

# Patterns in the temporal complexity of global chlorophyll concentration

Received: 2 August 2023

Accepted: 8 February 2024

Published online: 19 February 2024

 Check for updates

Vitul Agarwal <sup>1</sup>✉, Jonathan Chávez-Casillas <sup>2</sup>, Keisuke Inomura <sup>1</sup> & Colleen B. Mouw <sup>1</sup>

Decades of research have relied on satellite-based estimates of chlorophyll-*a* concentration to identify oceanographic processes and plan in situ observational campaigns; however, the patterns of intrinsic temporal variation in chlorophyll-*a* concentration have not been investigated on a global scale. Here we develop a metric to quantify time series complexity (i.e., a measure of the ups and downs of sequential observations) in chlorophyll-*a* concentration and show that seemingly disparate regions (e.g., Atlantic vs Indian, equatorial vs subtropical) in the global ocean can be inherently similar. These patterns can be linked to the regularity of chlorophyll-*a* concentration change and the likelihood of anomalous events within the satellite record. Despite distinct spatial changes in decadal chlorophyll-*a* concentration, changes in time series complexity have been relatively consistent. This work provides different metrics for monitoring the global ocean and suggests that the complexity of chlorophyll-*a* time series can be independent of its magnitude.

Satellites have been used to monitor global ocean color for decades, and satellite-derived products have been critical in determining the global trends of primary productivity<sup>1–3</sup>, coastal runoff<sup>4,5</sup>, sea ice extent<sup>6,7</sup>, and harmful algal blooms<sup>8</sup>. Of particular relevance is chlorophyll-*a* concentration ([chl-*a*]), the primary pigment used by phytoplankton to perform photosynthesis, which can be reliably estimated from the reflectance of blue and green light from the oceans<sup>9</sup>. Chlorophyll-*a* concentration estimates are used for various goals: estimating primary productivity<sup>3</sup>, developing ecological indicators<sup>10</sup>, monitoring long-term trends<sup>11</sup>, or testing earth-system-models<sup>12,13</sup>. Despite the necessity of using [chl-*a*] time series in global studies, it is unclear what role, if any, the time series complexity of [chl-*a*] time series plays in accounting for region-specific differences in global ocean color.

The roughness or complexity of a time series can strongly indicate multiple phenomena: high stochasticity in a target process, measurement error, long tails in the data, rapidly changing system states, or the confluence of all these factors and more. As phytoplankton populations display chaotic dynamics, non-linear behavior, and intermittent instability<sup>14–16</sup>, measurements of the complexity of [chl-*a*] time series might also capture large-scale patterns that structure global phytoplankton communities. There

are many ways of estimating the natural complexity of a time series, such as calculating the fractal or Hausdorff dimension<sup>17</sup>, permutation entropy<sup>18</sup>, or Lyapunov exponents<sup>19,20</sup>. These are all different measures of complexity for dynamic systems and can often be related to one another. Although some of these calculations have been performed at select locations<sup>21</sup>, a global analysis of [chl-*a*] time series complexity would provide a deeper understanding of the spatio-temporal characteristics of phytoplankton blooms as observed by satellite radiometers.

We used approximately 25 years (1998–2022) of global [chl-*a*] observations that were based on measurements made with the Sea-viewing Wide Field of View Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer on the Aqua satellite (MODIS), Medium Resolution Imaging Spectrometer (MERIS), Visible and Infrared Imaging/Radiometer Suite (VIIRS) and the Ocean and Land Color Instruments (OLCI-A and OLCI-B) sensors. Chlorophyll concentration estimates were from the Garver-Siegel-Maritorena Model merged data product<sup>22,23</sup> at 25 km resolution and daily temporal resolution, available from the Hermes GlobColour website (<https://hermes.acri.fr/>). Using a merged product allowed us to increase the spatial and temporal coverage of our analysis; however, we also tested our analysis on

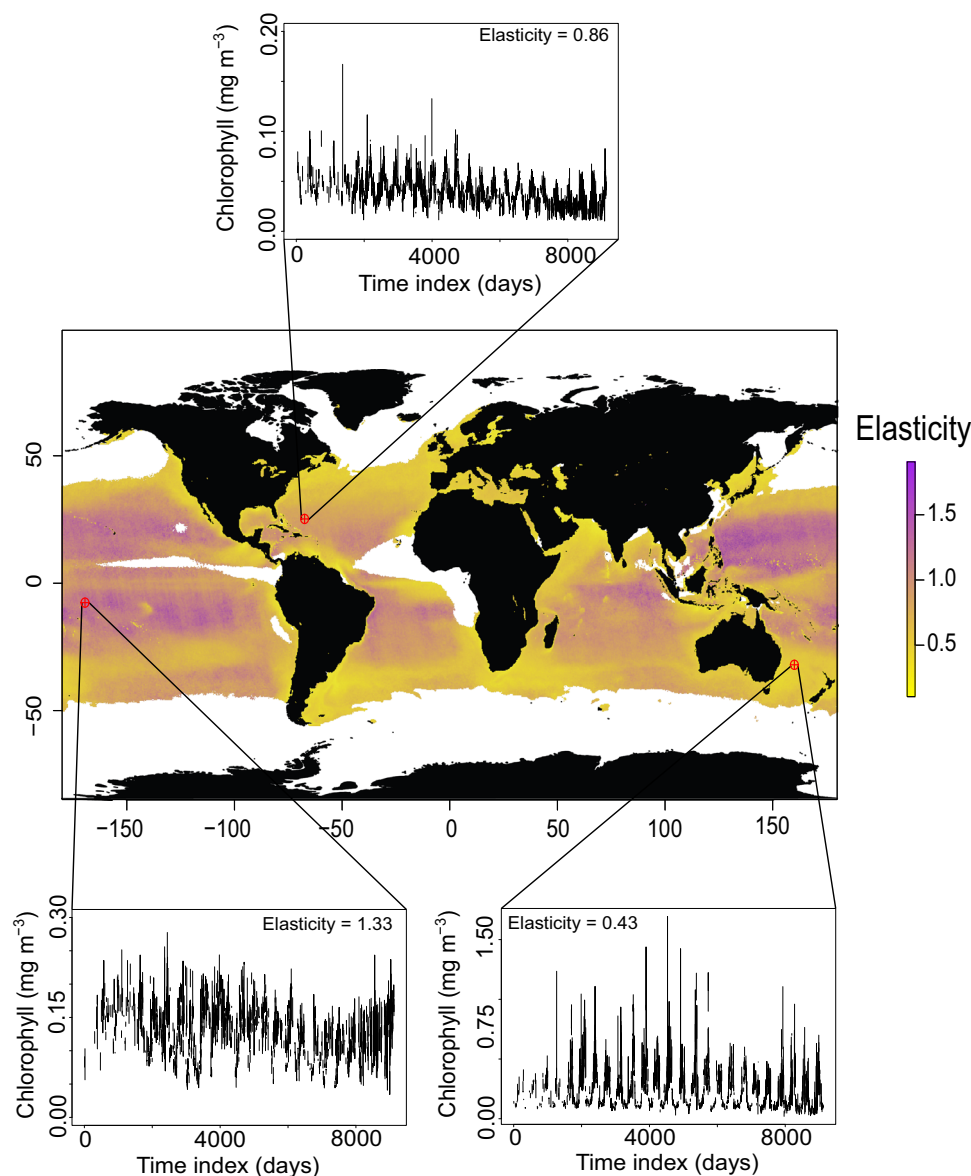
<sup>1</sup>Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA. <sup>2</sup>Department of Mathematics and Applied Mathematical Sciences, University of Rhode Island, Kingston, RI, USA. ✉e-mail: [vitulagarwal@uri.edu](mailto:vitulagarwal@uri.edu)

