Journal of the Meteorological Society of Japan

EARLY ONLINE RELEASE

The DOI for this manuscript is

DOI:10.2151/jmsj.2024-010

J-STAGE Advance published date: January 17th, 2024

The final manuscript after publication will replace the preliminary version at the above DOI once it is available.

Global historical reanalysis with a 60-km AGCM and surface pressure observations: OCADA

Masayoshi Ishii,

Hirotaka Kamahori,

Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Hisayuki Kubota,

Hokkaido University, Sapporo, Japan

Masumi Zaiki,

Seikei University, Tokyo, Japan

Ryo Mizuta,

Hiroaki Kawase,

Masaya Nosaka,

Hiromasa Yoshimura,

Naga Oshima,

Eiki Shindo,

Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Hiroshi Koyama,

Japan Agency for Marin-Earth Science and Technology, Yokohama, Japan

Masato Mori,

Research Institute for Applied Mechanics, Kyushu University, Kasuga, Japan

Shoji Hirahara,

Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Yukiko Imada,

University of Tokyo, Kashiwa, Japan
Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Kohei Yoshida,

Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Toru Nozawa,

Okayama University, Okayama, Japan

Tetsuya Takemi,

Disaster Prevention Research Institute, Kyoto University, Uji, Japan

Takashi Maki,

Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

and

Akio Nishimura

Climate Prediction Division, Japan Meteorological Agency, Tokyo, Japan

December 11, 2023

Corresponding author: Masayoshi Ishii , Meteorological Research Institute, Japan Meteorological Agency, 1-1, Nagamine, Tsukuba, Ibaraki, 305-0052, Japan.
E-mail: maish@mri-jma.go.jp

Abstract

A historical atmospheric reanalysis from 1850 to 2015 was performed using an atmospheric general circulation model assimilating surface pressure observations archived in international databases, with perturbed observational sea surface temperatures as a lower boundary condition. Posterior spread during data assimilation provides quantitative information on the uncertainty in the historical reanalysis. The reanalysis reproduces the evolution of the three-dimensional atmosphere close to those of the operational centers. Newly archived surface pressure observations greatly reduced the uncertainties in the present reanalysis over East Asia in the early 20th century. A scheme for assimilating tropical cyclone tracks and intensities was developed. The scheme was superior to the present several reanalyses in reproducing the intensity close to the observations and the positions. The reanalysis provides possible images of atmospheric circulations before reanalyses with full-scale observations become available, and opportunities for investigating extreme events that occurred before World War II. Incorporating dynamical downscaling with a regional model that includes detailed topography and sophisticated physics is an application of historical reanalysis to reveal the details of past extreme events. Some examples of past heavy rainfall events in Japan are shown using a downscaling experiment, together with dense rainfall observations over the Japanese islands.

Keywords climate; atmosphere; reanalysis; data rescue

1. Introduction

Atmospheric reanalyses provide long-term four dimensional atmospheric evolution using state-of-the-art atmospheric general circulation models (AGCMs) and well-archived observational records (*e.g.*, Kalnay et al., 1996; Uppala et al., 2005; Onogi et al., 2007). Because of the satisfactory quality and length of the available data, they have been used in various climate research and application studies. Meanwhile, reanalyses stretching back more than 100 years were conducted using sophisticated data assimilation schemes such as four-dimensional variational methods and Kalman filters and relatively abundant surface observations such as pressure and sea surface temperature. Whitaker et al. (2004) first implemented this idea using an ensemble Kalman filter and surface-pressure observations. Their reanalysis is called 20CR, and the latest 20CR version 3 stores assimilation results from the early 19th century onward. Their reanalysis has been continuously improved (Compo et al. 2006, 2011; Slivinski et al. 2019). The atmospheric reanalysis of this study is one of the counterparts of the 20CR. A similar reanalysis by the European Center for Medium-Range Weather Forecasts (ECMWF) was released as ERA-20C (Poli et al. 2016) in which marine wind observations were assimilated together with surface pressure observations. A historical reanalysis with a coupled atmosphere-ocean model was also attempted uniquely by ECMWF, which is called CERA-20C (Laloyaux et al. 2018). The model integration in these reanalyses except CERA-20C requires long-term radiative forcing and ocean-surface boundary conditions. The Coupled Model Intercomparision Project (CMIP; *e.g.*, Taylor et al., 2012) provided various natural and anthropogenic forcing suitable for historical climate simulations, and the forcing is frequently used in the historical reanalysis. Historical observational fields of sea surface temperature (SST) and sea ice concentration (SIC) provided by the Hadley Center (HadISST2; Rayner et al., 2006; Titchner and Rayner, 2014) were used in 20CR, ERA-20C, and CERA-20C, sometimes combining other observational SST and SIC data. Moreover, the above-mentioned reanalyses provide uncertainty information by conducting multiple data assimilation experiments with stochastic model physics and spatiotemporally varying background errors (ERA-20C and CERA-20C) and ensemble Kalman filter (20CR).

Atmospheric models can greatly affect on the quality of the four-dimensional atmospheric circulations presented in the above reanalyses. Many AGCMs have produced long-term climate simulations with observed boundary conditions and external forcing factors (Gates et al. 1999). Furthermore, AGCM is a major component of earth system models (ESMs) that simulate interactions between atmospheric, oceanic, chemical, and biological processes. Climates spanning hundreds of years into the past and into the future have been successfully simulated by ESMs typically under the recent CMIP experiments (Taylor et al., 2012; Eyring et al., 2016). As a result, the atmospheric models and external forcing factors are currently capable of hundreds-year-long reanalyses.

The global atmospheric observation network was established mainly for surface and upper air profile observations before the satellite era starting from the late 1970s. Surface observations such as surface pressure (Ps) and SST required in hundreds-year-long reanalyses have been archived in international databases called the International Comprehensive Ocean-Atmosphere Data Set (ICOADS version 3.0; Freeman et al., 2017) and the International Surface Pressure Databank (ISPD version 4.7; Compo et al., 2019). These observations cover the global region from the 19th century or earlier, more densely than the upper air observations. Since the early 2000s, world-wide atmospheric data rescue activities have been active under the International Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative (Allan et al. 2011). Historical observational data are generally subject to various types of bias, and Ps observations are not exceptional. Observations at high altitudes are sometimes challenging, because their values must be adjusted to match the levels between the model and the stations for the data assimilation. Despite this adjustment, the resultant Ps fields may still have residual errors even when considering the model biases (Slivinski et al. 2019).

Observed SSTs are given to the AGCM as a boundary condition at the sea surface when the AGCM is solely integrated. Several long-term gridded SST analyses were produced using ICOADS (Rayner et al. 2003; Hirahara et al. 2014; Huang et al. 2015) and these have been used for the atmospheric reanalyses. Considerable efforts have been made to reduce uncertainties in the analyzed SSTs caused by measurement and human biases and spatiotemporal changes in the observational data distributions (Folland and Parker 1995; Kennedy et al. 2011; Chan and Huybers 2019; Chan et al. 2019). These SST analyses have been available on a monthly calendar basis from the mid-19th century onward. Overall similarities have been found among the present SST analyses (Huang et al. 2015). The uncertainties are typically large before World War II, and decrease over time (Huang et al. 2016).

This study focuses on the reproducibility of East Asian climate and severe weather events from the mid-19th century to the mid-20th century (Okuda 1981; Fujibe 2008, 2010; Watanabe 2012; Fujibe and Matsumoto 2022). In many cases, the severe weather events were directly caused or influenced by tropical cyclones. Therefore, the reanalysis must represent observed typhoons well enough. There are many observational records from the Japan Meiji Restoration (1868) and weather observations at some stations from the early 19th century (Kubota et al. 2021). Some of these records are already stored in the latest ISPD, and data rescue activities are currently being conducted by some research groups cooperating with ACRE. These observations are useful for better representation of severe events in historical reanalyses.

The subsequent sections present the observations and methodology used in this study, and the verification results are compared with previous reanalyses and observations unused in the present reanalysis. Finally, concluding remarks are presented.

2. Data and Method

A historical atmospheric reanalysis was performed from 1850 to 2015, in which atmospheric circulations were updated by assimilating surface pressure observations every three hours with an ensemble Kalman filter using an atmospheric general circulation model (AGCM). The end of the reanalysis period was due to the availability of the observations used this time. In this study, the historical reanalysis is referred to as over-centennial atmospheric data assimilation (OCADA). This nickname is a play on the family name of Dr. Takematsu Okada (1874–1956) who was the fourth Director General of the Central Meteorological Observatory of Japan. He contributed to the modernization of Japan's early meteorological and maritime services and observation networks.

The spectral atmospheric model used in this study is a version of the Meteorological Research Institute AGCM: MRI-AGCM3.2 (Mizuta et al. 2012). The spectral model resolution was TL319 (approximately 60 km at the equator) with 64 vertical levels from the surface to 0.01 hPa. The model was configured for long-term integration with prescribed CMIP Phase 5 greenhouse gases, aerosols, and ozone concentrations (Taylor et al. 2012), and has been applied to past climate simulations and future climate projections (Mizuta et al. 2012, 2017). A database of climate simulations with MRI-AGCM3.2 is known as d4PDF (database for policy decision making for future climate changes; (Mizuta et al. 2017; Fujita et al. 2019; Nosaka et al. 2020). Many studies using d4PDF have reported that the simulation is superior in reproducing atmospheric phenomena in response to global SST variations due to the use of a 60-km atmospheric model and 100 ensemble members of model simulations (Ishii and Mori 2020). Therefore, the same model setup used in d4PDF was chosen for the present reanalysis. The AGCM boundary condition at the sea surface was given by the observed sea surface temperatures from COBE-SST2 (Hirahara et al. 2014) in the reanalysis, as in d4PDF.

Two major observational datasets relelated to surface pressure were used in OCADA,: the International Surface Pressure Databank version 4.7 (ISPD; Compo et al., 2019) and the International Best Track Archive for Climate Stewardship version 4 (IBTrACS; Knapp et al., 2018). IBTrACS was used to control the positions and intensities of model-simulated tropical storms during data assimilation, as described below. In addition, surface pressure observations over East and South-East Asia before World War II were newly digitized and were used first time in this study.

The data assimilation procedure is based on Hunt et al. (2007), which is referred to as local ensemble transform Kalman filter (LETKF). An LETKF-based reanalysis system has been developed from scratch in this study. A series of AGCM ensemble model integrations starting with different initial conditions were performed simultaneously, and then the model states were updated every three hours by LETKF. The system repeats the cycle of model integration and optimization for months or years of data assimilation in a single computational job. No model restarts are required at any time for the three-hourly LETKF procedure, making the computation efficient in reducing elapsed time. Through a preliminary reanalysis experiment with several different ensemble sizes, the size of LETKF was finally determined to be 80 which mostly minimized the differences in 2000–2004 between OCADA and the operational reanalysis of the Japan Meteorological Agency (JRA-55; Kobayashi et al., 2015; Harada et al., 2016).

2.1 Surface Pressure Observations

The surface pressure databank, ISPD ver. 4.7 contains observational data from 1806 to 2015 over global land and oceans (Compo et al. 2011). Worldwide data rescue efforts have collected many of these data under the Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative (Allan et al. 2011). The database was merged with the latest ICOADS 3.0 (Freeman et al. 2017) and was expanded by adding data over both land and oceans from new data rescue projects (Compo et al. 2019).

Figure 1 shows the monthly surface pressure (Ps) observation records and global data coverage. The number of records has increased over time, approaching 70 million by the end of 2015. Although the Ps observations were recorded at various time intervals such as three and four times a day, depending on the stations, ships, and dates of observation, many reports after 1960, that is, after the International Geophysical Year, were recorded every three hours at 00UTC, 03UTC, and so on. The global data coverage increases with time before 1960 over land and before 1980 over the oceans, and these values are nearly constant thereafter. In addition, the coverage over land increases monotonically before 1960, while the coverage over the oceans suffers from two world wars around the mid-1910s and early 1940s. The peaks of the ocean data coverage in the 1850s and the 1880s were due to the US Maury Collection and the US Marine Meteorological Journals

Collection, respectively (Woodruff et al. 1987, 2011). Fig. 1

2.2 Tropical Cyclone Best Track Data

The latest IBTrACS best-track archive (Knapp et al. 2018) stores the three-hourly track positions of tropical cyclones (TCs) over global regions from 1841 to 2021. The estimated or observed Ps and wind speeds at the TC centers are also included when they are available. The TC positions and central pressures were used in OCADA to reproduce past TC tracks and intensities as closely as observed. Because IBTrACS collected tropical cyclone data from all available sources, some of these "spurs" labeled in IBTrACS indicate the same TC with different values among them (Schreck et al. 2014). Therefore, all spurs have been discarded in the present historical reanalysis.

Many central pressure values in IBTrACS are undefined, accounting for as much as half around 1980. Prior to the reanalysis, missing central pressures in IBTrACS were replaced by Ps observations better archived in ISPD with observation type codes related to TCs. This merged TC dataset was suitable to perform data assimilation for this study. The assimilation scheme for the TC track and intensity is described in Section 2.10.

2.3 Additional Surface Observations

In addition to ISPD, Ps observations available in East and South-East Asia were used to improve atmospheric circulation around East Asia in the present reanalysis. These data have been collected from domestic data rescue activities in Japan, as listed in Table 1.

The monthly reports of the Central Meteorological Observatory (CMO), the predecessor to the Japan Meteorological Agency, recorded observations from more than 100 stations in Japan. This study used Ps observations from 66 stations (#1 in the table). Since most of the recorded Ps observations were instrument readings, gravity and sea level corrections were applied to them. From the late 19th century to the early 20th century, Japan experienced many severe meteorological phenomena that caused serious disasters. Therefore, Ps observations at data-available stations in Japan for five days around the dates of the selected individual events were intensively prepared to aim at better representations of the events in OCADA, digitized from the CMO monthly reports: 26 stations for the flash flood at Hikone in September 1896, 16 stations for the TC passing over the Tokyo Bay in September 1917, and 59 stations for Typhoon Muroto in September 1934. The severe weather events in 1896 and 1934 are presented in Section 4.

Japan began systematic weather monitoring by the CMO with a modernized observation network during the Meiji era (1868–1912). In the early stages of modernization, surface observations at lighthouses were dominant (#2). This reanalysis used digitized records from 1877 to 1882, although they were available from 1872 to 1930. Before the Meiji era, weather observations were made at limited stations at Edo (Tokyo), Dejima, Nagasaki, and Yokohama (#3, #4, #5, and #6). Except for those at Edo, these were the result of foreign labor in Japan. In addition, several observational reports at Philippine stations were available (Akasaka 2014; Kubota and Chan 2009; Kobayashi and Yamamoto 2013), and digitized Ps observations at 57 stations were used (#7).

The positions and intensities of TCs during 1892–1944 over the western North Pacific were reported by Wadachi (1952) and Hsu et al. (1973). A total of 778 TCs were archived and used in this study (#8). When the maximum 10-m wind speeds of the TC were not available, the sustained wind speeds were estimated from the central pressures using the Atkinson and Holliday (1977) relationship. The wind data were used to verify the assimilation results.

Further details of each item in Table 1 are presented in Appendix A.

Table 1

A relative increase in the number of additional Ps observations to the ISPD version 4.7 is shown in Fig. 2. This number gradually increased from the late 19th century, but began to decline in the 1930s. These observations improve the global land surface database by a maximum about 20 %. The large increase is partly due to frequent daily records at many stations, with additional stations located only in East Asia, the Philippines, and Micronesia. A spike in the 1880s was due to the inclusion of 3-hourly records per day in the lighthouse observations. Fig. 2

2.4 Land-Surface Ps Bias Corrections

Some surface pressure observations in ISPD suffer from severe biases that are often observed in other types of historical observations. As shown by the time series of Ps anomalies relative to the JRA-55 climatology at a specific station (Fig. 3a), a constant bias correction does not necessarily work for the entire observation period at the station. Note that the model climatology is assumed to be the same between OCADA and JRA-55, here. In fact, these AGCMs are from the same lineage, although some physics schemes differ between the two. Different offsets often appear, exceeding tens of hectopascals in the time series. Several reasons for these offsets include instrumental errors, wrong elevation records, and human errors (Allan and Ansell 2006; Ansell et al. 2006). Among these, sea level adjusted pressure values with ground level elevation or vice versa are frequently observed in the database. Such offsets must be removed prior to climate reanalysis. Since it is difficult to identify the reasons for the individual offsets, observational biases in Ps were defined as long-term averages of deviations from the monthly climatology interpolated to observation dates and locations. The elevations at the observation sites were adjusted to those of the interpolated JRA-55 grid points using a constant lapse rate. In addition, the bias correction scheme must work for sudden changes in the Ps anomaly time series as shown in Fig. 3a. Two methods have been introduced separately for land and ocean observations. Fig. 3

One for land observations was to detect long-term mean differences between the observations and the JRA-55 climatology, and the mean differences were defined as observational biases, assuming they were unchanged for certain ranges of days. Three pairs of averaging range and bias size thresholds were incorporated: if 365-, 182-, and 91-day averages of temporally persistent differences exceeded 3, 7, and 10 hPa, respectively, all Ps observations in those periods were considered biased. These thresholds were determined by comparing the maximum pressure changes for the three averaging ranges in the JRA-55 Ps time series. Observations remained unchanged when data samples were unavailable for 91 days. When the magnitude of the bias was less than 1.5 hPa, the bias was set to zero. The correction scheme was not applied to observations related to TC. The red line in Fig. 3a shows an example of detected biases. Significant changes in Ps anomalies occur several times in the 1850s, 1870s, and 1880s, while no biases lasting for several years or decades are seen in four parts of the time series. Figure 3b depicts a histogram of the maximum magnitudes of the correction amounts, that is bias $\times(-1)$, at each station. Biases were detected at more than 80 % of the stations, ranging from -400 hPa to 500 hPa. The variety of biases after 1950 increased compared to before 1950, reflecting the increase in diversity and frequency of the observations. In particular, not all Ps records counted may be problematic, since JRA-55 does not necessarily represent the truth. The biases of the additional observations (Section 2.3) were mostly detected within $\pm$ 3 hPa, and their mean differences from the JRA-55 climatology were within $\pm$ 1 hPa after the correction.

