ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 42, FEBRUARY 2025, 373–381

• **Original Paper** •

Discovering Climate Change during the Early 21st Century via Wasserstein Stability Analysis

Zhiang XIE[1], Dongwei CHEN[*2], and Puxi LI[*3]

[1]*Department of Earth and Space Sciences*, *Southern University of Science and Technology*, *Shenzhen* 518055, *Guangdong*, *China*

[2]*School of Mathematical and Statistical Sciences*, *Clemson University*, *Clemson* 29634, *SC*, *USA*

[3]*State Key Laboratory of Severe Weather*, *Chinese Academy of Meteorological Sciences*, *China Meteorological Administration*, *Beijing* 100081, *China*

(Received 7 November 2023; revised 21 May 2024; accepted 19 June 2024)

ABSTRACT

Climate change is an essential topic in climate science, and the accessibility of accurate, high-resolution datasets in recent years has facilitated the extraction of more insights from big-data resources. Nonetheless, current research predominantly focuses on mean-value changes and largely overlooks changes in the probability distribution. In this study, a novel method called Wasserstein Stability Analysis (WSA) is developed to identify probability density function (PDF) changes, especially the extreme event shift and nonlinear physical value constraint variation in climate change. WSA is applied to the early 21st century and compared with traditional mean-value trend analysis. The results indicate that despite no significant trend, the equatorial eastern Pacific experienced a decline in hot extremes and an increase in cold extremes, indicating a La Niña-like temperature shift. Further analysis at two Arctic locations suggests sea ice severely restricts the hot extremes of surface air temperature. This impact is diminishing as the sea ice melts. By revealing PDF shifts, WSA emerges as a powerful tool to re-examine climate change dynamics, providing enhanced data-driven insights for understanding climate evolution.

Key words: climate change, probability density function, extremes, Wasserstein distance

Citation: Xie, Z., D. W. Chen, and P. X. Li, 2025: Discovering climate change during the early 21st century via Wasserstein stability analysis. *Adv. Atmos. Sci.*, **42**(2), 373−381, https://doi.org/10.1007/s00376-024-3324-6.

Article Highlights:

- A novel method called Wasserstein Stability Analysis (WSA) is introduced to detect changes in probability distributions.
- Shifts in extreme events such as La Nina-like events can be identified by WSA.
- The nonlinear physical constraints imposed by the sea ice can be detected well by WSA.

1. Introduction

Understanding the evolution of the climate system over time stands as a cornerstone in the realm of climate science. This imperative is emphasized in the findings of the Intergovernmental Panel on Climate Change Sixth Assessment Report (IPCC AR6), which illuminates the persistent warming trend in global mean surface air temperature (SAT) since the 20th century, predominantly driven by human influence. This has sparked extensive discussions regarding the trend of annual mean SAT in the context of global warming (Delworth and Knutson, 2000; Delworth et al., 2015; Li et al., 2015; Huang et al., 2017).

Yet, navigating the intricacies of climate change in an era characterized by the proliferation of expansive and complex datasets—often referred to as "big data"—poses profound difficulties. While mean values represent significant statistical features of datasets, encapsulating the overall energy conditions and adeptly filtering noise within low-quality datasets, the advent of the big data era demands novel methodologies. As datasets become more precise and finer in higher resolution, the challenge lies in unraveling underlying patterns, correlations, and deep insights—a task where the richness of the information transcends mere mean-value changes.

The early 21st century is marked by a negative phase of the Pacific Decadal Oscillation (PDO) and cooling of the

* Corresponding authors: Dongwei CHEN, Puxi LI
Emails: dongwec@g.clemson.edu, lipx@cma.gov.cn

Pacific Ocean (Douville et al., 2015), accompanied by La Niña-like Pacific temperature patterns (Kosaka and Xie, 2013). Concurrently, the Arctic sea ice experienced a persistent decline at a relatively high rate (Cai et al., 2021), exerting a significant impact on SAT. However, conventional trend or mean-value analysis struggles to detect extreme patterns like La Niña events or distinguish whether the warming trend is due to human influence or other factors like sea ice. In this context, the emergence of a novel mathematical tool named Wasserstein distance (W-distance), which is often used to quantify the distance and similarities between two probability distributions from optimal transport theory (Figalli and Glaudo, 2021), offers fresh insights.

While W-distance has found applications in climate science, such as detecting changes in forced Lorenz attractors (Robin et al., 2017), model evaluation (Vissio et al., 2020), oceanographic data analysis (Hyun et al., 2022), and data assimilation (Tamang, 2022), its potential in the physical analysis of big data in the real world remains relatively unexplored. Here, we apply W-distance to the physical analysis of climate data during the early 21st century. By leveraging the W-distance, we introduce a a novel method—Wasserstein Stability Analysis (WSA)—dedicated to uncovering probability density function (PDF) variability in climate change. This method, embracing signals like extremes and physical value constraints, transcends the confines of conventional mean-value analysis. Through this new method, we extract insights from the realm of big data that exceed mere averages or trends, thereby enriching our understanding of climate change dynamics.

The paper is organized as follows: In section 2, we describe the dataset and WSA algorithm. Optimal mass transportation and W-distance are also briefly introduced. In section 3, we apply WSA to the early 21st century. Through the new method, we obtain a clear equatorial eastern Pacific signal as highlighted in previous studies, and an evident physical value constraint change in the Arctic that has not yet been fully explored. Finally, in section 4, we draw conclusions and discuss ideas for future work.