single missions (i.e., MODIS) and found only small differences on a global scale (Figure S1). Previous studies have used this product to understand phytoplankton bloom phenology based on change point algorithms<sup>24,25</sup>, cycles in chlorophyll dynamics<sup>26</sup>, the climate-driven components of [chl-*a*] variability<sup>27</sup>, and the contributions of citizen science programs<sup>28</sup>.

Here, we show how two different complexity metrics can indicate region-specific differences in several phenomena, such as the regularity of daily changes in [chl-*a*] or the likelihood of extreme values within a [chl-*a*] time series. Some of these differences are likely driven by differential rates of satellite spatial coverage, ocean color algorithm error, and naturally occurring dynamics. On a decadal time scale, we observe changes in both the mean magnitude of [chl-*a*] and the complexity properties of the time series. This suggests that our understanding of the global ocean may need to account for intrinsic variation within [chl-*a*] time series, as well as how we perceive and estimate [chl-*a*] from satellite radiometers.

## Results and Discussion

An important component in global [chl-*a*] studies is the definition of a phytoplankton bloom, often described as an anomalous or distinct increase in the biomass of a particular species (or set of species). Determining when a bloom occurs is not trivial, with varying strategies that can produce different results<sup>29,30</sup>. One example is defining the bloom start date as the point where [chl-*a*] rises a certain percentage above the median [chl-*a*]<sup>31–33</sup>. We defined an analogous term to what economists call elasticity to analyze the sensitivity of [chl-*a*] time series to thresholds, where the elasticity of a [chl-*a*] time series is a measure of the responsiveness of the percentage change between two variables. The variables measured were the number of times that there was an increase in daily [chl-*a*] above a certain threshold and the threshold itself (see Methods for further details). Higher elasticity describes a rougher time series and a greater sensitivity to thresholds, likely due to greater variability within day-to-day [chl-*a*] change (Fig. 1). As elasticity does not heavily weigh the magnitude of [chl-*a*], but instead the variability in the magnitude of [chl-*a*], it is not a direct measure of



**Fig. 1 | Global map of elasticity calculated from 25 years (1998–2022) of daily-scale chlorophyll-*a* concentration time series.** Darker colors (purple) indicate greater elasticity for the corresponding 25x25km pixel, whereas lighter colors (yellow) indicate lower elasticity. Some examples of time series with their elasticity

values are provided. Areas in white are pixels where greater than 80% of the time series were missing observations or less than 400 days with consecutive measurements. 'Time index' refers to the day of sampling starting from January 1<sup>st</sup>, 1998, to January 1<sup>st</sup>, 2023.

phytoplankton growth dynamics or productivity. It captures the regularity of change within a time series for each pixel, irrespective of the mechanisms involved in driving those changes.

When viewed globally, the elasticity of [chl- $a$ ] time series varied across broad ocean regions. The oligotrophic gyres, known for their low chlorophyll concentration and primary production rates<sup>34,35</sup>, had the highest elasticity values. In contrast, regions with high chlorophyll concentration and prominent seasonal cycles, such as the Patagonian Shelf, the Baltic Sea, the west coast of Central America, and the southwest Pacific Ocean<sup>36–39</sup>, showed lower elasticity values. As elasticity is sensitive to the regularity of [chl- $a$ ] change, regions with strong, periodic drivers of phytoplankton blooms (often manifested in seasonal or annual patterns) would naturally have lower values of elasticity in comparison to regions where the periodicity of [chl- $a$ ] change is irregular. Figure S2 highlights how elasticity can vary for three different model time series. If all daily changes within a time series were equal, its elasticity would equal to 0 (no sensitivity to thresholds). When evaluated for the relative frequency of daily changes, some regions were prominent in their dynamic behavior, with more days of significant [chl- $a$ ] increase regardless of any chosen threshold (Figure S3). Interestingly, the North Atlantic gyre had lower than average elasticity values than both Pacific gyres, despite similar values of [chl- $a$ ]. The difference is likely due to the stronger seasonal patterns in the North Atlantic and the influence of the annual spring bloom<sup>40,41</sup>, which would impart more regularity in the [chl- $a$ ] time series in the North Atlantic. Our results suggest that the elasticity metric tracks the relative strength of climatic and seasonal forcing of [chl- $a$ ] across broad ocean regions and might provide a quantitative measure of differentiating between them. Global Earth-system models may benefit by matching the relative strength of climatic drivers to the measured elasticity of [chl- $a$ ] time series on a global scale. For observational studies that rely on measurements of [chl- $a$ ], more dynamic behavior of time series in different ocean regions might necessitate an increase in the precision and number of samples taken to identify trends.