2.5 Maritime Ps Bias Corrections

Unlike the measurements at land stations, the positions of the ship Ps observations are generally moving. In addition, the ship call signs identifying individual ships are not well known in the observation databases (Ishii et al. 2005). Therefore, another method for the bias detection of ship observations has been introduced. A ship-observed value relative to the

JRA-55 climatology was compared separately with observations at stations within 110 km and 1 hr and those of the other ships within 220 km and 2 hr around the ship under inspection. The station data used were already bias-corrected, and those with elevations less than 25 m were chosen. An average of the differences sampled by either type of comparison is a candidate for ship observation bias. The comparison with land stations has a higher priority, and a median filter was applied before averaging to exclude outliers. An advantage of this procedure is that it does not require long-term records for comparison. As a result, biases were successfully obtained for more than half of the all ships after 1900, but only a few percent before 1900. Most of the biases were distributed within $\pm$ 100 hPa, and those in the 19th century were difficult to compute due to data sparseness. Because metadata such as ship call sign and barometer height were not considered this time, there is much room for improvement in this method. Ship observations are considered unbiased when the absolute biases are less than 1.5 hPa, which is the same as for station observations.

2.6 Evaluation of Bias Correction Methods

Before conducting the reanalysis with the bias-corrected observations, a statistical-based objective analysis of 6-hourly sea level pressure (SLP) from 1845 to 2015 was conducted on a global 1° longitude $\times$ 1° latitude grid. This is a sister product called COBE-SLP2 for sea surface temperature analysis, COBE-SST2 (Ishii et al. 2005; Hirahara et al. 2014), and is a counterpart of Haldley Center monthly historical mean sea level pressure (HadSLP2; Allan and Ansell, 2006) and a daily mean sea level pressure reconstruction over a European-North Atlantic region (Ansell et al. 2006). Two types of statistical analyses were performed using bias-corrected and uncorrected Ps observations, and the 6-hourly analysis was computed as the sum of the JRA-55 monthly climatology, the reconstructed 30-day mean anomalies relative to the climatology, and 6-hourly changes relative to the monthly anomalies. The 6-hourly statistical analysis mimics the historical reanalysis with respect to sub-daily intervals, and its computational cost is much lower than that of the reanalysis.

The bias-corrected SLP analysis is in good agreement with the operational reanalysis for a period after 1960 (Fig. 4). The anomaly correlation coefficients (ACCs) against ERA5 are mostly between 0.8 and 0.96 over the Northern Hemisphere (NH), while those for the Southern Hemisphere (SH) are slightly smaller. There is a strong (weak) seasonality in ACCs in NH (SH): high correlation in winter and low in summer. Here, ERA5 is assumed to be close to the truth. The similarity between COBE-SLP2 and ERA5 remains high over the period globally. The ACCs of JRA-55 in SH before 1980 is gradually worse backward in time, probably due to the use of an older observational database in JRA-55. Fig. 4

Figure 5 shows 100-yr trends in SLP and global and hemispheric mean SLP time series obtained from statistical analyses. The magnitudes of the local trends became small in the bias-corrected analysis. In case of the uncorrected data, there were large positive trends in mountainous regions such as the Rocky Mountains, southern Brazil, Australia, and some areas of Africa, and negative trends around Greenland, the coastal regions of Antarctica, Cuba, and the Philippines. There were significant positive biases over Eurasia in the early 20th century (not shown), and the Ps biases were substantially removed by bias correction, as observed in the northern hemisphere mean SLP time series. These results suggest that the biases over land are more severe than those over the oceans. Fig. 5

The global mean Ps represents the total atmospheric mass, and variations in the water vapor content and greenhouse gases in the atmosphere may cause atmospheric mass changes (Trenberth and Smith 2005). For example, observed column integrated water vapor increased by 5 ± 0.36 %/K at a 5% significance level between 1988–2014 (Allan et al. 2022). This corresponds to an increase of 0.06 hPa. However, this this change cannot be comfirmed by our statistical analysis because the global mean SLPs in Fig. 5 fluctuated within ±0.1 hPa, which is twice the observed trend. Meanwhile, it became clear that the hemispheric averages of one hemisphere compensated for the other around the global mean on decadal time scales when the biases were corrected. A similar compensation was confirmed in the seasonal cycle of Ps ranging at $\pm$ 1 hPa due to dry air and moisture changes caused by the tropical and subtropical monsoonal activities (Trenberth and Smith 2005).

In the present reanalysis, Ps observations with absolute biases larger than 150 hPa (100 hPa) over land (oceans) were not used. After subtracting the biases from the Ps observations, the corrected observations were inspected using quality control procedures: gross error check of the observation minus guess, buddy check, and data thinning, following Ishii et al. (2005) and Hirahara et al. (2014).

2.7 SST and Its Perturbations

The gridded SST observations given by COBE-SST2 (Hirahara et al. 2014) are based on an objective analysis of SST observations, and are defined on a monthly 1° longitude $\times$ 1° latitude grid. The reliability of SST is affected by the observation distribution in space and time. The uncertainty information was provided by COBE-SST2 as analysis errors at all grid points. The globally averaged analysis errors decrease approximately with time (Fig. 1). The reasons for the temporal peaks and troughs in the time series are similar to those presented in Section 2.1. In the historical analysis, COBE-SST2 was interpolated to daily, and SST perturbations with amplitudes proportional to the analysis errors were constructed and used in the LETKF.

The set of SST perturbations represents random fluctuations at grid points on assimilation dates between different members of the LETKF. Meanwhile, the spatiotemporal changes in each member of the perturbed SSTs are continuous. The perturbed SSTs were composed of variations depending on the uncertainty of COBE-SST2 plus those due to ocean eddies. These two parts are independent of one another. The reason for using eddy-related perturbations is that SST variations due to ocean eddies are poorly represented in COBE-SST2 which is based on in situ observations only. The sea ice concentration included in COBE-SST2 was also perturbed consistently with the SST perturbations at each grid point. The construction of the SST perturbations is detailed in Appendix B.

There are advantages to using of SST perturbations in SST-forced AGCM experiments. Actual atmospheric events regarded as natural variations probabilistically respond to the observed SSTs; therefore ensemble AGCM experiments with different initial conditions are often performed (*e.g.*, Watanabe et al., 2013). In such experiments, SST perturbations work effectively according to our experience (Appendix B). In addition, different initial conditions are not always necessary, as the perturbed SSTs alone excite a comparable internal variability in the AGCM. Similar SST perturbations have already been used in a large ensemble climate simulation aimed at future changes in atmospheric extremes (Mizuta et al. 2017; Fujita et al. 2019; Nosaka et al. 2020), known as d4PDF. Many studies using d4PDF simulations have been undertaken to understand the probabilistic behavior of natural variations such as typhoon activity, monsoons, blocking, and atmospheric rivers (Ishii and Mori 2020).

In OCADA, SST perturbations act as a source of observational uncertainties in the atmospheric circulations, and standardize the probabilistic AGCM responses to observed SSTs (Poli et al. 2016). The ensemble spread of OCADA reflects the SST perturbations over time (Fig. 1). Before the 1870s, the SST anomalies over the Niño3 region (150°W – 90°W, 5°S – 5°N) suffered from large uncertainties (Fig. 13 of Hirahara et al., 2014), and correspondingly large ensemble spreads appeared in OCADA. In this case, the ensemble mean states may be featureless. This point needs to be considered when defining the LETKF parameters, as discussed later.

2.8 Data Used for Verification

The historical reanalysis OCADA was verified with observations and reanalyses. The observations used were independent of the current reanalysis: surface air temperature, upper air temperature, and rain gauge precipitation. These gridded observations cover a period of more than 60 years. CRUTEM5 stores monthly surface air temperature anomalies relative to the 1961-1990 average (Osborn and Jones 2014), and the temperatures are defined only on land from 1850 onward on a 5° × 5° longitude-latitude grid. Monthly upper air temperature data, HadAT2 (Thorne et al. 2005), are based on radiosonde observations, and their anomalies from the 1966 – 1995 climatology at nine pressure levels from 850 hPa to 30 hPa were provided on a 10° × 5° longitude-latitude basis from 1958 to 2012. The observational precipitation dataset used in this study is provided by the Global Precipitation Climatology Center (GPCC) of the Deutscher Wetterdienst (Schneider et al. 2010). The dataset contains monthly precipitation amounts from 1901 to 2020, defined on a 0.5° × 0.5° longitude-latitude grid over the global land area, excluding the Antarctic continent.

Several previously published atmospheric reanalyses were used for comparison; conventional reanalyses used all available observations including satellite observations: JRA-55 (Kobayashi et al. 2015) and ERA5 (Hersbach et al. 2020), and one with only surface pressure observations: 20CRv3 (Slivinski et al. 2019). The present historical reanalysis is a counterpart to 20CR.

An observation network called "KUNAI-KANSOKU" or "KUNAI" was the predecessor of the present Automated Meteorological Data Acquisition

System (AMeDAS) maintained by the Japan Meteorological Agency. In the AMeDAS network, surface meteorological observations have been made at stations about 17 km apart over the Japan Islands, while a similar observation density to AMeDAS was maintained in KUNAI (Fujibe 2012). Precipitation observations from the latter were recorded once a day, and some of the data records were digitized with the support of domestic research funds in Japan (Fujibe 2008; Matsumoto 2013). The data were used to verify a severe atmospheric event in 1934.

2.9 Data Assimilation

The atmospheric model was integrated, assimilating the two-dimensional surface pressure at three-hourly intervals using an 80-member LETKF scheme. Simultaneously, the three-dimensional zonal and meridional winds, air temperature, and specific humidity are updated in the model consistently with the analyzed surface pressure field. The LETKF scheme follows Hunt et al. (2007) with an extension of a four-dimensional ensemble Kalman filter (Hunt et al. 2004). For instance, at 03UTC, observations for 3 hours after 00UTC are used to compare model states with spatiotemporally collocated observations. An optimal atmospheric state was obtained by adding the analyzed increments to the model background field, and the model was restarted from this state.

Localization and inflation are important LETKF parameters (Hunt et al. 2007). The localization parameters, which limit the spatial range of the impact of observations, were set to 3,000 km and 400 hPa in the horizontal and vertical directions, respectively. The inverse observational errors are multiplied by weights given by a function (Gaspari and Cohn 1999), ensuring zero weights at twice the localization scale. Inflation of the ensemble AGCM spread is necessary to compensate for the lack of spreads due to the limited ensemble size. In the historical reanalysis, the chosen inflation factors varied in space and time, considering the characteristics of atmospheric variations and changes in observational distributions (Whitaker et al. 2004; Compo et al. 2011). Considering the aforementioned SST perturbations, the inflation parameters were set to smaller values than in the case where the perturbed SSTs are not used: 1.1, 1.01, and 1.01 for 30 °N – 90 °N, 10°S – 10 °N, and 90°S – 30 °S, respectively. Those from 10° to 30 ° are linearly interpolated. In the vertical direction, the inflation beneath the 200 hPa level is the same as at the surface and is linearly reduced toward 30 hPa and zero above 30 hPa. This set of inflation factors was unchanged for years after 1980 because the global observational coverage shown in Fig. 1 saturates over this period. As the SST perturbations inflate the model ensemble spread, the LETKF inflation factors were set to zero before 1946. The inflation factors were linearly interpolated between 1946 and 1980. It was confirmed that the model ensemble spread is sufficiently represented by the SST perturbation alone in model runs without observational constraints. Therefore, it is expected that further inflation by using the factor will sometimes result in too much ensemble spread or too large increments. That is why the inflation factor before 1946 is set zero in this study.

The ratio of the error standard deviations of the surface pressure between the model and the observations was set to 1:4. To avoid spurious diffusion in the model integration for data assimilation, a fourth-order hyperdiffusion was applied to the optimal states. This acted as a low-pass filter in the horizontal sigma plane, while no filter was used in the vertical direction.

2.10 Tropical Cyclone Data Assimilation

Tropical storms cause severe damage to societies in East and South-East Asia. One of the purposes of the current climate reanalysis documented here is to understand the past severe weather systems. Therefore, this reanalysis incorporates data assimilation of tropical cyclone tracks. To accomplish this, model simulated cyclones are forced to maintain their observed position and intensity as much as possible by assimilating an additional set of 13 artificial Ps observations around the center of each observed TC. The artificial observations were placed at the center and at the eastern, northern, western, and southern points of three concentric circles separated by two geodetic degrees (Fiorino 2002; Onogi et al. 2007). Their pressure values were defined with respect to a cyclone in the ensemble mean background state spatially closest to the observation, approximating the structure of the model-generated TC by fitting the Schloemer (1954) formula with a Newton's method:

$$P(r) = P_c + |\Delta P| \exp(-r/R), \tag{1}$$

where P_c is the central pressure, ΔP is the amount of depression from the ambient pressure, r is the distance from the center, and R is a parameter that determines the horizontal structure of the TC. Using Eq. (1), the four observations along a concentric circle have the same values.

In most cases, it was possible to find a model-generated storm near the observed TC within four geodetic degrees from the center of the TC, as demonstrated in Section 3.1. To ensure that the assimilated TC is close to the observation in terms of position and intensity, a penalty depending on the distance between the model and observed TCs is imposed on the observation-minus-guess values and the inverse observation error variance; the penalty for the former is subtraction by $ad/2$, and that for the latter is multiplication by $1 + ad/2$, where a is a constant and d is the distance between the model-generated and observed locations. Here, a is set to 1, and d is upper bounded within two geodetic degrees. The error variances of the concentric observations were magnified by factors ranging from 1 to 3:

the observations were more credible the farther they were from the center of TC. This minimized the distortions of the analyzed fields, especially along the outermost circle.

If the central Ps observation is available in the ISPD-combined IBTrACS (Section 2.2), the artificial observations are modified, preserving their spatial structure, after replacing P_c with actual observations and preserving the TC structure given by Eq. (1). In the initial stage of TC, the depression 7.5 hPa and $R = 1$ in geodetic degree are given a priori only in the case of no pressure observations and no model-generated TCs in the vicinity of the observations.

A tropical cyclone generated in the model was grown or decayed by correcting its position and intensity at the surface using the above assimilation scheme. The TC structure is determined by the model physics and dynamics. In fact, the positions of the observed TCs were well maintained in the model mostly when the distances between the model-generated and observed TCs were within two geodetic degrees. In other cases, a new TC was placed in the model, or the model surface pressures were changed. The radial array of Ps observations used in the data assimilation was made in the shape of the model-simulated storm. Hereafter, they are referred to as pseudo-observations imitating model-simulated TC (POIMT). As described above, the present methodology using POIMT is much simpler than that with the prescribed dynamic and thermodynamic structures of a hurricane vortex, which has been widely used (*e.g.*, Kurihara et al., 1993; Zou and Xiao, 2000; Kobayashi et al., 2015), and non-symmetric structures of TC were ignored in POIMT.

3. Reanalysis Results

The reanalysis integration started in January 1845, and the results from 1850 to 2015 are presented in this section. The ensemble means and spreads of the 80-member ensemble data assimilation were computed on the TL319 Gaussian grid, while the outputs of all ensemble members were converted to a $1^{\circ} \times 1^{\circ}$ grid for convenience of data handling and storage. The model output variables are slightly more than, but mostly the same as those of d4PDF. The volume of all output from one member was approximately 2.7 terabytes.

3.1 Near Surface

Only surface pressure (Ps) observations were assimilated in MRI-AGCM3.2 with the lower boundary condition given by the perturbed observational SSTs. Model Ps necessarily approaches observations over the globe because of the use of data assimilation (Fig. 6). From 1979 to 2015, when the ERA5 used satellite observations on a full scale, the correlation coefficients of monthly Ps between OCADA and ERA5 exceeded 0.85 in most regions. Before the satellite era (1958–1978), OCADA still agrees satisfactorily with the ERA5 with correlation coefficients greater than 0.8 north of 60°S. Hatched areas spreading over the southern oceans appear less confident; "confidence" denoted by $1 - \sigma_{ens}/\sigma_{clim}$ is less than 0.35, where σ_{ens} and σ_{clim} indicate ensemble spreads and interannual standard deviations, respectively. The threshold of 0.35 was taken from Slivinski et al. (2021). Poor agreement in highland areas, such as the Tibetan Plateau, Rocky Mountains, and Central Africa, may be due to insufficient data samples or inadequate Ps bias correction, some of which originates from JRA-55 as seen in other reanalyses (Allan and Ansell 2006). The relatively low correlation coefficients in the southern oceans may be caused by relatively few observations there, and the amplitudes of the Ps anomalies were smaller by 1–3 hPa than in ERA5. Fig. 6

The surface air temperature (SAT) and precipitation of OCADA were compared with the observational data provided by the Climate Research Unit of the University of East Anglia. The CRUTEM5 for temperature and GPCC for precipitation used in this study are the latest and available over land. The comparison provides independent validation of the reanalysis because surface observations, except for Ps, were not used in OCADA.

