scientific reports

Check for updates

OPEN

Deep generative model super-resolves spatially correlated multiregional climate data

Norihiro Oyama[1]✉, Noriko N. Ishizaki[2], Satoshi Koide[1] & Hiroaki Yoshida[1]

Super-resolving the coarse outputs of global climate simulations, termed downscaling, is crucial in making political and social decisions on systems requiring long-term climate change projections. Existing fast super-resolution techniques, however, have yet to preserve the spatially correlated nature of climatological data, which is particularly important when we address systems with spatial expanse, such as the development of transportation infrastructure. Herein, we show an adversarial network-based machine learning enables us to correctly reconstruct the inter-regional spatial correlations in downscaling with high magnification of up to 50 while maintaining pixel-wise statistical consistency. Direct comparison with the measured meteorological data of temperature and precipitation distributions reveals that integrating climatologically important physical information improves the downscaling performance, which prompts us to call this approach πSRGAN (Physics Informed Super-Resolution Generative Adversarial Network). The proposed method has a potential application to the inter-regionally consistent assessment of the climate change impact. Additionally, we present the outcomes of another variant of the deep generative model-based downscaling approach in which the low-resolution precipitation field is substituted with the pressure field, referred to as ψSRGAN (Precipitation Source Inaccessible SRGAN). Remarkably, this method demonstrates unexpectedly good downscaling performance for the precipitation field.

The increase of greenhouse gases in the air composition due to human activities is now believed to have led to the rise in the frequency of unusual disasters[1–4]. To prevent an irreversible collapse of the current ecosystem and resulting impoverishment of human lives, many countries have set specific medium- and long-term goals for the reduction of greenhouse gas emissions, and similar paradigm shifts in decision making have occurred even at the private sector level.

Numerical approaches are regarded as the most powerful and reliable scientific option at the moment in quantitatively evaluating the efficacy of political or management plans that aim to tackle climatological issues. The Global Climate Model (GCM) is the prime example, which has accurately reproduced past and current climate changes, and its reliability of quantitative future estimates is sufficiently high[5]. Such future projections with high accuracies rely on the overall consideration of the global atmospheric and oceanic circulation (and even still more complicated ingredients such as chemical[6] and biological[7] processes)[8–12], and thus, the horizontal spatial resolution is sacrificed by the required computational costs; the typical resolution of the GCMs is only down to the order of 1° in longitude and latitude, corresponding to a grid size of more than a hundred kilometers on the equator. Therefore, to exploit the GCM outputs to assess the impact of climate change and to make proper decisions, it is obviously vital to super-resolve the coarse grid spacing of simulations and to reach the fine resolution of interest. Here, special attention should be given to reproducing the inherent spatial correlation of the meteorological variables, as well as the local statistics, in decision making by integrating multiregional information[13–17], such as transportation infrastructure development and sustainable energy networks, future urbanization, and agricultural intensification.

A variety of techniques to super-resolve GCM outcomes, which are referred to as the *downscaling* (DS) methods in meteorology and climatology, have been developed[18–25]. They are categorized roughly into two groups: dynamical[18–21] and statistical DS methods[22–25]. The dynamical downscaling method is based on physical footings: several coupled differential equations are numerically integrated with the results of the GCM (or any other crude-resolution simulation results) being used as the boundary conditions. However, the computational cost again creates a trade-off between the accuracy and the feasibility. In contrast, in the statistical approaches, we

[1]Toyota Central R &D Labs, Inc., Bunkyo-ku, Tokyo 112-0004, Japan. [2]Center for Climate Change Adaptation, National Institute for Environmental Studies, Tsukuba 305-8506, Japan. ✉email: Norihiro.Oyama.vb@mosk.tytlabs.co.jp

turn a blind eye to the physical laws behind the data. Instead, empirical links between the large- and local-scale climates are identified and applied to the crude-resolution climate model outputs. Since the systematic errors of the naively interpolated GCM output (referred to as the bias) are locally corrected such that the statistical properties are precisely reproduced, the spatial correlation, i.e., the information on the events occurring at distant places, is discarded[26–28]. The statistical downscaling methods overcoming the latter problem remain to be developed.

In this paper, we propose a machine learning approach that super-resolves the GCM outputs and reproduces both the local statistics and the instantaneous spatial correlations between distant regions. Among several options for improving the resolution of geophysical or climatological data[29–32], our method is based on the generative adversarial network (GAN) approach, which has been proven to be a very powerful downscaling tool through several previous studies[33,34]. To accurately reproduce the physical nature, we use auxiliary but climatologically important data, sea-level pressure distribution and topographic information, in addition to the target variables, temperature and precipitation distributions (see Fig. 1A and the next section for more details). Since this method falls within the criteria of the first-level physics informed super-resolution methods[35], we name our method π SRGAN (Physics Informed Super-Resolution Generative Adversarial Network). The direct comparison with the measured meteorological data shows that the local statistical properties are obtained using the practical output from the GCM simulations as accurately as the conventional statistical downscaling method that is focused on matching these properties. We then highlight that the spatial correlation of variables is accurately reproduced, which could not be achieved with conventional downscaling methods (see Fig. 1B). The present method is therefore the next generation downscaling method that has a potential application in climate change assessment considering both local-scale and interregional events. We also considered another variant of the SRGAN that projects the high-resolution temperature and precipitation field from the low-resolution information about only temperature and pressure (we call this variant the ψSRGAN: Precipitation-Source-Inaccessible SRGAN). With this special variant, we demonstrate the surprisingly robust ability of the SRGAN-based methods to express natural results.
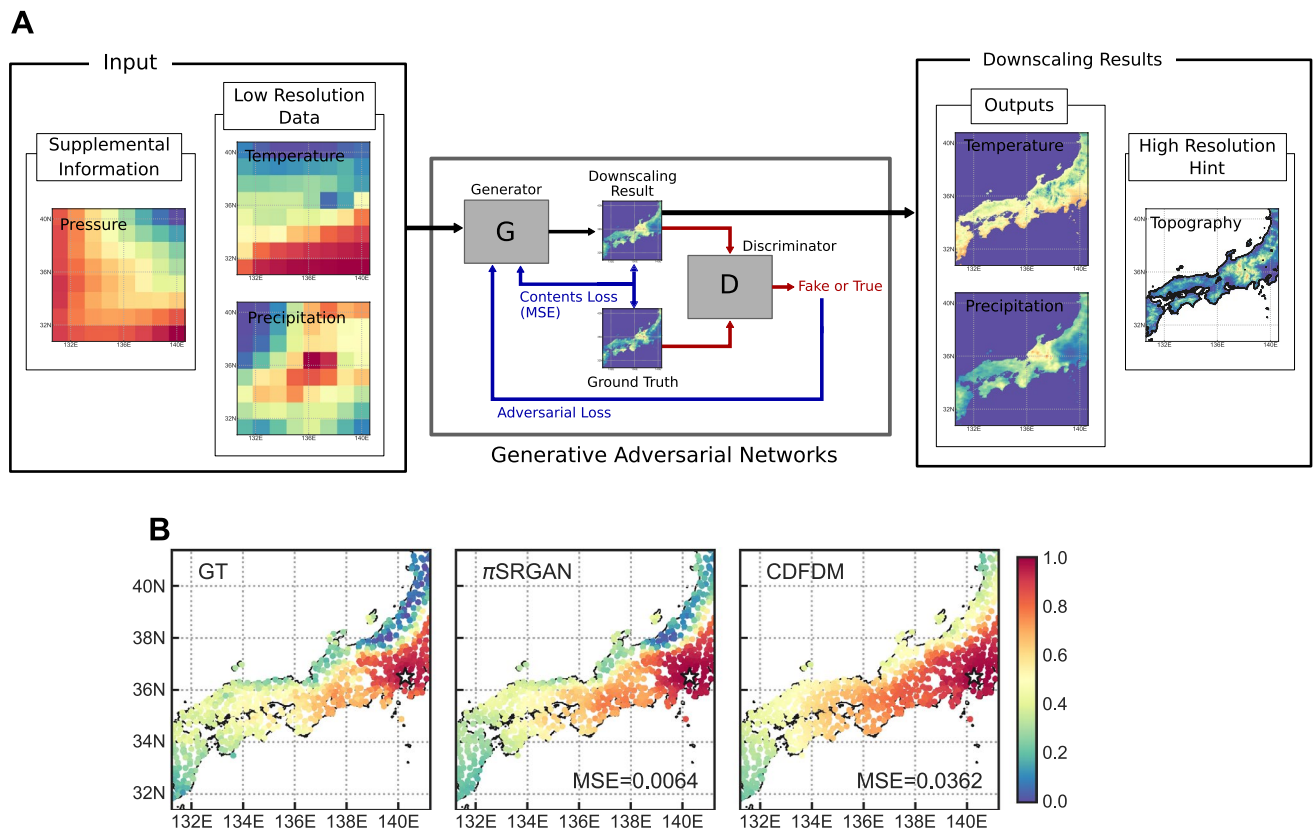

