Compare SKT and GAM

Mike Wessel

2023-04-17

Initial data analysis conducted using R script SKT GAM Comp Run.r which saves out a file called df_total2. That file contains the results of SKT and the GAM based method applied to 15 sites including fixed stations, and estuarine and lake probabilistic designs. All sites contained at least 60 nonmissing observations which is part of the standard PC inclusion criteria.

The analysis includes two types of GAM based runs, one on the raw data (i.e., not date averaged so e.g. W1 would have 4 observations for each date), and one using a date average which preserves the period of time when there were 9 sampling periods which were then reduced to 8.

The SKT uses sample period median values for its test based on the current 8 sample period assignments used for the Pinellas County Report Card.

The results below suggest there a total of 8 discrepancies which we will examine one at a time in alpha numeric order.

All the following plots and analysis use the “Raw” data to express the data in its most basic form

Trend Result Comparison

station	param	skt	GAM_Avg	GAM_Raw	skt_direction
08-03	Chl_a	Trend	Trend	Trend	Decreasing
08-03	TN	Trend	Trend	Trend	Decreasing
09-02	Chl_a	Trend	Trend	Trend	Increasing
09-02	TN	No Trend	Trend	Trend	Not
15-04	Chl_a	Trend	Trend	Trend	Increasing
15-04	TN	Trend	Trend	Trend	Decreasing
17-03	Chl_a	Trend	Trend	Trend	Increasing
17-03	TN	No Trend	No Trend	No Trend	Not
19-08	Chl_a	Trend	Trend	Trend	Decreasing
19-08	TN	Trend	No Trend	No Trend	Decreasing
SA	Chl_a	Trend	Trend	Trend	Decreasing
SA	TN	Trend	Trend	Trend	Decreasing
SB	Chl_a	Trend	Trend	Trend	Decreasing

station	param	skt	GAM_Avg	GAM_Raw	skt_direction
SB	TN	Trend	Trend	Trend	Decreasing
W1	Chl_a	Trend	No Trend	No Trend	Increasing
W1	TN	Trend	Trend	Trend	Increasing
W2	Chl_a	Trend	Trend	Trend	Increasing
W2	TN	Trend	Trend	Trend	Increasing
W3	Chl_a	No Trend	No Trend	No Trend	Not
W3	TN	Trend	Trend	Trend	Increasing
W4	Chl_a	No Trend	No Trend	No Trend	Not
W4	TN	Trend	Trend	No Trend	Increasing
W5	Chl_a	No Trend	No Trend	Trend	Not
W5	TN	No Trend	No Trend	No Trend	Not
W6	Chl_a	No Trend	No Trend	No Trend	Not
W6	TN	Trend	Trend	No Trend	Increasing
W7	Chl_a	Trend	No Trend	No Trend	Decreasing
W7	TN	Trend	Trend	Trend	Increasing
W8	Chl_a	No Trend	No Trend	Trend	Not
W8	TN	Trend	Trend	Trend	Increasing