Figure 7 shows good agreement between the ensemble means and the

SAT and precipitation observations over global, European, and East Asian regions. In the early period in panel (c), differences are likely due to better representation of East Asian surface temperature in OCADA by incorporating new historical Ps data, while CRUTEM5 may be less credible before 1890. The atmospheric models used in OCADA and JRA-55 have similar characteristics each other because some physics schemes are similar between them. However, excessive precipitation over the tropics observed in JRA-55 (Kobayashi et al. 2015) was not seen in OCADA, probably because a different cumulus convection scheme is used in MRI-AGCM3.2 (Mizuta et al. 2012). The confidence intervals for SAT given by the ensemble spreads are small in the later decades due to the increasing number of surface observations (Fig. 1). In contrast, those for precipitation are substantial even in recent years, reflecting discrepancies largely at local scales. Large global mean SAT differences are seen in the 1920s and around 2000, which are caused by unwanted mismatches mainly over Eurasia and North America, respectively (*cf.*, Fig. 5 and Fig. S-1 of Supplemental Materials). These errors seem to be related to biased observations in mountainous areas, probably due to the inclusion of observational pressure biases. The differences in global SAT (panel a) and East Asian precipitation (panel f) appear large around 1910–1930. These might be caused by some biases recently reported by Chan et al. (2019), which are not yet properly handled in COBE-SST2, or by temporally invariant aerosol and ozone concentrations externally given to the model (Mizuta et al. 2017). Fig. 7

The reproduction of realistic tropical cyclones is one of the main subjects of the present reanalysis because they play a critical role in natural disasters in East Asia. Figure 8 shows an example of the assimilated TC named Tokage that hit Japan on October 20, 2004. According to local measurement reports, the central surface pressure reached 955 hPa at 06UTC, while those of 966 hPa and 976 hPa were reached for OCADA and JRA-55, respectively. The difference of about 10 hPa in the central Ps between OCADA and JRA-55 seems remarkable, since the ensemble spread at the center of the TC is approximately 5 hPa. The TC reproducibility of MRI-AGCM has been investigated by previous studies in terms of occurrence frequency, spatiotemporal distribution, and intensity (Murakami et al. 2012; Mizuta et al. 2017; Yoshida et al. 2017). Overall, the model favorably captures the characteristics of actual TCs. In addition, the TC track data assimilation embedded in OCADA appears to work as expected. Although the 60-km AGCM is not good at representing terrain-trapped precipitation along the Japanese archipelago, the precipitation amounts of two reanalyses are comparable to those of GPCC and a local observation network in Japan (Fig. S-2 of Supplemental Material). Fig. 8

The LETKF procedure corrects the position of the model-predicted TC to that of the observations by assimilating POIMT (Section 2.10). The following results compare six-hourly TC track positions and intensities between OCADA and IBTrACS. In this comparison, reanalysis SLPs were searched for TCs within six geodetic degrees around the observed TCs in the IBTrACS. Since the track information and central TC pressures of IBTrACS were assimilated in the present reanalysis, the comparison shown below is not independent for these two variables. For the maximum 10-m wind speed, the comparison result will provide an incomplete verification, as many of them are estimated from corresponding pressure fields.

The mean position errors, which are the distances between the analyzed and observed TCs, were within 2.3 geodetic degrees for all TCs for the entire reanalysis period, and within 3 geodetic degrees between the predicted and observed TCs. The threshold of 2° adopted as the spatial scale of the penalty, d (Section 2.10), is close to these values. The correction amounts of the global TC positions obtained by the LETKF resulted in 0.7 – 1° over time. Approximately 50 % of all predicted TCs before 1950 and 25 % in recent decades had position errors larger than 2°. The decreasing errors with time imply that the model background fields become suitable for maintaining model-simulated TCs as observed (Fig. 1).

Figure 9 shows the effectiveness of OCADA in reproducing TC activities in recent decades, compared with JRA-55, ERA5, and 20CR, based on all

6-hourly TC samples from the genesis to the decay. The position errors are mostly within 2°, and the differences in central pressure and maximum 10-m wind speed are approximately within ± 10 hPa and ± 10 ms^{-1}, respectively. Table 2 shows the quantitative comparison of these errors. In particular, the errors of the OCADA's central pressure are smaller than those of the other reanalyses. The tropical cyclones represented by OCADA are generally close to observations. The TC intensities of OCADA tend to be weaker than those of IBTrACS. However, the absolute TC intensity distributions appear to be in good agreement with those of IBTrACS. Furthermore, severe TCs exist in OCADA with a frequency of more than 80 % of the observations, using severity thresholds of severance: 982.5 hPa and 27.5 m$s^{-1}$for central pressure and maximum 10-m wind, respectively (Table 2). The frequency of 20CR is also higher than in the conventional reanalyses. The reanalysis with only surface observations may have an advantage in reproducing TC activities near the surface. Baker et al. (2021) showed that TC activities are underrepresented in conventional reanalyses such as JRA-55 and ERA5. Similarly, TCs with central pressures below 990 hPa and maximum 10-m winds above 20 ms^{-1}appear weak in this comparison. In contrast, the position errors within 1° are more dominant in the two conventional reanalyses than in OCADA. Other statistics with minimum pressure and maximum 10-m wind samples of each TC (not shown) showed a similar result to Fig. 9. In JRA-55, unrealistic weakening trends in the analyzed tropical cyclone intensity were reported by Kobayashi et al. (2015), but no such thing was confirmed in OCADA. Note that no specific TC assimilation scheme was adopted in ERA5 and 20CR.

Fig. 9

Table 2

3.2 Upper Air

The three-hourly atmospheric upper air states were updated by assimilating only surface pressure observations. Therefore, the quality of the OCADA largely depends on the assimilation scheme and model performance adopted in the reanalysis experiment and the observation density varying in space and time.

Figure 10 displays anomaly correlation coefficients (ACCs) of air temperature between the reanalyses and HadAT for the whole period of HadAT (1958 – 2012). The variations in upper air temperatures are well represented in JRA-55, which used radio-sonde observations. In contrast, the two historical reanalyses of OCADA and 20CR reasonably present the variations at the 500 hPa level, but those at 200 hPa poorly agree with the observations, despite the data-rich period. This might be a limitation of the current historical reanalyses. OCADA Surface pressures agree well with ERA-5 on land rather than in the tropical and southern oceans (Fig. 6). Similar patterns are observed for upper air temperatures. HadAT is based only on radio-sonde observations which have been carefully inspected to minimize the effect of any pervasive biases (Thorne et al. 2005). However, none of the reanalyses considered here well represent the observed variations over India and the African Continent. Fig. 10

Figure 11 displays monthly root-mean-square differences (RMSDs) of zonal mean upper air temperature and zonal wind in d4PDF, OCADA, and 20CR, compared with ERA5. The same AGCM was used in OCADA and d4PDF, the latter of which is free from observations. Comparing OCADA to d4PDF, the improvement due to data assimilation is more obvious for the zonal wind than for the air temperature. In high latitudes, the RMSDs of two variables are slightly larger in OCADA than in 20CR. Large RMSDs in air temperature of OCADA around 60°N near the surface originated mainly from disagreement with ERA5 over land areas. These features are commonly seen in JRA-55; a possible reason is differences between the boundary layer schemes of the modeling centers. When the reference is changed from ERA5 to JRA-55, the situation between OCADA and 20CR is apparently reversed (not shown). Consequently, the quality of OCADA, particularly in the upper air and at high latitudes, substantially depends on the atmospheric model used. Fig. 11

The geopotential height at the 500 hPa isobar surface, which is a representative variable in the troposphere, is a good indicator of the skill of daily weather forecasts (*e.g.*, Magnusson and Källén, 2013). Figure 12 shows the RMSDs of the 500 hPa geopotential height between OCADA and JRA-55 for the northern (20°N – 80°N) and southern (80°S – 20°S) hemispheres. The annual mean RMSDs in the Northern Hemisphere were less than 30 m for most of the period, sometimes exceeded 30 m, and were mostly constant from 1958 to 2015. Seasonal dependencies appear in the 6-hourly RMSDs: larger RMSDs from winter to spring and smaller from summer to autumn in both hemispheres, although the seasonality is not so robust. The dominant daily weather systems in each season may have influenced this result. A previous study (Compo et al. 2006) reported that the RMSDs of 20CR correspond to those of the 3-4-day lead weather forecast at operational centers, and the latest 20CR maintains this quality level against the latest weather forecasts at the ECMWF (blue lines in Fig. 12 and 25–30 hPa shown in Fig. 5 of Slivinski et al., 2021). In the Southern Hemisphere, RMSDs are larger than those in the Northern Hemisphere, and they decrease in time from 90 m to 40 m. Compared with 20CR, RMSDs of OCADA are 20 %–30 % larger in the Northern Hemisphere. It is inferred from the smaller spreads than RMSDs in OCADA that the required inflation factors or the SST perturbations are insufficient (Section 2.9). Fig. 12

3.3 Validation of Reanalysis Uncertainties

Another approach was taken to validate the ensemble spreads of OCADA and to evaluate the impact of additional Ps observations in East and South-East Asia on the reanalysis for the period before ERA5 and JRA-55 became available. We assumed that the number of observations is the main factor for the uncertainties in OCADA. To estimate the uncertainties, pseudo-reanalysis experiments for 2001 were performed with collocated Ps observations corresponding to those of a past year. The years for the pseudo-reanalysis were chosen every 20 years from 1865 to 1965. The uncertainties of the reanalysis for these years were computed as RMSDs between the pseudo- and the original reanalysis with the full set of Ps observations, assuming that the original reanalysis is the truth. Similar approaches have been adopted to validate analysis errors in statistical analyses of oceanographic elements (*e.g.*, Hirahara et al., 2014; Ishii et al., 2017).

Figure 13 demonstrates the reduction of uncertainties in global mean geopotential height profiles and hemispheric and tropical-subtropical mean surface pressures by comparing RMSDs and spreads. The maximum geopotential RMSDs appeared at levels near the centers of the subtropical jets and the minimum RMSDs near the surface throughout the period. The RMSDs decrease over the period due to the increasing number of observations available in the database (Fig. 1). This test showed that the uncertainties represented by RMSD and the spread agree well with each other, and that the spreads are rather sensitive to changes in the number of observations. Notably, the years 1885 and 1945 are rather irregular in terms of data availability: the number of observations increases in the former and oppositely decrease in the latter owing to World War II (Section 2.1). The ensemble spread was slightly underestimated before 1920, and this underestimation appears severe in the Northern Hemisphere spread of surface pressure. In contrast, the spreads in low latitudes and the Southern Hemisphere are comparable to the corresponding RMSDs. Fig. 13

Figure 14 shows the relative decreases in RMSDs of 500 hPa geopotential height and SLP due to the additional observations in East and South-East Asia (Section 2.3) by comparing RMSDs of the pseudo-reanalyses with and without the additional observations in 1925. Sizable reductions of the uncertainties in geopotential height and SLP are observed in East Asia, and the areas of reduction extend eastward to the dateline. In the tropical region, the reduction is less impressive likely because the observed SSTs dominantly determine the atmospheric circulations there (Fig. 11). The zonal extent of the RMSD reduction areas in the tropics should be close to the scale of the atmospheric tides, which is approximately 90° of longitude. Such features are confirmed in the figure, although the signals are weak. In the meantime, RMSDs become larger particularly in West Asia, North

America, Alaska, and Europe. It is unlikely that the improvement over the East Asia prolongs along the subtropical jet on the 500 hPa isobaric surface. The reason is unclear, but direct observations may be needed to determine the upper air conditions over the western North America and some areas in Eurasia. Regarding TCs, no significant improvement was confirmed in these experiments. In fact, Ps observations near the TC centers are scarce, and hence, the accuracy of TC center positions is primarily important. In contrast, surface observations around TCs should be beneficial to realize the environmental conditions for TCs. However, the latter remains unsolved. Fig. 14

4. Severe Events before WW-II in Japan

The latest Intergovernmental Panel on Climate Change Assessment Report states that global warming brings with it the threat of heavy precipitation (Masson-Delmotte et al. 2021). However, historical records suggest possible candidates for extreme precipitation events more threatening than those observed after the International Geophysical Year. In Japan, one of them is a heavy precipitation event that occurred in Hikone, located around the center of Honshu Island. The event occurred in September 1896, and a daily rainfall amount of approximately 600 mm/day was reported in the observation records (Okuda 1981). In addition, the water level of Lake Biwa near Hikone and the largest in this country, reached approximately 4 m, according to a Japanese flood information site (Shiga Prefecture 2020). The return period of the 600 mm rainfall at Hikone was out of extreme statistics (Suzuki and Kikuchihara 1984) or was estimated as ten thousand years (Fujibe 2010). Internal variations in the atmosphere are rich in variety. Therefore, understanding such extreme events is essential for advancing research in meteorology and climatology.

The Hikone heavy rainfall event was plausibly reproduced in OCADA (Fig. 15), although the precipitation amount was approximately a quarter of that observed. Given this, we should consider the appropriateness of using POIMT around the observed tropical cyclone. As mentioned in Section 2.10, POIMT is based on model-simulated tropical cyclones that were used to maintain the track close to the observation. If POIMT had not been used, the event would not have been effectively simulated, because station and ship observations are scarcely available. This historical reanalysis was produced with the goal of its practical use for meteorological and climatological studies and various climate impact assessment studies for specific East Asian regions. Therefore, the assimilation of POIMT was adopted in OCADA. Note that the heavy precipitation northwest of the TC shown in the figure is probably unrealistic according to the literature (Watanabe 2012) and observational records maintained by the Japan Meteorological Agency. The imperfection of the TC simulation as well as the data assimilation may affect the present result. Fig. 15

Examining the appropriateness of reanalysis with POIMT in a meteorological sense might be possible using locally available observations. For example, it is inferred from the weather map and available precipitation observations for the Hikone case that a seasonal rain front located over Honshu Island, activated by a southern typhoon, brought heavy rainfall (Okuda 1981). To further understand the event, one approach is dynamical downscaling with a higher-resolution regional model equipped with more realistic topography and sophisticated physics than the global model. The regional model can then be used to quantitatively reproduce an event and how it occurs.

Figure 16 shows the result of downscaling for another significant event, Typhoon Muroto, in September 1934, for which KUNAI observations are available. A regional model NHRCM (Non-Hydrostatic Regional Climate Model; Sasaki et al., 2011) was used for it. This model is also the same as the one used in d4PDF, and the horizontal resolution of NHRCM is 20 km. The regional model started on July 1 of the same year from an initial condition given by the global model and was integrated constraining NHRCM along the computational domain boundary derived from OCADA. In addition, prognostic variables in NHRCM were nudged to OCADA of low wave numbers in the upper air. The downscaling is promising. As observed, the distribution of weak and strong precipitation areas was represented in the downscaling, and topographical effects on precipitation were confirmed. The precipitation in front of the TC is weaker in the global model than in NHRCM. Compared with KUNAI, the 20-km resolution of NHRCM appears to be insufficient to represent terrain trapped precipitation. Further downscaling up to 5 km or 2 km resolution is planned in future studies for more detailed information on the extreme events. Fig. 16

5. Concluding Remarks

This study presented a historical atmospheric reanalysis from 1850 to 2015 using a 60-km mesh AGCM and surface pressure (Ps) observations. The three-dimensional atmospheric states were updated by assimilating surface observations with an 80-member LETKF scheme constructed in this study. The pressure observations were taken from ISPD version 4.7, and additional East and South-East Asia observations were included here. A scheme specifically designed for assimilating the observed tropical cyclone tracks and intensities was incorporated into it. Sea surface temperatures from COBE-SST2 were used for the ocean surface boundary condition of AGCM during the assimilation integration, and 80 sets of perturbed COBE-SST2 were used to drive the individual members of the LETKF. Prior to the reanalysis, large Ps biases in the observations were removed so that the long-term averages were close to the JRA-55 climatology. In addition, newly archived observations, mostly in East and South-East Asia, provided frequent daily records that increased the total volume of the Ps database by 20 % in the early 20th century.

This new reanalysis, named OCADA, reproduced the atmospheric circulation and various extreme events with uncertainties equivalent to the LETKF ensemble spreads. The ensemble spread appears to mostly capture the actual uncertainty in the global means. However, those at local scales do not. In addition, OCADA had a lower performance in representing 500 hPa geopotential heights compared to 20CR. These results suggest room for improvement in the assimilation scheme and the LETKF inflation and localization parameters used.

It is difficult to perfectly remove biases from the historical observations. Its goal is also unclear, and there is no choice but to repeat bias corrections and objective analyses. The Ps bias correction schemes have been developed separately for land stations and ships. However, this methodology is still under development. The choice of JRA-55 for the base field needs to be reconsidered, and the efficient use of the metadata could overcome the incompleteness of the scheme due to missing ship call signs. Regarding COBE-SST2, observation biases before World War II may not have been sufficiently removed. However, the actual biases in SST observations before the 1940s seem too complicated to be thoroughly removed (Chan and Huybers 2019). In addition, the resolution of COBE-SST2, $1^\circ \times 1^\circ$, is insufficient to resolve the variability in the Kuroshio and Gulf Stream and coastal regions (Reynolds et al. 2007). Careful treatment of observations and appropriate boundary conditions are required to reproduce regional climates.