Figure 1. Schematic diagram of πSRGAN and the distribution of spatial correlation coefficients. (**A**) High-resolution topography and low-resolution sea level pressure, in addition to the low resolution data corresponding to the output, are supplied to the generative adversarial networks. (**B**) The reconstructed distributions of spatial correlation coefficients, indicating the correlation strength from the reference site at Tokyo [35.735° N, 139.6683° E], obtained with πSRGAN and a conventional CDFDM are compared to the ground truth (GT).

Results

Super-Resolution Generative Adversarial Networks with various data.

We employ a super-resolution method based on generative adversarial networks (Super-Resolution Generative Adversarial Networks: SRGAN, see Methods section for details) as the basic machine learning algorithm, which was proven to have potential in DS with a scale factor up to 50[33]. Although the original SRGAN was able to restore physical consistency in the turbulent wind velocity field, which was shown in terms of the well-known Kolmogorov 5/3 power-law[39], it was also reported that it showed a worse performance in reproducing the basic statistics, such as the pixelwise consistency like mean squared error, than a less sophisticated deep learning approach[33]. In this work, considering two distinct variants in addition to the standard SRGAN, we show that the integration of the low-resolution input with auxiliary information enables to overcome the drawback of relatively poor reproducibility of simple statistical properties and that the ability of SRGAN-based methods to downscale in a "physically natural" manner is quite robust against the change in the input low-resolution information.

There are a vast variety of LR information, as seen in several similar recent attempts[34,40–42]. Among them, we employed the sea-level pressure, one of the fundamental hydrodynamic (or aerodynamic) variables on which the various quantities of sub-models of GCMs are based, as a piece of key auxiliary information. Also, this variable is described with fewer assumptions in the models than other meteorological variables such as humidity. In the literature, strong links between synoptic-scale horizontal circulation and vertical motion are discussed in terms of the sea-level pressure[22,43–45]. In the first variant, we incorporate the low-resolution pressure field as an auxiliary physical information (Fig. 1A), which serves as guidance for the DS of the target variables, namely temperature and precipitation. In this method, moreover, we introduced the topographic information as another auxiliary information since it can be utilized in a high-resolution format only if we assume it is identical over the time window of interest (order of 10 to a 100 years). The topographic information is indirectly supplied as a part of teacher data during the training by adding to one of the output channels. In this way, we can provide both low-resolution and high-resolution auxiliary data in an unambiguous manner without any artificial operation (like resolution matching by interpolation or pooling). Since the use of supplemental physical information during learning is regarded as primary-level physics-informed machine learning[35], we call this method the Physics-Informed SRGAN (πSRGAN for short).

The second variant of SRGAN is designed to generate high-resolution temperature and precipitation fields using solely low-resolution data pertaining to the temperature and pressure fields. This variant is referred to as the Precipitation-Source-Inaccessible SRGAN (ψSRGAN) and demonstrates the surprisingly robust capability of SRGAN-based methods to describe "physically natural" precipitation fields.

The performances of three variants of SRGAN (standard SRGAN, ψSRGAN, and πSRGAN) are evaluated via direct comparisons among them and with a non-machine learning-based method: we summarize these methods in Table 1. The cumulative distribution function-based downscaling method (CDFDM) is the widely used conventional statistical DS method (see the method section for the details), and the SRGAN refers to the original SRGAN-based method presented in Ref.[33].

Data sets.

We use the climate model simulation outputs for the low-resolution input and the real observation data for the high-resolution ground truths in the case studies. As the low-resolution data, we used the Japanese 55-year reanalysis (JRA-55) data from 1980 to 2018[37] with data assimilation. The grid spacing is 1.25°. The daily data corresponding to the reference data (in Japanese local time) were created from 3-h simulation data. Specifically, data at 0Z, 3Z, 6Z, 9Z and 12Z on the target date and data at 15Z, 18Z and 21Z on the previous day of the target date were averaged to obtain the daily data in JST. The reference high-resolution data were the Agro-Meteorological Grid Square Data (AMGSD)[36]. The 1 km-meshed daily data over Japan are constructed using the in-situ observation network system of the Japan Meteorological Agency, which covers the entire land area over Japan from 122° to 146° east and from 24° to 46° north. Upon being fed into the networks, all of the data undergo a process of normalization and concatenation. For further information regarding the technical aspects of these procedures, please refer to the SI Appendix.

We use the data from 1980 to 2018 (14,245 days in total). These data are split into training, validation, and test datasets in a time-series manner as summarized in Table 2 for both low-resolution (JRA-55) and high-resolution (AMGSD) data. We emphasize that this time-series partitioning, characterized by a substantial volume of test data, represents a challenging task for downscaling mid-term future projections, and consequently, necessitates the incorporation of the climate change trend. The AMGSD data were adjusted such that the grid spacing was

Abbr. name	Explanation	Low-resolution data	High-resolution data
GT	Ground truth (observation results offered by AMGSD[36])	–	–
LR	Low resolution data of JRA-55[37]	–	–
SRGAN	Standard SRGAN-based method[33]	PRC, TMP	PRC, TMP
πSRGAN	Physics-informed SRGAN	PRC, TMP, SLP	PRC, TMP, TOPO†
ψSRGAN	Precipitation-source-inaccessible SRGAN	TMP, SLP	PRC, TMP
CDFDM	Cumulative distribution function-based downscaling method[38]	PRC, TMP	PRC, TMP

Table 1. Summary of protocols compared in this study. *PRC* precipitation, *TMP* temperature, *SLP* sea level pressure, *TOPO* topography. †In πSRGAN, TOPO is used as a hint, not targeted.

Training data	Validation data	Test data
1980–1997	1998–2000	2001–2018

Table 2. Year span for each data set.

0.025°/grid both in latitude and longitude. We extracted the data for the region from 130.625° to 140.625° east and from 30.625° to 40.625° north, which results in a 400 × 400 pixels square. The JRA-55 data of the corresponding region are 8 × 8 pixel squares, and thus, the scale factor for the DS tasks is 50.

Qualitative visualization. We first present typical qualitative visualizations for the temperature and the precipitation fields of one day in Fig. 2, which highlights the ambitious downscaling with the present large scaling factor of 50. Here, the high-resolution information of 2500 pixels is extracted from one single pixel in the low-resolution counterpart. We compare the results of different protocols (summarized in Table 1), along with the visualization of the original low-resolution JRA-55 and the high-resolution AMGSD data.

The difference in the downscaled temperature from the ground truth is not very large (the upper row of Fig. 2), and it is difficult to find any superiority or inferiority in performance from these qualitative plots. In contrast, the results for precipitation demonstrate rich information on the features of DS protocols (the lower row of Fig. 2). The CDFDM result shows an overly smoothed profile compared to the GT: high precipitation values (represented by red colors) are observed in a vaster area. On the other hand, SRGAN family finely reproduce the localized nature of the high precipitation areas, which the CDFDM fails to describe. Remarkably, even ψSRGAN also succeeded in reproducing the localized heavy rain event, although, in this method, the low-resolution precipitation field is not supplied as an input. The GAN-based methods[13–17] are recognized to be advantageous in reproducing such fine structures. The maximum precipitation values of the DS results are all very close to that of the GT. Please refer to Fig. S2 in the SI Appendix for the graphical depictions of the differences between the GT and DS outcomes, which offer a more direct and intuitive insight into the distinctions among the performances of different methods. We note that although the results for πSRGAN were excluded from Fig. 2 due to their substantial similarity with those for SRGAN and space limitations, they are included in Figure S2 of the SI Appendix.

