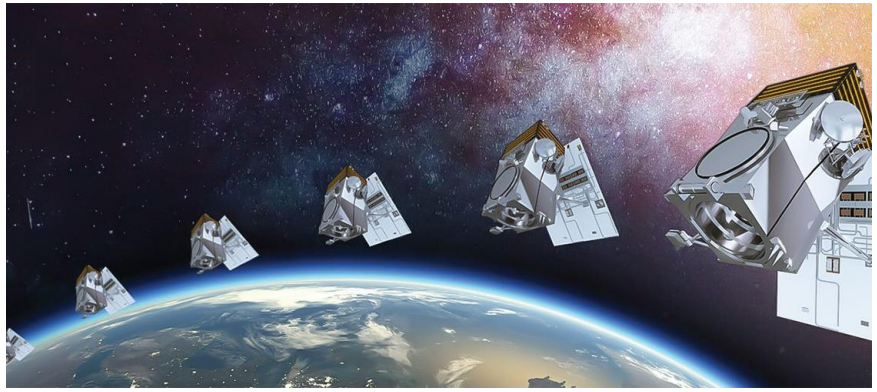
Statistical implications of temporal aggregation in inland water quality monitoring using multi-platform satellite missions

Megan Coffer
NOAA GST
megan.coffer@noaa.gov

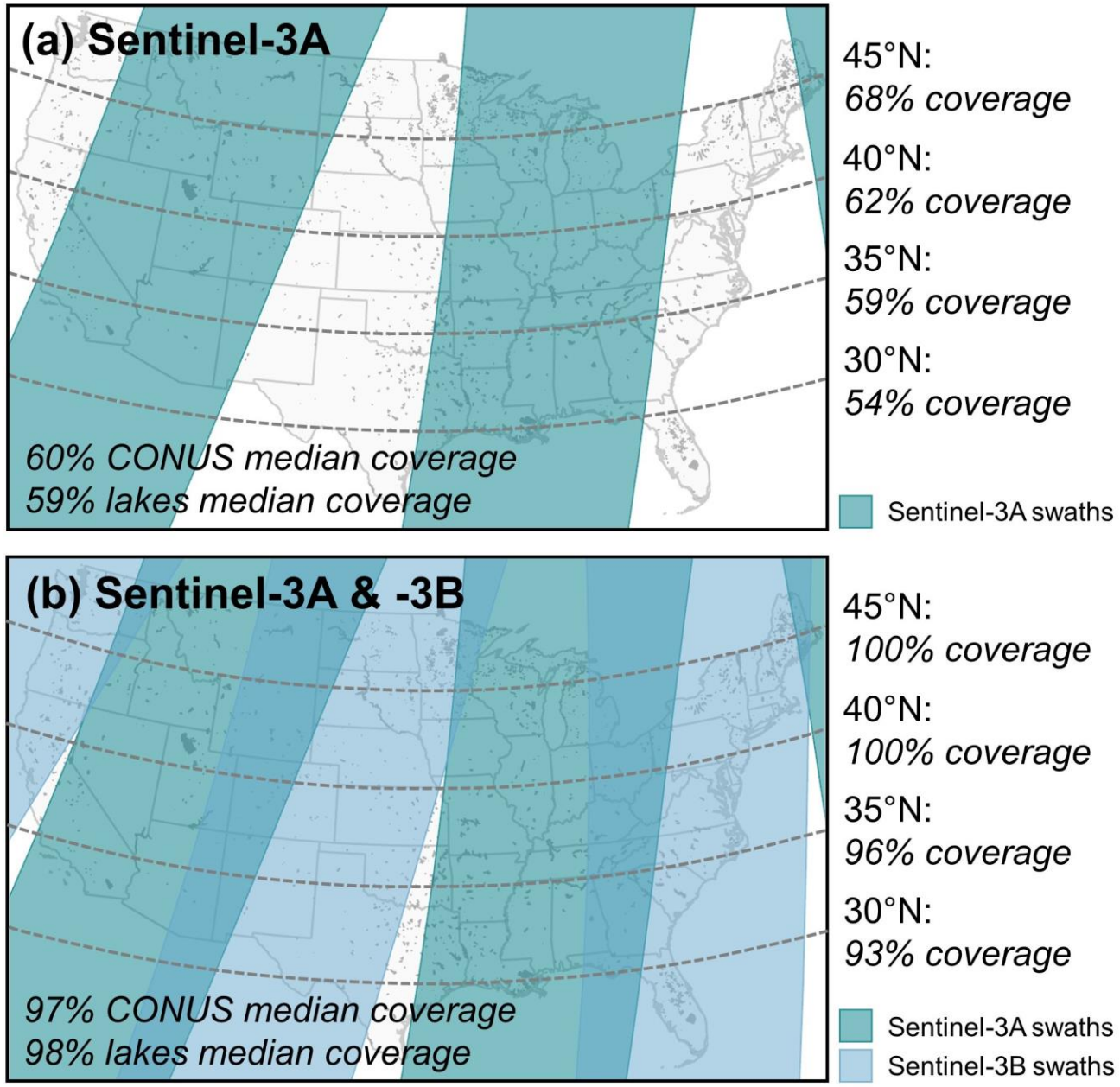
28 March 2024



*Upcoming Maxar **WorldView Legion** constellation to include at least **six platforms**, launched in pairs over the next several years*