Hundreds-year-long reanalyses are essential for a deep understanding of the internal variety of atmospheric circulation and weather systems. Severe past events causing the most severe social damage may occur under current or future climate conditions. Meanwhile, even well-known El Niño events vary in many aspects on decadal time scales (*e.g.*, Kleeman et al., 1999). Climate projections have been made to maintain our lifestyles in the future; however, global warming, which has been slowly progressing for more than 100 years, is poorly understood. Our knowledge based on observations is too short to answer these questions. Understanding past atmospheric and climatic events through such reanalyses is promising. Atmospheric data rescue is also essential. Fortunately, a number of observations are available in Japan, although not all of them have been digitized yet. Dynamical downscaling from the global historical reanalysis, OCADA, could help to understand past extreme events and long-term climate changes in Japan.

Acknowledgements

The authors thank Dr. G. Compo for providing the latest surface pressure data (ISPD version 4.7) and for his support. The digitized Japanese "KUNAI-KANSOKU" dataset was provided by Dr. F. Fujibe, Dr. J. Matsumoto, and Dr. H. Yamamoto. Their appreciation is extended to all the data providers of IBTrACS version 4, ERA5, CRUTEM5, HadAT, and GPCC. This study was supported in part by the Integrated Research Program for Advancing Climate Models (TOUGOU; from FY2017 to FY2021; Grant Number: JPMXD0717935561) funded by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan. The Earth Simulator maintained by JAMSTEC is extremely useful for conducting large-ensemble climate reanalysis. The tremendous help from Mr. H. Yamamoto, Mr. M. Kamei, Mr. S. Hayashi, and the JMA staff members are gratefully acknowledged.

H. Kubota was supported by the Grant-in-Aid for Scientific Research No. 25282085, 15KK0030 and 20K20328 and by the Young Scientific Research No. 21684028 of the MEXT. The author thanks Dr. Mong-Ming Lu of the Central Weather Bureau, Prof. Chih-Wen Hung of the National Taiwan Normal University, Prof. Chung-Hsiung Sui of the National Taiwan University, the Philippines Atmospheric Geophysical and Astronomical Services Administration, the University of Hawaii, and the Japan Meteorological Agency for sharing historical data and typhoon books.

Finally, the authors would like to thank the two anonymous reviewers for taking the time to read the manuscript and for their constructive comments.

Data Availability

The historical reanalysis data produced by this study are available at https://climate.mri-jma.go.jp/pub/archives/Ishii-et-al_OCADA/. The data server provides the ensemble means and spreads of the main variables defined on a $1^\circ \times 1^\circ$ grid in the NetCDF format. The other data of each ensemble member or the ensemble means and spreads on the original grid will be provided to the extent possible upon request. SST perturbations and COBE-SLP2 can be accessed from the same location. Access to the surface pressure observations used in this study, as well as KUNAI-KANSOKU, requires the permission from the individual owners, unless published. The historical reanalysis, 20CR, is available at https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html or https://doi.org/10.5065/H93G-WS83, and the operational reanalyses, JRA-55 and ERA5, at http://search.diasjp.net/en/dataset/JRA55 and https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, respectively. The University of East Anglia provides gridded air temperature observations called CRUTEM5 and HadAT at https://crudata.uea.ac.uk/cru/data/temperature/. Gridded precipitation observations are provided by the Global Precipitation Climatology Center of the Deutscher Wetterdienst, and are available at https://opendata.dwd.de/climate_environment/GPCC/html/download_gate.html.

Appendices

A. Additional Observations in East and South-East Asia

The sections below present the details of the additional observations shown in Table 1.

a. CMO Stations (1890–1945)

The Japan Meteorological Agency (JMA) manages the observational archives of the Central Meteorological Observatory (CMO), the predecessor of JMA, and those of the former governmental administrations and agencies. The observational data were recorded on paper or microfilms. The CMO data were published as monthly reports. Metadata for the use of altitude and gravity adjustments were included in the records of most stations. This study used observations measured at 66 CMO stations.

Hourly surface pressure observations at nine stations (Wakkanai, Sapporo, Hakodate, Niigata, Tokyo, Kobe, Shionomisaki, Fukuoka, Kagoshima) were digitized from the original CMO books under a support of the Innovative Program of Climate Change Projection for the 21st Century (FY2007–FY2011) funded by the Ministry of Education, Culture, Sports, Science, and Technology. These stations were the designated observatories of the

CMO, and weather observations were maintained by trained staff members.

As submarine telegraph lines connecting the Japanese main islands with the outlying islands became open in 1895, the CMO began to work in earnest on the outlying islands. Observations at seven stations in Southwest Islands, Kagoshima Prefecture, Hachijojima, Chichijima, and Iwoto in the CMO monthly reports from 1890 to 1945 were used in this study. Station pressure was recorded every one to eight hours depending on the stations, and altitude and gravity adjustments were made (Kubota et al. 2016).

A number of overseas Japanese stations were established after the Chinese-Japanese Peace Treaty of 1895 and the Russian-Japanese Peace Treaty of 1905. Overseas stations in Taiwan, Korea, Manchuria, China, Sakhalin, and the Kuril Islands reported in the monthly reports of the CMO of Japan were used from 1897 to 1941. Surface pressures at 34 stations were recorded every one to eight hours depending on the stations. The same adjustments as above were made.

Weather observations in the Micronesia Islands began in 1923 at the observatory of the South Seas Bureau of the old Japanese government. The CMO monthly reports include surface pressure observations made at 12 stations in Micronesia. These data are available from 1923 to 1945 (Kubota 2012). The pressure data were recorded every one to twelve hours depending on the stations, and altitude and gravity adjustments were made.

b. Lighthouse (1877–1882)

In 1868 immediately after the Edo era, the new Meiji government invited R. H. Branton from England to supervise the construction of lighthouses. Meteorological observations began at various lighthouses in 1872, and the British Meteorological Committee compiled the records of these observations (Nyomura 2002). The records from 1877 were microfilmed, and digitized into image files by JMA. Although Nyomura (2002) indicated that the microfilm records continue to 1953, the image files maintained by JMA are up to 1930.

The number of data-available lighthouses was 26 in January 1877 and then increased to 37 in 1882. The number of daily observations of pressure and temperature was two (9:00 and 21:00) per day until 1881, increasing to eight times (0:00, 3:00, 6:00, 9:00, 12:00, 15:00, 18:00, and 21:00) per day from 1882. The barometer readings were recorded in English inches. The local time was GMT + 9:00.

c. Edo (1838–1855)

Observations during the Edo era (1600–1868), whose capital was located in Tokyo, Japan, were made by the Tokugawa government bureau of Astronomy for Calendar Making of Edo. On April 1, 1842, this bureau was located 35.7°N, 139°E, and 20.0 m above mean sea level (MSL). After a fire on April 2, 1842 (Amano 1953), the bureau was reestablished on June 1, 1842 at 35.69°N, 139.75°E, 6.5 m above MSL, which is approximately 2 km south-southeast of the first location. The original handwritten lists of observations, stored at the National Astronomical Observatory in Japan, are not accessible, but contemporary handwritten copies of these lists are available from the National Archives of Japan (REIKEN-KOUBO, 1838–1855; handwritten documents available from the National Archives of Japan, 3-2 Kitanomaru Koen, Chiyoda-ku, 102-0091 Tokyo, Japan).

Meteorological observations of temperature (°F), pressure (English inch), and weather were made from December 17, 1838, to February 16, 1855. A blank period was due to the fire after April 1, 1842. Observations of atmospheric pressure were resumed on December 30, 1844. This suggests that meteorological instruments were severely damaged. No documentation of the meteorological instruments or their locations within the observatory was preserved (Zaiki et al. 2006).

Additionally, the original atmospheric pressure data from May 1845 to September 1848 are highly uncertain, and clearly show anomalous variations (Zaiki et al. 2006). Therefore, the pressure data for this period have not been used in this reanalysis.

d. Dejima (1848–1858)

Dejima (32.7°N, 129.9°E, 8m above MSL) was an artificial island in Nagasaki, where the Dutch were the only Westerners allowed to maintain a trading factory on the Japanese mainland during the 17th–19th centuries. The earliest systematic meteorological observation on Dejima was made by J. C. Blomhoff, who was the director of the Dutch Trading Post in Nagasaki in 1819. Philipp Franz von Siebold took over from the Blomhoff in 1825. Since then, the Dutch established a meteorological network in their colonies, and, fortunately, the Dutch were allowed to continue their activities after the country's opening in 1854.

The records used in this analysis were collected by O. G. J. Mohnicke (1845–1851) and Mohnicke's assistants J. A. G. A. L. Bassle (1845–1848) and F. C. Lucas (1848–1852). Mohnicke also participated in meteorological observations in the Dutch East Indies (present-day Indonesia). Later, J. K. van den Broek (1854–1857), J. L. C. Pompe van Meerdervoort (1857–1862) and others took over the observations.

The data from October 1848 to July 1856 and November 1857 to December 1858 were published in the Royal Netherlands Meteorological Institute (KNMI) Yearbook (1855, 1856, available from the KNMI Library, P.O. Box 201, 3730 AE De Bilt, Netherlands), and Pompe (1866a,b). The observation elements were temperature, pressure, humidity, precipitation, wind speed, wind direction, and cloud cover. The observation times were 6:00, 9:00, 15:30, and 22:00 local time (for barometric pressure and cloud cover only). A thermometer was placed outside the north wall of the second floor of the factory chief's house. It was placed 1 m from the wall and 8 m above mean sea level. The barometer was 8 m above mean sea level and was located in the room of the house, while the thermometer was placed outside. The atmospheric pressure was measured in millimeters with a mercury siphon barometer (Können et al. 2003).

e. Nagasaki Hospital (1871–1877)

The Nagasaki Dutch Hospital (37 m above MSL) was established on September 20, 1861 by physician Pompe van Meerdervoort on a hill approximately 500 m southeast of Dejima (now Nita-Sako Elementary School). After Pompe left Japan in November 1862, observations were resumed by pharmacist A. J. C. Geerts in November 1871, and then taken over by physician W. K. M. van Leeuwen van Duivenbode in November 1874 (Können et al. 2003).

Daily observation records from November 1871 to December 1877 were reported in the KNMI Yearbook (1875, 1876, 1877). The observation schedules were 07:00, 12:00, and 18:00 local time until October 1874 and 07:00, 12:00, and 19:00 local time thereafter. Atmospheric pressures were measured in millimeters, but no information about the instruments was available.

f. Yokohama (1864–1865)

The record for December 1864–November 1865, located at 35.45°N, 139.67°E, 3 m above MSL, was collected by P. Mourier, a French Navy doctor working at the Yokohama Ironworks. His published records (Mourier 1866) contain pressure in millimeters and temperature readings (maximum and minimum temperatures) for 7:00, 10:00, 16:00, and 22:00 local time, where Yokohama local time is 9:20 ahead of UTC, and humidity and precipitation were also recorded. It was reported that the instruments were calibrated in Paris in May 1864, and the observations were carefully collected at a place north of the house in the yard, 2.0 m above the ground, and shaded by double umbrellas (Zaiki et al. 2006).

g. Philippine Stations (1891–1944)

Historical station data were collected in the Philippines between 1891 and 1944. There were 40 stations from 1891 to 1941 and 17 from 1942 to 1944. Observations were recorded in the Boletin mensual del Observatorio Meteorológico de Manila published by the Manila Observatory from 1891 to 1901 (Akasaka 2014), the Monthly Bulletin of the Philippine Weather Bureau published by the Manila Central Observatory from 1901 to 1940 (Kubota and Chan 2009), the Monthly Meteorological Investigation in South Seas published by the CMO from 1940 to 1941, and the Monthly Bulletin of Philippine Weather Reports published by the 22nd meteorological field members of the Japanese Army from 1942 to 1944 (Kobayashi and Yamamoto 2013). Surface pressure data were recorded from 1 to 24 hours depending on the stations and adjusted to sea level pressure from 1891 to 1941, while station pressure data were recorded from 1942 to 1944.

h. Western North Pacific Typhoons (1892–1944)

Tropical cyclone tracks and intensities were recorded in Eighty years of typhoon tracks published by the Central Weather Bureau of Taiwan from 1892 to 1939 (Hsu et al. 1973), Geophysical Review published by the CMO from 1901 to 1939, and Trajectories of Tropical Cyclones (Wadachi 1952) from 1940 to 1944.

B. SST Perturbations

A COBE-SST2 perturbation has been introduced for use in various studies with climate models. As described below, there are several favorable characteristics in the perturbations: 1) each set of perturbed SSTs is randomly constructed, but the values are continuous in space and time; 2) the perturbed SSTs include uncertainties originating due to ocean eddies; 3) any number of perturbation members can be constructed; and 4) the sea ice concentration is also perturbed consistently with the SST perturbations. In addition, the magnitudes of the SST perturbations are comparable to the analysis errors estimated based on objective analysis theory. Hirahara et al. (2014) demonstrated that the theoretically computed analysis errors properly represent the uncertainties due to the spatiotemporal sampling density.

In the construction of the SST perturbation, it is assumed that the perturbed components can be decomposed by empirical orthogonal functions (EOFs), which are used to reconstruct the interannual components of the monthly SSTs of COBE-SST2 (Hirahara et al. 2014) as

$$\mathbf{t}(t_k) = \mathbf{s}(t_k)\mathbf{\Lambda}^{1/2}\mathbf{F}, \tag{B1}$$

where $\mathbf{t}(t_k)$ is a vector containing the global SSTs for the month t_k, $\mathbf{s}(t_k)$ is the normalized score, $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues, and $\mathbf{F}$ is a matrix of the eigenvectors. COBE-SST2 was obtained by using a variational minimization method, and the analysis errors, $\mathbf{P}^a$, were simultaneously obtained by

$$\mathbf{P}^a = \mathbf{F}(\mathbf{\Lambda}^{-1} + \mathbf{F}^t\mathbf{H}^t\mathbf{R}^{-1}\mathbf{H}\mathbf{F})^{-1}\mathbf{F}^t, \tag{B2}$$

where $\mathbf{H}$ is a bilinear interpolation operator from the grid point to an observation location, and $\mathbf{R}$ is a diagonal matrix of the observation error variances (Hirahara et al. 2014). A convenient mathematical characteristic of EOF is that EOFs are mutually independent vectors, and the degrees of freedom of the EOFs are generally much smaller than the number of spatial grid points. The leading 133 leading EOF modes were used in COBE-SST2, and, the diagonal components of $\mathbf{P}^a$, that is, analysis errors, were stored in the archive.

An official COBE-SST2 is defined on a monthly $1^\circ \times 1^\circ$ grid without the use of satellite observations; therefore, SST variations originating due to oceanic mesoscale eddies are introduced into the perturbation. The eddy component was also constructed continuously in space and time. The magnitudes of the eddy term are equivalent to the sigma of the differences between COBE-SST2 and NCEP OISST version 2 with in situ and satellite observations (Reynolds et al. 2002), averaged monthly from 1982 to 2008. The SST perturbation was added to the monthly SST analysis defined on a grid.

The SST perturbation $\boldsymbol{\delta}^i(t_k)$ of member $i = (1, 2, \cdots, N)$ for month t_k is defined by the following equation:

$$\boldsymbol{\delta}^i(t_k) \quad = \quad \mathbf{b}^i(t_k)\boldsymbol{\Gamma}^{1/2}\mathbf{G}\mathbf{F} + \mathbf{e}^i(t_k), \qquad \text{(B3)}$$

where $\mathbf{b}^i$ is a time-varying normalized score vector, and the $\boldsymbol{\Gamma}$ and $\mathbf{G}$ are matrices containing the eigenvalues and eigenvectors of $(\mathbf{A}^{-1} + \mathbf{F}^t\mathbf{H}^t\mathbf{R}^{-1}\mathbf{H}\mathbf{F})^{-1}$ in Eq. (B2), and $\mathbf{e}$ is the eddy component. Note that $\boldsymbol{\Gamma}$ is diagonal. The first term on the right side of Eq. (B3) indicates that the perturbation can be computed in proportion to the analysis errors. When mutually independent $\mathbf{b}^i$ and $\mathbf{e}_j$ are given, N perturbations can be obtained. Finally, the i-th perturbed SST is obtained by adding $\boldsymbol{\delta}^i(t_k)$ to COBE-SST2.

Matrix $(\mathbf{A}^{-1} + \mathbf{F}^t\mathbf{H}^t\mathbf{R}^{-1}\mathbf{HF})^{-1}$ was computed from SST observations in each calendar month, and this term is a core part of the COBE-SST2 analysis errors. In the case of d4PDF, where spatially fixed and temporally invariant analysis errors were used, the matrix was constructed using virtual observations uniformly distributed in space (Ishii and Mori 2020).

The time series of the score vector $\mathbf{b}^i$ is given by a m-th order autoregressive model:

$$\mathbf{b}^i(t_k) = \sum_{j=1}^{m} c_j^i \mathbf{b}^i(t_{k-j}) + \boldsymbol{\epsilon}^i(t_k). \tag{B4}$$

where c_j^i is the j-th weight for mode i, and $\boldsymbol{\epsilon}^i$ is a Gaussian noise vector. For spatiotemporal continuity of the perturbations, c_1^i should be close to 1. Assuming that SST uncertainties are related to natural variability and seasonality, a set of c_j was determined from past SST changes. In this study, the maximum of j is set to 12, but $c_j^i = 0$ for $i > 8$ and $j > 1$, considering the computational stability of the AR models and the estimation of c_j^i using the least squares method. The independence of the perturbations depends on the randomness of $\boldsymbol{\epsilon}^i$.