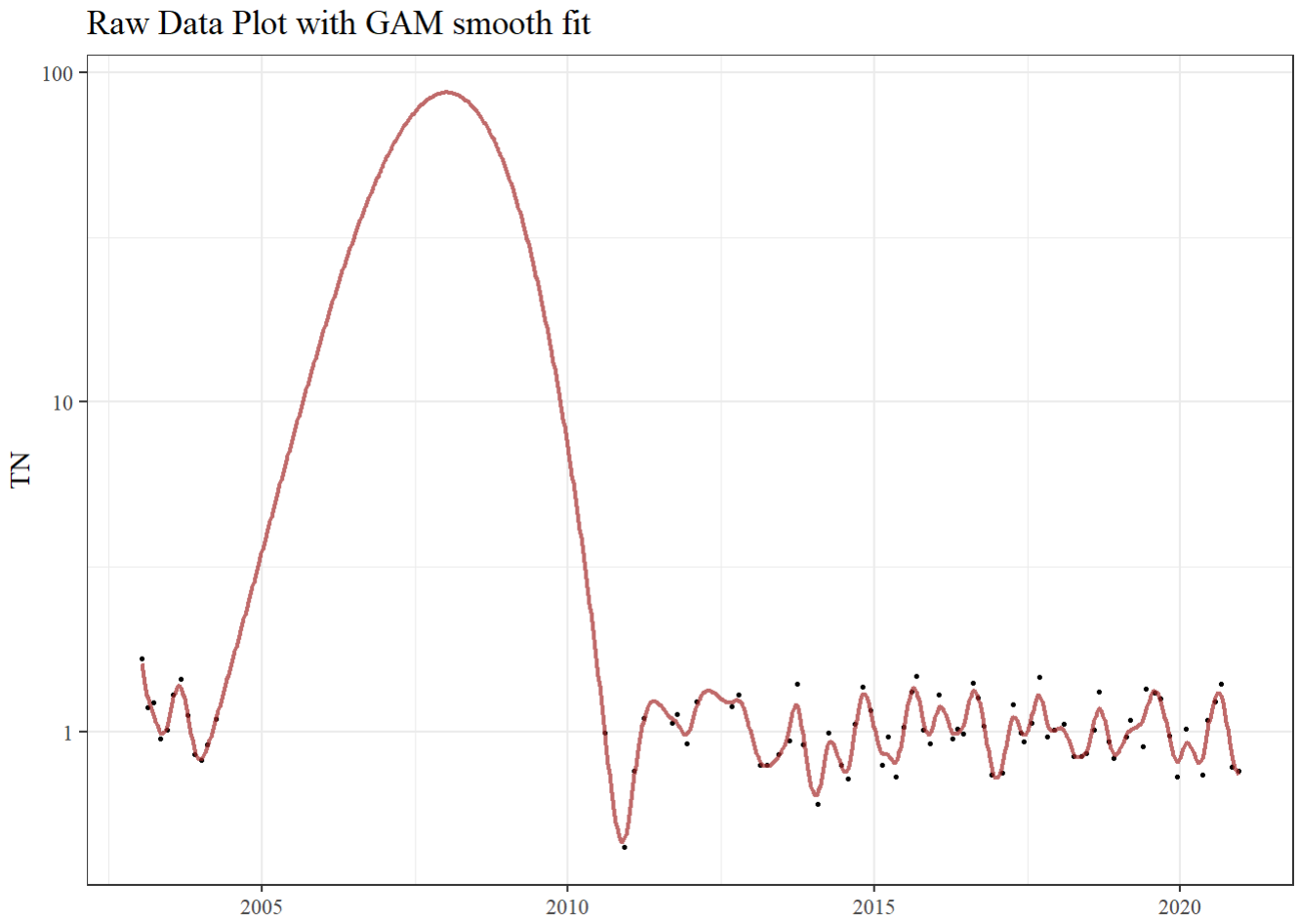
First Example

09-02 TN: The SKT suggests NO TREND while both GAM based tests suggest trend.

This site was not found to contain significant autocorrelation according to the SKT methodology.

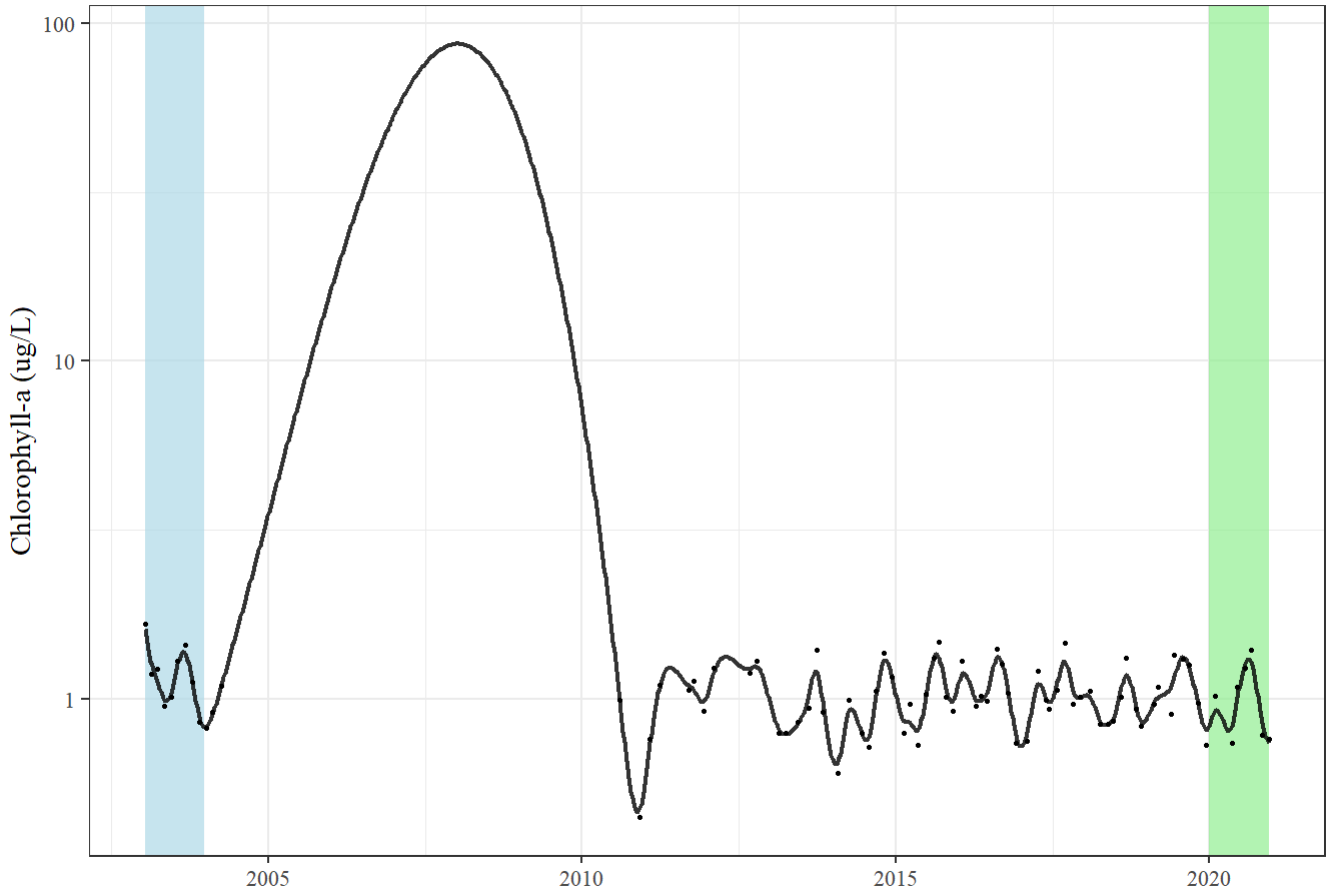
As seen in the plots below there is a data gap that may be throwing off the GAM method

```
## reducing knots for cont_year spline from 156
## reducing knots for cont_year spline from 143
## reducing knots for cont_year spline from 130
## reducing knots for cont_year spline from 117
## reducing knots for cont_year spline from 104
## reducing knots for cont_year spline from 91
```