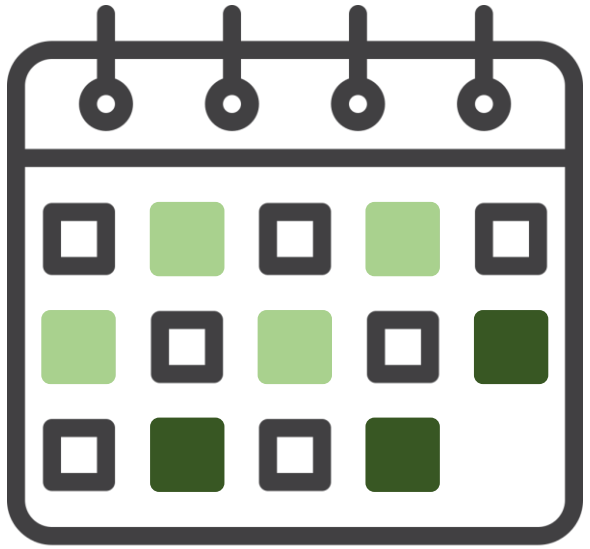
*Planet has launched **hundreds of CubeSats**, including their RapidEye and PlanetScope constellations*

*The Copernicus Programme launched **Sentinel-2A in 2015** followed by **Sentinel-2B in 2017** and **Sentinel-3A in 2016** followed by **Sentinel-3B in 2018***



Spatial coverage of Sentinel-3 increased from a daily median of 60% CONUS coverage with just Sentinel-3A to ***nearly 100% daily coverage after the inclusion of Sentinel-3B***

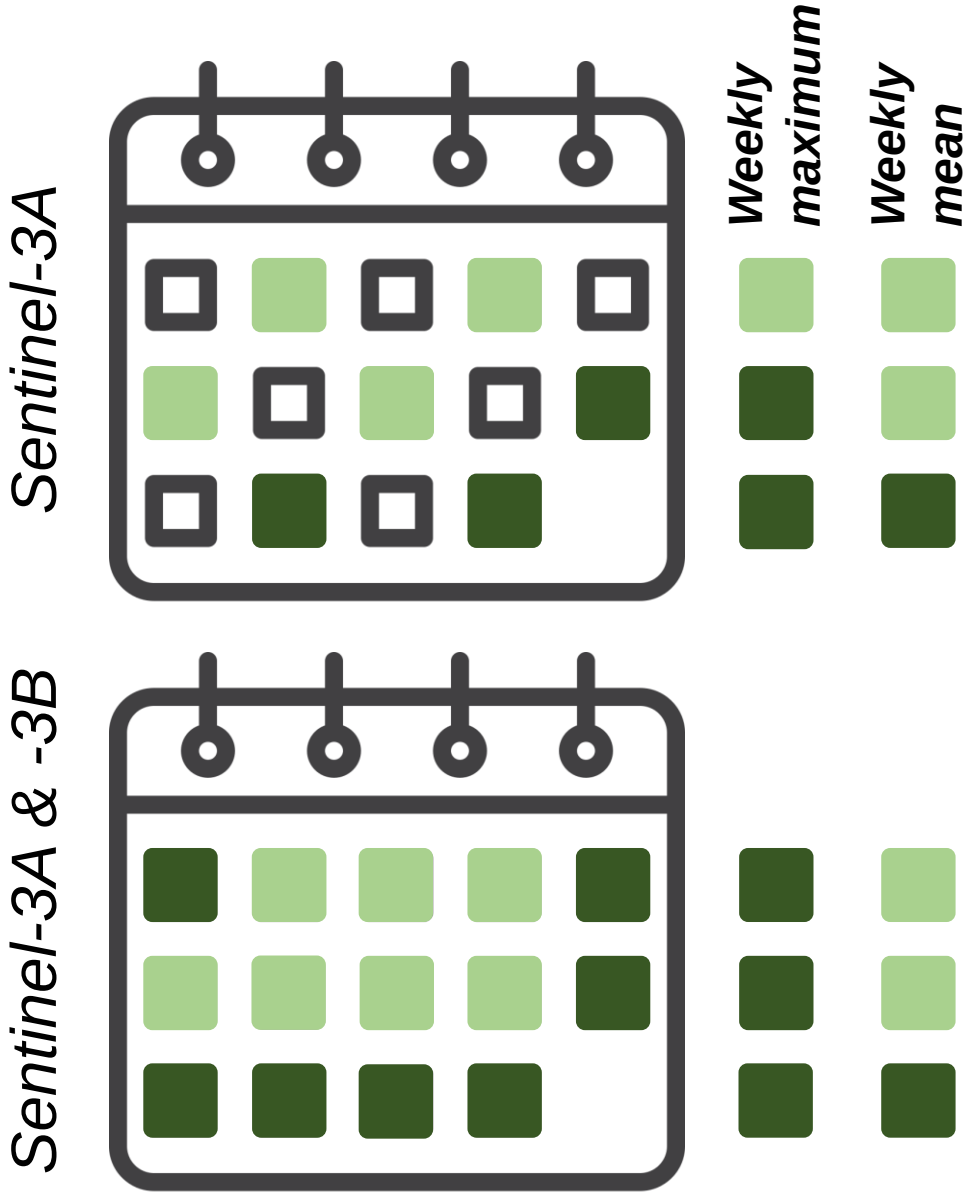
Sentinel-3A



Sentinel-3A & -3B

Increased temporal frequency provides the ***opportunity to observe additional, potentially more extreme water quality events;*** *change assessments may reflect changes in observational frequency rather than true environmental trends*

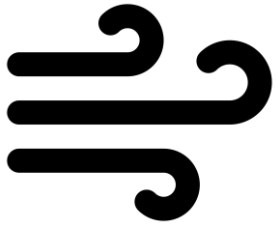
When aggregating over time (e.g., weekly or monthly composites), ***temporal aggregation strategies can impact results***



Maximum data value: *In cases where toxin data are unavailable, cyanobacterial biomass should use maximum concentrations* *(Ibelings et al., 2021)*

Measure of central tendency: *More robust representation as it minimizes outliers and fluctuations due to low-quality data, errors, or localized events* *(Piao et al., 2023)*