2. Data and methods

2.1. *Dataset*

As the description of PDFs necessitates a relatively large sample size, daily datasets are used for W-distance calculation in this work. Specifically, the surface 2-m air temperature data are obtained from the ERA5 dataset (0.25° × 0.25°, single layer, hourly) from 1998 to 2013 (Hersbach et al., 2020), and the spatial range is global. To better represent the large-scale signal and accelerate the data processing, we regrid the original dataset into daily data with a spatial resolution of 2.5° × 2.5° by spatial and temporal averaging. We also tested the results with another observed sea surface temperature dataset (Fig. S1 in the ESM) .

The sea-ice data used in the study come from Hadley Centre Sea Ice and Sea Surface Temperature monthly dataset (HadISST), which has a 1° × 1° spatial resolution (Rayner et al., 2003). The time range is from 1998 to 2013 and the spatial range is still global.

2.2. *Wasserstein Stability Analysis*

Let $c(x,y)$ be the cost function of transporting unit mass from x to y. The W-distance $[\mathcal{W}_p(\mu,\nu)]$ is induced by the study of how to transport mass from one probability distribution μ to another distribution ν on the Euclidean space $\mathbb{R}^d$ in the cheapest way by letting the cost function $c(x,y)$ be the p–distance $|x-y|^p$ between x and y where $p \in [1,\infty)$:

$$\mathcal{W}_p(\mu,\nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y)\, \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}}, \qquad (1)$$

where $\pi \in \Pi(\mu,\nu)$ means π is a joint distribution of μ and ν. More details about W-distance can be found in the Appendix. In this work, we use earth mover's distance, i.e. $p = 1$.

In the null hypothesis of statistical analyses, we usually assume that the random variable follows the same distribution. Following this logic, we claim that the random variable has an unstable PDF if the W-distance between its current state and a reference state is out of a specific confidence interval, representing a significant change from one climate stage to another. Therefore, the significant PDF change signal is detected by testing W-distance, and the climate change details can be further explored. In this section, we describe the development of the novel method of WSA to evaluate the magnitude of PDF change quantitatively. The new method has two steps: W-distance test and PDF analysis.

(1) W-distance test. Given two samples X and Y, both series are first standardized individually by the following scalar:

$$\frac{x-\hat{\mu}}{\hat{\sigma}}, \qquad (2)$$

where $\hat{\mu}$ is the mean value and $\hat{\sigma}$ is the standard deviation. Then, the W-distance $\mathcal{W}_1(X,Y)$ is obtained after standardization. Note that standardization is necessary since the geographic difference may lead to a large variation in W-distance. For example, the high surface temperature variation at high latitudes usually leads to a larger W-distance than its counterpart in tropical zones. Nevertheless, it does not necessarily mean the PDF change at high latitudes is stronger than that at low latitudes.

Our significance test is based on the Monte Carlo test (Haurwitz and Brier, 1981; Xie et al., 2017). The null hypothesis H_0 of WSA is that the PDF discrepancy between X and Y can be explained by the white noise. Subsequently, we generate new samples X_N and Y_N by adding white noise to X and Y, and calculate $\mathcal{W}_1(X_N,Y_N)$. Such an operation is performed 500 times. Under the null hypothesis with confidence level $1-\alpha$, $\mathcal{W}_1(X,Y)$ is within the confidence interval of these W-distances. If not, H_0 is rejected. The α in this study has been set as 0.01 and we perform the W-distance test for

all locations individually. A significance level of $\alpha = 0.05$ yields similar results (Fig. S2). More details about the W-distance test can be found in Algorithm 1 in the Appendix.

(2) PDF analysis. After detecting the significant PDF change zones where the W-distance is significant, we plot two PDFs and the corresponding change for each significant change zone, and then explore the climate change mechanism that makes W-distance significant. The next section will use two examples to demonstrate how it functions.

3. Results

During 1998–2013, the global mean surface temperature experienced a slower increase, which is referred to as the "global warming hiatus" in earlier studies (Kosaka and Xie, 2013). Recent research indicates that there was no true "hiatus" during that time, but the global surface temperatures were still rising, albeit more slowly (Simmons et al., 2017). This phenomenon suggests that physical processes other than the anthropogenic warming trend play a role in climate change, and the detection of PDF change may provide us with new information for this period. Therefore, by comparing the performance of WSA with traditional trend analysis, we can evaluate the similarities and differences between these two approaches, which could advance our understanding of climate science and inform future research in this area.

Figure 1a depicts the W-distance between SAT PDFs during 1998–2005 and 2006–13. There are two main large-scale significant W-distance signals appearing in Fig. 1a: the equatorial eastern Pacific and the Barents–Kara seas. The equatorial eastern Pacific has the greatest W-distance, with values ranging from 0.08 to 0.24; nevertheless, there is no discernible trend (Fig. 1b). Another noticeable W-distance signal can be seen over the Barents–Kara seas, which is located inside a significant warming trend zone in the Arctic (Fig. 1b). The following analysis will mainly focus on signals over the equatorial eastern Pacific and Barents–Kara seas to highlight the information that W-distance can provide about climate change. In practice, the Arctic region exhibits a significant temperature variation. As a result, the regional average may ignore some information from individual grid cells (see section 4 for detail). To avoid this, we selected three special sites to perform further PDF analysis:

(1) Site A (0°, 90°W): a site with a significant W-distance signal but without a significant linear trend;

(2) Sites B (79°N, 66°E) and C (79°N, 131°E): sites that both exhibit significant linear trends; however, Site B has a significant W-distance while C does not.