The uncertainty from transient SST changes such as mesoscale eddies,

$\mathbf{e}^i(t_k)$, in Eq. (B3) is modeled using a different first-order autoregressive model as follows:

$$\mathbf{e}^i(t_k) = \boldsymbol{\sigma}_e\{\alpha\boldsymbol{\eta}^i(t_k) + (1-\alpha)\boldsymbol{\eta}^i(t_{k-1})\}, \tag{B5}$$

where the vector $\boldsymbol{\sigma}_e$ is the standard deviation of the difference between COBE-SST2 and OISST, and $\boldsymbol{\eta}^i_{ik}$ is the spatially smoothed Gaussian noise with a horizontal scale of 450 km (Ebuchi and Hanawa 2000). The temporal continuity of $\mathbf{e}$ was ensured by $\alpha = 0.8$, which is a typical 1-month lag correlation of the SST differences. The westward migration of the eddy was not considered in Eq. (B5).

Sea ice concentration (SIC) perturbations were computed consistently with SST perturbations. The ICE-SST relationship given by statistically defined quadratic equations in COBE-SST2 (Hirahara et al. 2014) was used. Fig. 17

Examples of SST and SIC perturbations are shown in Fig. B1. In 1935, the observations were insufficiently sampled in contrast to recent decades. As a result, the perturbations are quite large, and those for SST are spatially between ±1 K and between ±0.2 for SIC. The spatial continuities of the perturbations are confirmed in the figure. Moreover, positive and negative perturbations extend to a few hundred or thousands of kilometers in space, and perturbed Niño3 SSTs (5°S – 5°N, 150°W – 90°W) often persist for a season or more in each member.

Thus far, several applications have demostrated the effectiveness of the SST and SIC perturbations. One experiment aimed to reproduce the 2010 Russian heat wave (Dole et al. 2011) by MRI-AGCM3.2 simulations forced by observed SSTs. The event is thought to result from natural variability (Dole et al. 2011; Watanabe et al. 2013); therefore, AGCM simulations can probabilistically reproduce the event. In the MRI-AGCM3.2 simulation with the SST perturbations, some of the 10-member runs starting from the same initial condition successfully reproduced the event, while all 10-member simulations with different initial conditions instead of the SST perturbations failed. Another experiment in which 100-member ensemble simulations with the perturbations were performed with the MIROC5 AGCM (Watanabe et al. 2010), yielded more unbiased distributions of atmospheric states than those starting the model integration with different initial conditions. From these experiments, it is inferred that AGCM simulations with SST perturbations standardize the atmospheric response to observed SSTs. Furthermore, SST-perturbation ensemble climate simulations have been applied to a large ensemble and high-resolution climate simulation database, known as d4PDF, which has been used for making policy decisions for future climate change in many centers and institutes across Japan (Ishii and Mori 2020).

References

Akasaka, I., 2014: Weather observation records from late 19th century to early 20th century in the Philippines. *Senshu University Institute of Humanities monthly bulletin*, **273**, 1–15.

Allan, R., and T. Ansell, 2006: A new globally complete monthly historical gridded mean sea level pressure dataset (hadslp2): 1850–2004. *Journal of Climate*, **19(22)**, 5816–5842.

Allan, R., P. Brohan, G. P. Compo, R. Stone, J. Luterbacher, and S. Brönnimann, 2011: The international atmospheric circulation reconstructions over the earth (ACRE) initiative. *Bull. Amer. Meteor. Soc.*, **92(11)**, 1421–1425.

Allan, R. P., K. M. Willett, V. O. John, and T. Trent, 2022: Global changes in water vapor 1979–2020. *J. Geophys. Res. Atmos.*, **127(12)**, e2022JD036728.

Amano, R., 1953: *Meteorological observations in Tokyo during the period 1838 -1855. Journal of the University of Mercantile Marine, A4, 167–194.* (in Japanese).

Ansell, T. J., P. D. Jones, R. J. Allan, D. Lister, D. E. Parker, M. Brunet, A. Moberg, J. Jacobeit, P. Brohan, N. Rayner, and others, 2006: Daily mean sea level pressure reconstructions for the european–north atlantic region for the period 1850–2003. *Journal of Climate*, **19(12)**, 2717–2742.

Atkinson, G. D., and C. R. Holliday, 1977: Tropical cyclone minimum sea level pressure/maximum sustained wind relationship for the western North Pacific. *Mon. Wea. Rev.*, **105(4)**, 421–427.

Baker, A. J., K. I. Hodges, R. K. Schiemann, and P. L. Vidale, 2021: Historical variability and lifecycles of north atlantic midlatitude cyclones originating in the tropics. *Journal of Geophysical Research: Atmospheres*, **126(9)**, e2020JD033924.

Chan, D., and P. Huybers, 2019: Systematic differences in bucket sea surface temperature measurements among nations identified using a linear-mixed-effect method. *Journal of Climate*, **32(9)**, 2569–2589.

Chan, D., E. C. Kent, D. I. Berry, and P. Huybers, 2019: Correcting datasets leads to more homogeneous early-twentieth-century sea surface warming. *Nature*, **571(7765)**, 393–397.

Compo, G. P., L. C. Slivinski, J. S. Whitaker, P. D. Sardeshmukh, C. McColl, P. Brohan, R. Allan, X. Yin, R. Vose, L. J. Spencer, L. Ashcroft, S. Bronnimann, M. Brunet, D. Camuffo, R. Cornes, T. A. Cram, R. Crouthamel, F. Dominguez-Castro, J. E. Freeman, J. Gergis,

B. S. Giese, E. Hawkins, P. D. Jones, S. Jourdain, A. Kaplan, J. Kennedy, H. Kubota, F. L. Blancq, T. Lee, A. Lorrey, J. Luterbacher, M. Maugeri, C. J. Mock, K. Moore, R. Przybylak, C. Pudmenzky, C. Reason, V. C. Slonosky, B. Tinz, H. Titchner, B. Trewin, M. A. Valente, X. L. Wang, C. Wilkinson, K. Wood, and P. Wyszynski, 2019: The International Surface Pressure Databank version 4. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. NOAA National Centers for Environmental Information. Accessed 12 April 2021.

Compo, G. P., J. S. Whitaker, and P. D. Sardeshmukh, 2006: Feasibility of a 100-year reanalysis using only surface pressure data. *Bull. Amer. Meteor. Soc.*, **87**, 175–190.

Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E. Gleason, R. S. Vose, G. Rutledge, P. Bessemoulin, and others, 2011: The twentieth century reanalysis project. *Quart. J. Roy. Meteor. Soc.*, **137(654)**, 1–28.

Dole, R., M. Hoerling, J. Perlwitz, J. Eischeid, P. Pegion, T. Zhang, X.-W. Quan, T. Xu, and D. Murray, 2011: Was there a basis for anticipating the 2010 russian heat wave? *Geophys. Res. Lett.*, **38(6)**.

Ebuchi, N., and K. Hanawa, 2000: Mesoscale eddies observed by TOLEX-ADCP and TOPEX/POSEIDON altimeter in the Kuroshio recirculation region south of Japan. *J. Oceanogr.*, **56(1)**, 43–57.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, **9(5)**, 1937–1958.

Fiorino, M., 2002: , Analysis and forecasts of tropical cyclones in the ECMWF 40-year reanalysis (ERA-40). Extended abstract of 25t Conference on Hurricanes and Tropical Meteorology. Technical report.

Folland, C. K., and D. E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, **121**, 319–367.

Freeman, E., S. D. Woodruff, S. J. Worley, S. J. Lubker, E. C. Kent, W. E. Angel, D. I. Berry, P. Brohan, R. Eastman, L. Gates, and others, 2017: ICOADS Release 3.0: a major update to the historical marine climate record. *Int. J. Climatol.*, **37(5)**, 2211–2232.

Fujibe, F., 2008: *A study on the frequency and long-term changes of extreme meteorological events.* available at https://kaken.nii.ac.jp/file/KAKENHI-PROJECT-18340145/18340145seika.pdf (in Japanese).

Fujibe, F., 2010: Assessment of uncertainty in estimating the return period of extreme rainfalls. *TENKI*, **57(7)**, 17–30. (Japanese text with English abstract).

Fujibe, F., 2012: Evaluation of background and urban warming trends based on centennial temperature data. *Papers in Meteorology and Geophysics*, **63**, 43–56.

Fujibe, F., and J. Matsumoto, 2022: Regional features related to timing of maximum and minimum precipitation during the warm season in japan. *Geographical review of Japan series B*, **95(2)**, 55–68.

Fujita, M., R. Mizuta, M. Ishii, H. Endo, T. Sato, Y. Okada, S. Kawazoe, S. Sugimoto, K. Ishihara, and S. Watanabe, 2019: Precipitation changes in a climate with 2-k surface warming from large ensemble simulations using 60-km global and 20-km regional atmospheric models. *Geophys. Res. Lett.*, **46(1)**, 435–442.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125(554)**, 723–757.

Gates, W. L., J. S. Boyle, C. Covey, C. G. Dease, C. M. Doutriaux, R. S. Drach, M. Fiorino, P. J. Gleckler, J. J. Hnilo, S. M. Marlais, and others, 1999: An overview of the results of the atmospheric model intercomparison project (amip i). *Bulletin of the American Meteorological Society*, **80(1)**, 29–56.

Harada, Y., H. Kamahori, C. Kobayashi, H. Endo, S. Kobayashi, Y. Ota, H. Onoda, K. Onogi, K. Miyaoka, and K. Takahashi, 2016: The jra-55 reanalysis: Representation of atmospheric circulation and climate variability. *J. Meteor. Soc. Japan*, **94(3)**, 269–302.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, and others, 2020: The era5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146(730)**, 1999–2049.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **32**, 57–75.

Hsu, J. H., C. J. Yao, P. Y. Wang, F. S. Yang, J. L. Shin, and S. C. Hsu, 1973: *Eighty years of typhoon tracks 1892-1977. Central Weather Bureau, Taiwan.*

Huang, B., V. F. Banzon, E. Freeman, J. Lawrimore, W. Liu, T. C. Peterson, T. M. Smith, P. W. Thorne, S. D. Woodruff, and H.-M.

Zhang, 2015: Extended reconstructed sea surface temperature version 4 (ERSST. v4). Part I: Upgrades and intercomparisons. *J. Climate*, **28(3)**, 911–930.

Huang, B., P. W. Thorne, T. M. Smith, W. Liu, J. Lawrimore, V. F. Banzon, H.-M. Zhang, T. C. Peterson, and M. Menne, 2016: Further exploring and quantifying uncertainties for extended reconstructed sea surface temperature (ERSST) version 4 (v4). *J. Climate*, **29(9)**, 3119–3142.

Hunt, B., E. Kalnay, E. Kostelich, E. Ott, D. Patil, T. Sauer, I. Szunyogh, J. Yorke, and A. Zimin, 2004: Four-dimensional ensemble Kalman filtering. *Tellus A: Dynamic Meteorology and Oceanography*, **56(4)**, 273–277.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230(1-2)**, 112–126.

Ishii, M., Y. Fukuda, S. Hirahara, S. Yasui, T. Suzuki, and K. Sato, 2017: Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, **13**, 163–167.

Ishii, M., and N. Mori, 2020: d4PDF: large-ensemble and high-resolution climate simulations for global warming risk assessment. *Prog. Earth Planet Sci.*, **7**, 58.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865–879.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, and others, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77(3)**, 437–472.

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, , and M. Saunby, 2011: Reassessing biases and other uncertainties in sea-surface temperature observations measured in situ since 1850, part 1: measurement and sampling uncertainties. *J. Geophys. Res. Oceans*, **116(D14103)**.

Kleeman, R., J. P. McCreary Jr, and B. A. Klinger, 1999: A mechanism for generating ENSO decadal variability. *Geophys. Res. Lett.*, **26(12)**, 1743–1746.

Knapp, K. R., H. J. Diamond, J. P. Kossin, M. C. Kruk, and C. J. Schreck, 2018: International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4. [CSV subset] . Accessed 9 May 2021.

Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, and others, 2015: The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48.

Kobayashi, S., and H. Yamamoto, 2013: Wartime meteorological observations of the Japanese military in East Asia and the collection of its neglected data. *Hist. Geogr.*, **267**, 82–98. (in Japanese).

Können, G., M. Zaiki, A. Baede, T. Mikami, P. Jones, and T. Tsukahara, 2003: Pre-1872 extension of the japanese instrumental meteorological observation series back to 1819. *J. Climate*, **16(1)**, 118–131.

Kubota, H., 2012: History of meteorological observation in the South Seas Bureau. *Chizujoho*, **32(3)**, 13–17. (in Japanese).

Kubota, H., and J. C. Chan, 2009: Interdecadal variability of tropical cyclone landfall in the Philippines from 1902 to 2005. *Geophys. Res. Lett.*, **36(12)**.

Kubota, H., Y. Kosaka, and S.-P. Xie, 2016: A 117-year long index of the Pacific-Japan pattern with application to interdecadal variability. *Int. J. Climatol.*, **36(4)**, 1575–1589.

Kubota, H., J. Matsumoto, M. Zaiki, T. Tsukahara, T. Mikami, R. Allan,

C. Wilkinson, S. Wilkinson, K. Wood, and M. Mollan, 2021: Tropical cyclones over the western north Pacific since the mid-nineteenth century. *Climate Change*, **164(3)**, 1–19.

Kurihara, Y., M. A. Bender, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.*, **121(7)**, 2030–2045.

Laloyaux, P., E. de Boisseson, M. Balmaseda, J.-R. Bidlot, S. Broennimann, R. Buizza, P. Dalhgren, D. Dee, L. Haimberger, H. Hersbach, and others, 2018: Cera-20c: A coupled reanalysis of the twentieth century. *Journal of Advances in Modeling Earth Systems*, **10(5)**, 1172–1195.

Magnusson, L., and E. Källén, 2013: Factors influencing skill improvements in the ECMWF forecasting system. *Mon. Wea. Rev.*, **141(9)**, 3142–3153.

Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, J. B. R. M. E. Lonnoy, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Z. (eds.), 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Matsumoto, J., 2013: *Long-term changes of precipitation intensity in monsoon Asia.* https://kaken.nii.ac.jp/file/KAKENHI-PROJECT-23240122/23240122seika.pdf (in Japanese).

Pompe van Meerdervoort, J. L. C., 1866a: *Verslag over de Gouvernements Geneeskundige Dienst op het eiland Desima en in Japan over 1857 en 1858. Geneeskundig Tijdschrift voor Nederlands Indie, Nieuwe Serie, 2, 495–572.* Available from National Museum of Ethnology Library, P.O. Box 212, 2300AE, Leiden, Netherlands.

Pompe van Meerdervoort, J. L. C., 1866b: *Weerkundige waarnemingen, gedaan op het eiland Desima in Japan gedurende het jaar 1858. Desima Publishers, 53 pp.* Available from KNMI Library, P.O. Box 201, 3730 AE De Bilt, Netherlands.

Mizuta, R., A. Murata, M. Ishii, H. Shiogama, K. Hibino, N. Mori, O. Arakawa, Y. Imada, K. Yoshida, T. Aoyagi, and others, 2017: Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bull. Amer. Meteor. Soc.*, **98(7)**, 1383–1398.

Mizuta, R., H. Yoshimura, H. Murakami, M. Matsueda, H. Endo, T. Ose, K. Kamiguchi, M. Hosaka, M. Sugi, S. Yukimoto, and others, 2012:

Climate simulations using MRI-AGCM3.2 with 20-km grid. *J. Meteor. Soc. Japan Ser. II*, **90**, 233–258.

Mourier, P., 1866: *Yokohama (Japon). Observations Méetéorologiques faites pendant l'anne 1865. Annuaire de la Société Météorologique de France 14, 105–119.* (in French).

Murakami, H., R. Mizuta, and E. Shindo, 2012: Future changes in tropical cyclone activity projected by multi-physics and multi-sst ensemble experiments using the 60-km-mesh mri-agcm. *Climate dynamics*, **39**, 2569–2584.

Nosaka, M., M. Ishii, H. Shiogama, R. Mizuta, A. Murata, H. Kawase, and H. Sasaki, 2020: Scalability of future climate changes for +1.5K, +2K, and +4K global warming in NHRCM large ensemble simulations. *Prog. Earth Planet Sci.*, **7(1)**, 1–13.

Nyomura, Y., 2002: Meteorological observation at lighthouses. *KISHO*, **(3)**, 36–42. (in Japanese).

Okuda, Y., 1981: On heavy rainfall and flood on september 4–11, meiji 29 (1896) —- heavy rainfall that brought the highest water level to lake biwa since the meiji era. *Annual review of Yashirogakuin University*, **(20)**, 34–65. (in Japanease).

Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, and others, 2007: The JRA-25 reanalysis. *J. Meteor. Soc. Japan*, **85(3)**, 369–432.

Osborn, T. J., and P. D. Jones, 2014: The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth Syst. Sci. Data*, **6**, 61–68.

Poli, P., H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. Tan, C. Peubey, J.-N. Thépaut, and others, 2016: Era-20c: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, **29(11)**, 4083–4097.

Rayner, N., P. Brohan, D. Parker, C. Folland, J. Kennedy, M. Vanicek, T. Ansell, and S. Tett, 2006: Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: The hadsst.

Rayner, N., D. E. Parker, E. Horton, C. K. Folland, L. V. Alexander, D. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res. Oceans*, **108(D14)**.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15(13)**, 1609–1625.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *Journal of climate*, **20(22)**, 5473–5496.

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

Schloemer, R. W., 1954: *Analysis and synthesis of hurricane wind patterns over Lake Okeechobee, Florida*. Number 31. US Department of Commerce, Weather Bureau.