Valid sub-weekly changes



Wind-driven blooms

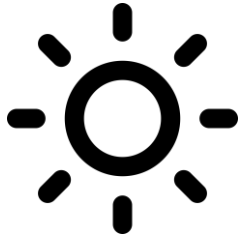
Short-lived events

Changes in buoyancy

Erroneous sub-weekly changes

Cloud & cloud shadow

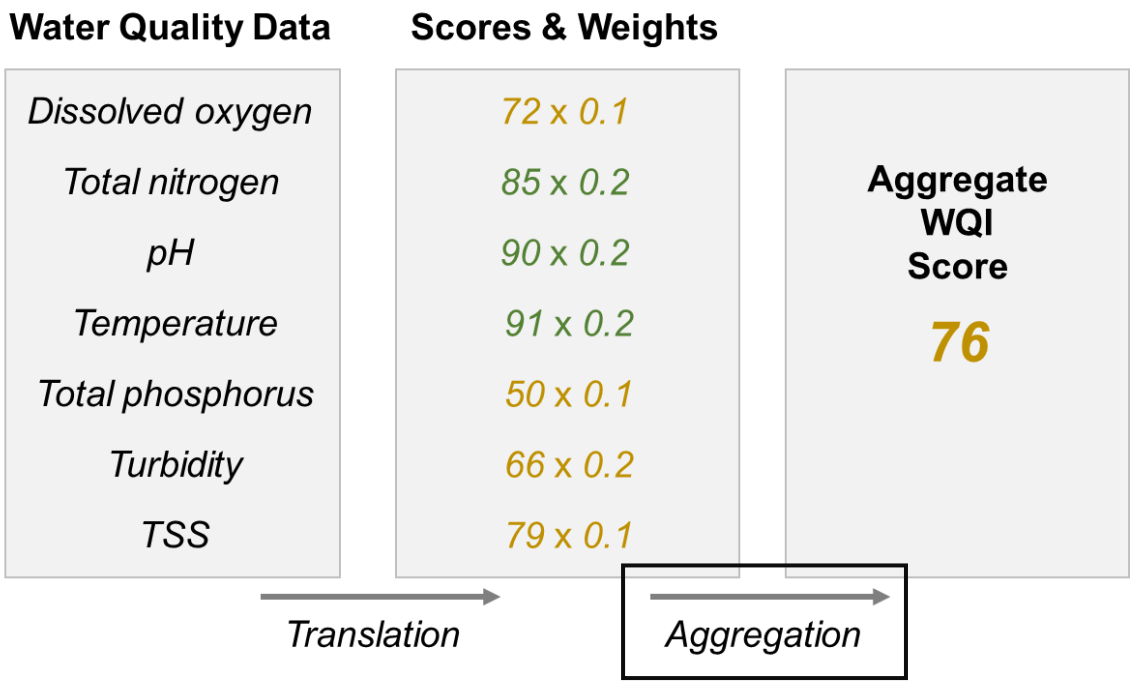
Adjacency effects



Sun glint

Intent is not to determine which aggregation strategy most closely matches true conditions, but instead to assess ***how each temporal aggregation approach can impact change statistics***

***Water Quality Index (WQI)** is a general method used to report water quality data for many parameters as a single number or score*



*Using water quality data, a study found that **aggregation method can impact benefit-cost analysis**, with total benefit estimates **varying from $82 million to $504 million nationally**.*

Walsh and Wheeler, 2012. *NCEE Working Paper.*

Assess impact of temporal aggregation methods on change and trend analyses from multi-platform satellite missions

Using simulated data values, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean (arithmetic)

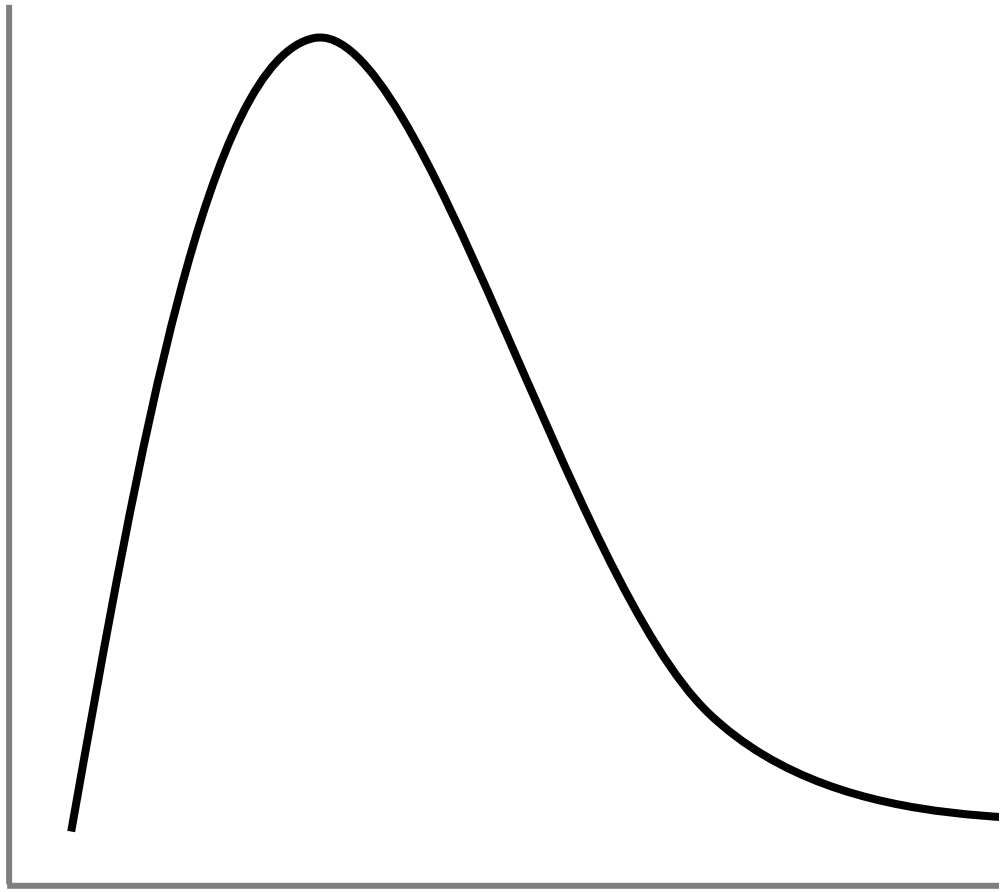
Using satellite-estimated cyanobacteria concentrations, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean

Evaluate differences in change assessments of satellite-estimated cyanobacteria concentrations aggregated via the maximum and mean

Using simulated data values, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean (arithmetic)

Using satellite-estimated cyanobacteria concentrations, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean

Evaluate differences in change assessments of satellite-estimated cyanobacteria concentrations aggregated via the maximum and mean

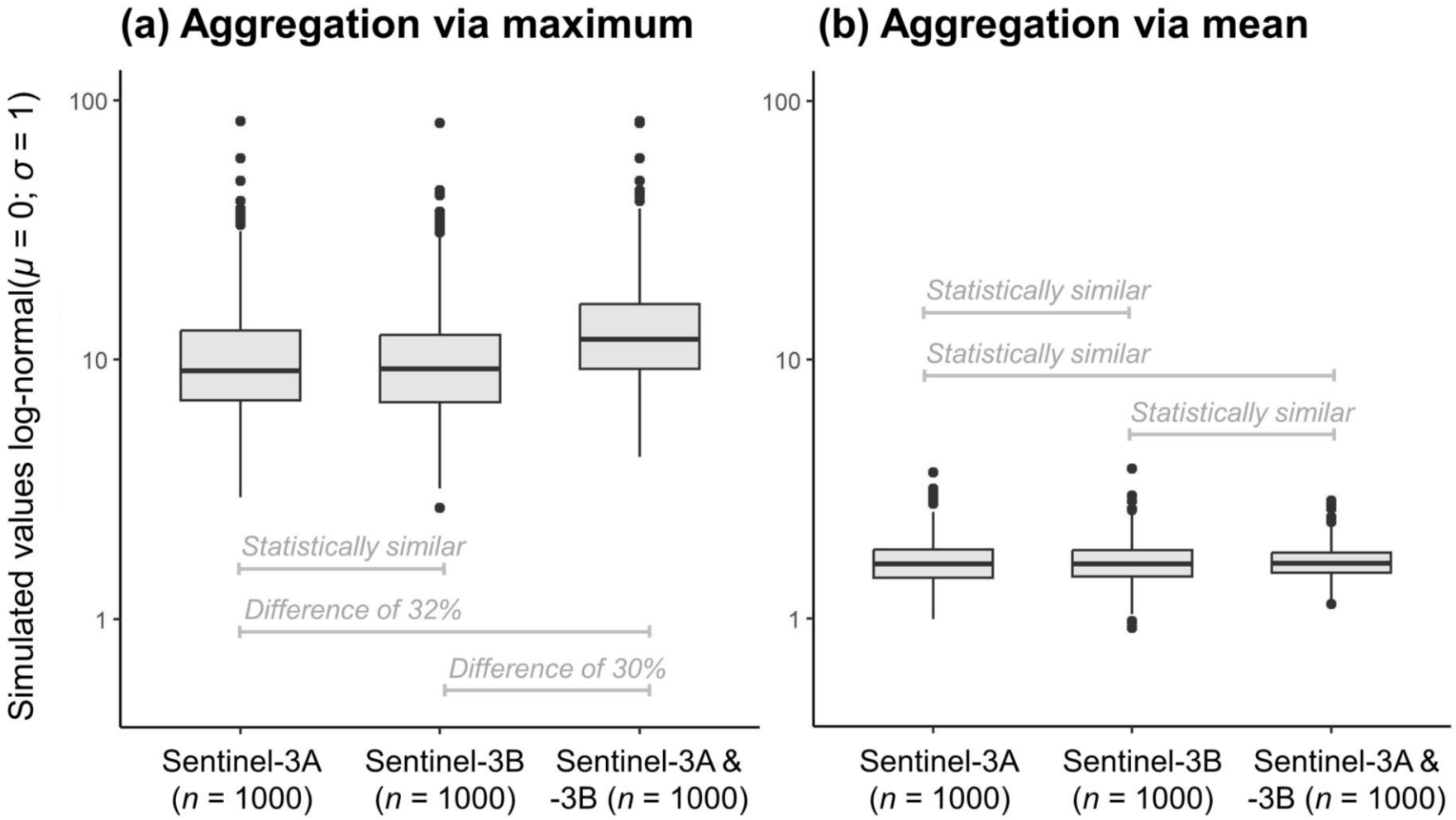


Without continuous, large-scale field observations, simulation study can ***guide interpretation of satellite-based results***

Water quality observations from Sentinel-3A, Sentinel-3B, and combined Sentinel-3A & -3B simulated using a ***log-normal distribution***

Distributions of aggregated values compared via ***Wilcoxon signed-rank test***

Wilcoxon, 1945. *Biometrics.*



When aggregating via the maximum, Sentinel-3A & -3B values were much higher than Sentinel-3A or Sentinel-3B individually; this ***increase is no longer evident when aggregating via the mean***

Using simulated data values, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean (arithmetic)

Using satellite-estimated cyanobacteria concentrations, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean

Evaluate differences in change assessments of satellite-estimated cyanobacteria concentrations aggregated via the maximum and mean

