arXiv reprint, arXiv:1901.03407.

Chong, Y. S., & Tay, Y. H. (2017). Abnormal event detection in videos using spatiotemporal autoencoder. In *International Symposium on Neural Networks* (pp. 189–196). Cham: Springer. Retrieved from https://arxiv.org/abs/1701.01546

Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J.-H., Dunne, K. A., et al. (2020). SPEAR: The next generation GFDL modeling system for seasonal to multidecadal prediction and projection. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001895. https://doi.org/10.1029/2019MS001895

Emori, S., & Brown, S. J. (2005). Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophysical Research Letters*, *32*, L17706. https://doi.org/10.1029/2005GL023272

Folland, C. K., Renwick, J. A., Salinger, M. J., & Mullan, A. B. (2002). Relative influences of the Interdecadal Pacific Oscillation and ENSO on the South Pacific Convergence Zone. *Geophysical Research Letters*, *29*(13), 1643. https://doi.org/10.1029/2001GL014201

Fujibe, F. (2013). Clausius–Clapeyron-like relationship in multidecadal changes of extreme short-term precipitation and temperature in Japan. *Atmospheric Science Letters*, *14*, 127–131. https://doi.org/10.1002/asl2.428

Harris, L. M., Lin, S.-J., & Tu, C. Y. (2016). High resolution climate simulations using GFDL HiRAM with a stretched global grid. *Journal of Climate*, *29*, 4293–4314. https://doi.org/10.1175/JCLI-D-15-0389.1

Hatsuzuka, D., & Sato, T. (2019). Monthly extreme precipitation in Japan using large-ensemble regional climate simulations. *Journal of Hydrometeorology*, *20*, 563–574.

Hatsuzuka, D., Sato, T., Yoshida, K., Ishii, M., & Mizuta, R. (2020). Regional projection of tropical-cyclone-induced extreme precipitation around Japan based on large ensemble simulations. *SOLA*, *16*, 23–29. https://doi.org/10.2151/sola.2020-005

IPCC. (2013). *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 1535). Cambridge: Cambridge University Press.

JMA. (2019). *Climate Change Monitoring Report 2019*. Retrieved from https://www.jma.go.jp/jma/en/NMHS/ccmr/ccmr2019.pdf

Kamiguchi, K., Arakawa, O., Kitoh, A., Yatagai, A., Hamada, A., & Yasutomi, N. (2010). Development of APHRO_JP, the first Japanese high-resolution daily precipitation product for more than 100 years. *Hydrological Research Letters*, *4*, 60–64. https://doi.org/10.3178/hrl.4.60

Kawano, T., & Kawamura, R. (2020). Genesis and maintenance processes of a quasi-stationary convective band that produced record-breaking precipitation in northern Kyushu, Japan on 5 July 2017. *Journal of the Meteorological Society of Japan*, *98*, 673–690.

Kawase, H., Imada, Y., Sasaki, H., Nakaegawa, T., Murata, A., Nosaka, M., & Takayabu, I. (2019). Contribution of historical global warming to local-scale heavy precipitation in Western Japan estimated by large ensemble high-resolution simulations. *Journal of Geophysical Research: Atmospheres*, *124*, 6093–6103. https://doi.org/10.1029/2018JD030155

Kharin, V., & Zwiers, F. W. (2000). Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *Journal of Climate*, *13*, 3760–3787.

Kikuchi, K., Wang, B., & Kajikawa, Y. (2012). Bimodal representation of the tropical intraseasonal oscillation. *Climate Dynamics*, *38*, 1989–2000. https://doi.org/10.1007/s00382-011-1159-1

Knapp, K., Kruk, M. C., Levinson, D. H., Diamond, H. J., & Neuman, C. J. (2010). The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, *91*, 363–376. https://doi.org/10.1175/2009BAMS2755.1

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*, 5–48. https://doi.org/10.2151/jmsj.2015-001

Kriegler, E., Bauer, N., Popp, A., Humpenöder, F., Leimbach, M., & Strefler, J. (2017). Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century. *Global Environment Change*, *42*, 297–315. https://doi.org/10.1016/j.gloenvcha.2016.05.015

Kusunoki, S., Mizuta, R., & Matsueda, M. (2011). Future changes in the East Asian rain band projected by global atmospheric models with 20-km and 60-km grid size. *Climate Dynamics*, *37*, 2481–2493.