Figure 2 shows the SAT PDFs and their changes from

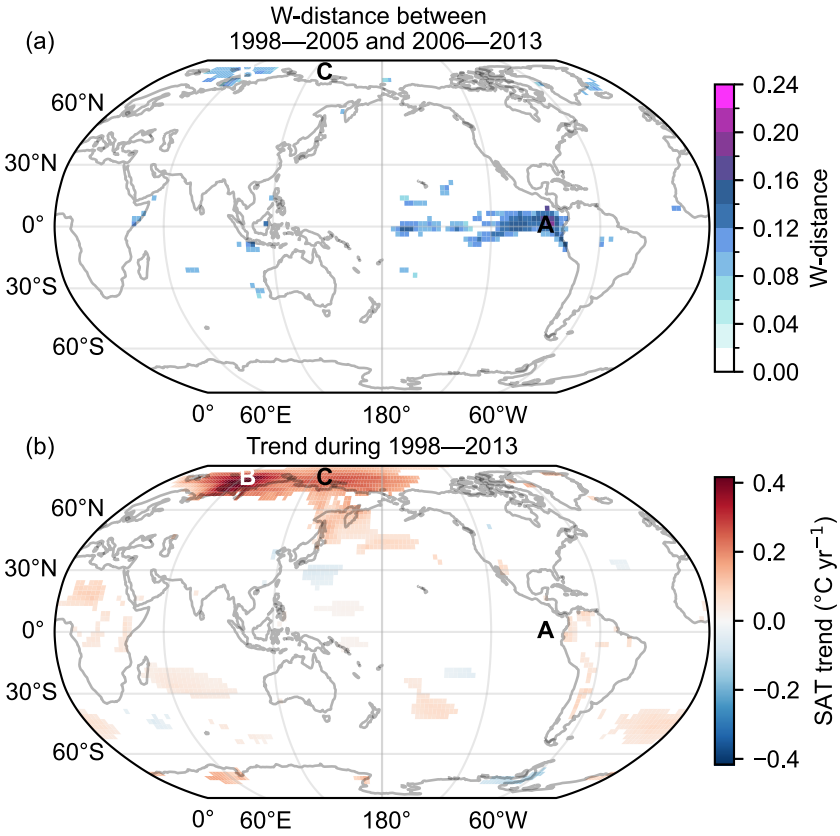

Fig. 1. (a) W-distance between SAT PDFs during 1998–2005 and 2006–13. (b) Linear trend during 1998–2013. Only statistically significant regions (99% significance level) are shown. Three sites, A (90°W, 0), B (66°E, 79°N), and C (131°E, 79°N), have been chosen for further analysis.

the period 1998–2005 to the period 2006–13 at Site A. The mean value and standard deviation stay relatively stable between the two periods (23.0°C vs. 23.2°C for the mean value; 2.2°C vs. 2.1°C for the standard deviation). However, in the second period, the hot extremes (more than two standard deviations, roughly 27°C) decrease by more than ten days per year, while the mild hot events (near one standard deviation) increase by roughly 30 days per year. Meanwhile, the cold extremes (less than −2 standard deviations, 19°C) increase by more than ten days per year, while the mild cold events (near −1 standard deviation) decrease by more than 30 days per year. The hot extreme decreasing and cold extreme increasing indicate a La Niña-like SAT shift.

Many early studies pointed out the cooling of the Pacific Ocean, which partially offset the global warming trend, correlating with the negative phase of the PDO in the early 21st century (Kosaka and Xie, 2013; England et al., 2014; Douville et al., 2015). Given the insignificant trend of observed SAT in the equatorial eastern Pacific, those studies hypothesized a potential impact of the equatorial eastern Pacific on the warming slowdown, and subsequently validated it through numerical modeling experiments to ascertain the significance of La Niña-like sea surface temperature changes. In our study, with WSA, we observe that in the equatorial eastern Pacific, while the overall PDF of SAT exhibits a warming shift, the decrease in hot extremes alongside the increase in cold extremes counteracts the main PDF shift (Fig. 2a), resulting in an insignificant mean-value change but a significant PDF shape change (indicated by high W-distance). The results demonstrate WSA's capability as an analytical tool to detect extreme shifts, enabling the rapid identification of crucial signals during the early 21st century.

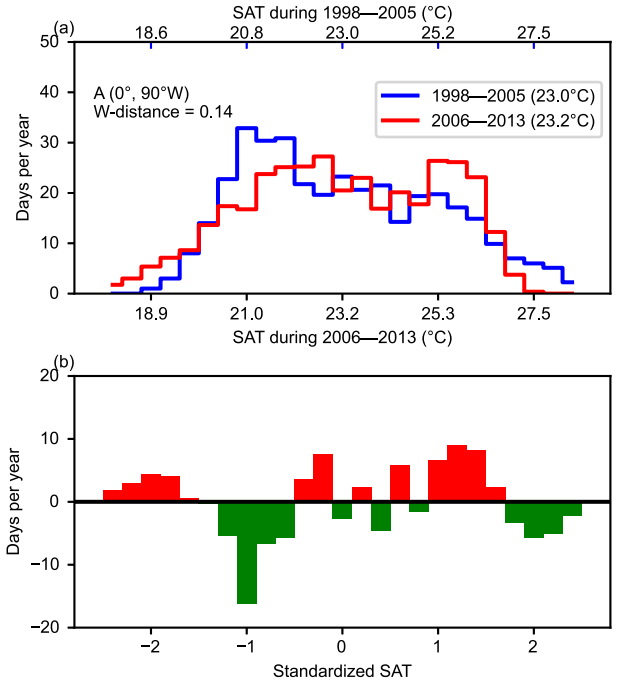

Fig. 2. (a) SAT PDFs for the periods 1998–2005 (blue) and 2006–13 (red). The bottom tick labels represent observed SAT values from 2006 to 2013, whereas the top tick labels represent observed SAT values from 1998 to 2005. (b) The difference between two PDFs of standardized SAT series (2006–13 minus 1998–2005) at site A (0°, 90°W). The frequency has been standardized as days per year. The mean value for the time period is shown by bracketed numbers in the legend.