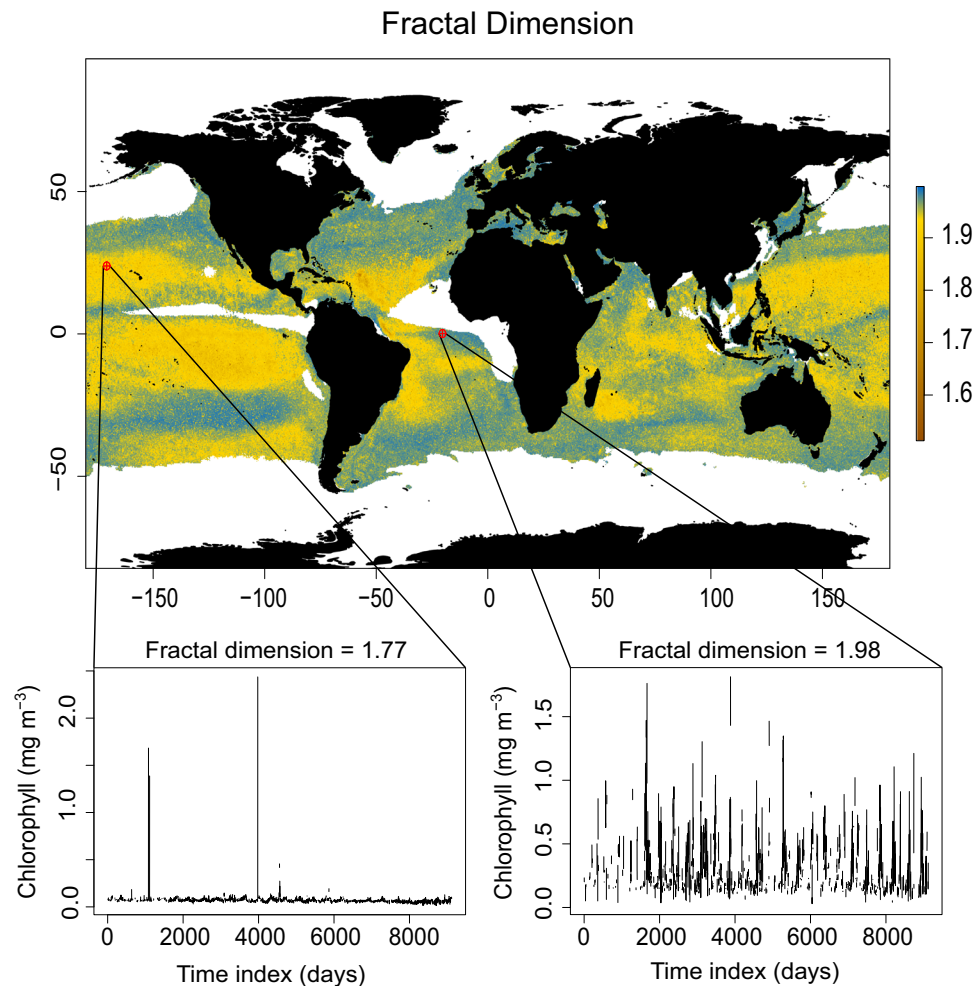
Another metric often used to compute the complexity of a time series is the fractal dimension. When considering a time series, we can think of it as the graph of function from  $\mathbb{R} \rightarrow \mathbb{R}$  where  $\mathbb{R}$  refers to the set of real numbers. As a graph, it then becomes a subset of  $\mathbb{R}^2$ . It is known that when the function is smooth (i.e., differentiable), its fractal dimension is 1, but when this function is not differentiable, the fractal dimension might be greater. This dimension cannot be larger than 2 as the graph is an embedding of  $\mathbb{R}^2$ . The fractal dimension is an indicator of how rough a curve is. It can also be understood as the autocorrelation of the time series as the lag period tends to 0. A higher fractal dimension implies higher instantaneous autocorrelation of the time series. For example, it is known that the fractal dimension of a Brownian motion, a non-differentiable continuous function, is 1.5. In this sense, the fractal dimension gives us a way to quantify how volatile a time series and how likely we are to observe extreme values on it.

When we computed the fractal dimension of every [chl- $a$ ] pixel time series, the values across the global ocean were remarkably consistent (typically above 1.85), with some key differences in particular regions (Fig. 2). The subtropical North and South Pacific, tropical North Atlantic, the Amazon plume region, and the Eastern coast of Madagascar had lower-than-average fractal dimensions. Part of the reason for this difference could be tied to the influence of physical ocean dynamics: such as the South Equatorial current and freshwater discharge from the Amazon River<sup>42,43</sup>, eddy formation from the Agulhas current near Madagascar<sup>44–46</sup>, the Gulf Stream and hurricane formation in the tropical North Atlantic<sup>47,48</sup> and the influence of tropical instability waves in the equatorial Pacific<sup>49</sup>. Another reason could be the episodic nature of elevated [chl- $a$ ] in many of these regions, which may lead to a greater number of outliers within the [chl- $a$ ] time series. As the box-counting method has been shown to return lower fractal

dimension values in time series with sharp changes<sup>50</sup>, we first remove most outliers from each time series (see Methods) before calculating the fractal dimension. This was done to minimize the error in our computations due to likely erroneous measurements. In this case, our global map of fractal dimension reveals the likelihood of anomalous spikes within each time series (long tails in the data), possibly due to naturally occurring dynamics, but also, atmospheric correction and ocean color algorithm error. Regions less prone to outliers have time series that return higher fractal dimensions due to fewer anomalous spikes.

We conducted several additional analyses to further test the sensitivity of elasticity and fractal dimension to our methodological choices. As the estimation of satellite chlorophyll using merged data can often contain bias due to the differences among individual sensors<sup>51–54</sup>, we calculated elasticity and fractal dimension for MODIS data alone (2002–2022) and compared it against the merged dataset for the entire time period. The results were largely similar on a global scale (Figure S1), with some minor differences in the magnitude of elasticity for the oligotrophic gyres. Similarly, when we tested for the influence of time series resolution (daily, weekly, monthly, etc.) on the metrics, our results showed that aggregated time series are typically smoother (i.e., possess lower fractal dimension). The elasticity also changes with coarser resolution for some regions and indicates geographically similar, but greater differences (Figure S4). To get an estimate of deviation for each complexity metric, we re-calculated the elasticity and fractal dimension for every time series based on random selected windows of 4000 days each. The standard error based on 30 trials suggests that elasticity can vary up to  $\pm 0.20$  in some regions such as the oligotrophic ocean, whereas fractal dimensions only vary up to  $\pm 0.04$ . Figure S5 highlights the deviation of both complexity metrics for the global dataset. Lastly, we also compared both complexity metrics to some traditional measures of variability for each time series (mean [chl- $a$ ], standard deviation, relative standard deviation, and average seasonal amplitude). The calculated fractal dimension of each time series showed a nonlinear positive relationship, whereas the elasticity showed a nonlinear negative relationship with each metric (Figures S6–7). When evaluated against each other on a similar scale, elasticity and fractal dimension did not present any strong relationships (Figures S8–9).