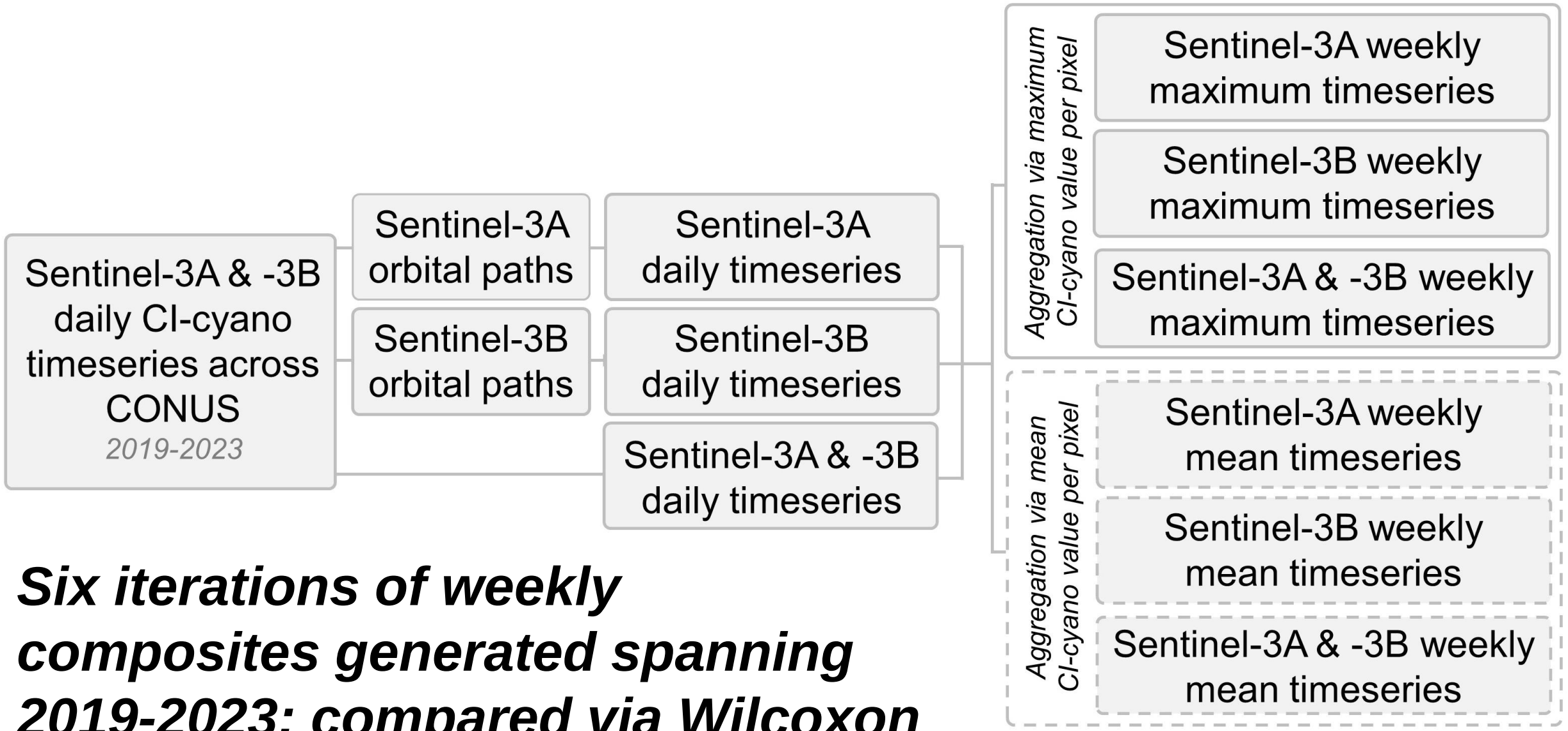
The **Cyanobacteria Assessment Network** (CyAN)

Interagency project *between NOAA, NASA, USGS, and USEPA*

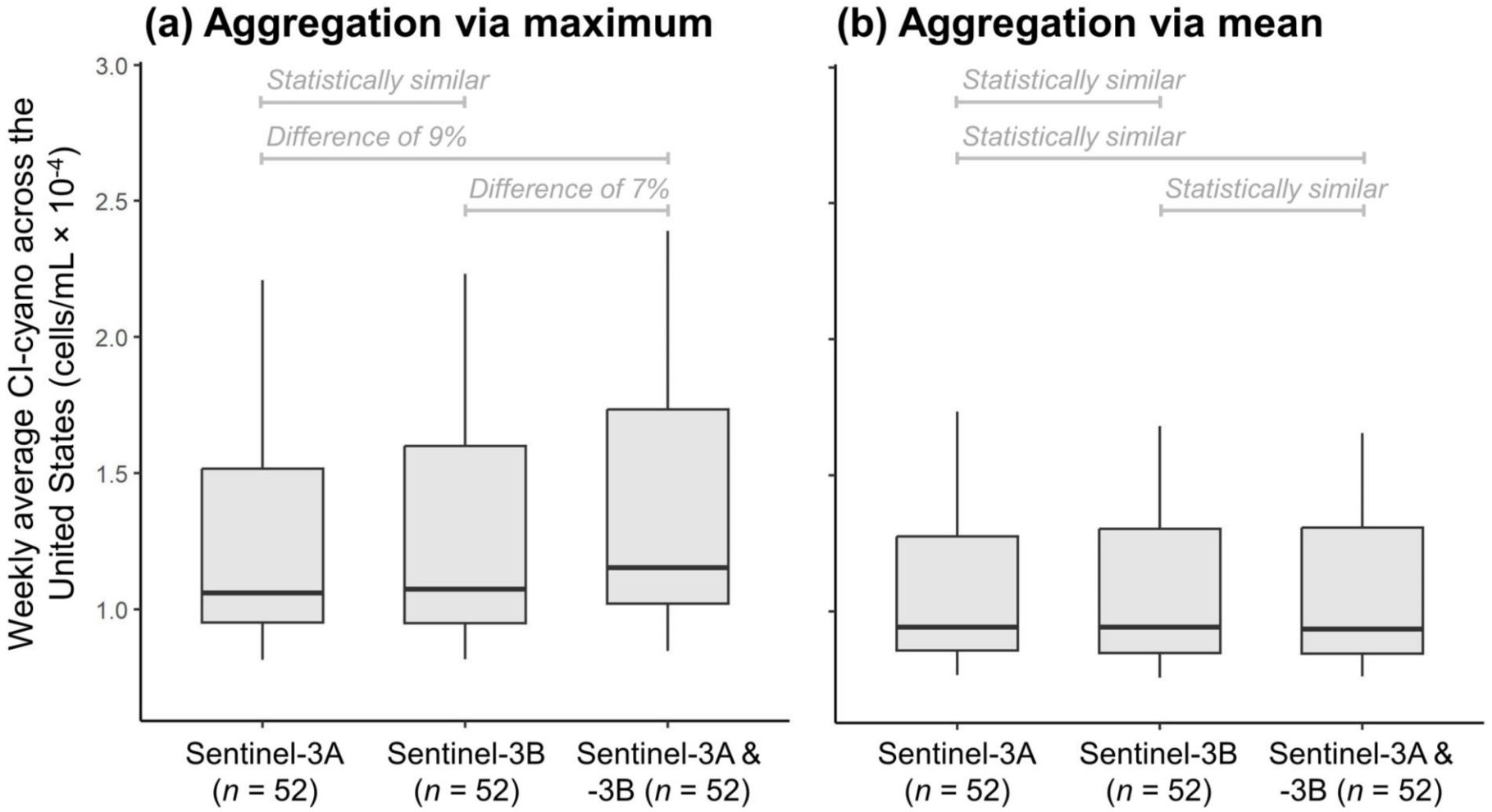
Provided ***national coverage*** *of* ***cyanobacteria*** *concentration using the* ***cyanobacteria index (CI-cyano)***