Site B is another region with a significant W-distance, suggesting a strong PDF change from the period 1998–2005 to the period 2006–2013 (Figs. 3a and c). At Site B, the SAT shows significant warming from −10.2°C to −6.4°C. Meanwhile, the standard deviation also decreases, indicating an elimination of SAT variation. Since Site B is located within a zone covered by sea ice, the maximum SAT is strongly constrained to being near the frozen line in both periods, and thus this causes a sharp frequency jump near the frozen line (Fig. 3a). However, as the mean value increases and the standard deviation decreases during the second period, the frozen line at Site B shifts significantly to the left. As a result, the frequency peak of hot events shifts to the left. Meanwhile, the slope near the frozen line becomes flatter, leading to a smoother PDF shape. In contrast, Site C exhibits a change that follows a similar pattern but is considerably smaller in extent in terms of mean value, standard deviation, and frozen line shift (Figs. 3b and d). Consequently, the W-distance at Site C is smaller than that at Site B and fails the 99% significance test. Actually, the existence of sea ice strongly limits the maximum surface temperature and the corresponding SAT near the frozen line. Therefore, the sharp peak shape of the SAT PDF appears in regions covered by sea ice. Nevertheless, regions with normal ocean conditions are expected to have a more Gaussian distribution. Therefore, the alteration at Site B and Site C suggests that the variability in sea-ice concentration may play a role in the SAT PDF change from the first to the second period.

By checking the full-time series in two periods, we can better understand the PDF change at Site B and Site C. From 1998–2005 to 2006–2013, the interval with ±1 standard deviation shrinks and moves upward at Site B because of the change in mean value and standard deviation (Fig. 4a). Clearly, this change leads to most of the SAT variabilities being below +1 standard deviation and also more cold extremes. For Site C, we can see a similar change but with a much smaller magnitude (Fig. 4b). As discussed, this difference can be attributed to the sea ice and its strong temperature constraint. Figure 5 shows the sea-ice concentration change between 1998–2005 and 2006–2013. We can see that Site B is in a zone where sea ice is melting the most. Hence, the PDF change at Site B is the most intense. As a comparison, Site C also experiences sea-ice loss but to a relatively small extent. As the temperature increases, the sea-ice concentration decreases, and the effect of sea-ice restriction gradually fades. This is why Site C also displays a similar PDF change but with a less significant level. In summary, the significant W-distance change at Site B implies the removal of physical temperature constraint imposed by sea ice that causes the sharp peak near the frozen line in the PDF. Actually, one of the most significant developments of IPCC AR6