Single-site statistics. Here and in the following subsections, we discuss the statistical features of downscaling results, focusing on the precipitation p, which is generally considered to be difficult to downscale accurately. In particular, we carefully examine the statistical consistency with the ground truth, which is crucial in actual usage of the DS results, e.g., in impact assessment of climate change in the future. Although the results presented in the main text are climatologically oriented indicators and not standard measures used in the field of image processing, we provide the values of pixel-wise mean squared error and corresponding peak signal-to-noise ratio in the SI Appendix.

We first measure the probability distribution functions (PDFs) of the precipitation data at 12 representative sites, $P_{\mathcal{S}}(p)$. Here, the PDFs are calculated using the set $\{p_k(l)|l \in \mathcal{S} \text{ and } k \in \mathcal{D}_{\text{test}}\}$, where $\mathcal{S}$ stands for the site of interest (each site includes 100 grid points: see Table S3 in the SI Appendix), $\mathcal{D}_{\text{test}}$ is the set of dates that are
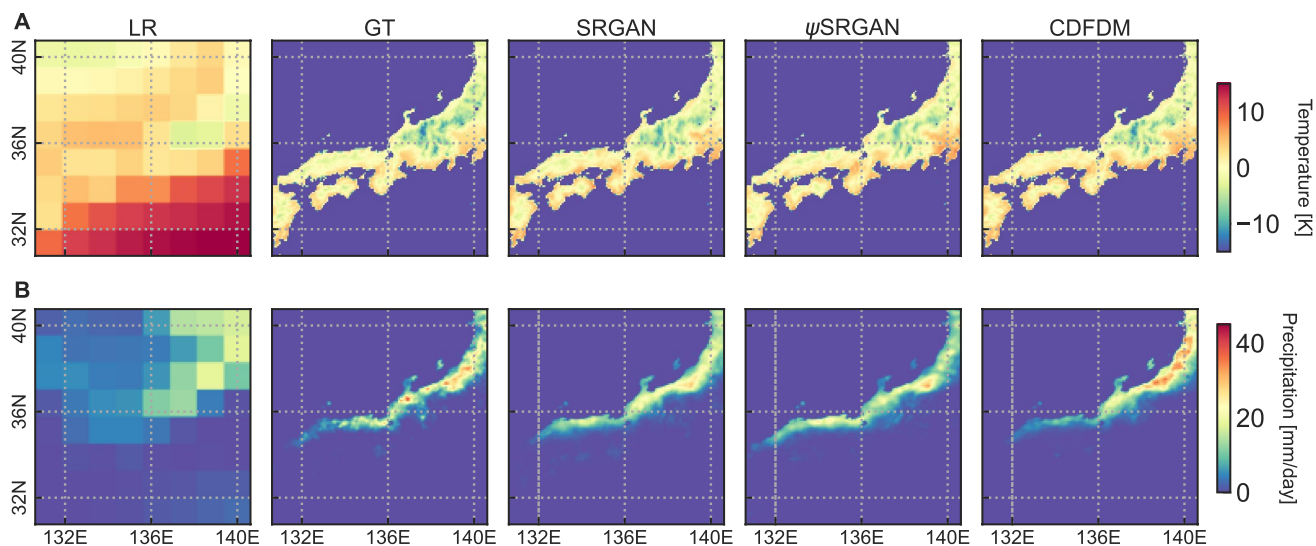

Figure 2. Downscaling results. Distributions of temperature (**A**) and precipitation (**B**) obtained with SRGAN, ψSRGAN, and CDFDM are compared with the corresponding low-resolution (8 × 8) inputs (LR) and the high-resolution (400 × 400) ground truth (GT). The resolution of the downscaled images is the same as that of the GT. The data from January 24, 2008, are displayed.

used for the test data (year span of 2001–2018; Table 2), and $p_k(l)$ is the value of the precipitation at the pixel l and for the date k (we omit the subscript unless necessary below). The results are shown in Fig. 3A–L. The 12 sites in Fig. 3 are chosen from the seaside areas within the system boundary of this study, as depicted in Fig. 3M. Table S3 in the SI Appendix provides more precise information (latitude, longitude, etc.) about these sites.

Overall, Fig. 3A–L shows that all methods express the regional dependence. Regarding each method, the CDFDM provides results matching the GT very well, including the heavy rainfall regime where $p > 50$mm per day up to the values at which $P^{GT}(p)$ becomes around 10^{-4}. This is expected because in the CDFDM the data are processed such that the resulting PDFs become completely consistent with the training data. If we shift our attention to the results of SRGAN family, we first notice that SRGAN and πSRGAN are as accurate as the CDFDM for most sites and most values of p. Moreover, surprisingly, even ψSRGAN succeeded in the projection of precipitation in the range $P^{GT}(p) > 10^{-3}$ at most sites although it was not provided with any direct information about the precipitation. In particular, we would like to stress that an extremely high accuracy has
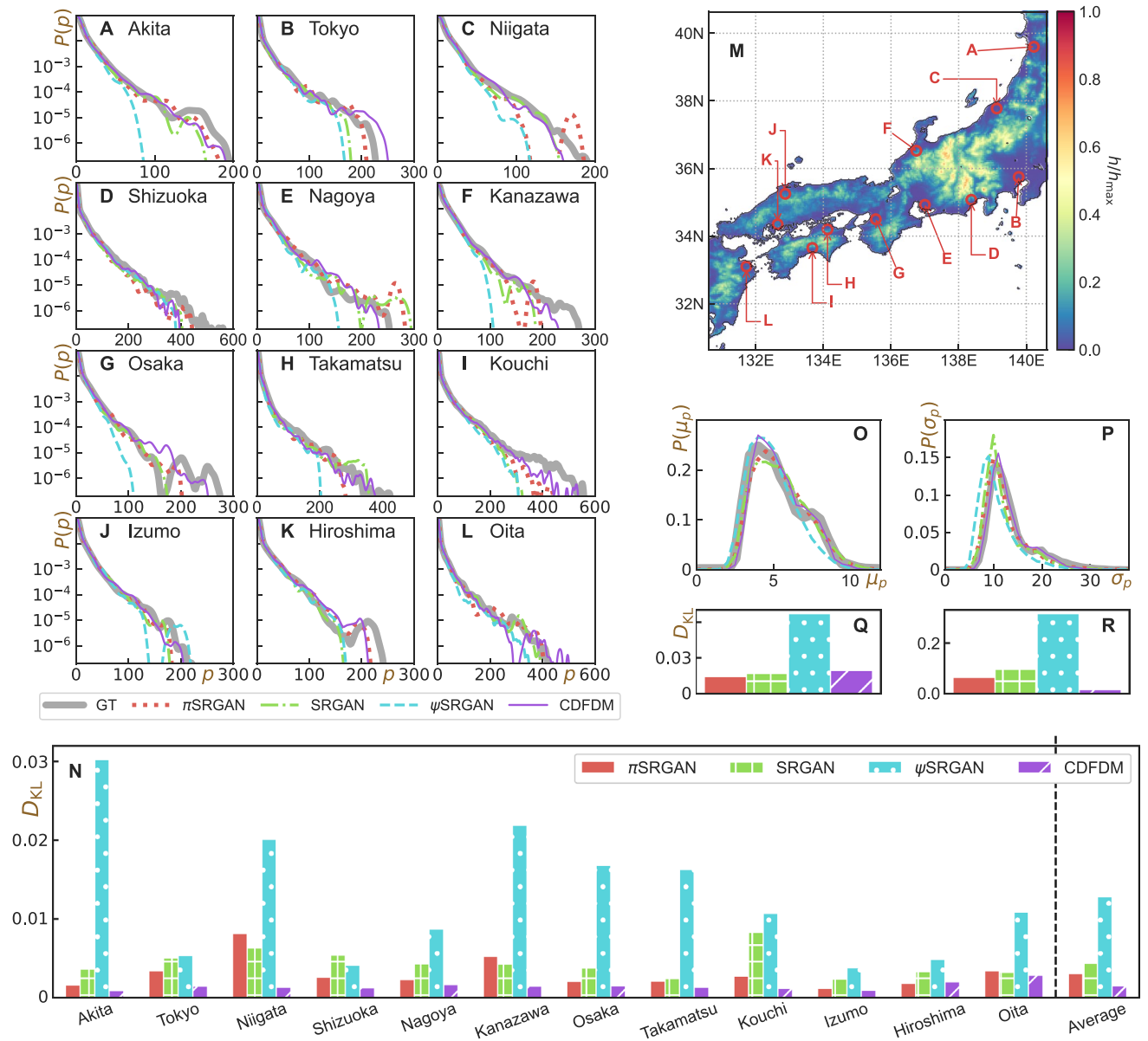