Schneider, U., A. Becker, P. Finger, E. Rustemeier, and M. Ziese, 2010: , GPCC Full Data Monthly Product Version 2020 at 2.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data. Technical report.

Schreck, C. J., K. R. Knapp, and J. P. Kossin, 2014: The impact of best track discrepancies on global tropical cyclone climatologies using IBTrACS. *Mon. Wea. Rev.*, **142(10)**, 3881–3899.

Shiga Prefecture, 2020: Overview of Lake Biwa Flood in 1896. `https://www.pref.shiga.lg.jp/suigaijyouhou/gaiyou/105764.html`. (in Japanese).

Slivinski, L., G. Compo, P. Sardeshmukh, J. Whitaker, C. McColl, R. Allan, P. Brohan, X. Yin, C. Smith, L. Spencer, and others, 2021: An evaluation of the performance of the twentieth century reanalysis version 3. *J. Climate*, **34(4)**, 1417–1438.

Slivinski, L. C., G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh, B. S. Giese, C. McColl, R. Allan, X. Yin, R. Vose, H. Titchner, and others, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the twentieth century reanalysis system. *Quart. J. Roy. Meteor. Soc.*, **145(724)**, 2876–2908.

Suzuki, A., and H. Kikuchihara, 1984: Method fbr estimating the return period of annual maximum daily precipitation considering about extreme values. *TENKI*, **31(3)**, 179–189. (in Japanease).

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93(4)**, 485–498.

Thorne, P. W., D. E. Parker, S. F. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res. Atmos.*, **110(D18)**.

Titchner, H. A., and N. A. Rayner, 2014: The met office hadley centre sea ice and sea surface temperature data set, version 2: 1. sea ice concentrations. *J. Geophys. Res. Atmos.*, **119(6)**, 2864–2889.

Trenberth, K. E., and L. Smith, 2005: The mass of the atmosphere: A constraint on global analyses. *J. Climate*, **18(6)**, 864–875.

Uppala, S. M., P. Kållberg, A. J. Simmons, U. Andrae, V. D. C. Bechtold, M. Fiorino, J. Gibson, J. Haseler, A. Hernandez, G. Kelly, and others, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131(612)**, 2961–3012.

Wadachi, K., 1952: *Trajectories of Tropical Cyclones, 1940–1950. The Central Meteorological Observatory Tokyo Japan, Tokyo.*

Watanabe, M., H. Shiogama, Y. Imada, M. Mori, M. Ishii, and M. Kimoto, 2013: Event attribution of the August 2010 Russian heat wave. *SOLA*, **9**, 65–68.

Watanabe, M., T. Suzuki, R. O'ishi, Y. Komuro, S. Watanabe, S. Emori, T. Takemura, M. Chikira, T. Ogura, M. Sekiguchi, and others, 2010:

Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity. *J. Climate*, **23(23)**, 6312–6335.

Watanabe, S., 2012: Flood damage in 1896 - system of meteorological observation and damage according to administrative divisions -. *Water Science*, **55(6)**, 111–133.

Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132(5)**, 1190–1200.

Woodruff, S. D., R. J. Slutz, R. L. Jenne, and P. M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, **68(10)**, 1239–1250.

Woodruff, S. D., S. J. Worley, S. J. Lubker, Z. Ji, J. Eric Freeman, D. I. Berry, P. Brohan, E. C. Kent, R. W. Reynolds, S. R. Smith, and others, 2011: ICOADS Release 2.5: extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967.

Yoshida, K., M. Sugi, R. Mizuta, H. Murakami, and M. Ishii, 2017: Future changes in tropical cyclone activity in high-resolution large-ensemble simulations. *Geophysical Research Letters*, **44(19)**, 9910–9917.

Zaiki, M., G. Können, T. Tsukahara, P. Jones, T. Mikami, and K. Matsumoto, 2006: Recovery of nineteenth-century Tokyo/Osaka meteorological data in Japan. *Int. J. Climatol.*, **26(3)**, 399–423.

Zou, X., and Q. Xiao, 2000: Studies on the initialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.*, **57(6)**, 836–860.

List of Figures

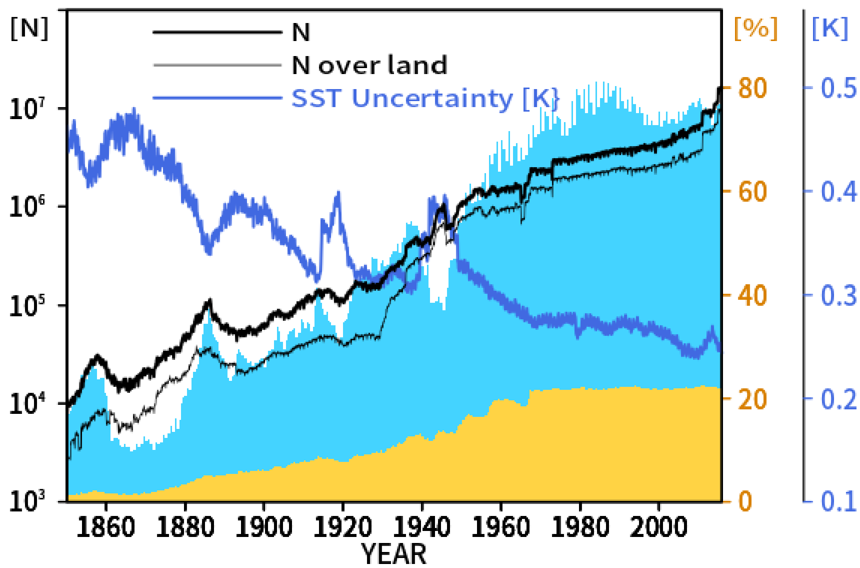

Fig. 1. Time series 1850 – 2015 of the monthly number of surface pressure observations over the globe (land and oceans; thick black) and over land (thin black), monthly global spatial data coverage (%; shades), and monthly SST analysis errors (K; blue). Data coverage is separately shown for the global land (orange) and oceans (light blue), and values at the upper side of the blue area indicate the total coverage. Coverage was estimated from the data distributions on a global 5° × 5° grid. Scales for each element are shown by colored vertical axes.

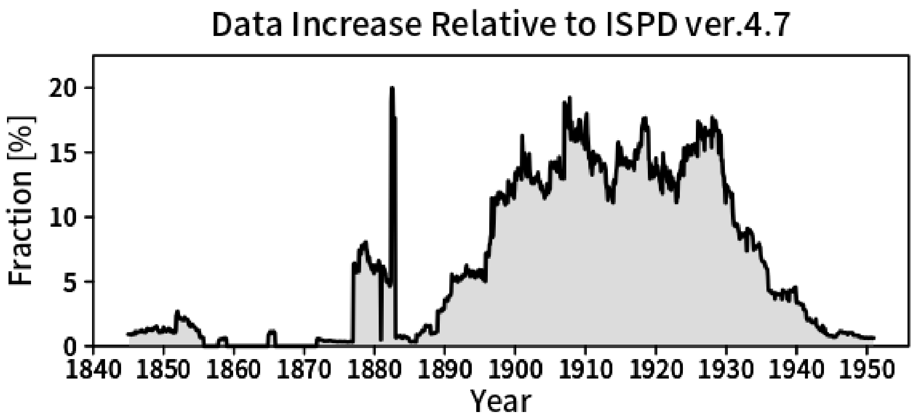

Fig. 2. Percentage increase in the monthly number of records of the Ps observations obtained through domestic data rescue in Japan, relative to the ISPD version 4.7.

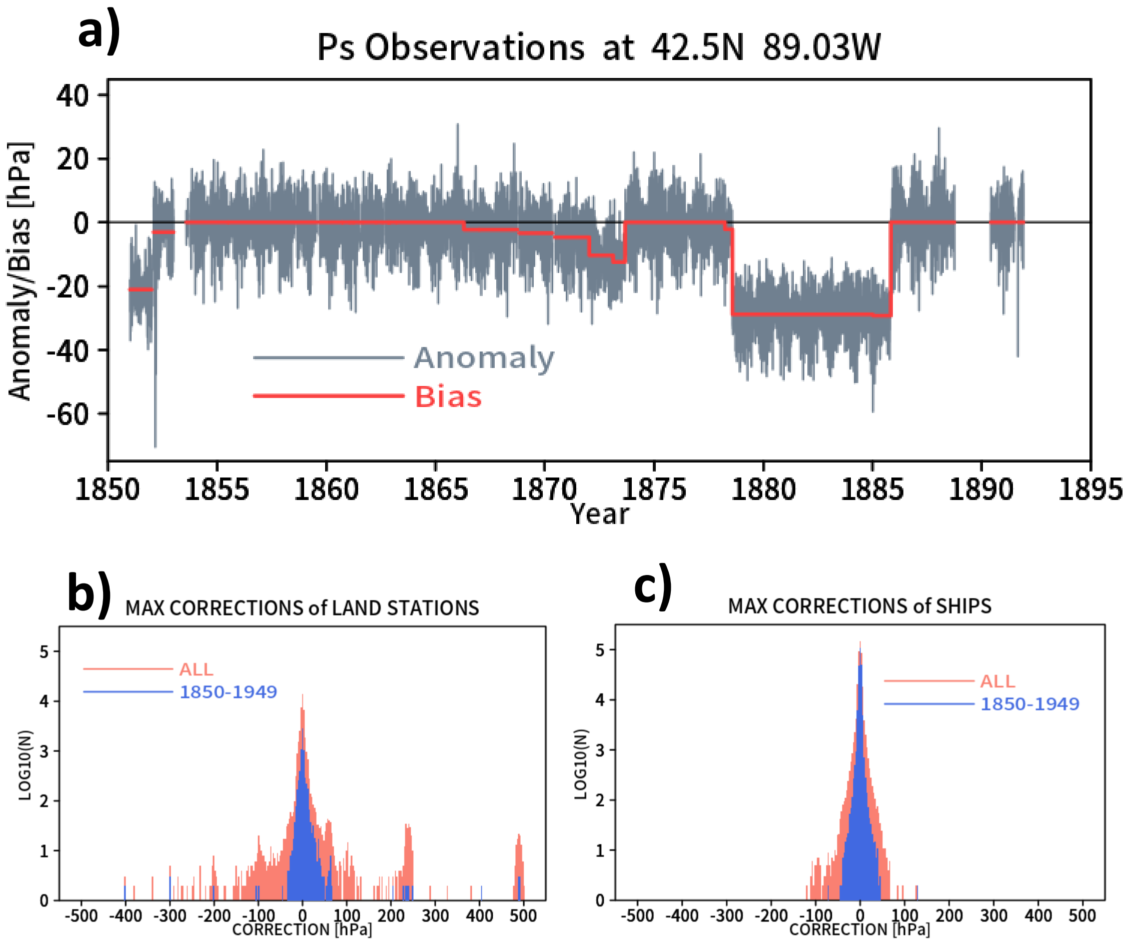

Fig. 3. a) Time series of sub-daily surface pressure anomalies (hPa; gray) relative to JRA-55 and identified biases (hPa; red) at a station located at (42.50°N, 89.03°W). The biases are defined as long-term mean deviations from the monthly JRA-55 climatology whose base period is 1961–2005. b) Histogram of the maximum bias correction amounts, that is, bias multiplied by -1, of the land surface observations for all and 1850–1949 periods, and c) those for ship observations. In b) and c), the bin width is 3 hPa. One maximum bias value per station and ship was counted.

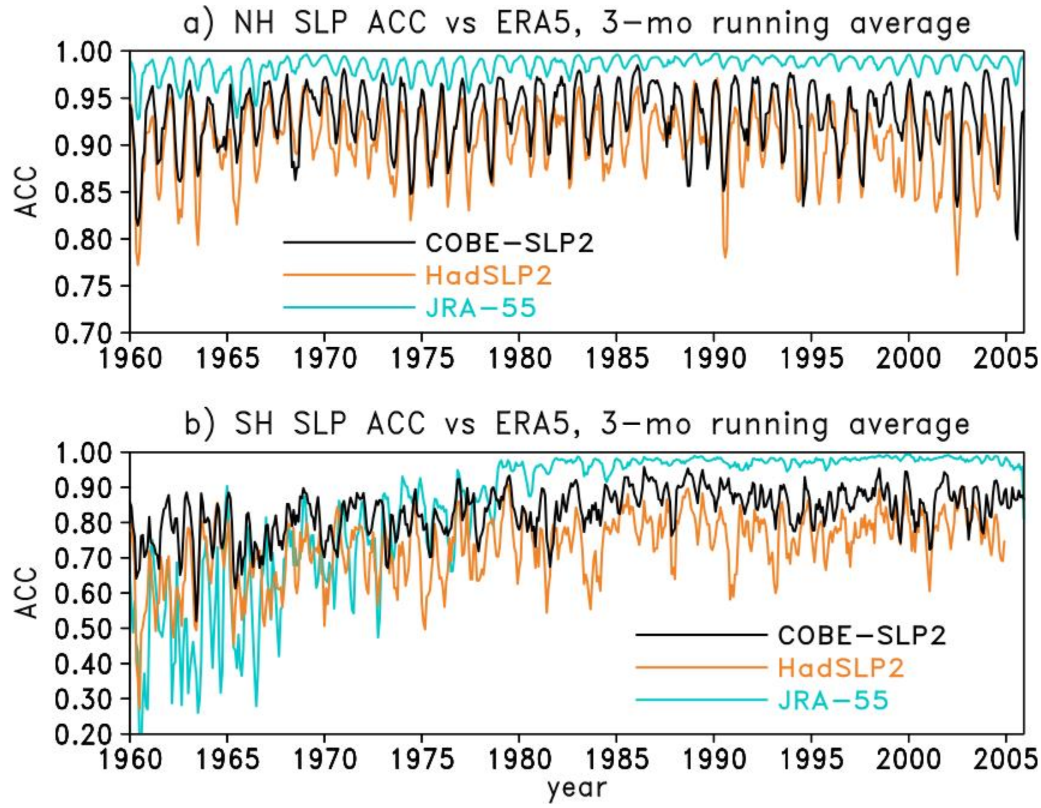

Fig. 4. a) Time series of the monthly SLP anomaly correlation coefficient (ACC) of COBE-SLP2 (black), HadSLP2 (orange), and JRA-55 (light blue) over the Northern Hemisphere, compared with ERA5. Three month running averaging is applied to the time series. b) Same as (a) but for the Southern Hemisphere.

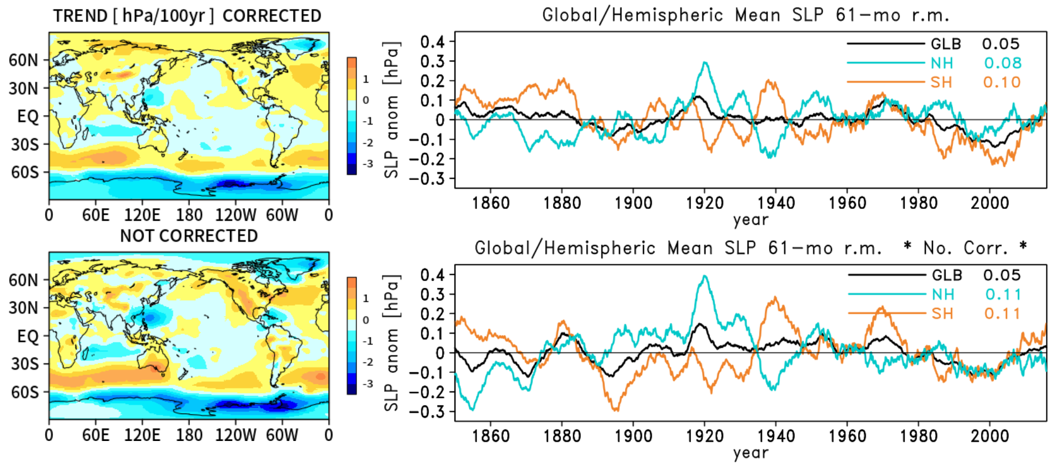

Fig. 5. Linear trends (hPa/100yrs; left; 1850–2005) and time series of 61-month running mean global (black) and hemispheric (NH: water blue, SH: orange) averages (hPa; right) of the SLP statistical analysis, COBE-SLP2, with (top) and without (bottom) the Ps bias correction. The legend in the right panels shows the root mean squares of each time series.

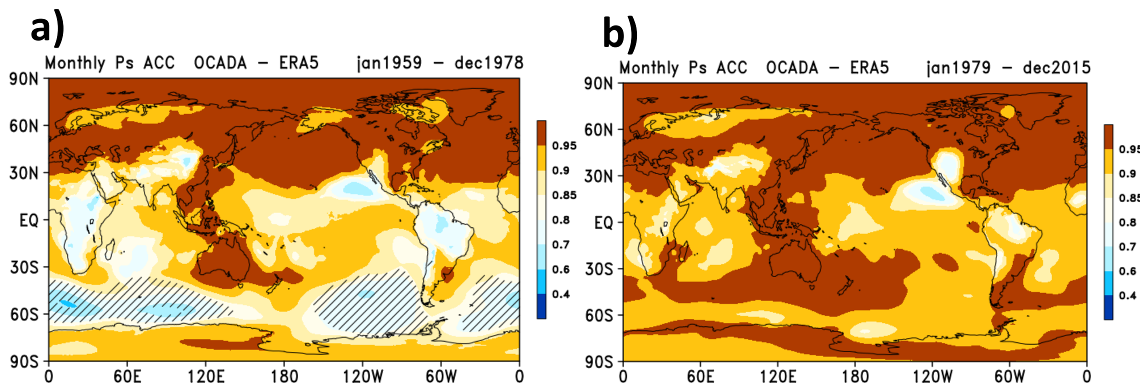

Fig. 6. Anomaly correlation coefficients of monthly surface pressure (Ps) between OCADA and ERA5 during a) 1959-1978 and b) 1979-2015. Hatched areas stand for “confidence” values less than 0.35, calculated according to Slivinski et al. (2021).