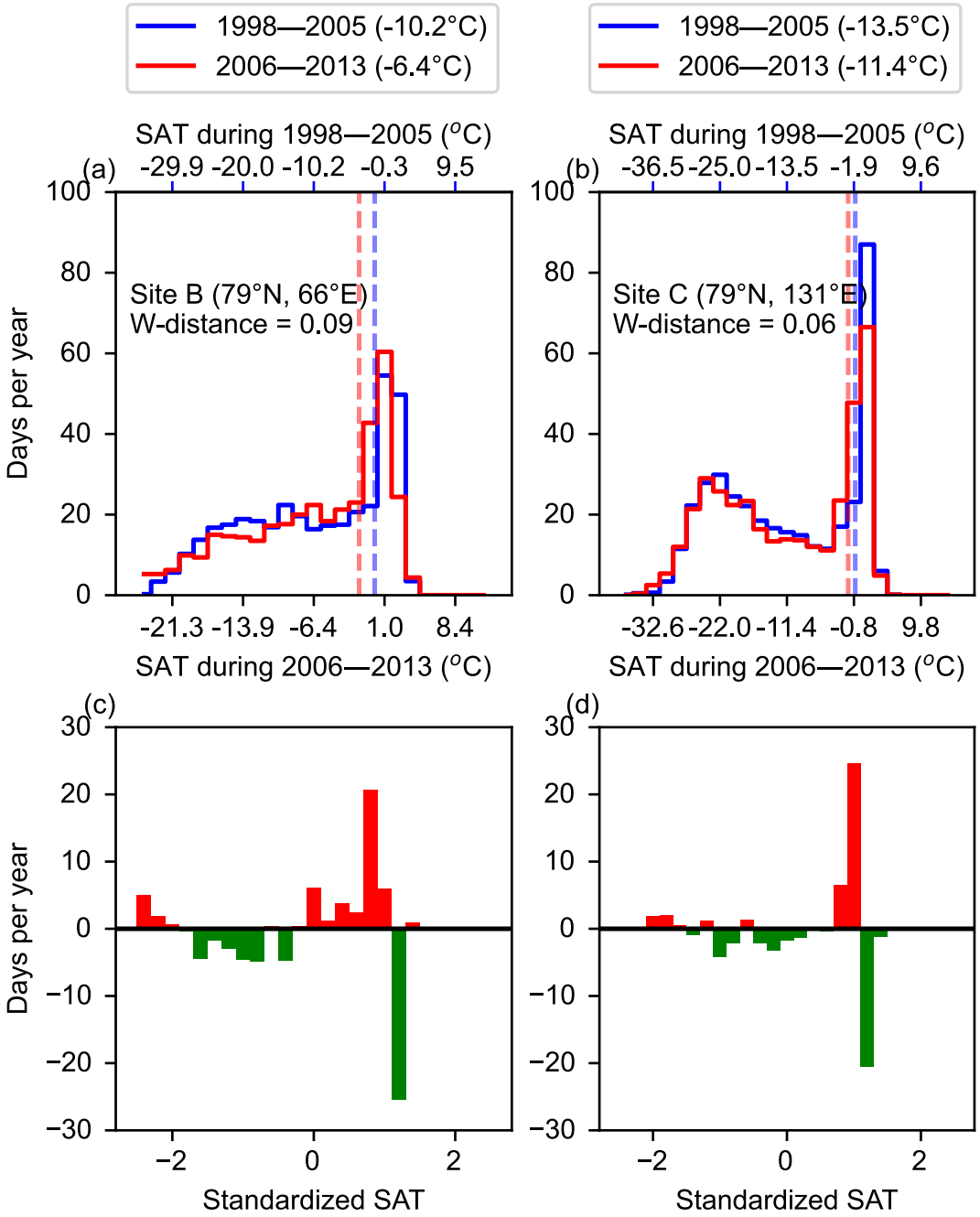

Fig. 3. As in Fig. 2 but (a) and (c) are for Site B, and (b) and (d) are for Site C. The dashed lines in (a) and (b) represent the ocean frozen temperature (−1.7 °C) in two distributions (blue for 1998–2005 and red for 2006–13).

compared with IPCC AR5 is the improved sea-ice modeling. Observational-based estimates of global mean surface temperature face a significant issue in areas where sea ice melts or grows, resulting in a switch between air temperature and sea surface temperature, which leads to reduced warming in IPCC AR5. According to Richardson et al. (2018), this underestimated warming in historical model simulations amounts to about 3% of the observed warming. Given that WSA is able to detect this sea-ice-induced change in PDF, it may be useful for future model development and data correction.

Overall, by detecting changes in PDFs between the early and later periods during the early 21st century, WSA is able to identify the two physical processes during the early 21st century: the La Niña-like tropical Pacific temperature shift and physical value constraint due to sea ice.

4. Conclusion and discussion

In this work, we develop a novel method called WSA to detect climate change signals. The results indicate that WSA shows some advantages compared with traditional

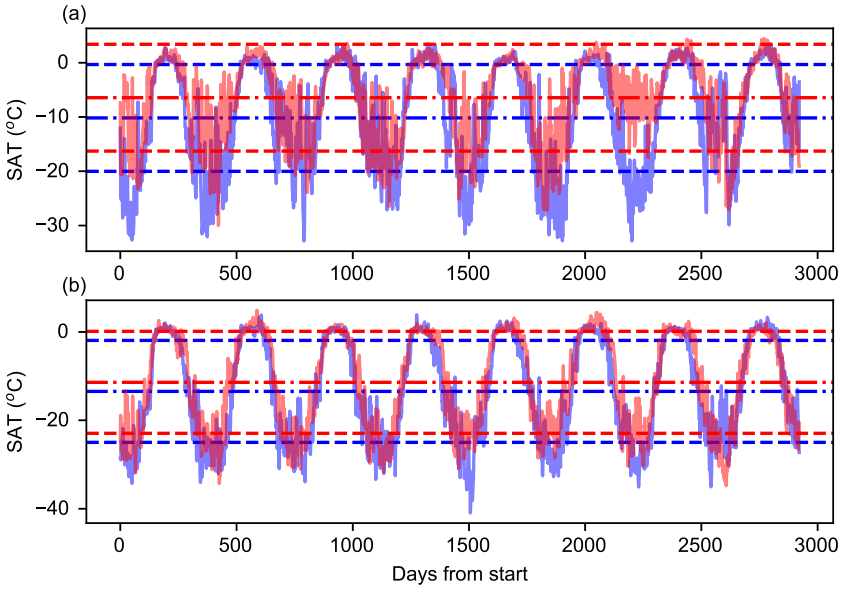

Fig. 4. Full daily SAT time series of 1998–2005 (blue) and 2006–13 (red) at Site B is plotted in Panel (a), and the counterpart at Site C is plotted in Panel (b). Dash-dot lines and dash lines represent mean value and ±1 standard deviation boundaries of 1998–2005 (blue) and 2006–13 (red), respectively.

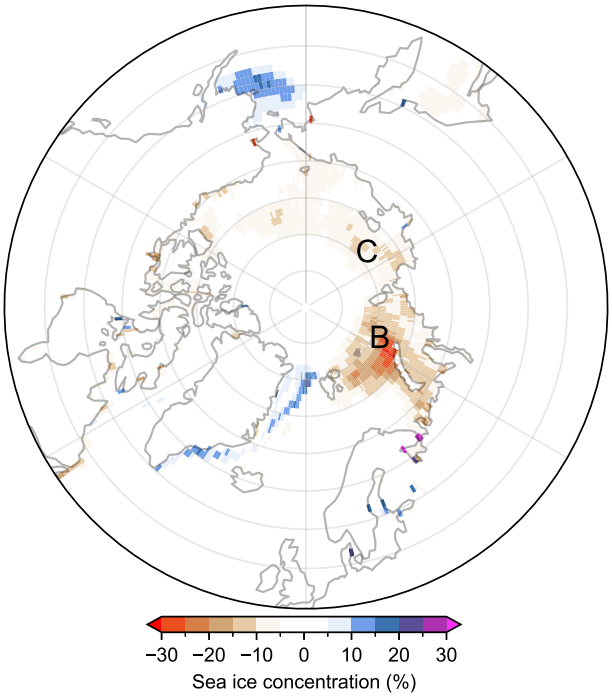

Fig. 5. The sea-ice concentration change (units: %) between 1998–2005 and 2006–13. The locations marked by letters correspond to Site B and Site C.

mean-value analysis in certain cases. Trend analysis based on mean values cannot detect the extreme event signals in the equatorial eastern Pacific Ocean in the early 21st century, while WSA can easily identify the La Niña-like temperature PDF shift. Additionally, due to the strong temperature constraint of sea ice, the sea-ice cover area shows a different PDF pattern compared to the open ocean. As a result, unlike mean-value analysis detecting significant trend regions, WSA can identify the strong sea-ice-loss area in the early 21st century, indicating a fundamental physical property change of the surface.

