

<https://doi.org/10.1038/s43247-024-01711-1>

A recent surge in global warming is not detectable yet



Claudie Beaulieu¹✉, Colin Gallagher², Rebecca Killick³, Robert Lund⁴ & Xueheng Shi⁵

The global mean surface temperature is widely studied to monitor climate change. A current debate centers around whether there has been a recent (post-1970s) surge/acceleration in the warming rate. Here we investigate whether an acceleration in the warming rate is detectable from a statistical perspective. We use changepoint models, which are statistical techniques specifically designed for identifying structural changes in time series. Four global mean surface temperature records over 1850–2023 are scrutinized within. Our results show limited evidence for a warming surge; in most surface temperature time series, no change in the warming rate beyond the 1970s is detected despite the breaking record temperatures observed in 2023. As such, we estimate the minimum changes in the warming trend required for a surge to be detectable. Across all datasets, an increase of at least 55% is needed for a warming surge to be detectable at the present time.

Global mean surface temperature (GMST) series are crucial for monitoring global warming. The warming can be quantified by a change from a base period (e.g. pre-industrial), or by the rate of change (the warming rate) over a time interval¹. The GMST naturally fluctuates in time, displaying short periods of accelerated or decelerated warming (Fig. 1). Considerable attention has focused on changes in the warming rate in the scientific literature and news media, with episodes of accelerated/decelerated warming (i.e., surges and slowdowns) being recently debated^{2–7}. These fluctuations may happen in the presence of long-term warming^{8,9} and can arise due to short-term variability (or noise) in the surface temperatures. Here, trend means the long-term change in mean temperatures and noise contains fluctuations about the trend.

Noise in temperature series is often characterized by a short-memory process such as an autoregression. In this and other short-memory models, the ocean and other slow climate component systems respond to random atmospheric forcing slowly, producing variability at time scales longer than that of white noise¹⁰. Short-memory fluctuations can be large enough to temporarily mask a long-term warming trend, creating the appearance of a slowdown. They can also exacerbate a warming trend, mimicking a surge¹¹. The key question is whether these fluctuations are occurring without any change in the underlying warming trend, or whether there has been an increase (warming surge) or decrease (warming slowdown) in the trend. To answer such questions, one needs to model the short-term variability in the GMST.

Several studies suggested that a slowdown in warming (the so-called hiatus) occurred in the late 1990s and investigated its causes⁵. The slowdown

was attributed to several factors, including large-scale variability in the Pacific Ocean^{12–16} and external forcings^{15,17,18}. However, studies focusing on the detection of this warming pause showed that the rate of change had not declined, and that this period (from approximately 1998–2012) was not unusual given the level of short-term variability present in the data^{11,19–22}. More specifically, studies analyzing GMST using changepoint detection methods, which are specifically designed to objectively detect the timing of trend changes, showed no warming rate changes circa 1998^{11,19,21}. Further, a study assuming that the changepoint time is known and took place in 1998 showed that the trends before and after 1998 were statistically indistinguishable²². Overall, evidence for a pause or slowdown circa 1998 lacked a sound statistical basis.

As per the Intergovernmental Panel on Climate Change (IPCC), detection of change refers to the “process of demonstrating that climate or a system affected by the climate has changed in some defined statistical sense without providing a reason for that change”^{23,24}. Typically, the process of attribution requires that a change is statistically detected²³.

The major agencies monitoring GMSTs all rank 2023 as the warmest year since the start of the instrumental record commencing in 1850^{25–28}. Clearly, global warming has not paused, and the current discussion about the rate of warming in the news media and literature has shifted to whether there has been a warming acceleration^{2,4,5,29,30}. For example²⁹, suggests that the warming rate has increased since 1990 due to a global increase in the Earth’s energy imbalance (the difference in incoming solar radiation and infrared radiation emitted to space)^{31–33}. Another recent study predicts an acceleration in the warming rate after 2010². With lessons learned from the

¹Department of Ocean Sciences, University of California, Santa Cruz, 1156 High Street, Santa Cruz, 95064 CA, USA. ²School of Mathematical and Statistical Sciences, Clemson University, 105 Sikes Hall, Clemson, 29634 SC, USA. ³School of Mathematical Sciences, Lancaster University, Bailrigg, Lancaster, LA1 4YF England, UK. ⁴Department of Statistics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, 95064 CA, USA. ⁵Departments of Statistics and Biological Systems Engineering, University of Nebraska-Lincoln, 340 Hardin Hall North, Lincoln, 68583 NE, USA. ✉e-mail: beaulieu@ucsc.edu

Fig. 1 | GMST anomalies from four datasets with superimposed piecewise linear model fitted trends. GMST anomalies are from NASA (blue), HadCRUT (yellow), NOAA (red), and Berkeley (grey) with fitted trends (thick lines) for **A** continuous models with changing autocorrelation, **B** discontinuous models with changing autocorrelation. Note: the model fits only show the trend in different regimes.

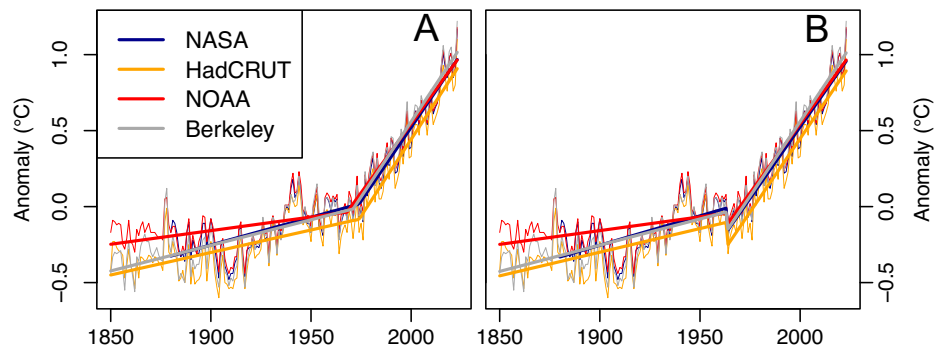


Table 1 | Changepoints detected in four global mean surface temperature datasets within continuous and discontinuous changepoint models

Dataset	Continuous	Discontinuous
NASA	1973 0.004, 0.020	1963 0.004, 0.019
HadCRUT	1973 0.003, 0.018	1963 0.003, 0.019
NOAA	1967 0.002, 0.017	1963 0.001, 0.018
Berkeley	1970 0.003, 0.019	1963 0.003, 0.020

Trends before and after the changepoint are presented below the changepoint year (in °C/year).

hiatus, debate still fresh, we assess whether a warming surge is statistically detectable.