Figure 3. Statistics of precipitation. (**A–L**) The probability distribution functions (PDFs) $P(p)$ as a function of the precipitation p at each site. 12 representative sites are chosen from all over Japan. The different ranges of the abscissa reflect the regional characteristics. See SI Appendix for the technical details in processing the PDFs. (**M**) The normalized topographic information and the locations of 12 sites of panels (**A–L**). (**N**) Bar plot of the values of the Kullback–Leibler divergence of DS methods for $P(p)$. (**O,P**) The PDFs of the mean μ_p and standard deviation σ_p of the precipitation over all the test data at each site; here the values of μ_p and σ_p at different sites serve as samples of the PDFs. (**Q,R**) Bar plot of the values of the Kullback–Leibler divergence for $P(\mu_p)$ and $P(\sigma_p)$.

been successfully obtained for Shizuoka (D), a representative site on the Pacific Ocean side (south side), where pressure-dominated summer-type precipitation events occur frequently. This indicates that the pressure field effectively serves as crucial information for the precipitation projection, such as the location of the typhoons. On the other hand, the accuracy is significantly lower at sites on the Sea of Japan side (north side), Akita, Niigata, and Kanazawa (A, C, F), which are less directly affected by typhoons. These trends are interestingly consistent with our knowledge, and it appears as if SRGAN is extracting physical laws from the data and making predictions, just as humans do. Then it is natural that this success of projection of the high-resolution precipitation from the low-resolution pressure drove us to believe the integration of the input information employed in πSRGAN would further improve the downscaling performance of SRGAN. However, since all SRGAN, πSRGAN, and CDFDM offer highly accurate results, it is difficult to visually judge from the graphs which one is better than the others: we make a quantitative comparison in the next paragraph. Before moving forward to the quantitative analysis, we remark on the discrepancies observed for tails in the large precipitation (small probability of $P^{\mathrm{GT}}(p) < 10^{-4}$) regime even in the cases of the CDFDM. These rare events corresponding to disaster-level torrential rains are very important from the perspective of disaster prevention but are beyond the limit of the current statistical DS methods, on which we provide an overview in Discussion section.

To investigate the difference in the performance of πSRGAN and SRGAN, we quantify the accuracy of each method using the Kullback–Leibler divergence D_{KL}:

$$D_{\mathrm{KL}}(P^{\mathrm{GT}}||P^{DS}) \equiv \int dp P^{\mathrm{GT}}(p) \log \frac{P^{\mathrm{GT}}(p)}{P^{DS}(p)}, \tag{1}$$

where $P^{\mathrm{GT}}(p)$ is the PDF of the GT and $P^{DS}(p)$ is that calculated using the downscaling results ($DS \in \{\pi\mathrm{SRGAN},$ $\mathrm{SRGAN}, \psi\mathrm{SRGAN}, \mathrm{CDFDM}\}$). Generally, the more different $P^{\mathrm{GT}}(p)$ and $P^{DS}(p)$ are, the larger D_{KL} becomes; D_{KL} vanishes when the two PDFs are exactly identical. Since the difference between two PDFs, $P^{\mathrm{GT}}(p)$ and $P^{DS}(p)$, is weighted by the ground truth distribution, the KL divergence places more importance on the frequently occurring events than on rare events. Technical details such as the data preprocessing employed are provided in SI Appendix. The KL divergence between the GT and DS results using distinct methods are shown by bar plots in Fig. 3N and summarized in Table 3, where the values averaged over the 12 sites are presented. The precise values of D_{KL} for each single site are provided in Table S4 in the SI Appendix. As expected from the fact that the CDFDM concentrates on matching these statistics for the training data, it gives the best values for most cases. However, it should be noted that, at Hiroshima (denoted by K), πSRGAN marks a better score than CDFDM. This result evidences the remarkable performance of πSRGAN concerning the basic statistical characteristics that the standard SRGAN can handle relatively inadequately. Indeed, among SRGAN family, πSRGAN marks the best performance if we compare them by the average value over 12 sites: $\bar{D}_{\mathrm{KL}}(P^{\mathrm{GT}}||P^{\pi\mathrm{SRGAN}})$ is smaller than $\bar{D}_{\mathrm{KL}}(P^{\mathrm{GT}}||P^{\mathrm{SRGAN}})$ by approximately 40% (the bars signify that the presented values represent the mean across 12 sites.). However, πSRGAN is not always better than SRGAN and it shows worse results than SRGAN at Niigata, Kanazawa, and Oita (C, F, L). It is noteworthy that these particular locations are precisely where the performance of ψSRGAN is significantly lacking. This observation suggests that the inclusion of low-resolution pressure fields may have led to undesired effects. We also note that, on the other hand, ψSRGAN exhibits a lower value of D_{KL} than that of the standard SRGAN at Shizuoka (Fig. 3D) where the pressure field is expected to play a crucial role in the determination of rainfall events. These findings about the effects of the introduction of auxiliary fields should be utilized for the future refinement of the method. To give a conclusion for this section, remarkably, even the standard SRGAN shows the same order of values of D_{KL} as those of CDFDM. Moreover, the provision of climatologically important auxiliary information can further improve the precision by 40%, evidenced by the results of πSRGAN.

Statistics over all sites. As another meteorologically important statistical point of view, we further measure the statistics over all sites: the PDFs of the mean μ_p and the standard deviation σ_p of the precipitation calculated over all test data on each pixel l:

$$\mu_p(l) \equiv \frac{1}{N_{\mathrm{test}}} \sum_{k}^{N_{\mathrm{test}}} p_k(l), \tag{2}$$

$$\sigma_p(l) \equiv \sqrt{\frac{1}{N_{\mathrm{test}}} \sum_{k}^{N_{\mathrm{test}}} \left(p_k(l) - \mu_p(l)\right)^2}, \tag{3}$$

where $k \in \mathscr{D}_{\mathrm{test}}$ is again the sample index, and N_{test} is the number of samples in $\mathscr{D}_{\mathrm{test}}$. The probability distribution of μ_p and σ_p, denoted by $P(\mu_p)$ and $P(\sigma_p)$, are shown in Fig. 3O,P. Note that here the values calculated on each

	πSRGAN	SRGAN	ψSRGAN	CDFDM
$\bar{D}_{\mathrm{KL}}$	3.06×10^{-3}	4.38×10^{-3}	1.28×10^{-2}	$\mathbf{1.50 \times 10^{-3}}$

Table 3. Average KL divergence of PDFs. Significant values are in [bold].

pixel serve as samples for these PDFs. The corresponding KL divergence D_{KL} for $P(\mu_p)$ and $P(\sigma_p)$ are presented in Fig. 3Q,R as well.