While the regional average is commonly used to depict large-scale features, its suitability varies, especially concerning PDF changes. In tropical regions like the equatorial eastern Pacific, the regional average aligns closely with sites due to uniform sea surface temperatures. However, at high latitudes, regional averages may not capture significant PDF shifts. For instance, in our analysis of the Barents–Kara seas (40°–80°E, 70°–80°N), the W-distance of the regional average SAT between the periods 1998–2005 and 2006–13 is not significant (in the electronic supplementary material), despite most grid cells within the region passing the 99% significance test (Fig. 1a). This discrepancy arises from differences in mean temperature and standard deviation among grid cells, highlighting that extreme values in one specific grid cells may not carry the same significance across grid cells, leading to an insignificant cold and warm shift from a regional perspective. Despite this, site-specific analyses, like Site B in our study, remain physically representative, uncovering a crucial mechanism: the weakening of physical constraints due to sea ice melting over time. As a result, to ensure robust results, we advocate for a more strict significance test level and caution in regional averaging.

It should be noted that the discussion of PDF change during the early 21st century is incomplete in this study. We just chose several typical regions or sites to illustrate the abundant information from the significant PDF change and verify the capability of the WSA algorithm. A more detailed discussion on scattered significant signals in the Indian Ocean and South America is another interesting topic and worthy of

investigation in a separate study. Meanwhile, we only applied the WSA method to SAT in this study. Similarly, we can also apply the method to precipitation, wind speed, and pressure, which may also offer more insights into climate change. Furthermore, while our current work is based on ERA5 reanalysis data, due to its widespread availability, sufficient spatial coverage (i.e., at the global scale), high temporal resolution and fine horizontal resolution, it is important to note that the WSA method is designed to be versatile and adaptable to different datasets. The choice of ERA5 in this context serves as a demonstration, showcasing the capabilities of the method. Researchers have the flexibility to substitute other observational datasets, including any other satellite retrievals or ground-based observations, in order to suit their specific needs and preferences.

It should be emphasized that despite the slowdown in global mean SAT during the early 21st century, global warming has not ceased; rather, it is manifested in various other ways. Previous studies documented that the global ocean heat content in the upper 2000 m has not stopped and has continuously increased in response to the increased atmospheric greenhouse gases from human activities (Cheng et al., 2024). Furthermore, from another perspective, the WSA method can offer new insights into climate change. We used it to successfully detect changes in PDFs over the equatorial eastern Pacific Ocean and sea-ice melting areas in the Arctic, and the possible underlying physical mechanisms have also been discussed. However, the current WSA method can only detect whether the distribution has changed or not, and the direction of change is not well depicted. For example, Fig. 1 can only give us a clue to discover PDF change signals, but cannot directly tell us whether most of the regions have more cold extremes or warm extremes. Therefore, how to measure the direction of PDF change is still a challenging topic and needs to be explored in the future.

Data and Software Surface 2-meters air temperature data is from ERA5 reanalysis products of the European Centre for Medium-Range Weather Forecast (ECMWF) (Hersbach et al., 2020), which can be accessed at the following link https://doi.org/10.24381/cds.adbb2d47. Sea ice data is from the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST) of Met Office in British (Rayner et al., 2003), which could be accessed via the following link http://www. metoffice.gov.uk/hadobs/hadisst/. The code of Wasserstein Stability Analysis can be found online via the link https://doi.org/10.5281/zenodo.7839648.

Acknowledgements. The authors greatly appreciate the comments of Dr. Zhaolu HOU from the Ocean University of China and Dr. Shuailei Yao from the Institute of Atmospheric Physics, Chinese Academy of Sciences. This work was supported by the National Key Research and Development Program of China (Grant No. 2021YFC3000904), the National Natural Science Foundation of China (42005039), and the Science and Technology Development Fund of CAMS (Grant No. 2024KJ013).

APPENDIX

This appendix explains the W-distance in optimal transport and W-distance test in WSA. In the first section, we introduce the W-distance, and especially we list the open source code to compute the W-distance. Then, in the second section, we give a detailed description of the W-distance test of WSA, which is formulated in Algorithm 1.

1. Introduction to W-distance and optimal transport

This section introduces the optimal transport and W-distance. In practice, W-distance can quantify the distance and similarities between two probability distributions. Compared to other similar metrics like Kullback–Leibler divergence, W-distance is rigorously defined and satisfies the metric axioms (Figalli and Glaudo, 2021). W-distance provides a powerful metric to quantify the similarities and discrepancies between probability distributions.