When calculated annually, both the elasticity and the average fractal dimension of the global ocean had a significant shift in 2003 and a gradual return to 1998 levels thereafter (Fig. 3a–c), whereas the mean [chl- $a$ ] showed a general decrease till 2019. Interestingly, a second transition around 2019 shows a rise in mean [chl- $a$ ], a sharp increase in fractal dimension and a drop in elasticity. The error was larger prior to 2003 for every metric, likely due to fewer observations and the sole data source being SeaWiFS. The satellite data record from 2003 onwards contained merged observations and included data from MODIS-Aqua and MERIS. As data collection by SeaWiFS stopped in 2010 and by MERIS in 2012, while data collection by VIIRS began in 2012, it is possible that the sharp changes around the years 2010–2012 can be attributed to changes in the temporal and spatial coverage of the time series during this period. Similarly, although previous studies have noted little to no inter-annual differences in [chl- $a$ ] over the 2003 transition<sup>26</sup>, changes in mean complexity over this period are likely tied to the increase in spatial coverage due to a greater number of sensors in orbit. Even though we observed a decline in [chl- $a$ ] since 2003 in agreement with the existing literature<sup>55,56</sup>, the true estimates of changes in [chl- $a$ ] likely depend on the methods used to determine trends and the composition of the spatial dataset<sup>57,58</sup>. We only considered part of the global ocean based on data availability (i.e., fewer high-latitude regions). Consequently, the mean [chl- $a$ ] only provides context for the changes in elasticity and fractal dimension. Our results suggest a decline in global [chl- $a$ ] and a concurrent change in the structure of [chl- $a$ ] time series in the satellite record.



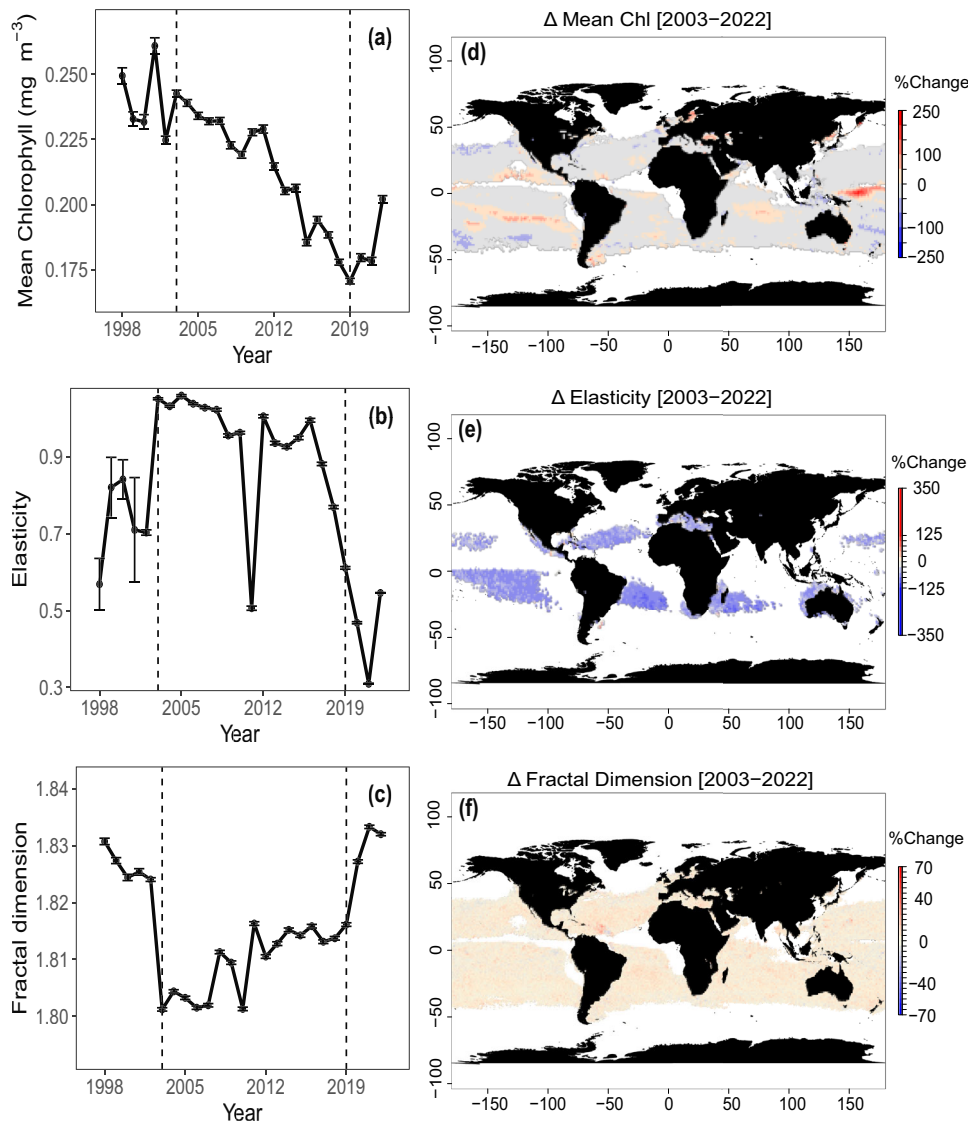
**Fig. 2 | Global map of fractal dimension calculated from 25 years (1998–2022) of daily-scale chlorophyll-*a* concentration time series with random resampling of any missing observations.** Blue and yellow indicate a higher fractal dimension for the corresponding 25x25km pixel, whereas brown indicates lower fractal dimensions. Most of the global ocean has relatively high fractal dimensions (global mean

= 1.95). Some examples of time series with their fractal dimension values are provided. Areas in white are pixels where greater than 80% of the time series were missing observations. ‘Time index’ refers to the day of sampling starting from January 1<sup>st</sup>, 1998, to January 1<sup>st</sup>, 2023. Note: the color scale is not linear and has been chosen to highlight spatial differences.