Remarkably, regarding the statistics of pixelwise average over all dates in the test dataset $P(\mu_p)$, πSRGAN (and moreover, SRGAN as well) achieves a better score than CDFDM. However, on the other hand, regarding $P(\sigma_p)$, CDFDM is the best and it shows almost identical results as GT. The small shifts of the whole curve of $P(\sigma_p)$ to the left of SRGAN-based methods are consequences of the underestimation of the high-precipitation events shown in Fig. 3A–L. These results suggest that SRGAN-based methods exhibit a bias towards typical values in downscaling results, as opposed to presenting bold projections of extreme events, compared to CDFDM. This is actually an anticipated tendency considering the design of the standard training scheme employed in machine learning-based methods.

Spatial correlation.

Next, we examine in detail the spatial correlation of the downscaled results. The importance of the spatial correlation of the meteorological variables, i.e., the relation between two distant sites, has been realized very recently[13–17], e.g., in the context of impact assessment of climate change. However, conventional DS methods such as CDFDM have proven to overestimate the correlation even though the statistical consistency with the GT is maintained[26–28]. Such a tendency is actually seen in the qualitative visualizations in Fig. 2, where the overly smoothed profiles are obtained. We thus systematically evaluate the accuracy in expressing the spatial correlation of the precipitation by measuring the Pearson's correlation coefficients of the precipitation $C_M^R(l, l')$ between two sites, l and l', which is defined as:

$$C_M^R(l, l') = \left\langle \frac{\frac{1}{N_M} \sum_k^{N_M} (\delta p_k^R(l) \delta p_k^R(l'))}{\sqrt{\frac{1}{N_M} \sum_k^{N_M} (\delta p_k^R(l))^2} \sqrt{\frac{1}{N_M} \sum_k^{N_M} (\delta p_k^R(l'))^2}} \right\rangle_M \tag{4}$$

where $\delta p_k^R(l) \equiv p_k^R(l) - \bar{p}_M^R(l)$ is the deviation of the k-th sample at site l from its reference average value $\bar{p}_M^R(l)$. The subscript M indicates that the average is taken over the data of month M, the superscript $R \in \{GT, \pi SRGAN, SRGAN, \psi SRGAN, CDFDM\}$ distinguishes the datasets and N_M represents the total number of test data samples belonging to month M. Since the distribution of the correlation coefficients is known to have features specific to each month, we measure the monthly values of the coefficients. Below, we focus on the results for $M =$ January, for which a previous work has pointed out the existence of a distinguished spatial pattern of precipitation correlation[28].

Figure 4A–C show the spatial distribution of the correlation coefficients $C_{Jan}^R(l, l')$, with Nagoya, Niigata, and Hiroshima being the reference points l (the locations of the reference points are marked by the star symbols). The correlations measured for the CDFDM are too high compared to the GT at almost all sites, as shown in Fig. 4A. This is mainly because the 2500 grid points extracted from the corresponding single low-resolution pixel tend to have similar values. In contrast, the results of the SRGAN family exhibit much sharper spatial contrast, e.g., the contrast between the north and south sides of the Chugoku area (around [36° N, 135° E]) is well captured. The differences in performance among these SRGAN-based methods are very subtle and a precise quantification is necessary to rank them: we will get back to this issue in the next paragraph. In Fig. 4B,C, we qualitatively observe the same difference in the accuracy among the methods. In particular, the SRGAN family, even including ψSRGAN, successfully reproduce the nonmonotonic nature of the correlation as a function of the distance from the reference site: e.g., in the results of the GT and SRGAN-based methods in Fig. 4B, along the north side coastline (see the arrow in the figure), the correlation decays quickly near the reference point and then grows again around the Noto peninsula (around [38° N, 137.5° E]). The CDFDM, on the other hand, merely exhibits the monotonic decay of the correlation along the same line. Please see also the SI Appendix for the difference plots between the GT and DS results.

To quantify the accuracy of $C_{Jan}^{DS}(l, l')$ for the different methods, we measure the mean square error (MSE) of the spatial distribution of the correlation coefficient defined as:

$$MSE_M^{DS}(l) = \frac{1}{N_{OS}} \sum_{l'}^{N_{OS}} (C_M^{DS}(l, l') - C_M^{GT}(l, l'))^2, \tag{5}$$

where l is the reference site and $N_{OS} \equiv 630$ is the number of observation stations (see SI Appendix for a detailed explanation). The values of $MSE_{Jan.}^{DS}$ measured based on each reference site are compared in Fig. 4D, and the average values are listed in Table 4 (the values for each site are shown in Table S4 in the SI Appendix). All SRGAN family exhibit much better results than those of the CDFDM for all sites considered here and even ψSRGAN offers twice better results. Specifically, the best one, πSRGAN, achieves 3.6 times better accuracy than the CDFDM for the average value over 12 sites. This result of the SRGAN-based methods being advantageous in achieving the "naturalness" of the spatial pattern is consistent with the report in Ref.[33]. If we further compare the results of SRGAN-based methods, although πSRGAN offers the best performance in terms of the mean value over all 12 sites, the standard SRGAN has the best values at the majority of locations, albeit by only small margins as shown in Fig. 4D (and Table S4 in the SI Appendix). We interpret this result as meaning that both πSRGAN and SRGAN demonstrate comparable performance in relation to the statistical characteristics of spatial correlation. Together with the discussion in the previous subsections, the results presented in this section enable us to conclude that in the present πSRGAN, the auxiliary fields enhance the reproducibility of the simple statistics (such as $P(p)$) while maintaining the expression ability of the natural spatial expanse. Such a strong downscaling ability highlights the applicability to local-scale and interregional assessments of climate change.
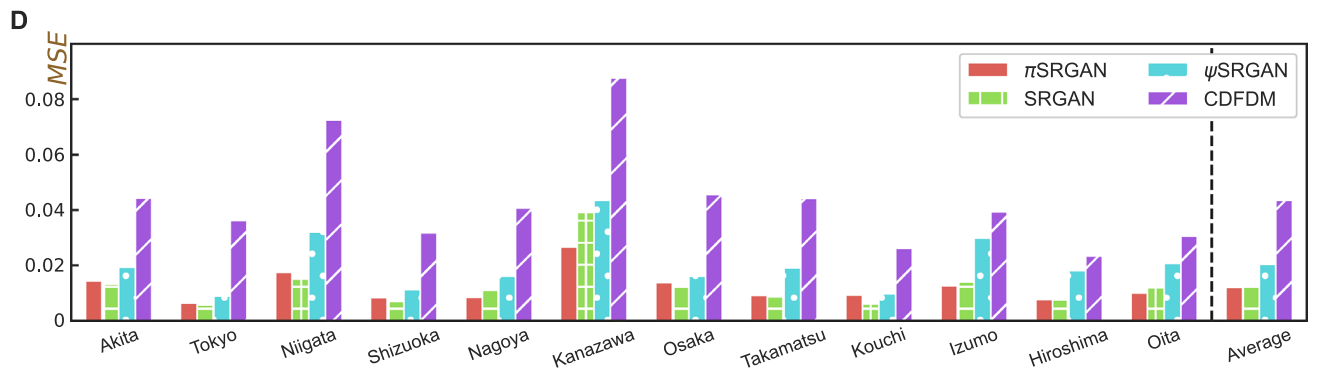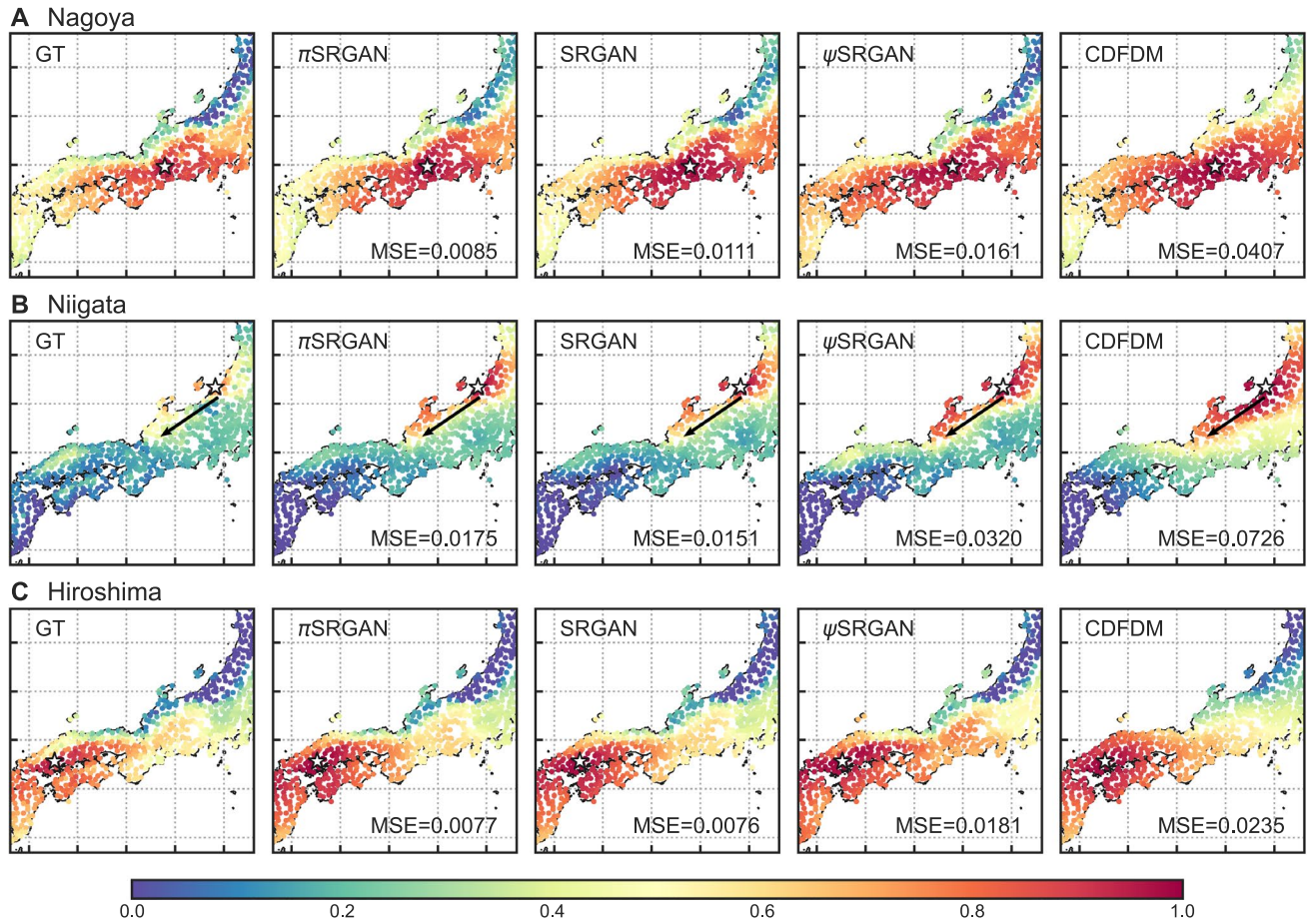
Figure 4. Spatial distribution of the correlation coefficients for precipitation. The distributions of January obtained with the πSRGAN, SRGAN, ψSRGAN, and CDFDM are compared against the ground truth (GT) in the case of the reference point of correlation at Nagoya [35.1667° N, 136.965° E] (**A**), Niigata [37.9133° N, 139.0483° E] (**B**), and Hiroshima [34.365° N, 132.4333° E] (**C**). The dot color indicates the values of $C_{Jan}^{R}(l, l')$ between the location of the dots and the reference site. The reference points are represented by star symbols. (**D**) The mean square error (MSE) of the correlation coefficients of the downscaled precipitations from those of the ground truth. Although in panels (**A–C**), the longitude and latitude values are omitted for reasons of space, they correspond to the same values as in Fig. 3M.