In 1781, Gaspard Monge proposed the concept of optimal transport from one practical situation: if one uses soil to build fortifications, what is the cheapest way to transport the soil? Let μ and ν be two probability measures (distributions) on $\mathbb{R}^d$, and this scenario leads to the following Monge formulation of optimal transport:

$$\inf_{T_{\#}\mu=\nu} \int_{\mathbb{R}^d} c(x, T(x))\, \mathrm{d}\mu(x)\,, \tag{A1}$$

where $c(x, T(x))$ is the cost of transporting unit mass from x to $T(x)$, and $\nu = T_{\#}\mu$ means that for any (measurable) set $A \subset \mathbb{R}^d$,

$$\nu(A) = \mu(T^{-1}(A))\,. \tag{A2}$$

Such T is said to be a transport map, and since T is a map, the mass at x could only be transported to one destination, which means Monge formulation does not allow splitting the mass.

In the 1940s, Leonid Kantorovich revisited Monge's problem and gave relaxation to Monge's formulation by splitting the mass. Kantorovich relaxation leads to the following formulation of optimal transport:

$$\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\,\mathrm{d}\pi(x,y), \tag{A3}$$

where $\pi\in\Pi(\mu,\nu)$ means π is a joint distribution with marginals μ and ν (also called a transport coupling with marginals μ and ν), i.e.,

$$\Pi(\mu,\nu)=\{\pi\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d):\pi(A\times\mathbb{R}^d)=\mu(A),\\ \pi(\mathbb{R}^d\times B)=\nu(B),A,B\subset\mathbb{R}^d\} \tag{A4}$$

When the cost function is the p–distance, i.e., $c(x,y)=|x-y|^p$, the metric side of Kantorovich formulation makes it valid to quantify the similarities between probability measures μ and ν via p-W-distance, defined as below:

$$\mathcal{W}_p(\mu,\nu)=\left(\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^p\;d\pi(x,y)\right)^{\frac{1}{p}}, \tag{A5}$$

where $p\in[1,\infty)$. In this work, we use the following earth mover's distance, i.e. $p=1$,

$$\mathcal{W}_1(\mu,\nu)=\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|\;d\pi(x,y). \tag{A6}$$

Earth mover's distance can be seen as the minimum amount of "work" required to transform mass from the probability measure (distribution) μ into another probability measure (distribution) ν, where the "work" is measured as the amount of distribution weight that must be moved, multiplied by the distance it has to be moved.

The code to calculate the earth mover's distance between two 1D distributions has been integrated into the *wasserstein_distance* function of the *Scipy* package. The input of *wasserstein_distance* is just the two sequences of observed values in the empirical distributions, then the function returns the computed distance between these two distributions. The details can be found in Scipy (2023).

Another option to compute W-distance is to use the *ot.emd2* function in the Python optimal transport package POT (Flamary et al., 2021). The *ot.emd2* function can compute the earth mover distance in 1D as long as the cost matrix is given by the corresponding l_1 distance. While POT is one of the most efficient exact optimal transport solvers, it has not been designed to handle large-scale optimal transport problems. Therefore, if one needs to solve the optimal transport with a large number of samples, we do not recommend using POT. Interested readers may refer to the POT homepage at https://PythonOT.github.io/ for more information.

2. W-distance test in Wasserstein Stability Analysis

Algorithm 1 provides details for the first step *W-distance test* in WSA. Given two samples X and Y, both series are first standardized individually by the scalar $(x-\hat{\mu})/\hat{\sigma}$, where $\hat{\mu}$ is the mean value and $\hat{\sigma}$ is the standard deviation. Then, the W-distances $\mathcal{W}_1(X,Y)$ between X and Y are obtained after standardization.

The null hypothesis H_0 of WSA is that the PDF discrepancy between two samples X and Y can be explained by the white noise. Clearly, the alternate hypothesis (H_1) is that the discrepancy of PDFs between X and Y results from factors other than white noise.

To conduct the Monte Carlo test, after obtaining the W-distance $\mathcal{W}_1(X,Y)$, we generate two new samples, X_N and Y_N, by adding white noise $N(0,1)$ to X and Y, where $N(0,1)$ is the standard Gaussian distribution. Then, one can obtain the W-distance $\mathcal{W}_1(X_N,Y_N)$ between the new samples X_N and Y_N. Such an additional operation of white noise is performed 500 times. Therefore, one can obtain a sequence of W-distances $\{\mathcal{W}_1(X_N,Y_N)\}_{N=1}^{500}$, and then obtain a confidence interval of W-distances. Under the null hypothesis H_0 that the distribution discrepancy between X and Y can be explained by the white noise $N(0,1)$, the W-distance between the original samples X and Y must be indistinguishable from the counterpart distance between the perturbed samples X_N and Y_N, i.e., $\mathcal{W}_1(X,Y)$ must be indistinguishable from $\mathcal{W}_1(X_N,Y_N)$. Under the null hypothesis H_0 with the confidence level $1-\alpha$, one would expect that $\mathcal{W}_1(X,Y)$ is within the confidence interval of $\{\mathcal{W}_1(X_N,Y_N)\}_{N=1}^{500}$. If not, H_0 is rejected, and one could claim that the distribution discrepancy between X and Y is not from the white noise $N(0,1)$ with the confidence level of $1-\alpha$, where $\alpha=0.01$.

In this study, a higher confidence level of $1-\alpha=99\%$, rather than the common 95%, is used to enhance the result interpretability, since W-distance is very sensitive to the PDF change. Furthermore, when adding white noise to the standardized samples of X and Y, we use the standard Gaussian distribution $N(0,1)$. Since a large size of samples is used, according to the central limit theorem, the amplitude of white noise is set as one standard deviation individually.

After performing the W-distance test for all locations, we can detect the places where the W-distance is significant, which are regarded as the significant PDF change zones. Then, we can carry out *PDF analysis* for the significant PDF change zones.