These results have several implications. First, an increase in the fractal dimension and a decrease in elasticity imply a lower frequency of anomalous events in many regions and an increase in stronger, repeatable cycles of [chl-*a*]. Part of this change can likely be attributed to the greater number of satellite sensors in orbit (and merged in the [chl-*a*] product) and a change in the magnitude of error-related outliers. Errors may decrease due to the greater availability of data and possibility of quality control but may also increase over time as the hardware of individual satellites age. Further improvements in satellite capabilities may allow future studies to increase the precision of their conclusions and enhance the predictability of [chl-*a*] time series by quantifying more accurately the impact of stochasticity and the measurement error<sup>59</sup>. While we did not evaluate the chlorophyll-*a* product uncertainty in this study, future studies may consider accounting for the difference between [chl-*a*] algorithms, different uncertainty calculations, and their relationship to time series complexity<sup>60–62</sup>. One possibility is that temporal complexity is driven by high variability within certain regions of the visible spectrum; for example, slight day-to-day differences in the green spectrum may propagate across [chl-*a*] calculations and lead to greater roughness within the time series. This variability could be caused by several factors – such as rapid differences in the taxonomic composition of phytoplankton or the influence of terrestrial discharge.

Second, when evaluated for differences across space, we also found that changes in mean [chl-*a*] and both the complexity metrics from 2003–2022 were inconsistent for different ocean regions (Fig. 3d–f). For example, areas such as the Black Sea, parts of the Western equatorial Pacific, and the subtropical Indian Ocean had an overall increase in [chl-*a*] in 2022 compared to 2003. In contrast, changes in the elasticity and fractal dimension were more spatially consistent across the global ocean. These patterns indicate that both metrics are insensitive to changes in [chl-*a*] magnitude and, thus, independent of the processes that control [chl-*a*] magnitude in different regions.

Third, although there is a long history and multiple benefits to categorizing marine ecological provinces based on environmental, geophysical, and biogeochemical data sources<sup>63–66</sup>, measurements of time series complexity offer a valuable addition to monitoring global ocean change on a larger scale. Our results could form the basis of further refinement in the definitions of marine ecological provinces. Future studies may also consider how disparate regions in the global ocean are similar in time series complexity and whether the patterns of observations by satellite radiometers are tied to specific environmental or ecological drivers of [chl-*a*]. Examples of possible drivers include mixed layer stability, eddy formation, temperature change and plankton growth and grazing rates.



**Fig. 3 | Inter-annual trends in time series complexity.** (a) mean chlorophyll-*a* concentration ( $\text{mg m}^{-3}$ ), (b) mean elasticity and (c) mean fractal dimension calculated yearly (1998–2022) from daily-scale time series data. Black points indicate the mean value and error bars are the 95% confidence intervals ( $1.96 \times S.E.$ ). The sample sizes for each point can be found in Supplementary Table 1. Gridlines mark the years of 2003 and 2019. Global maps of (d) the percentage change in mean chlorophyll-*a* concentration between the years 2003 and 2022, (e) the percentage change in mean elasticity between the years 2003 and 2022 and (f) the percentage

change in mean fractal dimension between the years 2003 and 2022. Blue values indicate negative change, whereas red values indicate positive change. White pixels are the locations where greater than 80% of the entire time series were missing observations, or where there were less than 100 consecutive days of observation in 2003 and 2022 each. Each pixel is  $25 \times 25 \text{ km}$ . There are more white pixels in subplot (e) because of less than 100 days of consecutive samples in 2003. We need more instances of repeated daily coverage to calculate elasticity (see Methods).

## Methods

We generated a 25-year daily time series (1998–2022) of [chl-*a*] by extracting a merged data product at 25 km resolution from <https://hermes.acri.fr/>. The merged data product used the Garver-Siegel-Maritorena Model<sup>22,23</sup>, which combined data from the Sea-viewing Wide Field of View Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer on the Aqua satellite (MODIS), Medium Resolution Imaging Spectrometer (MERIS), Visible and Infrared Imaging/Radiometer Suite (VIIRS) and the Ocean and Land Color Instruments (OLCI-A and OLCI-B) sensors.

Every time series on a daily scale contains some gaps due to insufficient satellite coverage and quality flags that were removed (Figure S10). To proceed with the analysis and remove any blank pixels, we instituted a cut-off of 20% to filter through the entire global time series dataset. Only the time series with greater than 20% of days

sampled over 25 years were considered in further analysis. This ensured a minimum of 1800 days had available [chl-*a*] estimates for every time series. These thresholds were chosen to strike a balance between data availability and spatial coverage. In our analysis, we wanted to maximize global spatial coverage without compromising on our ability to accurately estimate complexity due to sampling gaps.

We used a multi-step process to calculate the elasticity of every pixel time series. First, we created a 1-day-lagged time series and subtracted it from the original time series, giving us estimates of daily [chl-*a*] change. Any missing samples in the original time series were carried forward to the time series of daily change. Only the time series that had at least 400 samples of daily [chl-*a*] change proceeded to the next step in the calculation. Next, we created two uniform distributions of thresholds ( $\tau_1$  and  $\tau_2$ ; 1000 values each) that ranged from 1% to 50% of the median of the individual [chl-*a*] time series. For every time series,

we then calculated the number of days where daily [chl-*a*] change exceeded the threshold ( $X_1$  and  $X_2$ ). This calculation occurred independently for both sets of thresholds. Third, we calculated the relative difference in the threshold ( $\varepsilon$ ) used to find  $X_1$  and  $X_2$  according to Eq. 1, and the relative difference in daily change ( $N$ ) between  $X_1$  and  $X_2$  according to Eq. 2.

$$\varepsilon = \frac{T_2}{T_1} - 1 \quad (1)$$

$$N = \frac{X_2}{X_1} - 1 \quad (2)$$

Lastly, a linear model was created to find the slope between  $\log_e \varepsilon$  and  $\log_e N$  according to Eq. 3. The absolute value of the slope ( $\beta_1$ ) of this linear model was termed as the elasticity of the time series (as there are no non-negative values). The greater the value of  $\beta_1$ , the stronger the negative relationship between  $\log_e \varepsilon$  and  $\log_e N$ .

$$\log_e N = -\beta_1 \log_e \varepsilon + \beta_2 \quad (3)$$

It is known from the classical theory of linear models<sup>67</sup>, that the slope coefficient of a linear regression has the interpretation of being the average amount by which the response changes when the predictor changes by a unit. However, when considering a regression like the one in Eq. 3, where both predictor and response variable are measured in logarithmic scale, the coefficient of the predictor can be computed as

$$\frac{\partial \log_e N}{\partial \log_e \varepsilon} = \frac{\left(\frac{\partial N}{N}\right)}{\left(\frac{\partial \varepsilon}{\varepsilon}\right)} \quad (4)$$

which corresponds to what the economists denote as *elasticity*. Some examples of model time series and their respective elasticity values can be found in Figure S2 in the Supplementary material. As elasticity is based on relative change and is not sensitive to the magnitude of specific points, we did not remove any outliers from any time series for this part of the analysis. We also did not attempt to measure statistical significance, nor remove any pixels based on significance thresholds. Future studies may consider detailed statistical analyses to quantify the false discovery rate, the chosen number of threshold combinations, or the applicability of linear assumptions in the estimation of elasticity.