	π**SRGAN**	**SRGAN**	ψ**SRGAN**	**CDFDM**
MSE_{Jan}	$\mathbf{1.20 \times 10^{-2}}$	1.27×10^{-2}	2.04×10^{-2}	4.35×10^{-2}

Table 4. Average MSE of the correlation coefficients. Significant values are in [bold].

Discussion

We have developed a machine learning-based statistical downscaling (DS) method with a large scale-factor of 50, while maintaining both the basic statistical properties and the spatial correlation. We employed a physics-informed type approach[35] on the basis of the SRGAN-based method, and specifically, we developed a framework to use the proper auxiliary physical information along with the low-resolution input to attain large improvements in the DS performance as summarized in Fig. 1 and Tables 3, 4. High accuracy comparable to the CDFDM, a conventional method in actual use, was demonstrated by directly comparing the climatological statistical properties with the real data. More importantly, our approach exhibited the highly accurate reconstruction given in Fig. 4 of the natural spatial distribution of the precipitation correlation coefficient, which was a serious issue for the conventional statistical DS methods, including CDFDM[26–28]. Since the importance of the multiregional spatial correlation has recently been recognized[13–17], the present method is a promising new-generation alternative to conventional statistical DS methods, particularly in situations where the integration of the multiregional information is necessary.

The detection and prediction of rare events are vital issues *inter alia* in the context of climate change assessments. The methods including the present πSRGAN indeed have yet to accurately capture the low probability but significant rainfalls, as shown in Fig. 3. Here, we discuss possible directions to ameliorate the problem. First, we could raise the level of physics-informed machine learning in terms of the classification proposed in Ref.[35]. If we succeeded in directly incorporating some part of the governing equations into the learning process while maintaining the computational efficiency, local phenomena such as heavy rains would be predicted with high reliability. Another direction is to take measures to reform the basic machine learning architecture itself. Following the GAN-based approach, flow-based and diffusion model-based methods have attracted public attentions as powerful next-generation tools for general super-resolution tasks[46,47]. The main feature of these approaches is to generate multiple image candidates from a single input. Therefore, probabilistic information is expected to be drawn from the multiple super-resolved images, which would enable us to tackle the rare event predictions.

Another perspective concerns the use of machine learning techniques to improve the efficiency of dynamical downscaling, i.e., developing a high-speed machine-learning-based solver for the governing equations of climate models. Here we refer to an example of a speed up of multiscale simulations; in Ref.[48] the Gaussian process is used to reduce the computational burden of multiscale simulation for polymeric liquid to achieve a reduction by a factor of 30-100 without loss of accuracy. Breakthroughs driven by similar approaches are expected once the complexity of the governing equations for the climate models is overcome.

Finally, we refer to the generalization ability of SRGAN. Here, we have selected SRGAN instead of πSRGAN due to the anticipated lack of high generalization ability of the latter (πSRGAN relies on topographic information that is specific to the training area). In the SI Appendix, we present the results of the generalization test, in which we tried to execute downscaling computations for samples derived from a different area than the one employed for training. Specifically, the test area encompasses the region spanning from 135.625° to 145.625° east and from 35.625° to 45.625° north, with a 5° shift in both the eastward and northward directions from the original region used for the training. The findings of the examination demonstrate considerably inferior performance compared to those reported in the main text, exposing the deficient generalization capability. This suboptimal performance of the generalization ability is a somewhat predictable attribute since the training data are all from a specific same region. Even though we did not explicitly provide information about the topography in SRGAN, it is plausible that the network learned it indirectly through the temperature field, which exhibits a strong correlation with topography. We stress that we observe large errors even for Niigata and Kanazawa, which were part of the original computational domain. To enhance the generalization ability, we would need to incorporate samples from a more extensive range of areas. The exploration of such an approach is left for future research.

Methods

CDFDM.   Among a variety of statistical methods, we use, as a reference, the cumulative distribution function-based downscaling method (CDFDM) with quantile mapping that is in actual use.

If we simply map the low-resolution GCM simulation results onto the point at which the observations are available, we generally see a systematic difference, defined as bias, which comes from the systematic error of the model prediction and/or from the interpolation error. Removing this inherent bias is especially important in applying the downscaling results to the impact assessments. In the CDFDM, bias is corrected via an empirical transfer function constructed in advance using measured data of distributional variables and the corresponding simulation results. The detailed procedure of constructing the transfer function is described as follows[38].