Liu, Y., Racah, E., Prabhat, Correa, J., Khosrowshahi, A., Lavers, D., et al. (2016). Application of deep convolutional neural networks for detecting extreme weather in climate datasets. In *ABDA'16-International Conference on Advances in Big Data Analytics* (pp. 81–88). Retrieved from https://arxiv.org/abs/1605.01156

Min, S., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more intense precipitation extremes. *Nature*, *470*, 378–381. https://doi.org/10.1038/nature09763

Murakami, H., Delworth, T. L., Cooke, W. F., Zhao, M., Xiang, B., & Hsu, P.-C. (2020). Detected climatic change in global distribution of tropical cyclones. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(20), 10706–10714.

Murakami, H., Vecchi, G. A., Underwood, S., Delworth, D. L., Wittenberg, A. T., Anderson, W. G., et al. (2015). Simulation and prediction of Category 4 and 5 hurricanes in the high-resolution GFDL HiFLOR coupled climate model. *Journal of Climate*, *28*, 9058–9079. https://doi.org/10.1175/JCLI-D-15-0216.1

Murakami, H., Vecchi, G. A., Villarini, G., Delworth, D. L., Gudgel, R., Underwood, S., et al. (2016). Seasonal forecasts of major hurricanes and landfalling tropical cyclones using a high-resolution GFDL coupled climate model. *Journal of Climate*, *29*, 7977–7989.

Neena, J. M., Jiang, X., Waliser, D., Lee, J.-Y., & Wang, B. (2014). Eastern Pacific intraseasonal variability: A predictability perspective. *Journal of Climate*, *27*, 8869–8883.

NOAA. (2021). The Saffir-Simpson hurricane wind scale. Retrieved from https://www.nhc.noaa.gov/pdf/sshws.pdf

Ohba, M., Kadokura, S., Yoshida, Y., Nohara, D., & Toyoda, Y. (2015). Anomalous weather patterns in relation to heavy precipitation events in Japan during the Baiu season. *Journal of Climate*, *16*, 688–701.

Racah, E., Beckham, C., Maharaj, T., Kahou, S. E., Prabhat., & Pal, C. (2017). Extreme weather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. *Advances in Neural Information Processing Systems*, *30*, 3405–3416.

Reichstein, M., Campps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhals, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*, 195–204.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., & Fujimori, S. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Schemm, S., Rudeva, I., & Simmonds, I. (2015). Extratropical fronts in the lower troposphere–global perspectives obtained from two automated methods. *Quarterly Journal of the Royal Meteorological Society*, *141*, 1686–1698.

Shaw, S. B., Royem, A. A., & Riha, S. J. (2011). The relationship between extreme hourly precipitation and surface temperature in different hydroclimatic regions of the United States. *Journal of Climate*, *12*, 319–329.

Tsuji, H., Yokoyama, C., & Takayabu, Y. N. (2020). Contrasting features of the July 2018 heavy rainfall event and the 2017 Northern Kyushu rainfall event in Japan. *Journal of the Meteorological Society of Japan*, *98*, 859–876.

Tuan, B. M. (2019). Extratropical forcing of submonthly variations of rainfall in Vietnam. *Journal of Climate*, *32*, 2329–2348.

Utsumi, N., Kim, H., Kanae, S., & Oki, T. (2017). Relative contributions of weather systems to mean and extreme global precipitation. *Journal of Geophysical Research: Atmospheres*, *122*, 152–167. https://doi.org/10.1002/2016JD025222

van der Wiel, K., Kapnick, S. B., Vecchi, G. A., Cooke, W. F., Delworth, T. L., Jia, L., et al. (2016). The resolution dependence of contiguous US precipitation extremes in response to CO_2 forcing. *Journal of Climate*, *29*, 7991–8012.

Yu, R., & Li, J. (2012). Hourly rainfall changes in response to surface air temperature over eastern contiguous China. *Journal of Climate*, *25*, 6851–6861. https://doi.org/10.1175/JCLI-D-11-00656.1

Zhang, W., Vecchi, G. A., Murakami, H., Delworth, T., Wittenberg, A. T., Rosati, A., et al. (2016). Improved simulation of tropical cyclone responses to ENSO in the western North Pacific in the high-resolution GFDL HiFLOR coupled climate model. *Journal of Climate*, *29*(4), 1391–1415.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018a). The GFDL global atmosphere and land model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*, 691–734. https://doi.org/10.1002/2017MS001208

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018b). The GFDL global atmosphere and land model AM4.0/LM4.0: 2. Model description, sensitivity studies, and tuning strategies. *Journal of Advances in Modeling Earth Systems*, *10*, 735–769. https://doi.org/10.1002/2017MS001209

Zhao, W. Z., & Du, S. H. (2016). Learning multiscale and deep representations for classifying remotely sensed imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *113*, 155–165.