For the second part of the analysis, we used the R package `fractaldim`<sup>68</sup> to calculate the fractal dimension of every [chl-*a*] time series. First, we cleaned every time series by removing outliers that exceeded 3 standard deviations from the mean [chl-*a*] value. This was done to remove any aberrant spikes of [chl-*a*] that were likely incorrect and were missed in quality control. On average, there were about 0.004% of all pixel time series that were removed as outliers (-1–2 days every year across all locations). Given the large number of sampling gaps in some cases, we also performed bootstrapping for every pixel time series, which is a resampling technique. To be precise, the gaps in each time series were filled in by randomly selecting values (with replacement) from within the measurements of the same time series. This method of resampling was chosen to ensure we had a complete time series for every pixel without adding bias. We then calculated the fractal dimension of the time series using the box-counting method. The number of missing values and the range of [chl-*a*] (max - min) for every pixel time series did not show any relationship with the calculated fractal dimension (Figure S11). Examples of model time series and their calculated fractal dimension can be found in Figure S12 in the Supplementary material.

For calculating the yearly elasticity and fractal dimension, the process was similar to the one described above, except the time series

length was limited to the sampling points in each year from 1998–2022. To ensure enough data, we only calculated the fractal dimension and the mean [chl-*a*] of the time series (before random resampling) if at least 100 days were sampled for that year. Similarly, we only calculated the elasticity if there were at least 100 consecutive samples for that year. The results were reported in Fig. 3 as numerical averages for every pixel with error bars representing the 95% confidence intervals ( $1.96 \times S.E.$ ). Table T1 in the supplementary material reports the number of samples used to calculate the annual [chl-*a*] mean, fractal dimension, and elasticity. Percent change for either metric (represented as  $\rho$ ) was calculated according to Eq. 5:

$$\% \text{ change} = \frac{\rho_{2022} - \rho_{2003}}{\rho_{2003}} \times 100 \quad (5)$$

All the analyses and plotting were conducted in R<sup>69</sup> using the R packages “`raster`”<sup>70</sup>, “`maps`”<sup>71</sup>, “`ggplot2`”<sup>72</sup>, “`cowplot`”<sup>73</sup>, “`fractaldim`”<sup>68</sup>, “`MASS`”<sup>74</sup>, “`viridis`”<sup>75</sup> and “`ncdf4`”<sup>76</sup>.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Remote sensing data are available from <https://hermes.acri.fr/>. Data can be downloaded for specific locations and time periods under “GlobColour data search”. Sample datasets (processed) to reproduce part of the analysis can be found at <https://doi.org/10.5281/zenodo.10498362><sup>77</sup>. Source data are provided with this paper.

## Code availability

The code required for the analysis is available at <https://github.com/vitul-agarwal-1/chl-complexity> (<https://doi.org/10.5281/zenodo.10498357>). All the analyses and plotting were conducted in R<sup>69</sup> using the R packages `raster`<sup>70</sup>, `maps`<sup>71</sup>, `ggplot2`<sup>72</sup>, `cowplot`<sup>73</sup>, `fractaldim`<sup>68</sup>, `MASS`<sup>74</sup>, `viridis`<sup>75</sup> and `ncdf4`<sup>76</sup>.

## References

- Behrenfeld, M. J. et al. Climate-driven trends in contemporary ocean productivity. *Nature* **444**, 752–755 (2006).
- McClain, C. R. A decade of satellite ocean color observations. *Ann. Rev. Mar. Sci.* **1**, 19–42 (2009).
- Lee, Z., Marra, J., Perry, M. J. & Kahru, M. Estimating oceanic primary productivity from ocean color remote sensing: A strategic assessment. *J. Mar. Syst.* **149**, 50–59 (2015).
- Kudela, R. M. & Chavez, F. P. The impact of coastal runoff on ocean color during an El Niño year in Central California. *Deep. Res. Part II Top. Stud. Oceanogr.* **51**, 1173–1185 (2004).
- Chérubin, L. M., Kuchinke, C. P. & Paris, C. B. Ocean circulation and terrestrial runoff dynamics in the Mesoamerican region from spectral optimization of SeaWiFS data and a high resolution simulation. *Coral Reefs* **27**, 503–519 (2008).
- Brown, Z. W. & Arrigo, K. R. Contrasting trends in sea ice and primary production in the Bering Sea and Arctic Ocean. *ICES J. Mar. Sci.* **69**, 1180–1193 (2012).
- Arrigo, K. R., van Dijken, G. & Pabi, S. Impact of a shrinking Arctic ice cover on marine primary production. *Geophys. Res. Lett.* **35**, L19603 (2008).
- Stumpf, R. P. Applications of satellite ocean color sensors for monitoring and predicting harmful algal blooms. *Hum. Ecol. Risk Assess.* **7**, 1363–1368 (2001).
- Siegel, D. A., Maritorena, S., Nelson, N. B., Behrenfeld, M. J. & McClain, C. R. Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere. *Geophys. Res. Lett.* **32**, L20605 (2005).