The crude low-resolution data obtained from the GCM are first mapped onto a 2 km mesh using simple bilinear interpolation. At each mesh point, an empirical cumulative distribution function (CDF) is then constructed using the interpolated data of the variable of interest over a specified time window. The transfer function is defined as a map of a variable onto the one at which the corresponding CDF of the observation falls within the same quantile level. This preconstructed transfer function is applied under the assumption that the error-percentile relation is conserved over time. In the present study, the time window of a month is employed, while the original time window is over a half-year[38], to more sensitively capture the seasonal trend[49,50].

Note that while this CDFDM is a nonparametric method, the corrected CDF perfectly matches the corresponding CDF of the observation (for the training data); the statistical properties of the downscaling results are expected to reproduce the observation well. The bias-corrected climate scenario obtained with this method has been widely used in climate change impact studies[49,51,52].

SRGAN.   We employ a generative adversarial networks-based (GAN-based) method as the basic machine learning architecture, which is called Super-Resolution Generative Adversarial networks (SRGAN)[53]. The termi-

nology super-resolution (SR; or, in particular, single-image super-resolution) refers to a method of restoring a high-resolution image from the corresponding low-resolution data and is the counterpart of the downscaling in the realm of the general image processing. The GAN-based methods are capable of generating realistic images by pitting a discriminator network against a generator network that generates samples (see Fig. 1A). The discriminator network takes the real data (ground truths) and the fake data (output of the generator network) as inputs and identifies the authenticity of the input samples. The generator network tries to deceive the discriminator while the discriminator tries to judge with high accuracy. As a result, both networks spontaneously learn the "realistic" information. The SRGAN can reproduce fine textures that cannot be achieved by normal convolutional neural network-based variants and offers substantially improved realistic super-resolution images.

Such network-based super-resolution techniques have recently been used for the DS tasks of climatological data. In a representative report by Stengel and coworkers, Ref.[33], the authors compared the performances of SRGAN-based downscaling methods with previous methods (SRCNN: Super-Resolution Convolutional Neural-Networks). Although the SRCNN-based method appeared to be superior in evaluating the performance in terms of the simple pixelwise MSE, the SRGAN-based method provided *realistic* results satisfying the important physical requirements, e.g., the energy spectrum of the wind velocity field satisfied the Kolmogorov 5/3 scaling law[39] with remarkable accuracy. The network architecture in our πSRGAN is mostly the same as the original SRGAN introduced in Ref.[53], although the batch normalization layers are removed obeying Ref.[33]: the explanation of the precise architecture is presented in SI Appendix. We also summarize other technical details, such as the precise learning protocol, hyperparameter tuning, and the normalization of the data there. We note that the representative method compared to the πSRGAN referred to as "SRGAN" in our implementation is a slightly upgraded version including the high-resolution topography, which makes possible the decomposition of elements producing the improvement.

Data availibility

The datasets used and analysed in during this study are available from the corresponding author on reasonable request.

References

1. Min, S.-K., Zhang, X., Zwiers, F. W. & Hegerl, G. C. Human contribution to more-intense precipitation extremes. *Nature* **470**, 378–381. https://doi.org/10.1038/nature09763 (2011).
2. Pall, P. *et al.* Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature* **470**, 382–385. https://doi.org/10.1038/nature09762 (2011).
3. Kawase, H. *et al.* Contribution of historical global warming to local-scale heavy precipitation in Western Japan estimated by large ensemble high-resolution simulations. *J. Geophys. Res. Atmos.* **124**, 6093–6103. https://doi.org/10.1029/2018JD030155 (2019).
4. Imada, Y. *et al.* Advanced risk-based event attribution for heavy regional rainfall events. *NPJ Clim. Atmos. Sci.* **3**, 37. https://doi.org/10.1038/s41612-020-00141-y (2020).
5. Masson-Delmotte, V. *et al.* Climate change 2021: The physical science basis. In *Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* 2 (2021).
6. Sudo, K., Takahashi, M., Kurokawa, J.-I. & Akimoto, H. CHASER: A global chemical model of the troposphere 1. Model description. *J. Geophys. Res. Atmos.* **107**, ACH 7-1-ACH 7-20. https://doi.org/10.1029/2001JD001113 (2002).
7. Sato, H., Itoh, A. & Kohyama, T. SEIB-DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecol. Modell.* **200**, 279–307. https://doi.org/10.1016/j.ecolmodel.2006.09.006 (2007).
8. Watanabe, M. *et al.* Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Clim.* **23**, 6312–6335. https://doi.org/10.1175/2010JCLI3679.1 (2010).
9. Yukimoto, S. *et al.* A new global climate model of the meteorological research institute: MRI-CGCM3-model description and basic performance. *J. Meteorol. Soc. Jpn. Ser. II* **90A**, 23–64. https://doi.org/10.2151/jmsj.2012-A02 (2012).
10. Schmidt, G. A. *et al.* Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.* **6**, 141–184. https://doi.org/10.1002/2013MS000265 (2014).
11. Stott, P. A. *et al.* External control of 20th century temperature by natural and anthropogenic forcings. *Science* **290**, 2133–2137. https://doi.org/10.1126/science.290.5499.2133 (2000).
12. Scoccimarro, E. *et al.* Effects of tropical cyclones on ocean heat transport in a high-resolution coupled general circulation model. *J. Clim.* **24**, 4368–4384. https://doi.org/10.1175/2011JCLI4104.1 (2011).
13. Onat, N. C. & Kucukvar, M. Carbon footprint of construction industry: A global review and supply chain analysis. *Renew. Sustain. Energy Rev.* **124**, 109783. https://doi.org/10.1016/j.rser.2020.109783 (2020).
14. Ivanova, D. *et al.* Quantifying the potential for climate change mitigation of consumption options. *Environ. Res. Lett.* **15**, 093001. https://doi.org/10.1088/1748-9326/ab8589 (2020).
15. Fu, X., Lahr, M., Yaxiong, Z. & Meng, B. Actions on climate change, reducing carbon emissions in China via optimal interregional industry shifts. *Energy Policy* **102**, 616–638. https://doi.org/10.1016/j.enpol.2016.10.038 (2017).
16. Zhao, X. *et al.* Linking agricultural GHG emissions to global trade network. *Earth's Future* https://doi.org/10.1029/2019EF001361 *(2020)*.
17. Koks, E. E. & Thissen, M. A multiregional impact assessment model for disaster analysis. *Econ. Syst. Res.* **28**, 429–449. https://doi.org/10.1080/09535314.2016.1232701 (2016).
18. Giorgi, F. & Bates, G. T. The climatological skill of a regional model over complex terrain. *Mon. Weather Rev.* **117**, 2325–2347. https://doi.org/10.1175/1520-0493(1989)117<2325:TCSOAR>2.0.CO;2 (1989).
19. Wang, Y. *et al.* The climatological skill of a regional model over complex terrain. *J. Meteorol. Soc. Jpn. Ser. II* **82**, 1599–1628. https://doi.org/10.2151/jmsj.82.1599 (2004).
20. Déqué, M. *et al.* Global high resolution versus Limited Area Model climate change projections over Europe: quantifying confidence level from PRUDENCE results. *Clim. Dyn.* **25**, 653–670. https://doi.org/10.1007/s00382-005-0052-1 (2005).
21. Kawase, H. *et al.* Downscaling of the climatic change in the Mei-yu rainband in east asia by a pseudo climate simulation method. *SOLA* **4**, 73–76. https://doi.org/10.2151/sola.2008-019 (2008).
22. von Storch, H., Zorita, E. & Cubasch, U. Downscaling of global climate change estimates to regional scales: An application to iberian rainfall in wintertime. *J. Clim.* **6**, 1161–1171. https://doi.org/10.1175/1520-0442(1993)006<1161:DOGCCE>2.0.CO;2 (1993).