Changepoint techniques are used here to assess whether there has been a warming surge since the 1970s, and if so, to estimate when the surge(s) started (see Methods). Piecewise linear regression models that allow for trend changes are fitted to GMST datasets and assessed via changepoint techniques. When one does not a priori know the time of any changes, which is the case here, changepoint methods account for the number of possible different places where a new regime can begin, preventing overstating statistical significance by “cherry-picking” the location of the changepoint times. This is a critical statistical point. Alongside this, two types of changepoint models are used: discontinuous models where process means in the different regimes of the linear regression do not necessarily connect, and continuous models where the process means in successive regimes connect. While a continuous model is more physically realistic for a globally averaged GMST²¹, discontinuous fits are also provided as a sensitivity to model choice; a discontinuous model may approximate a continuous change in annual measurements. Our model fits are assessed by verifying whether the residual assumptions are met (see Supplementary Information). Four GMST records are analyzed at the annual time scale (see Methods). The Berkeley, HadCRUT, and NOAA series range from 1850 to 2023, while the NASA series covers 1880 to 2023. Since little evidence of a surge is concluded, a simulation study investigates how many additional years of GMST observations will be needed before a change in the warming rate becomes detectable.

Results

Can we detect a warming surge yet?

Continuous and discontinuous models were fitted to all annual GMST series (see Methods). Model fits and timings of any found changepoints are listed in Table 1 and illustrated in Fig. 1. For the continuous model, a single changepoint is detected near 1970 in all datasets (Fig. 1a). Similarly, we find one changepoint in all datasets for the discontinuous models (Fig. 1b).

While the timings detected are slightly earlier for the discontinuous models, both cases do not indicate any changes in trend after the 1970s. We allow the first-order autocorrelation parameter to vary between segments in the fits presented in Fig. 1a, b because a previous study that analyzed global surface temperature time series with changepoint detection suggested a reduction in autocorrelation after the 1960s¹¹. A changing autocorrelation allows us to better capture the larger serial autocorrelation in surface temperatures in the earlier part of the record. To assess sensitivity of our results to this choice, we also include continuous and discontinuous models fitted with fixed autocorrelation parameters (same in all regimes) (Supplementary Note 1). While there is variability in the number and timings of changes in the earlier part of the record, no warming surge is detected beyond the 1970s (Supplementary Figs. 1 and 3). With a changing autocorrelation, assumptions on the residuals seem valid both for the continuous and discontinuous models, but there is more leftover autocorrelation in the residuals when imposing a fixed AR(1) autocorrelation (Supplementary Note 2; Supplementary Fig. 2; Supplementary Table 1).

To illustrate how a change in assumptions can yield false detections, we also fitted a discontinuous model that does not take into account autocorrelation (assuming independent errors) on the HadCRUT dataset (see Fig. 2). We include this model here to emphasize that ignoring autocorrelation can lead to spurious detection of changepoints. In fact, multiple spurious changes in the rate of warming are detected after 1970 with this model. A warming acceleration in the 1970s is detected, followed by a different regime with a similar trend starting in 2000, and finally an acceleration in warming in 2012. Similar results are found in the other datasets (Supplementary Fig. 4). However, these fits are not valid as residuals are all strongly autocorrelated (Supplementary Fig. 5; Supplementary Table 1). This illustrates how changepoint analyses can produce substantially different results if autocorrelation is ignored^{11,34,35}. Furthermore, false detection issues are exacerbated with discontinuous fits that tend to enhance the impression of a change in trend²¹.

We focus on whether there has been a change in the warming trend, but here we also quantify how unusual the 2023 observed temperature is. In Fig. 3, we fit a continuous model with changing autocorrelation (same model as in Fig. 1a) withholding 2023. We then use this model to make a one step prediction for 2023, representing the expected anomaly considering a continued trend in warming and observed autocorrelation in the four respective datasets. The observed 2023 anomalies are much larger than the predicted values (Fig. 3). More specifically, the observed anomalies for 2023 are all in the 99th percentile when compared to the statistical predictions and their associated standard errors, indicating a large departure from the ongoing warming trend.

How many years are needed to detect a surge?

The fitted models suggest that no changepoints (surges or pauses) have occurred after the 1970s in the GMSTs analyzed. However, it would be somewhat naive to categorically conclude that no surge has occurred since it is possible that the change in trend is too small or that there is not yet enough data for statistical detection. In this section, we consider how far into the

Fig. 2 | An example of GMST time series with a spurious fit. A HadCRUT GMST anomalies (black) with fitted superimposed discontinuous piecewise linear trends (red) calculated assuming independent errors and **B** scatterplot of the errors to illustrate a positive correlation. A hypothesis test provides strong evidence that the independent errors assumption is not valid with a p -value < 0.000008 (see Supplementary Table 1).

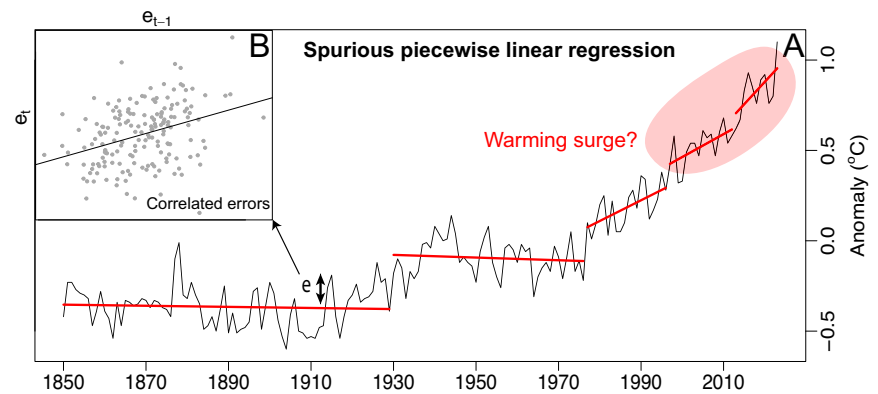
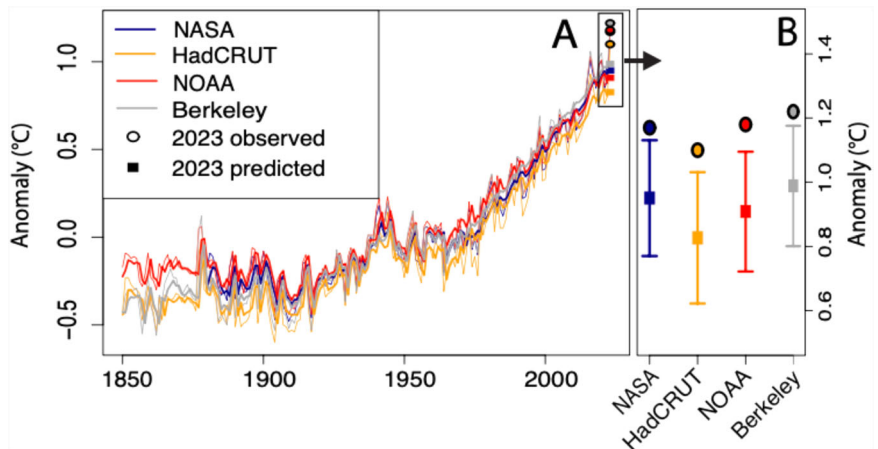