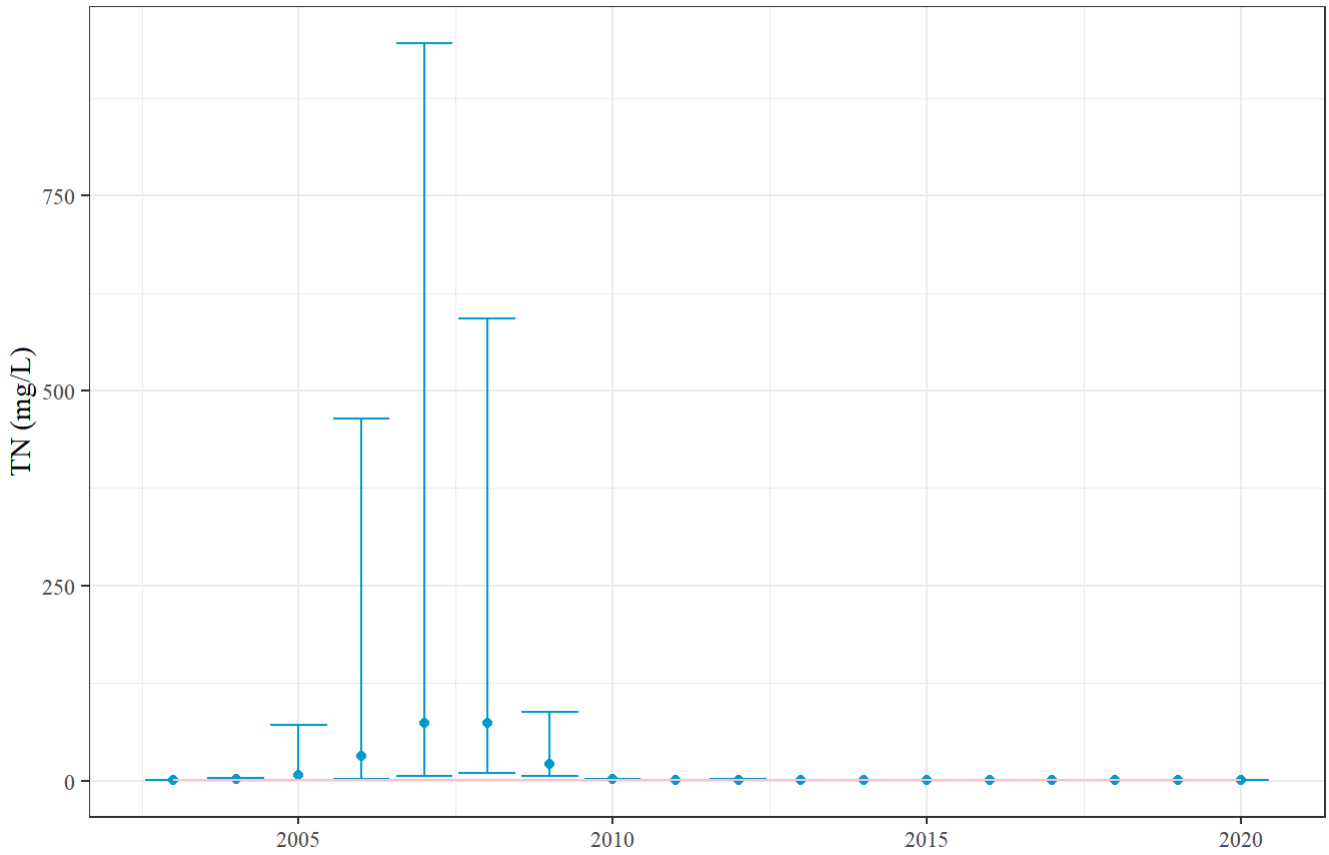

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 1.15, Test: 0.96, Change: -16.75%, p < 0.01