Data are delivered as daily data and as weekly composites aggregated via the ***maximum CI-cyano value***

Schaeffer et al., 2015. *Eos Transactions.*



Six iterations of weekly composites generated spanning 2019-2023; compared via Wilcoxon signed-rank test

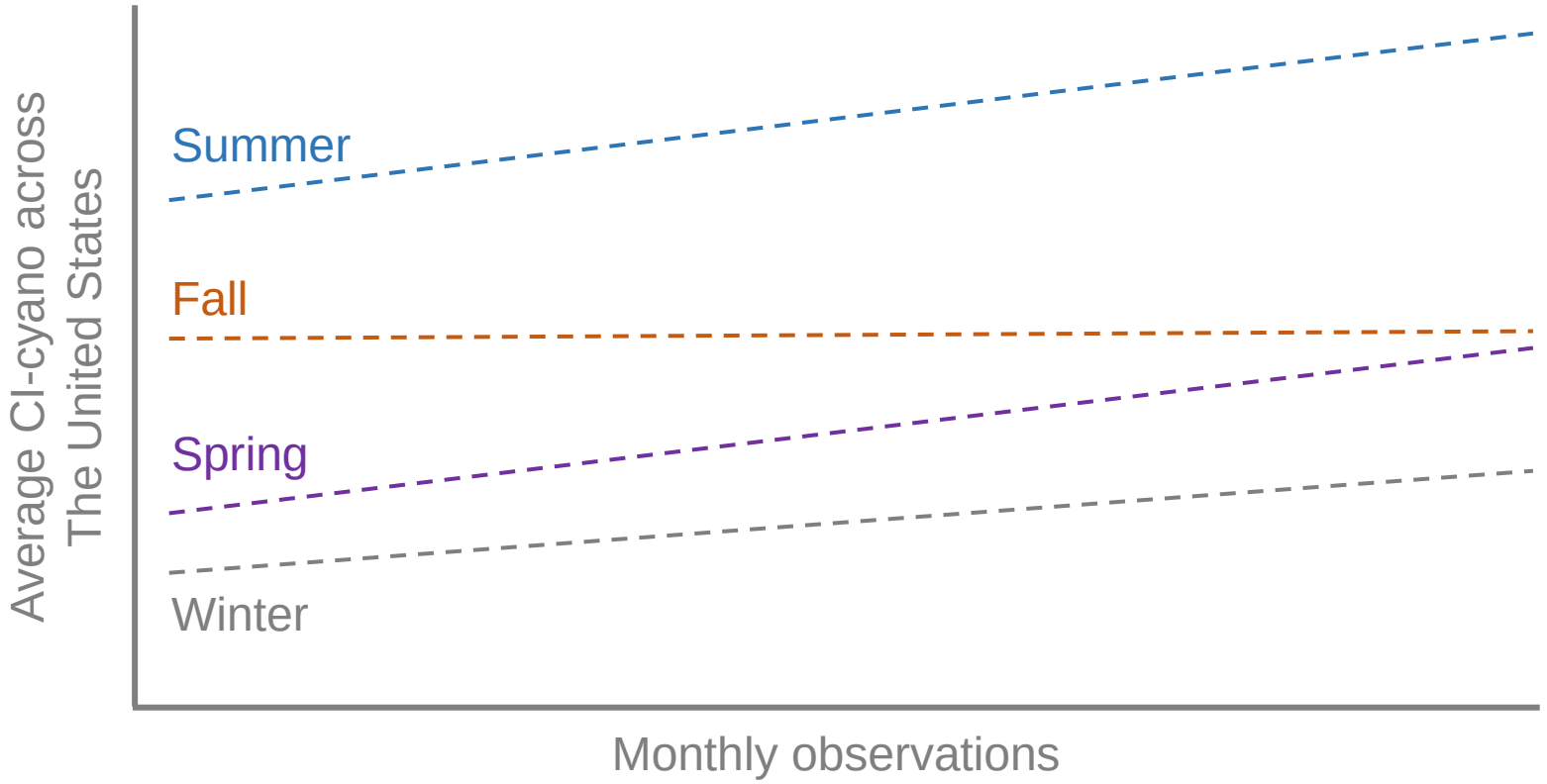


CI-cyano results followed the simulation study, where aggregation via ***maximum generated higher CI-cyano across CONUS*** *for combined Sentinel-3A & -3B;* ***distributions similar when aggregating via mean***