Fig. 3 | Observed vs predicted GMST anomalies for 2023. A GMST anomalies from the NASA (blue), HadCRUT (yellow), NOAA (red), and Berkeley (grey) series with superimposed piecewise linear model (as in Fig. 1a). The fits displayed here include the autocorrelation. The model is fitted until 2022, and used to predict the 2023 data point (square) for each dataset based on the fitted trend and autocorrelation. The actual observations for 2023 are also presented (circle). **B** Zoom in over 2023 with observations (circle) and predictions (square). The 2023 predictions include 95% prediction intervals (vertical bars). The observed 2023 anomalies are all outside the 95% intervals, indicating a large departure from the expected mean.



future GMST must be observed and how large a surge is needed to identify a statistically significant change at the current warming rate.

With the HadCRUT GMST from 1970–2023, we computed how large a surge would need to be to become statistically detectable at the $\alpha = 0.05$ significance level (see Methods). Elaborating, during the 1970–2023 period, the maximum difference in trends occurs in 2012, with the estimated segments being 1970–2012 and 2013–2023, respectively. Enforcing continuity between the two regimes, the estimated trends are 0.019 °C/year (first segment) and 0.029 °C/year (second segment), a 53% increase. Accounting for the short-term variability in the HadCRUT GMST over 1970–2023 and the added uncertainty for the changepoint location, the second segment (2013–2023) would need a slope of at least 0.039 °C/year (more than a 100% increase) to be statistically different than 0.019 at the $\alpha = 0.05$ significance level right now. The estimated slope of 0.029 °C/year falls far short of this needed increase. While it is still possible there was a change in the warming rate starting in 2013, the HadCRUT record is simply not long enough for the surge to be statistically detectable at this time.

Figure 4 shows the magnitude of trend change required for different potential changepoint locations from 1990 to 2015 and extending the time series from 2024 until 2040. The changepoint times considered encompass surge timings suggested in the scientific literature and media^{2,4,29}. For example, to detect a warming surge that starts in 1990 over the period of 1970–2024, the magnitude of the surge needs to be at least 67% relative to the 1970–1990 trend. This is equivalent to a change of trend from 0.018 °C/year over 1970–1990 to 0.030 °C/year over 1991–2024. If observations are extended into the future until 2030, the minimum surge detectable is 61%, becoming 55% by 2040.

To detect a surge starting in 2008 (as suggested in ref. 4) with a 2024 vantage year, the magnitude of the surge increase needs to be at least 75%.

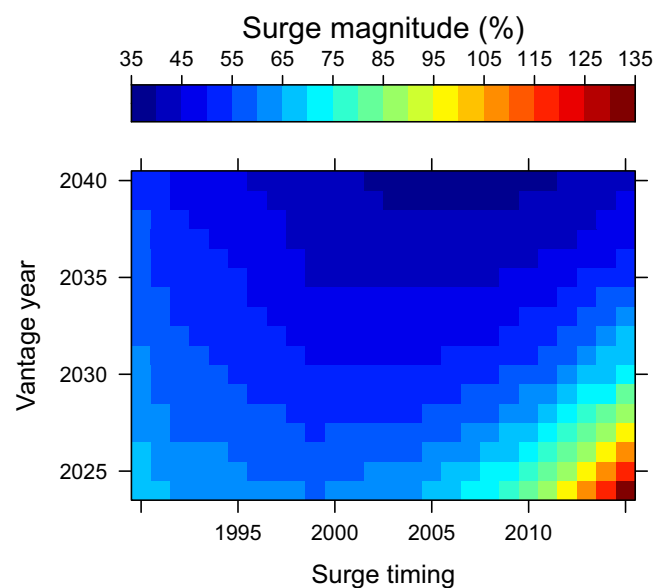


Fig. 4 | Minimum magnitude of a detectable (5% critical level) warming surge (%) given a range of potential timings for the start of the surge and different timings for the end of the time series. Surge magnitude estimates are based on the HadCRUT global mean surface temperature observed trend, variability and autocorrelation. Assuming a starting point in 1970, we consider a potential changepoint in trend for all years between 1990 to 2015 and potential vantage year from 2024 to 2040. The colorbar indicates the minimum magnitude of a surge to be detectable given the timing of the surge and the vantage year.

Extending to the vantage year 2040, a surge would need to increase by at least 39% to be detectable.

To detect a warming surge starting in 2010 and ending in 2024, the trend needs to have changed by 84% (equivalent to a trend of 0.034 °C/year from 2010–2024). If the time series extends to 2030, the surge would need to change by at least 58% (a magnitude of 0.028 °C/year from 2010–2030) to be detectable. If the time series is further extended to 2040, a surge of at least a 39% change (corresponding to a magnitude of 0.026 °C/year from 2010–2040) could be detectable.

We present this simulation based on time series properties of the HadCRUT data here; however, patterns are the same across all GMST datasets (Supplementary Notes 3–4; Supplementary Table 2; Supplementary Figs. 6–8). These estimates include the added uncertainty of an unknown changepoint location (see Methods). If the timing of a surge was known *a priori* from separate observational platforms (e.g., from satellites) or from climate model simulations, the minimum detectable surge magnitude would reduce (given the same short-term variability). In this case, the changepoint location is known and the problem reduces to testing whether the trends are the same before and after the hypothesized change time. Across all datasets, an increase of at least 55% is needed for a warming surge to be detectable in 2024.