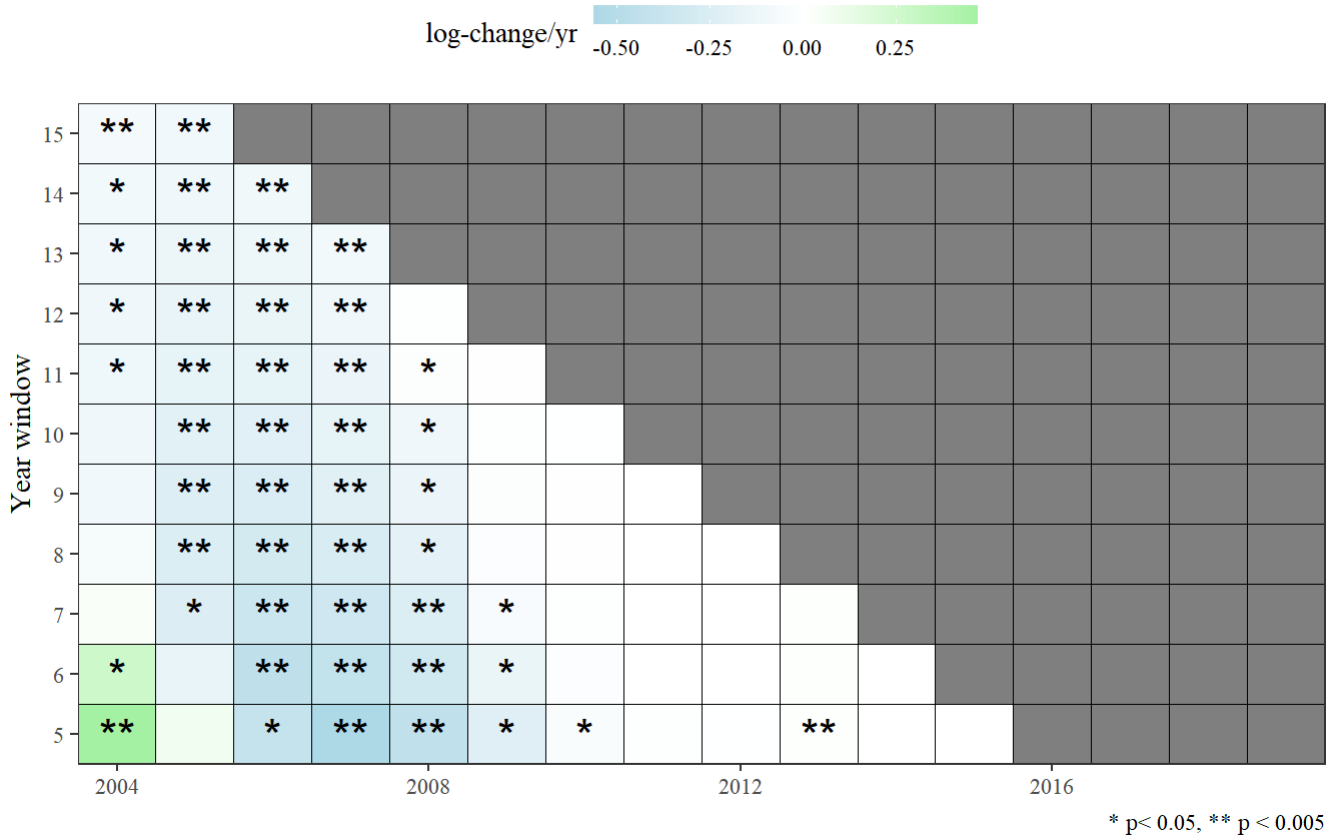


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope -0.02, log-slope -0.01 (-0.01, 0), p < 0.05



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



Second Example

19-08 TN: The SKT suggests a statistically significant DECREASING TREND while both GAM based tests suggest NO TREND.

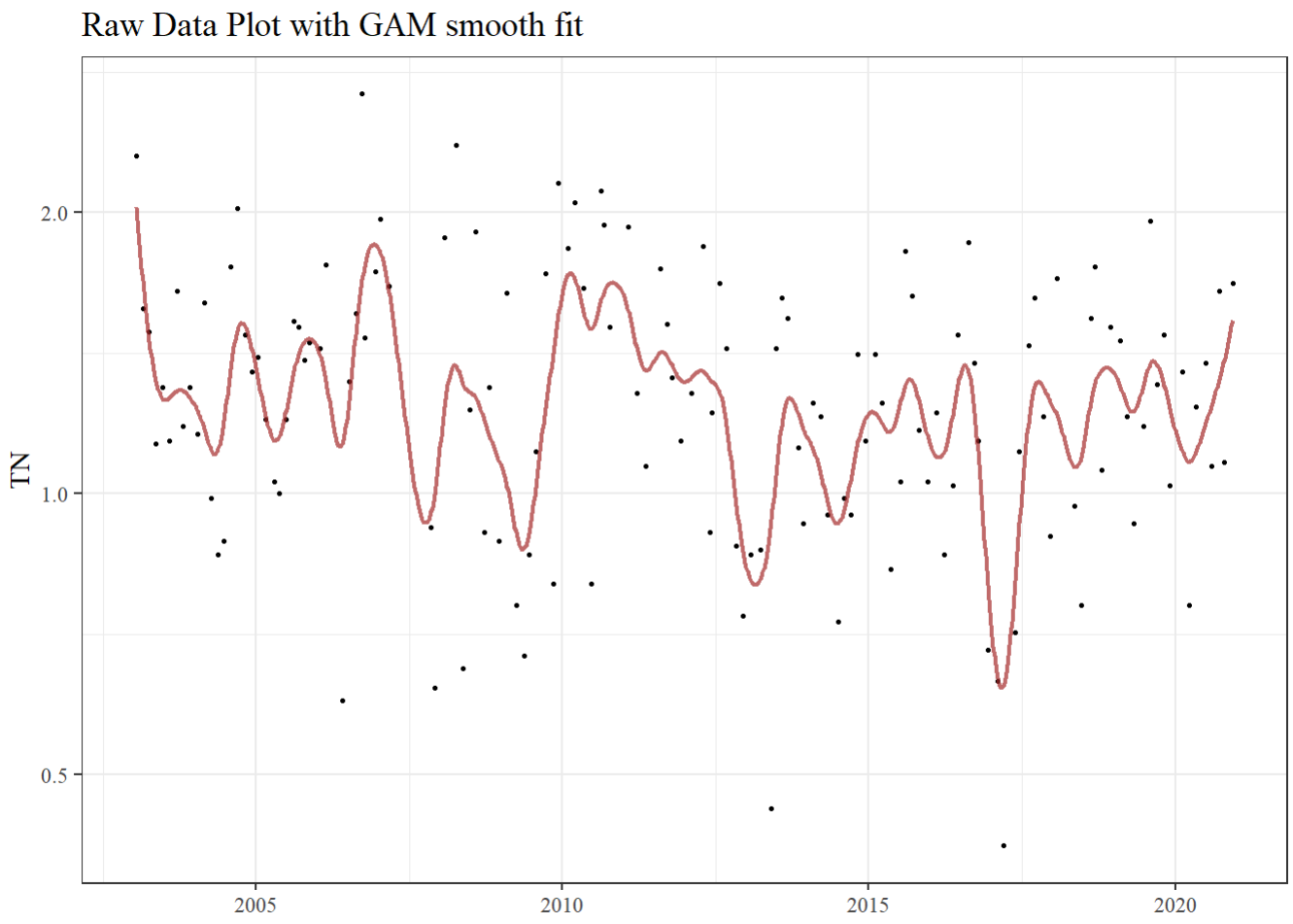
This site was not found to contain significant autocorrelation according to the SKT methodology.

GAM produced a warning that the Transformation introduced infinite values in continuous y-axis.

Again,neither the Raw or Averaged GAM methods suggest significant results.