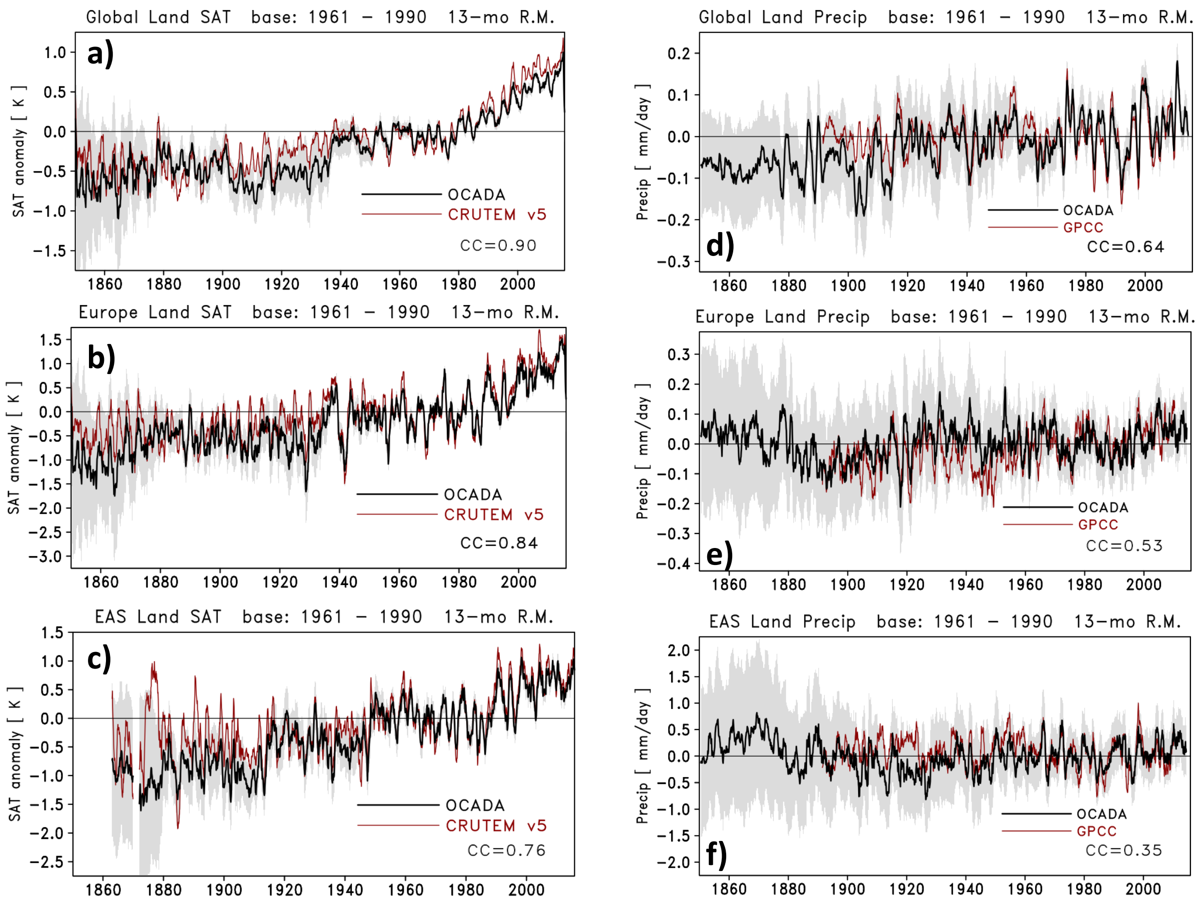

Fig. 7. Time series of monthly mean temperature (left) and precipitation (right) of OCADA (black) averaged over the globe (a, d), Europe (30°W – 60°E, 35°N – 70°N; b, e), and East Asia (130°E – 150°E, 25°N – 45°N; c, f). Station observations (red) of temperature (CRUTEM5) and precipitation (GPCC) are superimposed in left and right panels, respectively. Only land data are used for averaging. CRUTEM5 contains missing grids, and hence SATs collocated between CRUTEM5 and OCADA were compared. GPCC is available since 1891. Two-sigma spreads of the reanalysis are shown by gray shading. Temperature and precipitation anomalies are defined as deviations from the 1961–1990 averages. Thirteen-month running averaging is applied to all time series. CC in each panel denotes the correlation coefficient between the two curves.

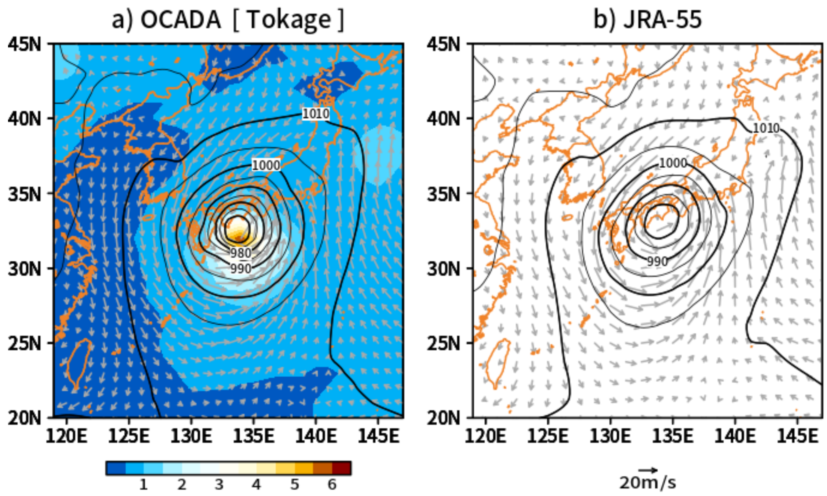

Fig. 8. Sample of TC Tokage in OCADA at 06UTC on October 20, 2004 (left), compared with that of JRA-55 (right). Contours indicate sea level pressure (hPa), and the interval is 5 hPa. Shading in the left panel indicates spreads (hPa) of the LETKF ensemble members. Wind vectors (ms^{-1}) at the 950 hPa isobar surface are shown. There is a small difference in the horizontal resolution between the two reanalyses: 1° × 1° in OCADA and 1.25° × 1.25° in the JRA-55.

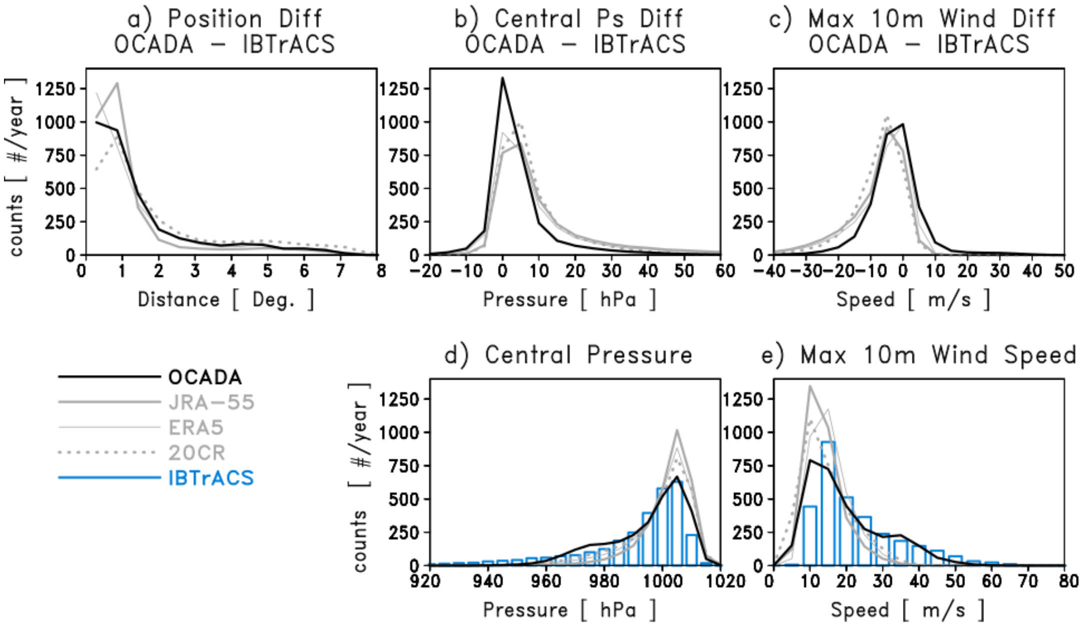

Fig. 9. Histograms of differences in a) TC positions (geodetic degree), b) TC central pressures (hPa), and c) maximum 10-m wind speed differences (ms^{-1}) between reanalyses and observations, d) central pressures, and e) 10 wind speeds. Black line shows OCADA defined on the $1^{\circ} \times 1^{\circ}$ grid, and the gray lines indicate the other reanalyses of different resolution: JRA-55 with $1.25^{\circ} \times 1.25^{\circ}$ horizontal resolution, ERA5 with $0.25^{\circ} \times 0.25^{\circ}$, and 20CR with $1^{\circ} \times 1^{\circ}$. Blue bars in d) and e) denote IBTrACS. Bin sizes for position, pressure, and wind errors are 1 °, 5 hPs, and 5 ms^{-1}, respectively. The statistics were based on 6-hourly TC data from 1979 to 2015 in the global region.

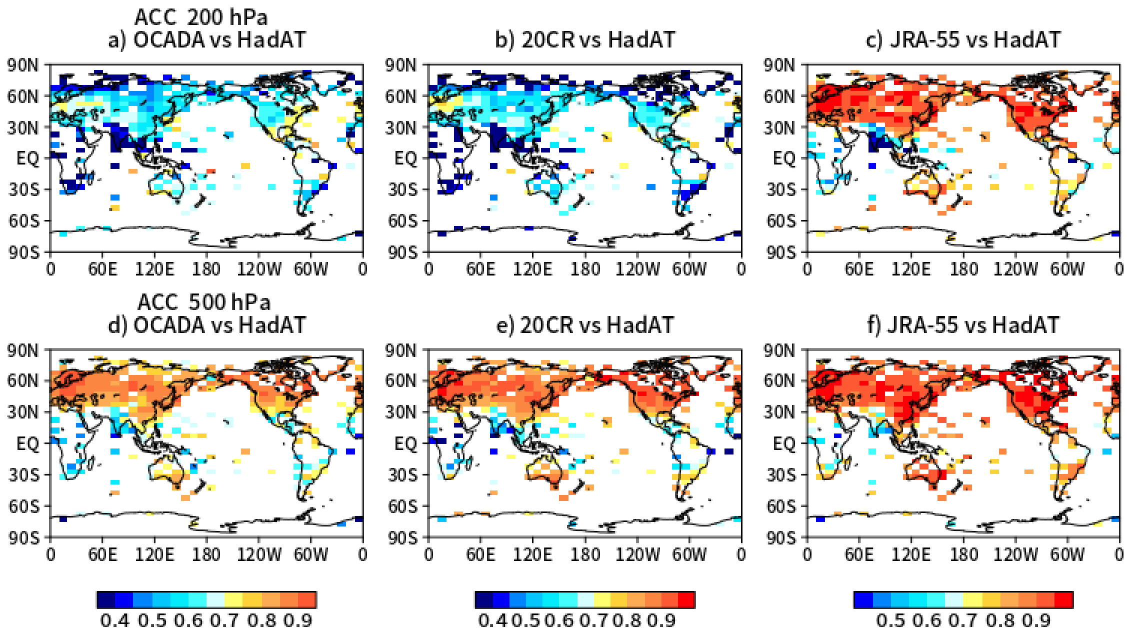

Fig. 10. Anomaly correlation coefficients of OCADA (a, d), 20CR (b, e), and JRA-55 (c, f) against HadAT at 200 hPa (upper) and 500 hPa (lower) isobar surfaces. The statistical period is 1958 – 2012. The reference period is 1966–1995, the same as for HadAT.

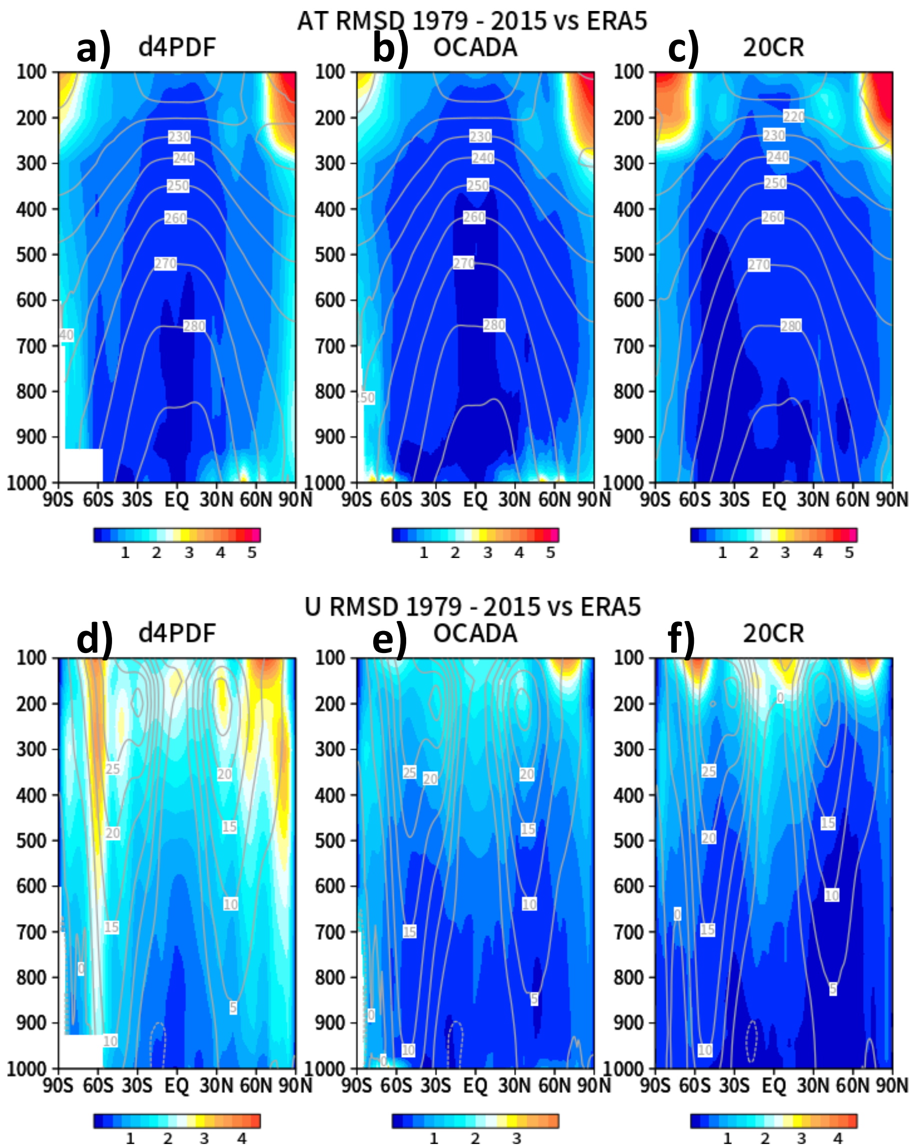

Fig. 11. RMSDs of monthly zonal mean upper air temperatures (K; a, b, c) and zonal winds (ms^{-1}; d, e, f) of d4PDF (a, d), OCADA (b, e), and 20CR (c, f) compared with ERA5. Contours shown in each panel denote the respective climatology. The averaging period is from 1979 to 2015. In the case of d4PDF, the vertical resolution near the surface is coarser than the others.

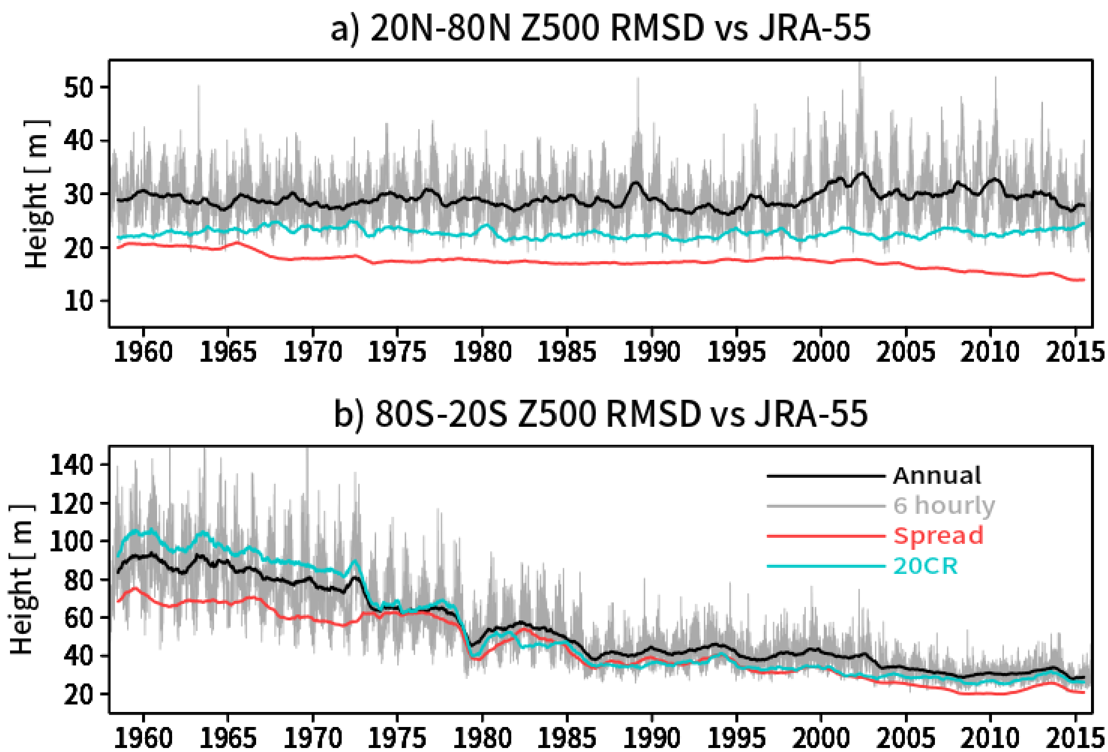

Fig. 12. Time series of annual-mean (black) and 6-hourly (gray) RMSDs of geopotential heights at 500 hPa averaged in a) 20°N – 80°N and in b) 80°S –20°S between OCADA and JRA-55. Red line indicates annual-mean spreads of the 80 ensemble members. Annual mean RMSDs for 20CR are also shown by blue lines.