Algorithm 1: Wasserstein Stability Analysis (WSA)

1: **procedure** (H_0: the discrepancy in probability distributions between X and Y is from white noise, with confidence level $1-\alpha$ where $\alpha=0.01$).
2: X and Y are standardized individually by the $(x-\hat{\mu})/\hat{\sigma}$ scalar.
3: Get $\mathcal{W}_1(X,Y)$
4: $N\leftarrow 1$
5: **while** $N\leqslant 500$ **do**
6: $X_N\leftarrow X+random(0,1)$ ▷ add white noise $N(0,1)$ to X and Y
7: $Y_N\leftarrow Y+random(0,1)$
8: Get $\mathcal{W}_1(X_N,Y_N)$ ▷ X_N and Y_N are series with white noise
9: $N\leftarrow N+1$
10: **end while**

11: **if** $\mathcal{W}_1(X,Y)$ is out of 99% confidence interval of $\{\mathcal{W}_1(X_N,Y_N)\}_{N=1}^{500}$ **then**
12: reject H_0 ▷ The discrepancy is not from $N(0,1)$
13: **else**
14: H_0 is not rejected
15: **end if**
16: **end procedure**

REFERENCES

Cai, Q. Q., J. Wang, D. Beletsky, J. Overland, M. Ikeda, and L. Y. Wan, 2021: Accelerated decline of summer Arctic sea ice during 1850–2017 and the amplified Arctic warming during the recent decades. *Environmental Research Letters*, **16**, 034015, https://doi.org/10.1088/1748-9326/abdb5f.

Cheng, L. J., and Coauthors, 2024: New record ocean temperatures and related climate indicators in 2023. *Adv. Atmos. Sci.*, **41**, 1068–1082, https://doi.org/10.1007/s00376-024-3378-5.

Delworth, T. L., and T. R. Knutson, 2000: Simulation of early 20th century global warming. *Science*, **287**, 2246–2250, https://doi.org/10.1126/science.287.5461.2246.

Delworth, T. L., F. R. Zeng, A. Rosati, G. A. Vecchi, and A. T. Wittenberg, 2015: A link between the hiatus in global warming and North American drought. *J. Climate*, **28**, 3834–3845, https://doi.org/10.1175/JCLI-D-14-00616.1.

Douville, H., A. Voldoire, and O. Geoffroy, 2015: The recent global warming hiatus: What is the role of Pacific variability. *Geophys. Res. Lett.*, **42**, 880–888, https://doi.org/10.1002/2014GL062775.

England, M. and Coauthors, 2014: Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change*, **4**, 222–227, https://doi.org/10.1038/nclimate2106.

Figalli, A., and F. Glaudo, 2021: *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows*. EMS Press.

Flamary, R., and Coauthors, 2021: Pot: Python optimal transport. *Journal of Machine Learning Research*, **22**(78), 1–8.

Haurwitz, M. W., and G. W. Brier, 1981: A critique of the superposed epoch analysis method: Its application to solar-weather relations. *Mon. Wea. Rev.*, **109**, 2074–2079, https://doi.org/10.1175/1520-0493(1981)109<2074:ACOTSE>2.0.CO;2.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**(730), 1999–2049, https://doi.org/10.1002/qj.3803.

Huang, J. B., and Coauthors, 2017: Recently amplified arctic warming has contributed to a continual global warming trend. *Nature Climate Change*, **7**(12), 875–879, https://doi.org/10.1038/s41558-017-0009-5.

Hyun, S., and Coauthors, 2022: Ocean mover's distance: Using optimal transport for analysing oceanographic data. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **478**(2262), 20210875, https://doi.org/10.1098/rspa.2021.0875.

Kosaka, Y., and S.-P. Xie, 2013: Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, **501**(7467), 403–407, https://doi.org/10.1038/nature12534.

Li, Q. X., and Coauthors, 2015: China experiencing the recent warming hiatus. *Geophys. Res. Lett.*, **42**, 889–898, https://doi.org/10.1002/2014GL062773.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. 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. Atmos.*, **108**, 4407, https://doi.org/10.1029/2002jd002670.

Richardson, M., K. Cowtan, and R. J. Millar, 2018: Global temperature definition affects achievement of long-term climate goals. *Environmental Research Letters*, **13**, 054004, https://doi.org/10.1088/1748-9326/aab305.

Robin, Y., P. Yiou, and P. Naveau, 2017: Detecting changes in forced climate attractors with Wasserstein distance. *Nonlinear Processes in Geophysics*, **24**, 393–405, https://doi.org/10.5194/npg-24-393-2017.

Scipy, 2023: Earth mover's distance code. Available from https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.wassersteindistance.html.

Simmons, A. J., P. Berrisford, D. P. Dee, H. Hersbach, S. Hirahara, and J.-N. Thépaut, 2017: A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. *Quart. J. Roy. Meteor. Soc.*, **143**, 101–119, https://doi.org/10.1002/qj.2949.

Tamang, S., 2022: Advancing the predictability of the earth system processes: Insights from optimal mass transport theory. PhD dissertation, University of Minnesota.

Vissio, G., V. Lembo, V. Lucarini, and M. Ghil, 2020: Evaluating the performance of climate models based on Wasserstein distance. *Geophys. Res. Lett.*, **47**, e2020GL089385, https://doi.org/10.1029/2020GL089385.

Xie, Z. A., A. M. Duan, and Q. Tian, 2017: Weighted composite analysis and its application: An example using ENSO and geopotential height. *Atmospheric Science Letters*, **18**, 435–440, https://doi.org/10.1002/asl.786.