Using simulated data values, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean (arithmetic)

Using satellite-estimated cyanobacteria concentrations, assess the impact of improved temporal coverage on data values aggregated via the maximum and mean

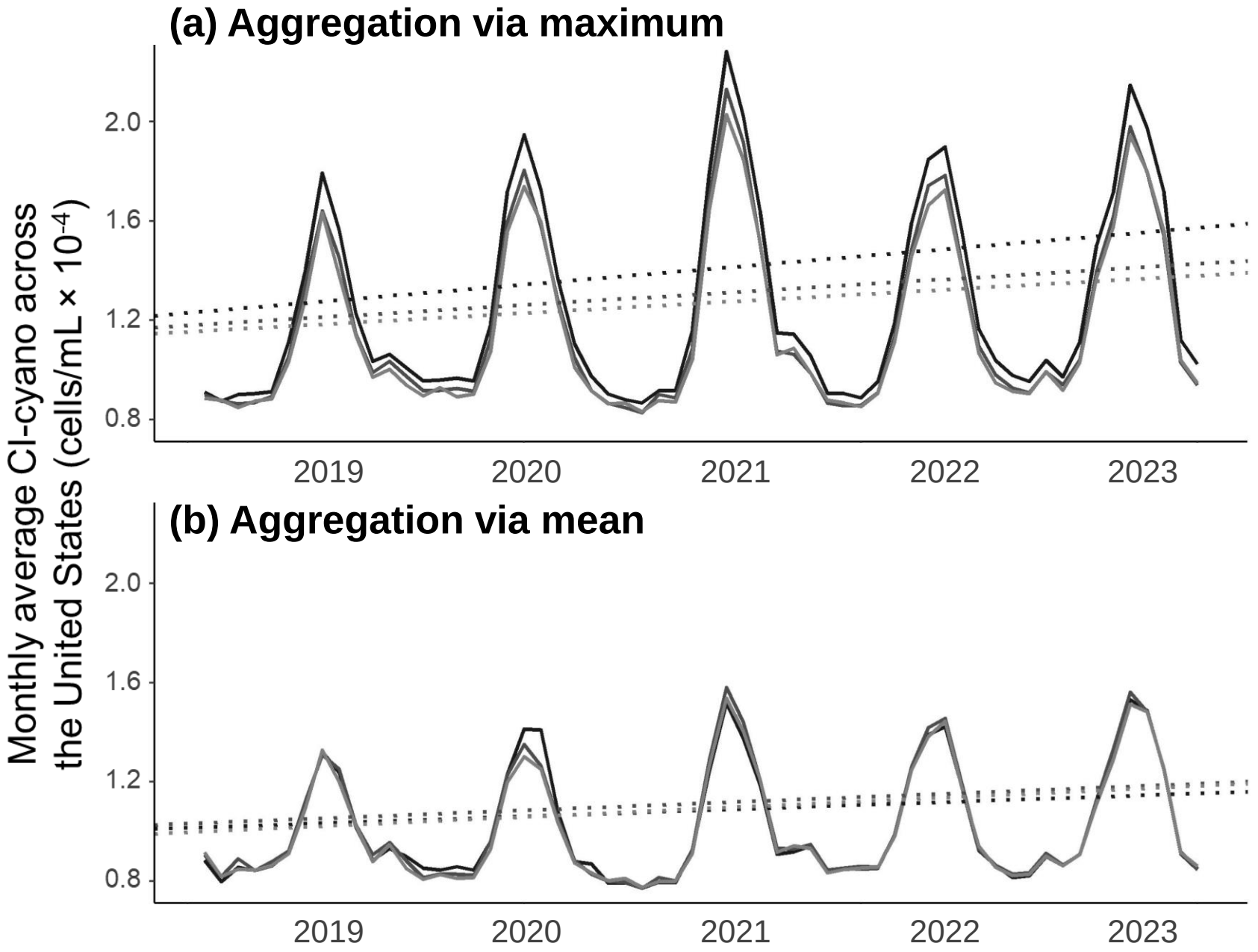
Evaluate differences in change assessments of satellite-estimated cyanobacteria concentrations aggregated via the maximum and mean



*Timeseries of **monthly average CI-cyano** generated for each of the six iterations*

*Trends assessed using the nonparametric **seasonal Mann-Kendall** test for trend and associated **Thiel-Sen slope***

Kendall, 1955; Mann, 1945; Sen, 1968; Theil, 1992.



Sentinel-3A & -3B:
Moderate increase (16%)