Over the different periods considered in the literature, the hardest surge timing to detect is 2015 when observations end in 2024. In this case, there are only nine years of observations after the change. The trend over those nine years would need to be 133% larger (0.044 °C/year) to become detectable. Based on Fig. 4, it is harder to detect a surge when it occurs close to the series' end. This behavior is also observed in a simulation study on the detection power for an increase in warming given different vantage years and surge timings (Supplementary Figs. 9–10). Detection power is lost with shorter time series (early vantage years) and for a late surge.

Discussion

GMST series fluctuate in time due to short-term variability, often creating the appearance of surges and/or slowdowns in warming. While these fluctuations may mimic an increase/decrease in the warming trend, they can simply arise from random noise in the series. This is important considering the warming hiatus discussion over the last decade and the more recent alleged warming acceleration. Formal detection of surges and pauses should account for noise (or short-term variability) and the additional uncertainty of identifying the changepoint times (unless the timing of a changepoint is suggested by independent model/theory/observations).

Here, several changepoint models were used to assess whether an acceleration in warming has occurred since 1970. Different changepoint model types were considered to assess sensitivity to model choice. After accounting for short-term variability in the GMST (characterized by an autoregressive process), a warming surge could not be reliably detected anytime after 1970. This holds regardless of whether the changepoint models impose continuity of mean responses between regimes or autocorrelation is fixed or time-varying. We further demonstrate that an acceleration is detected with a discontinuous model that assumes independent errors, which is not a statistically valid model choice. Model fits should be assessed for overall goodness of fit and produce residuals with a zero-mean and no autocorrelation (white noise). In the Supplementary Information, this is done by analyzing residuals from the model fits and testing them for residual autocorrelation.

While our focus is on whether there has been a continued acceleration in the rate of global warming, our analysis recognizes how unusual surface temperature anomalies were in 2023^{36–38}. Our model fit (in the continuous mean response model with changing autocorrelation, Fig. 1a) shows that the 2023 anomaly is larger than the 99th percentile of the expected mean, indicating a large departure from the current warming trend. One could consider including exogenous variables (such as ENSO) in the model, which would reduce variability in the residuals and enable statistically detection of any changes sooner³⁹. The fact that trend changes in GMST records were not detected after the 1970s does not rule out that some small changes may have

occurred; indeed, the records may be too short (or changes not large enough) to be detectable amidst the short-term variability. As such, a simulation study was conducted to assess when a warming surge will become detectable in the future. A change in the warming rate on the order of 35% around 2010 becomes detectable circa 2035. This is the case for both an acceleration or a slowdown in warming. Our simulations allow for either an increase or decrease in the trend (two-sided Student's test). Detection lengths would reduce with one-sided testing (say warming only), but this is not deemed justifiable given the recent discussion about a pause. Indeed, testing for a warming increase because the same observations suggest an increase will tend to overstate significance. Finally, our conclusion that an acceleration is not detectable at the global level yet may not apply at regional levels and a rigorous detection of regional warming surges should be the focus of future work.

Our conclusions are based on piecewise linear models. While piecewise linear models provide a good first-order approximation of any nonlinearities and prevent overfitting the data, no model will perfectly describe our scenario. The assumption that global surface temperatures contain first-order autocorrelation, which describes the dependency in the year-to-year noise values, is short-memory and geometrically decays in year. Other types of models used to describe the noise in surface temperature observations include a long-memory model (where the decay is a power law)^{40–42}. However, we do not consider these models here as identifying long-memory requires long time series and its presence tends to be more prominent in sea surface temperatures^{43,44}. Stochastic trend models where the GMST trend cointegrates with the trend in radiative forcings have also been considered^{45–47}. That said, these models have been mainly used for detection and attribution studies, and we focus on detecting a change in the warming trend here.

Method

Data

The following four GMST time series were analyzed in this study:

- The Hadley Centre/Climatic Research Unit, Version 5 (HadCRUT), surface temperature²⁶. This series is available at <https://www.metoffice.gov.uk/hadobs/hadcrut5/data/current/download.html>. The annual anomalies from 1850–2023 were used. Anomalies are relative to the 1961–1990 period.
- The Merged Land-Ocean Surface Temperature Analysis from the National Oceanic and Atmospheric Administration (NOAA GlobalTemp v5.1.0.) of²⁸. This series is available at https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series/globe/land_ocean/1/9/1850-2023. The annual anomalies from 1850–2023 were used. Anomalies are with respect to the 1901–2000 period.
- The Berkeley Earth Surface Temperatures (Berkeley) of ref. 25. This series is available at <http://berkeleyearth.org/data>. Anomalies are computed from the 1961–1990 baseline and cover the period 1850–2023.
- The Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) at the National Aeronautic Space Administration (NASA)²⁷. This series is available at <https://data.giss.nasa.gov/gistemp/> and spans 1880–2023. Anomalies are scaled to the 1951–1980 period.

Changepoint models

Our work entails fitting several changepoint time series models that partition the GMST into regimes with similar trends using piecewise linear regression models. This work is most concerned with changes in the trend of the series.

Changepoint analyses partition the data into different segments at the changepoint times. To describe this mathematically, our model allows for m changepoints during the data record $t \in \{1, \dots, N\}$, which occur at the times τ_1, \dots, τ_m , where the ordering $0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_m < N = \tau_{m+1}$ is imposed. The time t segment index $r(t)$ takes the value of unity for $t \in \{1, \dots, \tau_1\}$, two for $t \in \{\tau_1 + 1, \dots, \tau_2\}$, ..., and $m + 1$ for $t \in \{\tau_m + 1, \dots, N\}$. Hence, the m changepoint times partition the series into $m + 1$ distinct segments. The model for the whole series is

$$X_t = E[X_t] + \epsilon_t,$$

where $E[X_t]$ is the regression function. The regression functions considered in this manuscript include a continuous (joinpin) model, where we impose process means to meet at the changepoint times, and its discontinuous counterpart. The model errors $\{\epsilon_t\}$ all have a zero mean and allow for autocorrelation; more about this component is said below.