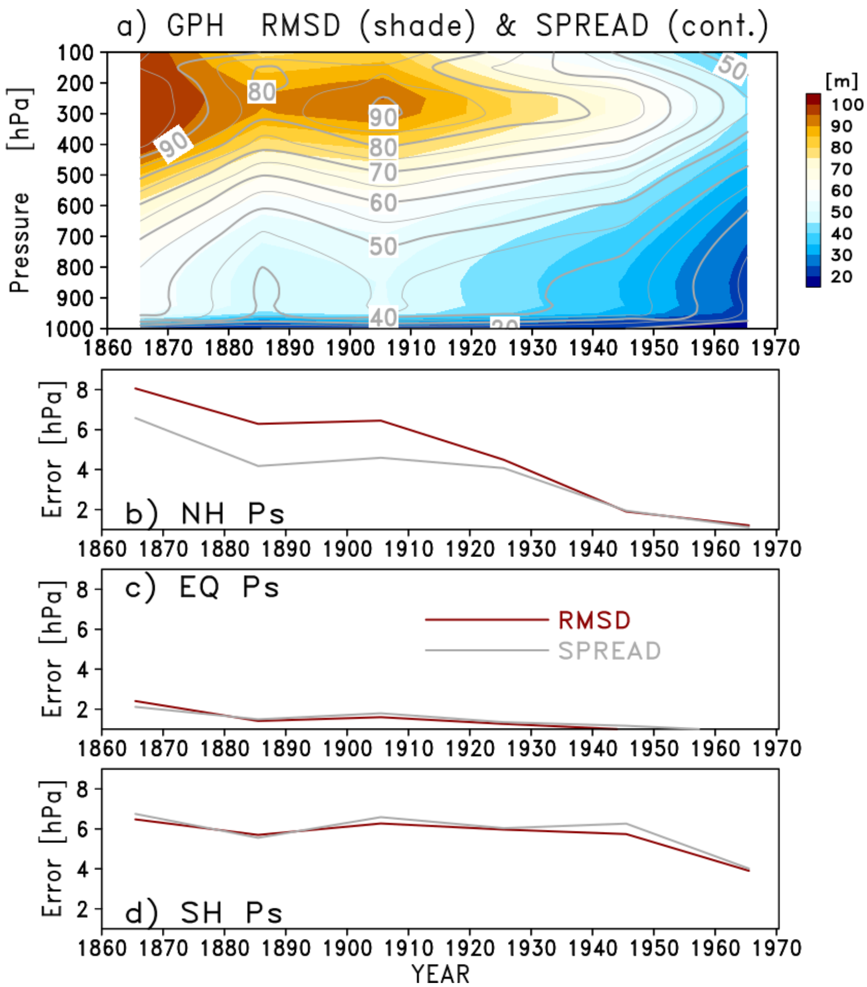

Fig. 13. Comparison between root-mean-square differences (RMSDs) and ensemble spreads for a) 6-hourly geopotential height profiles (GPH; m), b) 6-hourly surface pressures (Ps; hPa) averaged over the Northern Hemisphere (0° – 90°N) as a function of time, c) same as (b) but for low latitudes (30°S – 30°N), and d) same as b) but for the Southern Hemisphere (90°S – 0°). The time series are drawn with RMSDs and spreads defined every 20 years from 1865 to 1965. Shading in (a) and red line in (b, c, d) indicate RMSDs, and gray contour in (a) and gray line in (b, c, d) denote the ensemble spreads.

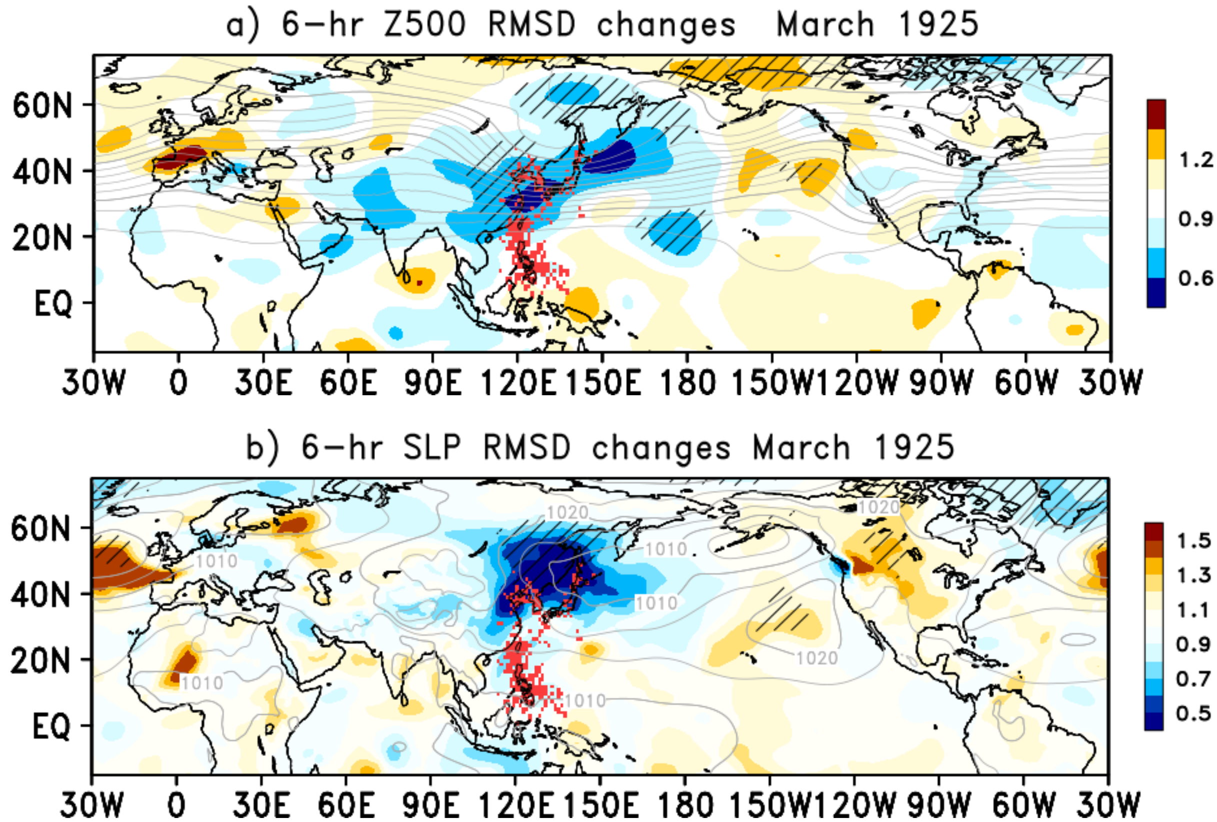

Fig. 14. Relative changes in RMSDs (shade) of a) geopotential height at 500 hPa and b) SLP in March 1925 between the pseudo-reanalyses with and without the additional observations in East and South-East Asia. Values less than one means the RMSDs are reduced due to the additional SLP observations. Red dots represent the positions of the additional observations. Hatched areas indicate that RMSD changes due to the additional observations are greater than one sigma of 6-hourly deviations from the 5-day running averages of OCADA in March 2001. The monthly averages in March 2001 are shown by gray contours.

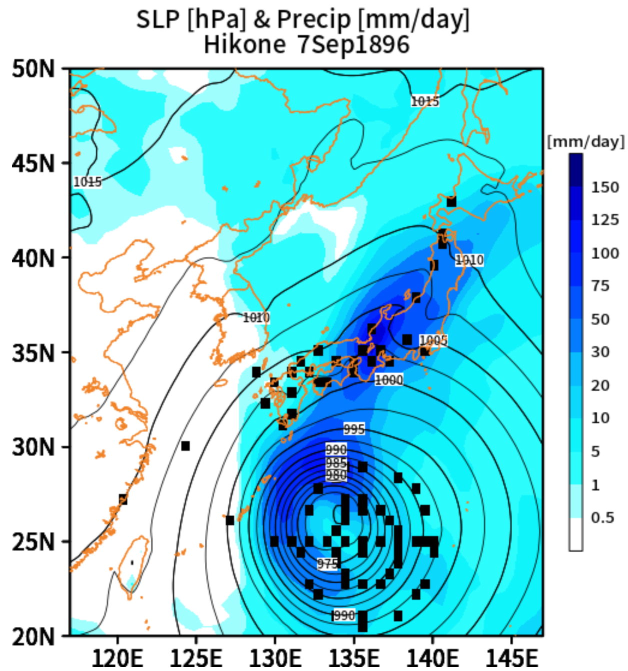

Fig. 15. A heavy precipitation event at Hikone (35°N, 136°E) on September 7, 1896. Shading denotes precipitation amount (mm/day). Contour indicates the daily mean SLP, and the interval is 5 hPa. Black square means positions of Ps observations available during the day. All POIMT observations at eight time steps during the day are plotted around the TC center.

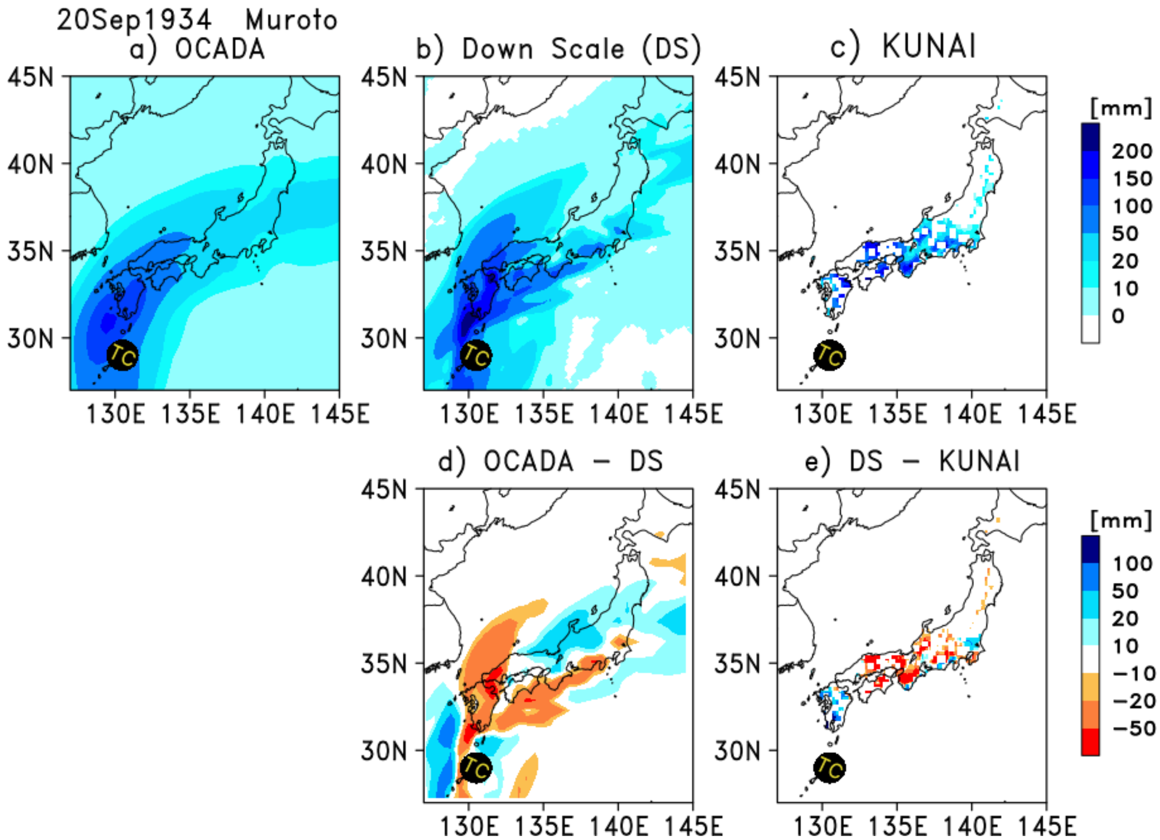

Fig. 16. Daily precipitation amounts (mm/day) of a) OCADA, b) the downscaling experiment, and c) KUNAI observations on September 20, 1934. Mark TC indicates the position of Typhoon Muroto at 12UTC of the day. In (c), KUNAI observations were averaged in daily $0.2^\circ \times 0.2^\circ$ boxes, and blank areas (white) denote no data. Differences between (a) and (b) and between (b) and (c) are shown in d) and e), respectively. The precipitation amount is that accumulated during 24 hours from 00JST of the day.

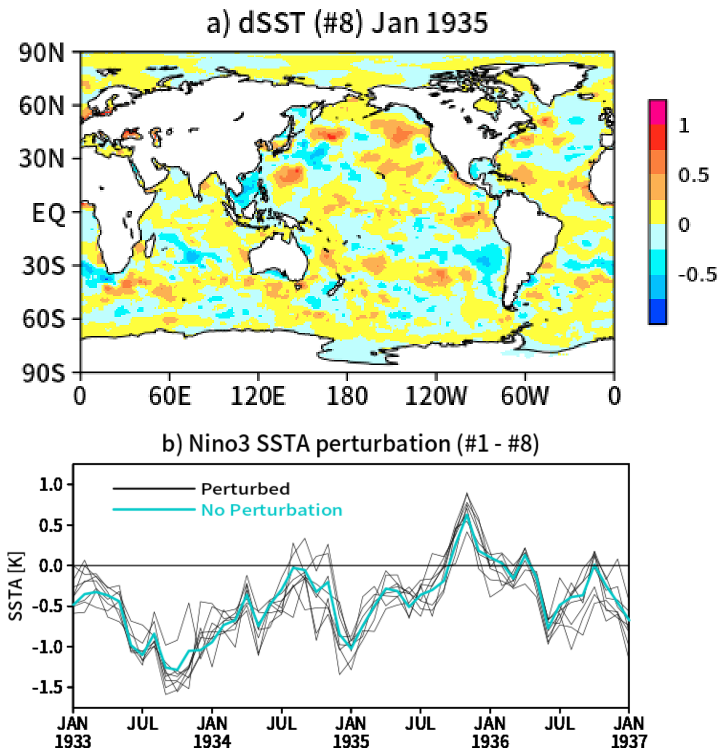

Fig. B1. Sample No. 8 of a) SST perturbation (K) and b) time series of perturbed (black) and unperturbed (blue) Niño3 SST anomalies (K) from eight members of the SST perturbation.

List of Tables

Table 1. Ps and TC observations added to ISPD and IBTrACS, respectively. Column "Stn.s" contains the number of stations and individual TCs.

No.	Station name	Elem.	Stn.s	Period	Agency/Institute	References
1	CMO stations	Ps	66	1890–1945	CMO	CMO original books and Monthly Reports
2	Lighthouses	Ps	37	1877–1882	Meiji government	Nyomura (2002)
3	Edo	Ps	1	1838–1855	Tokugawa Shogunate	REIKEN-KOUBO
4	Dejima	Ps	1	1848–1858	Dutch Trading Post	KNMI Yearbooks
5	Nagasaki Hospital	Ps	1	1871–1877	Nagasaki hospital	KNMI Yearbooks
6	Yokohama	Ps	1	1864–1865	Yokohama ironworks	Mourier (1866)
7	Philippine stations	Ps	57	1891–1944	Spain, USA, Japan	Akasaka (2014); Kubota and Chan (2009); Kobayashi and Yamamoto (2013)
8	WNP Typhoons	TC	778	1892–1944	Central Weather Bureau of Taiwan	Wadachi (1952), Hsu et al. (1973)

Table 2. Averages and standard deviations (SD) of the position (deg.), central pressure (hPa), and maximum 10-m wind(ms^{-1}) errors shown in Fig. 9a, b, and c, respectively. The columns of "severe" show the number of assimilated severe TCs relative to that of the observations (%). The thresholds for severity are 982.5 hPa and 27.5 ms^{-1} for central pressure and maximum 10-m wind, respectively.

	Position	Central Pressure			Max 10-m Wind		
Reanalysis	SD	mean	SD	severe	mean	SD	severe
OCADA	1.8	0.9	9.8	85	-0.5	8.6	81
JRA-55	1.6	2.3	14.9	15	-1.7	9.8	18
ERA5	1.8	1.8	13.6	38	-1.3	9.5	28
20CR	2.4	1.6	10.0	50	-1.6	7.9	36