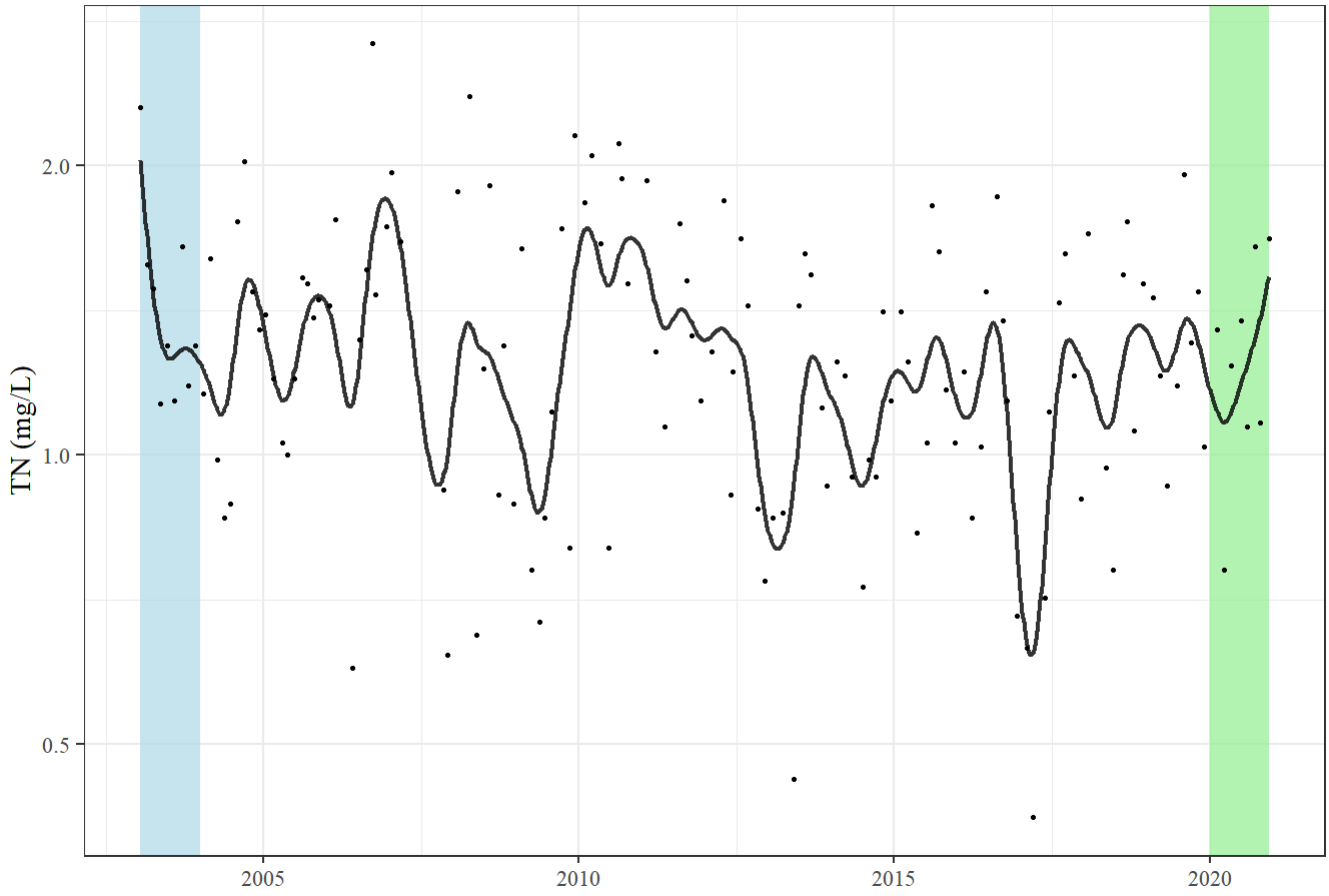
The regression plot with the box and whiskers below suggests a DECREASING slope.