10. Platt, T. & Sathyendranath, S. Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* **112**, 3426–3436 (2008).
11. Henson, S. A. et al. Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences* **7**, 621–640 (2010).
12. Follows, M. J., Dutkiewicz, S., Grant, S. & Chisholm, S. W. Emergent biogeography of microbial communities in a model ocean. *Science* **315**, 1843–1846 (2007).
13. Park, J., Stock, C. A., Dunne, J. P., Yang, X. & Rosati, A. Seasonal to multiannual marine ecosystem prediction with a global Earth system model. *Sci. (80-.)* **365**, 284–288 (2019).
14. Rogers, T. L., Munch, S. B., Matsuzaki, S., Ichiro, S. & Symons, C. C. Intermittent instability is widespread in plankton communities. *Ecol. Lett.* **26**, 470–481 (2023).
15. Giron-Nava, A. et al. Quantitative argument for long-term ecological monitoring. *Mar. Ecol. Prog. Ser.* **572**, 269–274 (2017).
16. Benincá, E. et al. Chaos in a long-term experiment with a plankton community. *Nature* **451**, 822–825 (2008).
17. Gneiting, T., Ševčíková, H. & Percival, D. B. Estimators of fractal dimension: Assessing the roughness of time series and spatial data. *Stat. Sci.* **27**, 247–277 (2012).
18. Bandt, C. & Pompe, B. Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* **88**, 174102 (2002).
19. Bryant, P., Brown, R. & Abarbanel, H. D. I. Lyapunov exponents from observed time series. *Phys. Rev. Lett.* **65**, 1523–1526 (1990).
20. Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. Determining Lyapunov exponents from a time series. *Phys. D. Nonlinear Phenom.* **16**, 285–317 (1985).
21. He, J. Application of generalized cauchy process on modeling the long-range dependence and self-similarity of sea surface chlorophyll using 23 years of remote sensing data. *Front. Phys.* **9**, 750347 (2021).
22. Maritorena, S., d’Andon, O. H. F., Mangin, A. & Siegel, D. A. Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues. *Remote Sens. Environ.* **114**, 1791–1804 (2010).
23. Maritorena, S. & Siegel, D. A. Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sens. Environ.* **94**, 429–440 (2005).
24. Friedland, K. D. et al. Asymmetry in the rate of warming and the phenology of seasonal blooms in the Northeast US Shelf Ecosystem. *ICES J. Mar. Sci.* 775–786 (2023).
25. Friedland, K. D. et al. Phenology and time series trends of the dominant seasonal phytoplankton bloom across global scales. *Glob. Ecol. Biogeogr.* **27**, 551–569 (2018).
26. Bashmachnikov, I., Belonenko, T. V. & Koldunov, A. V. Attribution and predictability of climate-driven variability in global ocean color. *Remote Sens. Environ.* **137**, 55–68 (2013).
27. Lim, H. G., Dunne, J. P., Stock, C. A. & Kwon, M. Attribution and predictability of climate-driven variability in global ocean color. *J. Geophys. Res. Ocean.* **127**, e2022JC019121 (2022).
28. Kirby, R. R., Beaugrand, G., Kleparski, L., Goodall, S. & Lavender, S. Citizens and scientists collect comparable oceanographic data: measurements of ocean transparency from the Secchi Disk study and science programmes. *Sci. Rep.* **11**, 15499 (2021).
29. Brody, S. R., Lozier, M. S. & Dunne, J. P. A comparison of methods to determine phytoplankton bloom initiation. *J. Geophys. Res. Ocean.* **118**, 2345–2357 (2013).
30. Isles, P. D. F. & Pomati, F. An operational framework for defining and forecasting phytoplankton blooms. *Front. Ecol. Environ.* **19**, 443–450 (2021).
31. Siegel, D. A., Doney, S. C. & Yoder, J. A. The North Atlantic spring phytoplankton bloom and Sverdrup’s critical depth hypothesis. *Sci. (80-.)* **296**, 730–733 (2002).
32. Henson, S. A., Robinson, I., Allen, J. T. & Waniek, J. J. Effect of meteorological conditions on interannual variability in timing and magnitude of the spring bloom in the Irminger Basin, North Atlantic. *Deep. Res. Part I Oceanogr. Res. Pap.* **53**, 1601–1615 (2006).
33. Sapiano, M. R. P., Brown, C. W., Schollaert Uz, S. & Vargas, M. Establishing a global climatology of marine phytoplankton phenological characteristics. *J. Geophys. Res. Ocean.* **117**, C08026 (2012).
34. Signorini, S. R., Franz, B. A. & McClain, C. R. Chlorophyll variability in the oligotrophic gyres: Mechanisms, seasonality and trends. *Front. Mar. Sci.* **2**, 1 (2015).
35. Behrenfeld, M. J. & Falkowski, P. G. Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.* **42**, 1–20 (1997).
36. Chiswell, S. M., Bradford-Grieve, J., Hadfield, M. G. & Kennan, S. C. Climatology of surface chlorophyll a, autumn-winter and spring blooms in the southwest Pacific Ocean. *J. Geophys. Res. Ocean.* **118**, 1003–1018 (2013).
37. Brando, V. E. et al. Phytoplankton bloom dynamics in the baltic sea using a consistently reprocessed time series of multi-sensor reflectance and novel chlorophyll-a retrievals. *Remote Sens.* **13**, 3071 (2021).
38. McClain, C. R. et al. Satellite ocean-color observations of the tropical Pacific Ocean. *Deep. Res. Part II Top. Stud. Oceanogr.* **49**, 2533–2560 (2002).
39. Rivas, A. L., Dogliotti, A. I. & Gagliardini, D. A. Seasonal variability in satellite-measured surface chlorophyll in the Patagonian Shelf. *Cont. Shelf Res.* **26**, 703–720 (2006).
40. Yoder, J. A., McClain, C. R., Feldman, G. C. & Esaias, W. E. Annual cycles of phytoplankton chlorophyll concentrations in the global ocean: A satellite view. *Glob. Biogeochem. Cycles* **7**, 181–193 (1993).
41. Pershing, A. J. & Stamieszkin, K. The North Atlantic Ecosystem, from Plankton to Whales. *Ann. Rev. Mar. Sci.* **12**, 339–359 (2020).
42. Smith, W. O. & Demaster, D. J. Phytoplankton biomass and productivity in the Amazon River plume: Correlation with seasonal river discharge. *Cont. Shelf Res.* **16**, 291–319 (1996).
43. Ciotti, A. M., Garcia, Ca. E. & Jorge, D. S. F. Temporal and meridional variability of Satellite-estimates of surface chlorophyll concentration over the Brazilian continental shelf. *Panamjas* **5**, 64–81 (2010).
44. Quartly, G. D. & Srokosz, M. A. Satellite observations of the Agulhas Current system. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **361**, 51–56 (2003).
45. Lutjeharms, J. R. E. & Machu, E. An upwelling cell inshore of the East Madagascar Current. *Deep. Res. Part I Oceanogr. Res. Pap.* **47**, 2405–2411 (2000).
46. Quartly, G. D., Buck, J. J. H., Srokosz, M. A. & Coward, A. C. Eddies around Madagascar - The retroflexion re-considered. *J. Mar. Syst.* **63**, 115–129 (2006).
47. Babin, S. M., Carton, J. A., Dickey, T. D. & Wiggert, J. D. Satellite evidence of hurricane-induced phytoplankton blooms in an oceanic desert. *J. Geophys. Res. Ocean.* **109**, C03043 (2004).
48. Gaube, P. & McGillicuddy, D. J. The influence of Gulf Stream eddies and meanders on near-surface chlorophyll. *Deep. Res. Part I Oceanogr. Res. Pap.* **122**, 1–16 (2017).
49. Strutton, P. G., Ryan, J. P. & Chavez, F. P. Enhanced chlorophyll associated with tropical instability waves in the equatorial Pacific. *Geophys. Res. Lett.* **28**, 2005–2008 (2001).
50. Mohammadi, S. Effect of outliers on the fractal dimension estimation. *Fractals* **19**, 233–241 (2011).
51. Morozov, E., Korosov, A., Pozdnyakov, D., Pettersson, L. & Sychev, V. A new area-specific bio-optical algorithm for the Bay of Biscay and assessment of its potential for SeaWiFS and MODIS/Aqua data merging. *Int. J. Remote Sens.* **31**, 6541–6565 (2010).
52. Staehr, S. U., Van der Zande, D., Staehr, P. A. U. & Markager, S. Suitability of multisensory satellites for long-term chlorophyll