The trend model regression structure we use has the simple piecewise linear form

$$E[X_t] = \alpha_{r(t)} + \beta_{r(t)}t,$$

where $\beta_{r(t)}$ and $\alpha_{r(t)}$ are the trend slope and intercept, respectively, of the linear regression in force during regime $r(t)$. An equivalent representation is

$$E[X_t] = \begin{cases} \alpha_1 + \beta_1 t, & 0 = \tau_0 < t \leq \tau_1, \\ \alpha_2 + \beta_2 t, & \tau_1 < t \leq \tau_2, \\ \vdots & \vdots \\ \alpha_{m+1} + \beta_{m+1} t, & \tau_m < t \leq \tau_{m+1} = N. \end{cases} \quad (1)$$

If continuity of the regression response $E[X_t]$ is imposed at the changepoint times, the restrictions

$$\alpha_i + \beta_i \tau_i = \alpha_{i+1} + \beta_{i+1} \tau_i, \quad 1 \leq i \leq m,$$

are imposed. These restrictions result in a model having m changepoints and $m + 2$ free regression parameters. Writing the model in terms of the free parameters $\alpha_1, \beta_1, \dots, \beta_{m+1}$ only gives

$$X_t = \alpha_1 + \sum_{i=1}^{r(t)-1} (\beta_i - \beta_{i+1})\tau_i + \beta_{r(t)}t + \epsilon_t.$$

The model errors $\{\epsilon_t\}_{t=1}^N$ are a zero mean autocorrelated time series. For a first-order autoregression (AR(1)), such a process obeys the difference equation

$$\epsilon_t = \phi \epsilon_{t-1} + Z_t, \quad (2)$$

where $\{Z_t\}$ is independent and identically distributed Gaussian noise with mean $E[Z_t] \equiv 0$ and variability $\text{Var}[Z_t] \equiv \sigma^2$, and $\phi \in (-1, 1)$ is an autocorrelation parameter representing the correlation between consecutive errors. It is important to allow for autocorrelation in the model errors in climate changepoint analyses^{11,34,48,49}; failure to account for autocorrelation can lead one to conclude that the estimated number of changepoints, \hat{m} , is larger than it should be. Higher-order autoregressions are easily accommodated should a first-order scheme be deemed insufficient. We will also consider cases where the autoregressive parameter changes at each changepoint time; these essentially let ϕ depend on time t via the regime index $r(t)$. The changepoint times for the autocorrelation and mean structure are constrained to be the same.

Estimation of the model parameters proceeds as follows. For a given number of changepoints m and their occurrence times τ_1, \dots, τ_m , one first computes maximum likelihood estimators of the regression parameters. These produce maximum likelihood estimators of all α_i and β_i . This fit gives a model likelihood, which we base on the Gaussian distribution since the series are globally and annually averaged. This likelihood is denoted by $L(m; \tau_1, \dots, \tau_m)$.

The hardest part of the estimation scheme lies with estimating the changepoint configuration. This is done via a Gaussian penalized likelihood. In particular, the penalized likelihood objective function O of form

$$O(m; \tau_1, \dots, \tau_m) = -2 \ln(L(m; \tau_1, \dots, \tau_m)) + C(m; \tau_1, \dots, \tau_m)$$

is minimized over all possible changepoint configurations. The penalty $C(m; \tau_1, \dots, \tau_m)$ is a charge for having m changepoints at the times τ_1, \dots, τ_m in a model. As the number of changepoints in the model increases, the

Table 2 | Model penalties for fitting in trend models

	Continuous	Discontinuous
Global ϕ	$p = 2m + 4$	$p = 3m + 4$
Piecewise ϕ	$p = 4m + 4$	$p = 5m + 4$

model fit becomes better and $-2 \ln(L)$ correspondingly decreases. However, eventually, adding additional changepoints to the model does little to improve its fit. The positive penalty term counteracts “overfitting” the number of changepoints, balancing likelihood improvements with a cost for having an excessive number of model parameters (changepoints and linear parameters within each segment). Many penalty types have been proposed to date in the statistics literature⁴⁹. One that works well in changepoint problems is the Bayesian Information Criterion (BIC) penalty

$$C(m; \tau_1, \dots, \tau_m) = p \ln(N),$$

where p is the total number of free parameters in the model. Table 2 lists values of p for the various model types encountered in this paper. For example, for a continuous model with a global AR(1) structure, there are $2m + 4$ free regression parameters in a changepoint configuration with m changepoints and $m + 1$ segment parameters. Also contributing to the parameter total are ϕ and σ^2 .

Finding the best m and τ_1, \dots, τ_m can be accomplished via a dynamic programming algorithm called PELT⁵⁰ or a genetic algorithm search as in refs. 51,52. PELT is computationally rapid, performs an exact optimization of the penalized likelihood, and was used here.

Given a specified model with trend and autocorrelation structure, a prediction interval can be computed for the last regime. Assuming that the forecast errors are normal, a 95% prediction interval for a h -step ahead forecast is given by

$$\hat{X}_{t+h} \pm 1.96 \hat{\sigma}_h \quad (3)$$

where h represents the number of years ahead, \hat{X}_{t+h} is the predicted temperature anomaly and $\hat{\sigma}_h$ is an estimate of the standard deviation of the h -step forecast distribution. The prediction intervals are computed using the model fit from the final segment in each series.

Testing trend differences

How can one determine the statistical significance of a potential surge in the warming rate at some point since 1970? To address this question, a model is needed. Whilst our running example here considers the HadCRUT series since 1970, other GMST datasets are easily analyzed (see Supplementary Material).

A simplification of (1) for a single continuous changepoint is

$$E[X_t] = \begin{cases} \alpha_1 + \beta_1 t, & 0 < t \leq \tau, \\ \alpha_1 + \beta_1 \tau + \beta_2(t - \tau), & \tau < t \leq N. \end{cases} \quad (4)$$

There are other good single changepoint slope change techniques for this task — two are HAC tests of ref. 53 and the two phase regression models of ref. 54 (the latter would have to be modified for autocorrelation). We develop a simple procedure here that is very accessible.