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
## reducing knots for cont_year spline from 144
```



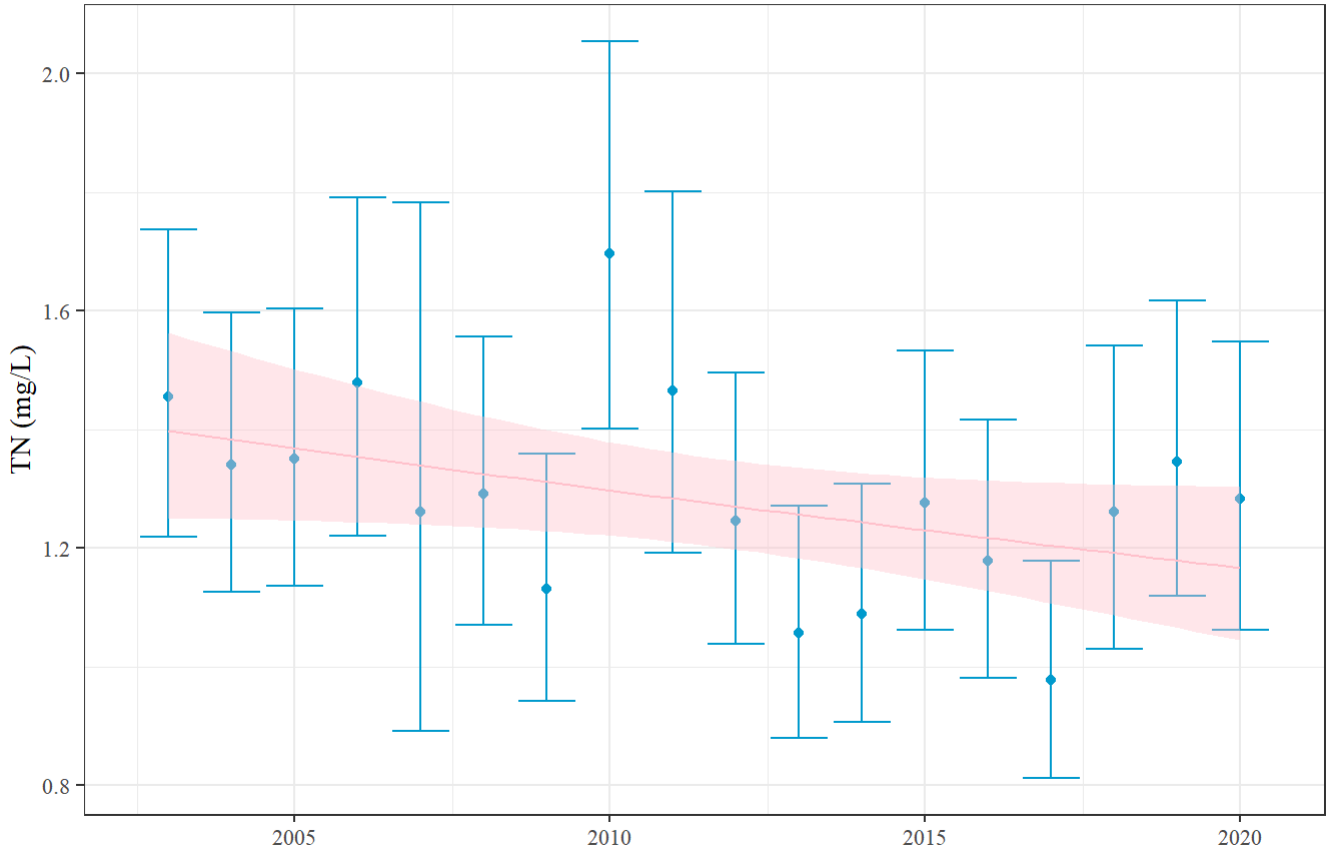

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 1.46, Test: 1.28, Change: -11.94%, ns

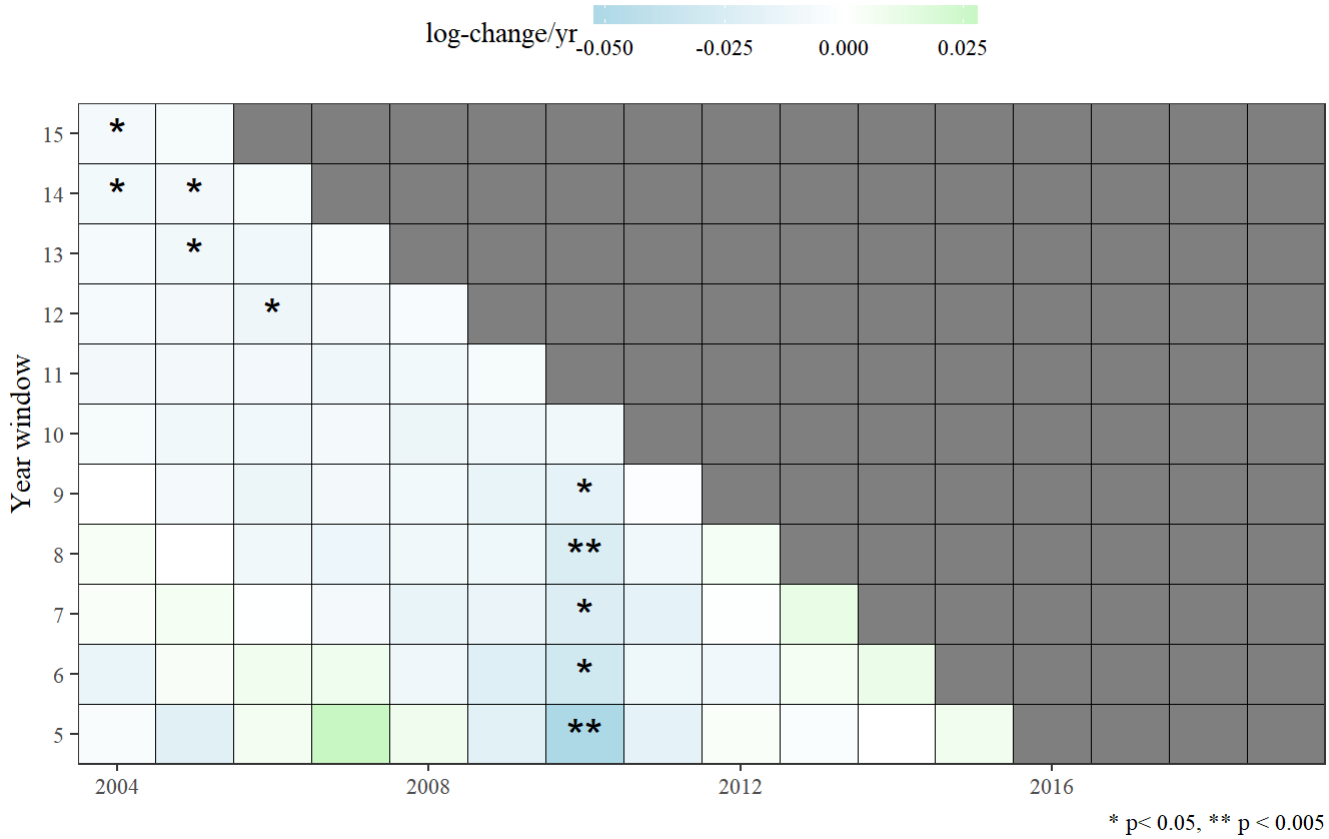


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope -0.01, log-slope 0 (-0.01, 0), ns



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



Third Example

W1 Chlorophyll a: The SKT suggests a statistically significant INCREASING TREND while both GAM based tests suggest NO TREND.