23. Wilby, R. L. *et al.* Guidelines for use of climate scenarios developed from statistical downscaling methods. In *Supporting Material of the Intergovernmental Panel on Climate Change, available from the DDC of IPCC TGCIA* **27** (2004).

24. Piani, C., Haerter, J. O. & Coppola, E. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoret. Appl. Climatol.* **99**, 187–192. https://doi.org/10.1007/s00704-009-0134-9 (2010).

25. Iizumi, T., Nishimori, M., Dairaku, K., Adachi, S. A. & Yokozawa, M. Evaluation and intercomparison of downscaled daily precipitation indices over Japan in present-day climate: Strengths and weaknesses of dynamical and bias correction-type statistical downscaling methods. *J. Geophys. Res.* **116**, D01111. https://doi.org/10.1029/2010JD014513 (2011).

26. Maraun, D. *et al.* Towards process-informed bias correction of climate change simulations. *Nat. Clim. Change* **7**, 764–773. https://doi.org/10.1038/nclimate3418 (2017).

27. Widmann, M. *et al.* Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment. *Int. J. Climatol.* https://doi.org/10.1002/joc.6024 *(2019).*

28. Ishizaki, N., Shiogama, H., Hanasaki, N., Takahashi, K., & Nakaegawa, T. Evaluation of the spatial characteristics of climate scenarios based on statistical and dynamical downscaling for impact assessments in Japan. *International Journal of Climatology* **43**(2), 1179–1192. https://doi.org/10.1002/joc.7903 (2023).

29. Kaur, H., Pham, N. & Fomel, S. Improving the resolution of migrated images by approximating the inverse Hessian using deep learning. *Geophysics* **85**, WA173–WA183. https://doi.org/10.1190/geo2019-0315.1 (2020).

30. Kaur, H., Sun, J., Aharchaou, M., Baumstein, A. & Fomel, S. Deep learning framework for true amplitude imaging: Effect of conditioners and initial models. *Geophys. Prospect.* https://doi.org/10.1111/1365-2478.13234 *(2022).*

31. Sachindra, D., Ahmed, K., Rashid, M. M., Shahid, S. & Perera, B. Statistical downscaling of precipitation using machine learning techniques. *Atmos. Res.* **212**, 240–258. https://doi.org/10.1016/j.atmosres.2018.05.022 (2018).

32. Baño-Medina, J., Manzanas, R. & Gutierrez, J. M. Configuration and intercomparison of deep learning neural models for statistical downscaling. *Geosci. Model Dev.* **13**, 2109–2124. https://doi.org/10.5194/gmd-13-2109-2020 (2020).

33. Stengel, K., Glaws, A., Hettinger, D. & King, R. N. Adversarial super-resolution of climatological wind and solar data. *Proc. Natl. Acad. Sci.* **117**, 16805–16815. https://doi.org/10.1073/pnas.1918964117 (2020).

34. Cheng, J. *et al.* Deepdt: Generative adversarial network for high-resolution climate prediction. *IEEE Geosci. Remote Sens. Lett.* **19**, 1–5. https://doi.org/10.1109/LGRS.2020.3041760 (2022).

35. Onishi, R., Sugiyama, D. & Matsuda, K. Super-resolution simulation for real-time prediction of urban micrometeorology. *SOLA* **15**, 178–182. https://doi.org/10.2151/sola.2019-032 (2019).

36. Ohno, H., Sasaki, K., Ohara, G. & Nakazono, K. Development of grid square air temperature and precipitation data compiled from observed, forecasted, and climatic normal data. *Clim. Biosphere* **16**, 71–79. https://doi.org/10.2480/cib.J-16-028 (2016).

37. Harada, Y. *et al.* The JRA-55 reanalysis: Representation of atmospheric circulation and climate variability. *J. Meteorol. Soc. Jpn Ser. II* **94**, 269–302. https://doi.org/10.2151/jmsj.2016-015 (2016).

38. Iizumi, T., Nishimori, M., Ishigooka, Y. & Yokozawa, M. Introduction to climate change scenario derived by statistical downscaling. *J. Agric. Meteorol.* **66**, 131–143. https://doi.org/10.2480/agrmet.66.2.5 (2010).

39. Frisch, U. *Turbulence* (Cambridge University Press, 1995).

40. Vandal, T. *et al.* Deepsd: Generating high resolution climate change projections through single image super-resolution. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '17*, 1663–1672. https://doi.org/10.1145/3097983.3098004 (Association for Computing Machinery, New York, NY, USA, 2017).

41. Vandal, T. *et al.* Generating high resolution climate change projections through single image super-resolution: An abridged version. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, 5389–5393. https://doi.org/10.24963/ijcai.2018/759 (International Joint Conferences on Artificial Intelligence Organization, 2018).

42. Yasuda, Y., Onishi, R., Hirokawa, Y., Kolomenskiy, D. & Sugiyama, D. Super-resolution of near-surface temperature utilizing physical quantities for real-time prediction of urban micrometeorology. https://doi.org/10.48550/ARXIV.2108.00806 (2021).

43. Maraun, D. & Widmann, M. *Statistical Downscaling and Bias Correction for Climate Research* (Cambridge University Press, 2018).

44. Huth, R. Statistical downscaling in central Europe: Evaluation of methods and potential predictors. *Clim. Res.* **13**, 91–101. https://doi.org/10.3354/cr013091 (1999).

45. Dayon, G., Boé, J. & Martin, E. Transferability in the future climate of a statistical downscaling method for precipitation in France. *J. Geophys. Res. Atmos.* **120**, 1023–1043. https://doi.org/10.1002/2014JD022236 (2015).

46. Lugmayr, A., Danelljan, M., Van Gool, L. & Timofte, R. SRFlow: Learning the Super-Resolution Space with Normalizing Flow. 715–732. https://doi.org/10.1007/978-3-030-58558-7_42 (SRFlow, 2020).

47. Li, H. *et al.* SRDiff: Single image super-resolution with diffusion probabilistic models. *Neurocomputing* **479**, 47–59. https://doi.org/10.1016/j.neucom.2022.01.029 (2022).

48. Seryo, N., Sato, T., Molina, J. J. & Taniguchi, T. Learning the constitutive relation of polymeric flows with memory. *Phys. Rev. Res.* **2**, 33107. https://doi.org/10.1103/PhysRevResearch.2.033107 (2020).

49. Yokohata, T. *et al.* Projections of surface air temperature required to sustain permafrost and importance of adaptation to climate change in the Daisetsu Mountains, Japan. *Sci. Rep.* **11**, 15518. https://doi.org/10.1038/s41598-021-94222-4 (2021).

50. Ishizaki, N. N., Shiogama, H., Hanasaki, N. & Takahashi, K. Development of cmip6-based climate scenarios for japan using statistical method and their applicability to impact studies. *Earth Sp. Sci. Open Arch.* https://doi.org/10.1002/essoar.10511571.1 *(2022).*

51. Hiruta, Y., Ishizaki, N. N., Ashina, S. & Takahashi, K. Regional and temporal variations in the impacts of future climate change on Japanese electricity demand: Simultaneous interactions among multiple factors considered. *Energy Conversion Manage. X* **14**, 100172. https://doi.org/10.1016/j.ecmx.2021.100172 (2022).

52. Hiruta, Y., Ishizaki, N. N., Ashina, S. & Takahashi, K. Hourly future climate scenario datasets for impact assessment of climate change considering simultaneous interactions among multiple meteorological factors. *Data Brief* **42**, 108047. https://doi.org/10.1016/j.dib.2022.108047 (2022).

53. Ledig, C. *et al.* Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings—30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* **2017-Janua**, 105–114. https://doi.org/10.1109/CVPR.2017.19 (2017). 1609.04802.

Acknowledgements

Author contributions

N.O. conducted numerical experiment, N.O., N.N.I., and H.Y. analyzed the data, N.O., S.K. and H.Y. invented the method, N.N.I., S.K., and H.Y. designed the work. All authors wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-32947-0.

Correspondence and requests for materials should be addressed to N.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.