If the single changepoint is known to occur at time τ , then the Student’s test based statistic

$$T_\tau = \frac{\hat{\beta}_2 - \hat{\beta}_1}{\widehat{\text{Var}}(\hat{\beta}_2 - \hat{\beta}_1)^{\frac{1}{2}}} \quad (5)$$

can be used to make inferences. Here, $\hat{\beta}_1$ and $\hat{\beta}_2$ are the estimated trends of the two segments before and after time τ . One concludes a surge in warming if T_τ is too large to be explained by chance variation (as gauged by a Student

Table 3 | Estimated parameters for simulating annual global surface temperature anomalies over 1970–2023 (null model)

Dataset	$\hat{\phi}$	$\hat{\alpha}_1$	$\hat{\beta}_1^{1970,2023}$	$\hat{\sigma}$
NASA	0.149	-0.076	0.019	0.095
HadCrut	0.087	-0.170	0.020	0.097
NOAA	0.190	-0.036	0.019	0.092
Berkeley	0.102	-0.064	0.020	0.096

distribution with $N - 3$ degrees of freedom); a change in the warming rate (negative or positive) is suggested when $|T_\tau|$ is too large to be explained by chance variation. In computing $\text{Var}(\hat{\beta}_2 - \hat{\beta}_1)$, the AR(1) parameters ϕ and σ^2 are needed. Estimates of the two slopes, AR(1) parameters and their standard deviations are provided in time series fitting software (such as the *arima* function in R⁵⁵).

Extreme care is needed when τ is unknown. Should the time τ be selected visually among the many possibilities where it can occur without accounting for this, statistical mistakes can ensue. This is why changepoint techniques are needed. For example, $\tau = 43$, which corresponds to 2012, has been suggested as a time when warming accelerated⁴. For the 1970–2023 data, ignoring the selection of the changepoint, at the 0.05 significance level, a $|T_\tau|$ of 2.007 or more indicates warming rate changes; a two-tailed Student's test was employed to allow for either an increase or a decrease in the warming rate. For our specific example, $|T_{43}| = |\hat{\beta}_1^{1970,2012} - \hat{\beta}_2^{2013,2023}|/\hat{\sigma} = |0.0286 - 0.0187|/0.0065 = 1.5281$ is not statistically significant at the 0.05 significance level.

If the time of the warming rate change is unknown (as is common), statistical significance is determined based on the null hypothesis distribution of

$$T_{\max} = \max_{\ell \leq \tau \leq u} |T_\tau|, \quad (6)$$

where ℓ and u are values that truncate the admissible changepoint times near the data boundaries for numerical stability. The T_{\max} statistic has significantly different statistical properties (more tail area) than $|T_\tau|$ for a fixed τ . A common truncation requirement, and one that we follow, is to truncate 10% at the series boundaries: $\ell = 0.1N$ and $u = 0.9N$. If the calculated T_{\max} statistic exceeds the threshold Q_N , where Q_N is the 0.95 quantile of the null hypothesis distribution of T_{\max} , then a statistically significant rate change is declared with confidence 95%. The most likely changepoint time, $\hat{\tau}$, is estimated as the τ at which $|T_\tau| = T_{\max}$ is maximal. Statistical tests of this type are discussed in ref. 56 and ref. 57. There, large sample distributions are derived to determine Q_N . However, due to the relatively short series since 1970, we use a Monte Carlo method with Gaussian AR(1) errors to determine statistical significance.

Elaborating, our Monte Carlo approach simulates many series using parameter estimates from the current data under the null hypothesis. For example, with the 1970–2023 HadCRUT data, the null hypothesis parameters are estimated as $\hat{\alpha}_1^{1970-2023} = -0.17$, $\hat{\beta}_1^{1970,2023} = 0.0199$, $\hat{\beta}_2 = 0$ (there is no second segment under the null hypothesis), $\hat{\phi} = 0.0865$, and $\hat{\sigma} = 0.097$ (Table 3). One hundred thousand time series were then simulated, T_{\max} was computed for each series, and the 0.95 quantile of these values was identified to estimate Q_N .

The simulated 95th percent quantile for the HadCRUT series is $Q_N = 3.1082$. The largest $|T_\tau|$ statistic occurs in 2012 and is $|T_{43}| = 1.5281 (= T_{\max})$, which is far from the required threshold of 3.1082. Hence, there is little evidence for a statistically significant change in the warming rate from 1970–2023 in the annual HadCRUT series; this conclusion holds for all GMST datasets considered in this paper.

So how large would the slope need to be in the second segment to declare a significant surge? We answer this for a baseline segment of 1970–2012 and a second segment from 2013–2023. To answer this, note that $\hat{\beta}_1$ and the numerator of T_τ do not depend on $\hat{\beta}_2$; thus, we can set $T_{\max} = T_{43} = 3.1082 = Q_N$ and solve for $\hat{\beta}_2$. This results in $\hat{\beta}_2 = 0.0388$. We see that a change in surge magnitude of $100(0.0388 - 0.0187)/0.0187 = 107\%$ between the two segments is required for 95% statistical confidence.

The same logic can be used to determine future estimated rates necessary for statistical significance. Using the HadCRUT series up to 2023, there is no statistical evidence of a surge starting in 2012 relative to the 1970–2012 segment. Will this still be the case in 2025? How about 2040? For any potential surge starting in 1990–2015, the data from 1970–2023 was used to estimate the warming trend slope, intercept, and AR(1) structure. We then simulated cutoff quantiles for 95% statistical significance as above for several considered vantage years, pushing out to 2040. Since the estimated standard deviation of the slope differences depends only on the segment lengths and the AR(1) parameters, the above procedure can be solved as above for the minimal slope necessary to induce statistical significance.

Using the HadCRUT series, one hundred thousand Gaussian series were simulated up to 2040 under our best working model (no surge, $\hat{\beta}_1^{1970,2023} = 0.0199$, $\alpha = -0.17$, $\phi = 0.0865$, and $\sigma = 0.097$). This gives the Monte Carlo quantile estimate $Q_{71} = 2.9877$. The numerator of the T_τ -statistic corresponding to a change starting in 2012 is estimated and solved for the minimally significant slope for the 2013–2040 segment: $\hat{\beta}_1^{1970,2012} + \sqrt{\text{Var}(\hat{\beta}_2^{2013,2023} - \hat{\beta}_1^{1970,2012})^{1/2}} Q_{71} = 0.0187 + (2.9877) 0.0025 = 0.0262$. In short, a 40% increase in the 2013–2040 warming rate relative to the 1970–2012 rate will be needed to declare a significant warming surge by 2040.