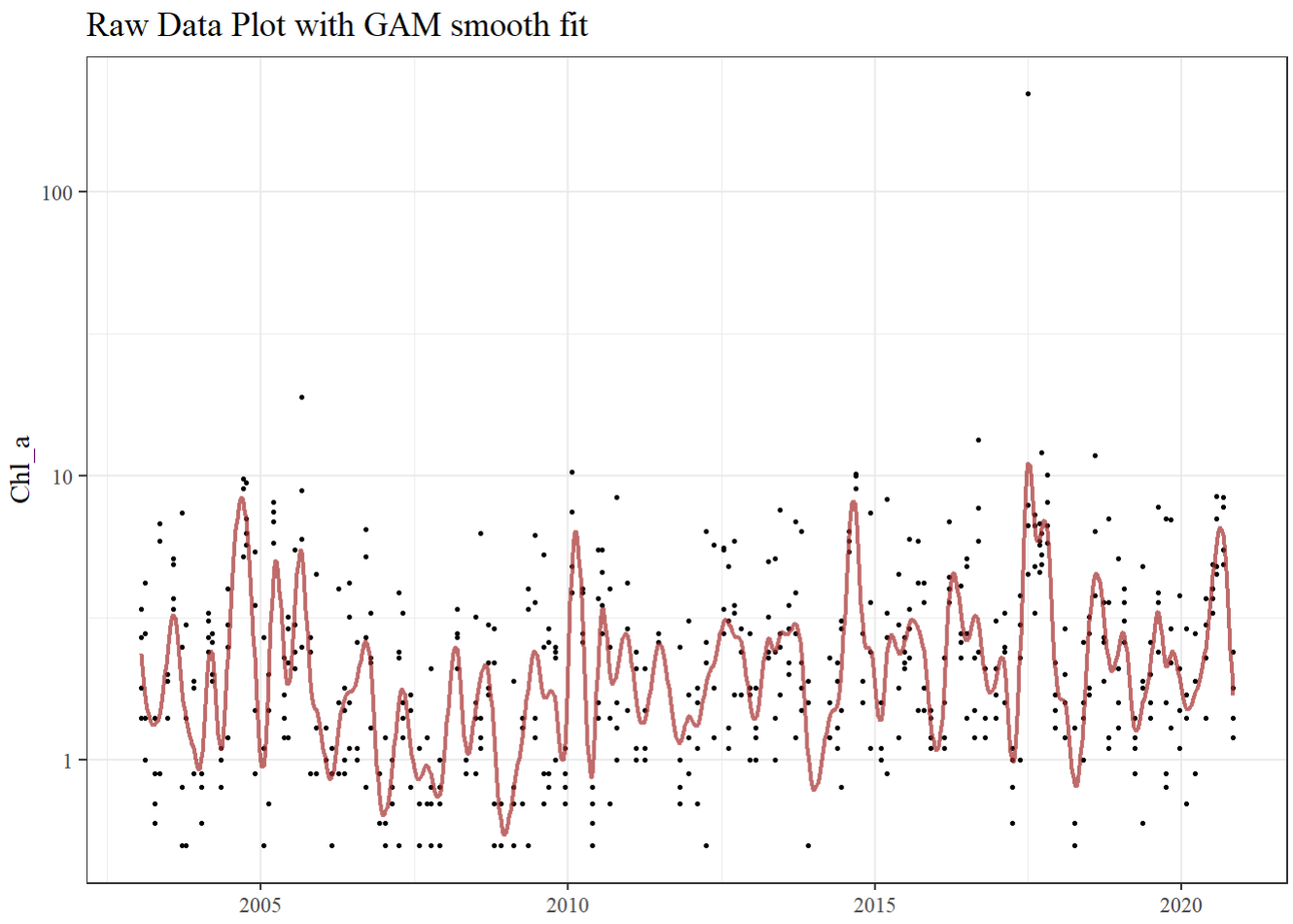
This site was not found to contain significant autocorrelation according to the SKT methodology.

GAM produced a warning that the Transformation introduced infinite values in continuous y-axis.

Again neither the raw or averaged GAM methods suggest significant results.

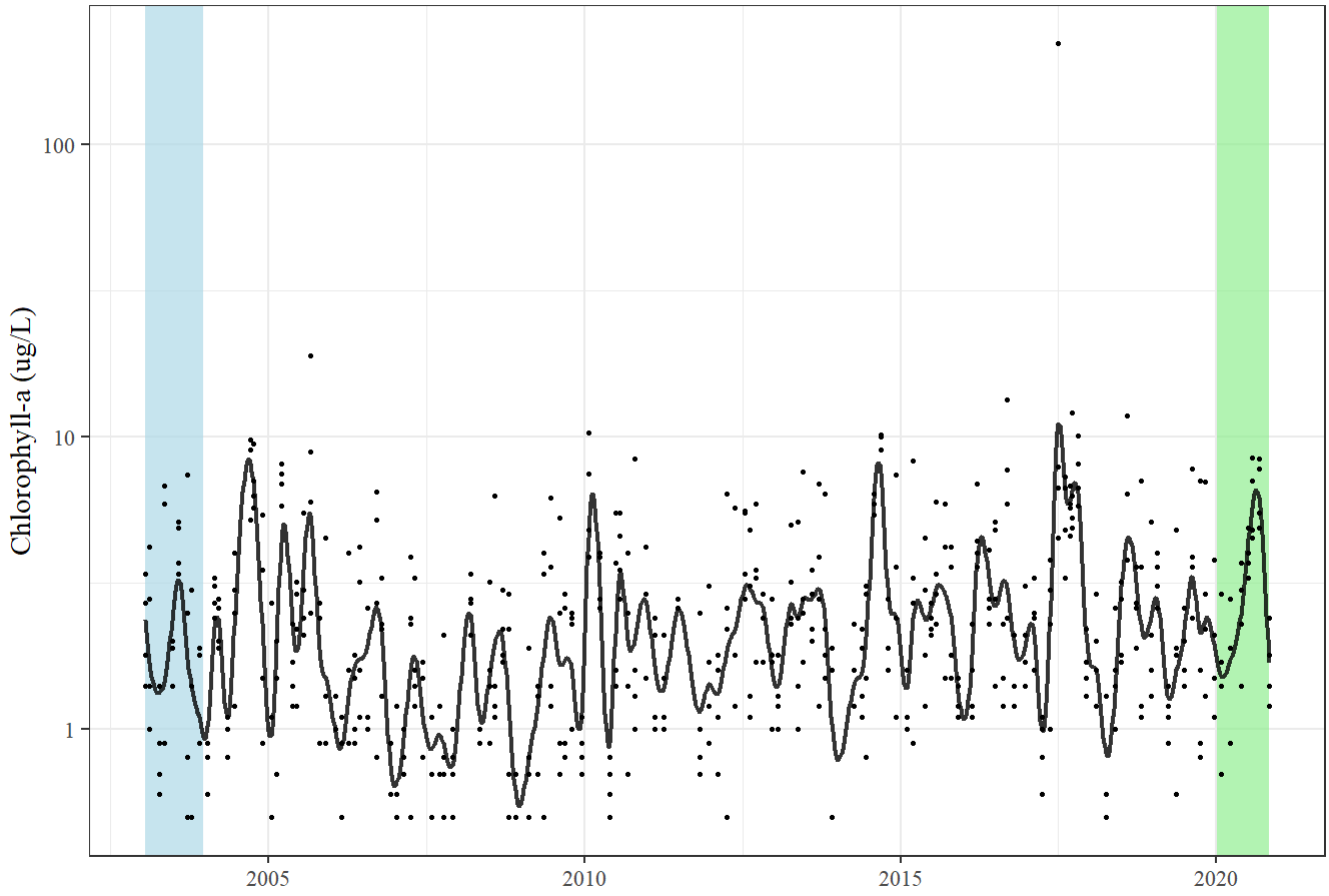
The regression plot with the box and whiskers below suggest a INCREASING slope.

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
## reducing knots for cont_year spline from 144
```