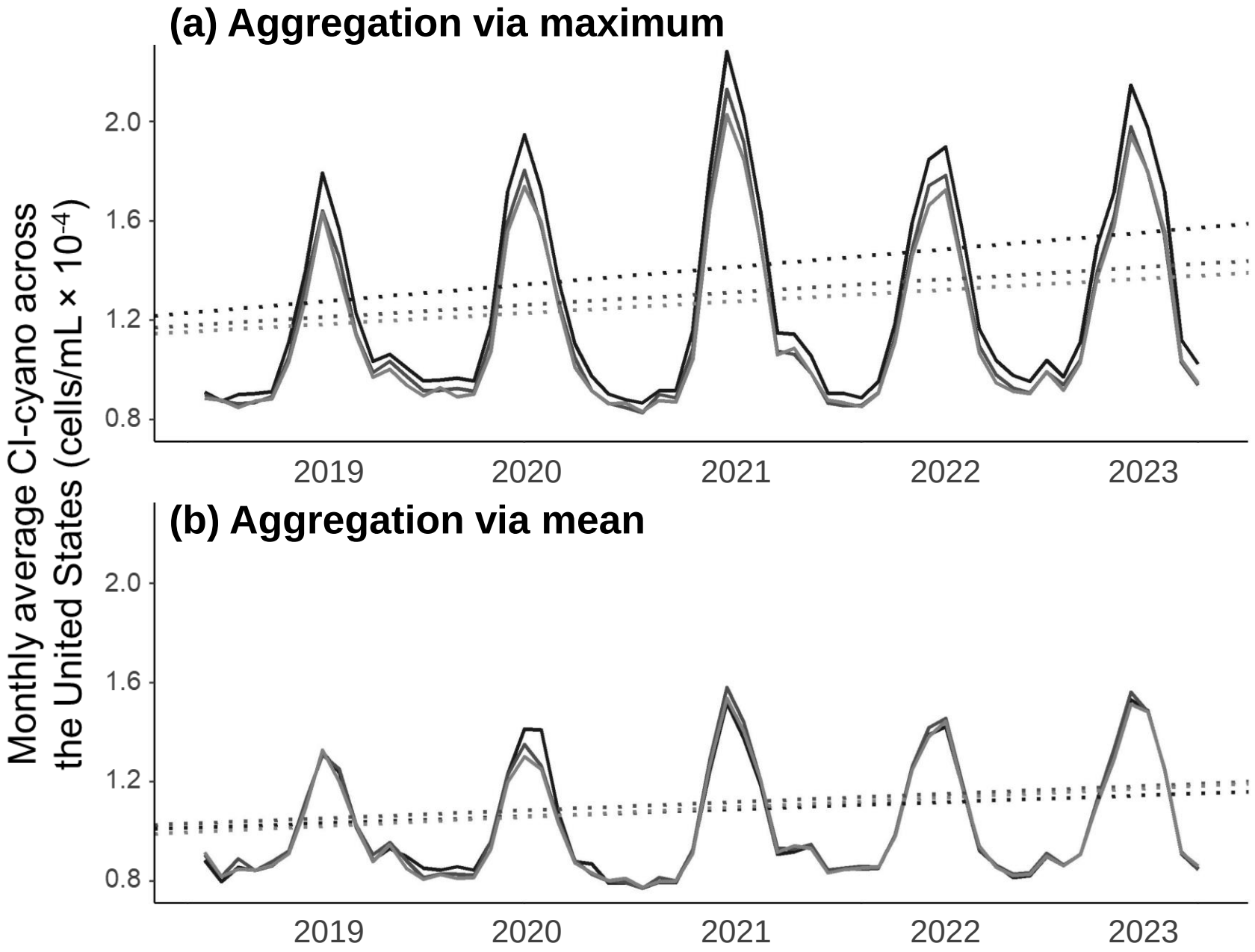
Sentinel-3A:
Moderate increase (12%)

Sentinel-3B:
Moderate increase (12%)

Sentinel-3A & -3B:
Moderate increase (8%)

Sentinel-3A:
Moderate increase (9%)

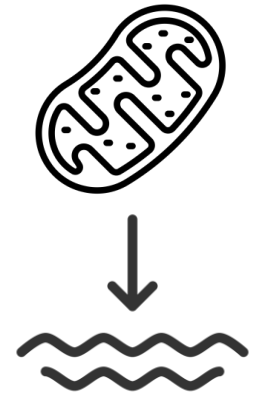
Sentinel-3B:
Moderate increase (9%)



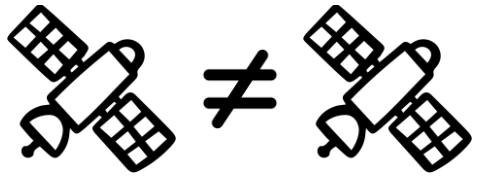
***Strength of trend unchanged**, consistent with previous results (Schaeffer et al., 2022; Hammond et al., 2020)*

*Percentage change **higher** for combined Sentinel-3A & -3B when **aggregating via maximum**; slightly **lower** when **aggregating via mean***

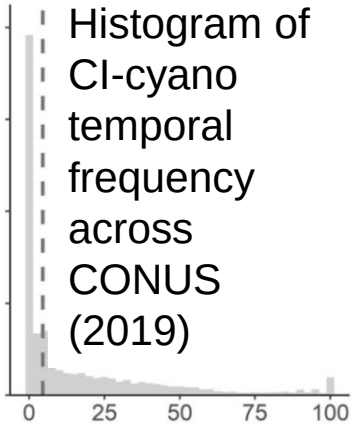
Movement within the water column: *Cyanobacteria migrate vertically at diurnal timescales (Cameron et al., 2024); collection times were within 20-45 minutes for Sentinel-3A and -3B (Schaeffer et al., 2022)*



Sensor differences: *Sentinel-3A reflectances higher than from Sentinel-3B (Lamquin et al., 2020)*



Skew observation distribution: *CI-cyano skewed toward high values (Coffer et al., 2021); can impact measures of central tendency*



Aggregation via mean

Aggregate via a measure of central tendency (such as the mean) likely more appropriate for analyses that span inconsistent or imbalanced sampling frequency

Temporal aggregation via the mean produced more consistent data distributions over for a simulation study and using Sentinel-3 data

Aggregation via maximum

Observed changes may reflect changes in observational frequency rather than true environmental changes

Next steps

Include observations from 2016 through 2018 (when only Sentinel-3A was available) in trend assessment

Perform a similar analysis using red-edge reflectance values from Sentinel-2A and -2B

Simulate upcoming impacts from the inclusion of additional satellite platforms in each mission, including Sentinel-3C & -3D and Sentinel-2C & -2D

Incorporate results for other measures of central tendency (*e.g.*, median data value)