The above process was repeated for each surge year, from 1990–2015, and each vantage year from 2024–2040. For each surge year start τ , the minimum statistically significant slope is calculated assuming that $T_{\max} = |T_\tau|$. For each τ , this minimally significant slope is compared to the estimated slope from the 1970–(1969+ τ) series segment to calculate its associated percent change. The results for the HadCRUT series are displayed in Fig. 4. Overall, one sees that either significantly increased warming or many more years of observations will be required before declaring any warming surge with a reasonable degree of confidence. This process is repeated for other GMST datasets based on the null hypothesis parameters listed in Table 3. Results are presented in the Supplementary Information.

Data availability

All the datasets used here are publicly available. Global mean surface temperature anomalies from the Hadley Centre/Climatic research Unit are available at <https://www.metoffice.gov.uk/hadobs/hadcrut5/data/current/download.html>. The Merged Land-Ocean Surface Temperature Analysis from the National Oceanic and Atmospheric Administration (NOAA GlobalTemp v5.1.0.) is available at https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series/globe/land_ocean/1/9/1850-2023. The Berkeley Earth Surface Temperatures time series is available at <http://berkeleyearth.org/data>. The Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) at the National Aeronautic Space Administration (NASA) is available at <https://data.giss.nasa.gov/gistemp/>.

Code availability

Codes are available at <https://doi.org/10.5281/zenodo.13340187>.

Received: 12 April 2024; Accepted: 19 September 2024;

Published online: 14 October 2024

References

1. Risbey, J. S., Grose, M. R., Monselesan, D. P., O'Kane, T. J. & Lewandowsky, S. Transient response of the global mean warming rate and its spatial variation. *Weather Clim. Extremes* **18**, 55–64 (2017).
2. Hansen, J. E. et al. Global warming in the pipeline. *Oxf. Open Clim. Change* **3**, kgad008 (2023).
3. Medhaug, I., Stolpe, M. B., Fischer, E. M. & Knutti, R. Reconciling controversies about the 'global warming hiatus'. *Nature* **545**, 41–47 (2017).
4. Hausfather, Z. I study climate change. The data is telling us something new. *New York Times* <https://www.nytimes.com/2023/10/13/opinion/climate-change-excessive-heat-2023.html>.