- assessment in coastal waters: A case study in optically-complex waters of the temperate region. *Ecol. Indic.* **134**, 108479 (2022).
53. Sathyendranath, S., Brewin, R. J. W., Jackson, T., Mélin, F. & Platt, T. Ocean-colour products for climate-change studies: What are their ideal characteristics? *Remote Sens. Environ.* **203**, 125–138 (2017).
54. Gómez-Jakobsen, F., Ferrera, I., Yebra, L. & Mercado, J. M. Two decades of satellite surface chlorophyll a concentration (1998–2019) in the Spanish Mediterranean marine waters (Western Mediterranean Sea): Trends, phenology and eutrophication assessment. *Remote Sens. Appl. Soc. Environ.* **28**, 100855 (2022).
55. Boyce, D. G., Lewis, M. R. & Worm, B. Global phytoplankton decline over the past century. *Nature* **466**, 591–596 (2010).
56. Mouw, C. B., Ciocchetto, A. B. & Yoder, J. A. A Satellite Assessment of Environmental Controls of Phytoplankton Community Size Structure. *Glob. Biogeochem. Cycles* **33**, 540–558 (2019).
57. Hammond, M. L., Beaulieu, C., Henson, S. A. & Sahu, S. K. Regional surface chlorophyll trends and uncertainties in the global ocean. *Sci. Rep.* **10**, 15273 (2020).
58. Gregg, W. W., Rousseaux, C. S. & Franz, B. A. Global trends in ocean phytoplankton: a new assessment using revised ocean colour data. *Remote Sens. Lett.* **8**, 1102–1111 (2017).
59. Agarwal, V., James, C. C., Widdicombe, C. E. & Barton, A. D. Intra-seasonal predictability of natural phytoplankton population dynamics. *Ecol. Evol.* **11**, 15720–15739 (2021).
60. Moore, T. S., Campbell, J. W. & Dowell, M. D. A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sens. Environ.* **113**, 2424–2430 (2009).
61. Mélin, F., Sclép, G., Jackson, T. & Sathyendranath, S. Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets. *Remote Sens. Environ.* **177**, 107–124 (2016).
62. Tilstone, G. H. et al. Performance of Ocean Colour Chlorophyll a algorithms for Sentinel-3 OLCI, MODIS-Aqua and Suomi-VIIRS in open-ocean waters of the Atlantic. *Remote Sens. Environ.* **260**, 112444 (2021).
63. Reygondeau, G. et al. Dynamic biogeochemical provinces in the global ocean. *Glob. Biogeochem. Cycles* **27**, 1046–1058 (2013).
64. Oliver, M. J. & Irwin, A. J. Objective global ocean biogeographic provinces. *Geophys. Res. Lett.* **35**, L15601 (2008).
65. Devred, E., Sathyendranath, S. & Platt, T. Delineation of ecological provinces using ocean colour radiometry. *Mar. Ecol. Prog. Ser.* **346**, 1–13 (2007).
66. Sonnewald, M., Sonnewald, M., Dutkiewicz, S., Hill, C. & Forget, G. Elucidating ecological complexity: Unsupervised learning determines global marine eco-provinces. *Sci. Adv.* **6**, eaay4740 (2020).
67. Frees, E. W. *Regression modeling with actuarial and financial applications*. Cambridge University Press <https://doi.org/10.1017/CBO9780511814372> (2009).
68. Sevcikova, H., Percival, D. & Gneiting T. fractalDIM: Estimation of Fractal Dimensions. R package version 0.8-5. (2021).
69. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> (2023).
70. Hijmans, R. raster: Geographic Data Analysis and Modeling. R package version 3.6-20, <https://rspatial.org/raster> (2023).
71. Becker, R. A., Minka, T. P. & Deckmyn, A. maps: Draw Geographical Maps. R package version 3.4.1. (2022).
72. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, <https://ggplot2.tidyverse.org> (2016).
73. Wilke, C. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 1.1.1, <https://wilkelab.org/cowplot/> (2020).
74. Venables, W. N. & Ripley, B. D. Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0. <http://www.stats.ox.ac.uk/pub/MASS4/>.
75. Simon, G. et al. Rvision - Colorblind-Friendly Color Maps for R. R package version 0.6.2. <https://github.com/sjmgarnier/viridis/> (2021).
76. Pierce, D. ncd4f: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. R package version 1.21, <https://cirrus.ucsd.edu/~pierce/ncdf/> (2023).
77. Agarwal, V. vitul-agarwal-1/chl-complexity: v1 [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10498362> (2024).

## Acknowledgements

VA and CBM acknowledge support from the Rhode Island Sea Grant (Grant number: NA22OAR4170123 to CBM). CBM acknowledges support from CINAR (Cooperative Institute for the North Atlantic Region; Grant number: NA19OAR4320074 to CBM). This work was supported by a grant from the Simons Foundation (LS-ECIAMEE-00001549 to KI). We would like to acknowledge the team at the Unity high-performance computing cluster located at the Massachusetts Green High Performance Computing Center (MGHPCC).

## Author contributions

VA conducted the analysis and wrote the first draft of the manuscript. JC, KI and CBM assisted with all aspects of this work.

## 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/s41467-024-45976-8>.

**Correspondence** and requests for materials should be addressed to Vitul Agarwal.

**Peer review information** *Nature Communications* thanks Jesús Mercado and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. 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