5. Mooney, C. & Osaka, S. Is climate change speeding up? Here's what the science says. *The Washington Post* <https://www.washingtonpost.com/climate-environment/2023/12/26/global-warming-accelerating-climate-change/>.
6. Millman, O. Global heating is accelerating, warns scientist who sounded climate alarm in the 80s. *The Guardian* <https://www.theguardian.com/environment/2023/nov/02/heating-faster-climate-change-greenhouse-james-hansen>.
7. Harvey, C. Global warming 'pause' never happened, scientists say. *The Washington Post* <https://www.washingtonpost.com/news/energy-environment/wp/2015/09/17/new-statistical-studies-dismantle-the-notion-of-a-global-warming-pause/>.
8. Easterling, D. R. & Wehner, M. F. Is the climate warming or cooling? *Geophys. Res. Lett.* **36**, L08706 (2009).
9. Thompson, D. W., Barnes, E. A., Deser, C., Foust, W. E. & Phillips, A. S. Quantifying the role of internal climate variability in future climate trends. *J. Clim.* **28**, 6443–6456 (2015).
10. Hasselmann, K. Stochastic climate models Part I. Theory. *Tellus* **28**, 473–485 (1976).
11. Beaulieu, C. & Killick, R. Distinguishing trends and shifts from memory in climate data. *J. Clim.* **31**, 9519–9543 (2018).
12. Kosaka, Y. & Xie, S.-P. Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* **501**, 403–407 (2013).
13. Trenberth, K. E. & Fasullo, J. T. An apparent hiatus in global warming? *Earth's Future* **1**, 19–32 (2013).
14. Kaufmann, R. K., Kauppi, H., Mann, M. L. & Stock, J. H. Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proc. Natl Acad. Sci.* **108**, 11790–11793 (2011).
15. Schmidt, G. A., Shindell, D. T. & Tsigaridis, K. Reconciling warming trends. *Nat. Geosci.* **7**, 158–160 (2014).
16. England, M. H. et al. Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nat. Clim. Change* **4**, 222–227 (2014).
17. Fyfe, J. C., von Salzen, K., Cole, J. N. S., Gillett, N. P. & Vernier, J.-P. Surface response to stratospheric aerosol changes in a coupled atmosphere–ocean model. *Geophys. Res. Lett.* **40**, 584–588 (2013).
18. Santer, B. D. et al. Volcanic contribution to decadal changes in tropospheric temperature. *Nat. Geosci.* **7**, 185–189 (2014).
19. Cahill, N., Rahmstorf, S. & Parnell, A. C. Change points of global temperature. *Environ. Res. Lett.* **10**, 084002 (2015).
20. Lewandowsky, S., Risbey, J. S. & Oreskes, N. The “pause” in global warming: Turning a routine fluctuation into a problem for science. *Bull. Am. Meteorol. Soc.* **97**, 723–733 (2016).
21. Rahmstorf, S., Foster, G. & Cahill, N. Global temperature evolution: recent trends and some pitfalls. *Environ. Res. Lett.* **12**, 054001 (2017).
22. Rajaratnam, B., Romano, J., Tsiang, M. & Diffenbaugh, N. Debunking the climate hiatus. *Climatic Change* **133**, 129–140 (2015).
23. Hegerl, G. et al. Good practice guidance paper on detection and attribution related to anthropogenic climate change. Tech. Rep., IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland (2010).
24. Bindoff, N. et al. 2013: Detection and attribution of climate change: From global to regional. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, Eds., Cambridge University Press, pp. 867–952.
25. Rohde, R. A. & Hausfather, Z. The Berkeley Earth land/ocean temperature record. *Earth Syst. Sci. Data* **12**, 3469–3479 (2020).
26. Morice, C. P. et al. An updated assessment of near-surface temperature change from 1850: The HadCRUT5 data set. *J. Geophys. Res.: Atmos.* **126**, e2019JD032361 (2021).
27. Lenssen, N. J. L. et al. Improvements in the GISTEMP uncertainty model. *J. Geophys. Res.: Atmos.* **124**, 6307–6326 (2019).
28. Vose, R. S. et al. Implementing full spatial coverage in NOAA's global temperature analysis. *Geophys. Res. Lett.* **48**, e2020GL090873 (2021).
29. Sanset, B. et al. Steady global surface warming from 1973 to 2022 but increased warming rate after 1990. *Commun. Earth Environ.* **4**, 400 (2023).
30. Forster, P. M. et al. Indicators of global climate change 2022: annual update of large-scale indicators of the state of the climate system and human influence. *Earth Syst. Sci. Data* **15**, 2295–2327 (2023).
31. Loeb, N. G. et al. Satellite and ocean data reveal marked increase in earth's heating rate. *Geophys. Res. Lett.* **48**, e2021GL093047 (2021).
32. von Schuckmann, K. et al. Heat stored in the Earth system 1960–2020: where does the energy go? *Earth Syst. Sci. Data* **15**, 1675–1709 (2023).
33. Minière, A., von Schuckmann, K., Sallée, J.-B. & Vogt, L. Robust acceleration of Earth system heating observed over the past six decades. *Sci. Rep.* **13**, 22975 (2023).
34. Lund, R. B., Beaulieu, C., Killick, R., Lu, Q. & Shi, X. Good practices and common pitfalls in climate time series changepoint techniques: A review. *J. Clim.* **36**, 8041 – 8057 (2023).
35. Shi, X., Beaulieu, C., Killick, R. & Lund, R. Changepoint detection: An analysis of the Central England temperature series. *J. Clim.* **35**, 6329 – 6342 (2022).
36. Rantanen, M. & Laaksonen, A. The jump in global temperatures in September 2023 is extremely unlikely due to internal climate variability alone. *npj Clim. Atmos. Sci.* **7**, 34 (2024).
37. Esper, J., Torbenson, M. & Büntgen, U. 2023 summer warmth unparalleled over the past 2,000 years. *Nature* **631**, 94–97 (2024).
38. Cheng, L. et al. New record ocean temperatures and related climate indicators in 2023. *Adv. Atmos. Sci.* **41**, 1068–1082 (2024).
39. Foster, G. & Rahmstorf, S. Global temperature evolution 1979–2010. *Environ. Res. Lett.* **6**, 044022 (2011).
40. Vyushin, D. I., Kushner, P. J. & Zwiers, F. Modeling and understanding persistence of climate variability. *Journal of Geophysical Research: Atmospheres* **117**, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD018240> (2012).
41. Franzke, C. Nonlinear trends, long-range dependence, and climate noise properties of surface temperature. *J. Clim.* **25**, 4172–4183 (2012).
42. Rybski, D., Bunde, A., Havlin, S. & von Storch, H. Long-term persistence in climate and the detection problem. *Geophys. Res. Lett.* **33**, L06718 (2006).
43. Poppick, A., Moyer, E. J. & Stein, M. L. Estimating trends in the global mean temperature record. *Adv. Stat. Climatol., Meteorol. Oceanogr.* **3**, 33–53 (2017).
44. Beaulieu, C., Killick, R., Ireland, D. & Norwood, B. Considering long-memory when testing for changepoints in surface temperature: A classification approach based on the time-varying spectrum. *Environmetrics* **31**, e2568 (2020).
45. Kaufmann, R. K., Kauppi, H., Mann, M. L. & Stock, J. H. Does temperature contain a stochastic trend: linking statistical results to physical mechanisms. *Climatic Change* **118**, 729–743 (2013).
46. Cummins, D. P., Stephenson, D. B. & Stott, P. A. Could detection and attribution of climate change trends be spurious regression? *Clim. Dyn.* **59**, 2785–2799 (2022).
47. Dergiades, T., Kaufmann, R. K. & Panagiotidis, T. Long-run changes in radiative forcing and surface temperature: The effect of human activity over the last five centuries. *J. Environ. Econ. Manag.* **76**, 67–85 (2016).
48. Beaulieu, C., Chen, J. & Sarmiento, J. L. Change-point analysis as a tool to detect abrupt climate variations. *Philos. Trans. R. Soc. A: Math., Phys. Eng. Sci.* **370**, 1228–1249 (2012).
49. Shi, X., Gallagher, C., Lund, R. & Killick, R. A comparison of single and multiple changepoint techniques for time series data. *Comput. Stat. Data Anal.* **170**, 107433 (2022).
50. Killick, R., Fearnhead, P. & Eckley, I. A. Optimal detection of changepoints with a linear computational cost. *J. Am. Stat. Assoc.* **107**, 1590–1598 (2012).

51. Davis, R. A., Lee, T. C. M. & Rodrigues-Yam, G. A. Structural break estimation for nonstationary time series models. *J. Am. Stat. Assoc.* **101**, 223–239 (2006).
52. Li, S. & Lund, R. Multiple changepoint detection via genetic algorithms. *J. Clim.* **25**, 674–686 (2012).
53. McKittrick, R. R. & Vogelsang, T. J. Hac robust trend comparisons among climate series with possible level shifts. *Environmetrics* **25**, 528–547 (2014).
54. Lund, R. & Reeves, J. Detection of undocumented changepoints: A revision of the two-phase regression model. *J. Clim.* **15**, 2547–2554 (2002).
55. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/> (2023).
56. Gallagher, C., Lund, R. & Robbins, M. Changepoint detection in climate time series with long-term trends. *J. Clim.* **26**, 4994–5006 (2013).
57. Robbins, M. W., Gallagher, C. M. & Lund, R. B. A general regression changepoint test for time series data. *J. Am. Stat. Assoc.* **111**, 670–683 (2016).

Acknowledgements

The authors thank NOAA, NASA, HadCRUT, and Berkeley for making surface temperature observations publicly available. Claudie Beaulieu acknowledges National Science Foundation Grant AGS-2143550; Robert Lund thanks National Science Foundation Grant DMS-2113592; Xueheng Shi thanks UNL Startup (ARD) 2162251011.

Author contributions

Claudie Beaulieu led the study's conceptualization and design, managed the project, performed data analyses, wrote and edited the manuscript and critically reviewed the findings. Colin Gallagher led the surge detectability analyses, wrote and edited the manuscript and critically reviewed the findings. Rebecca Killick led the changepoint analyses, wrote and edited the manuscript and critically reviewed the findings. Robert Lund wrote and edited the manuscript and critically reviewed the findings. Xueheng Shi wrote and edited the manuscript and critically reviewed the findings.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43247-024-01711-1>.

Correspondence and requests for materials should be addressed to Claudie Beaulieu.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Alireza Bahadori. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024