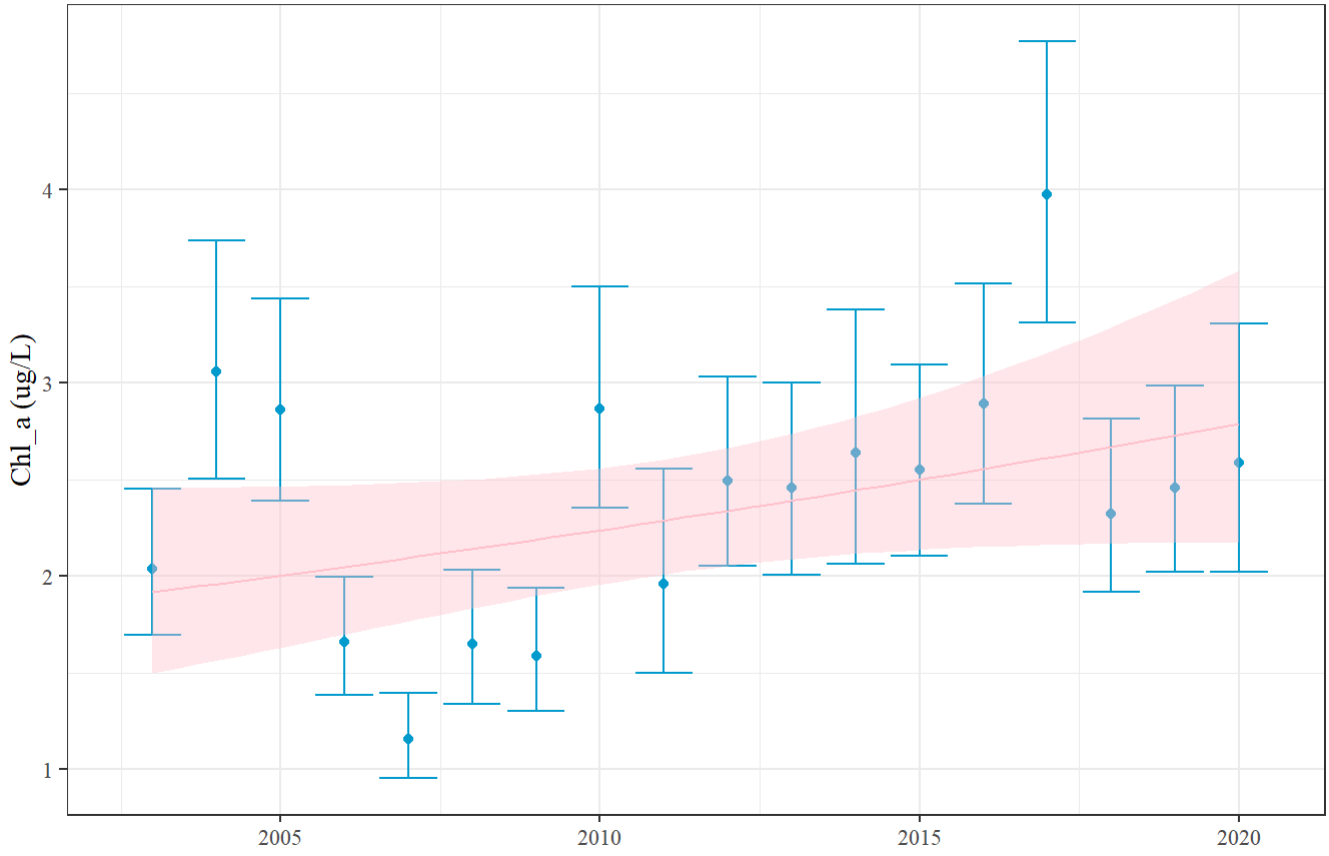

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 2.03, Test: 2.57, Change: 26.38%, ns

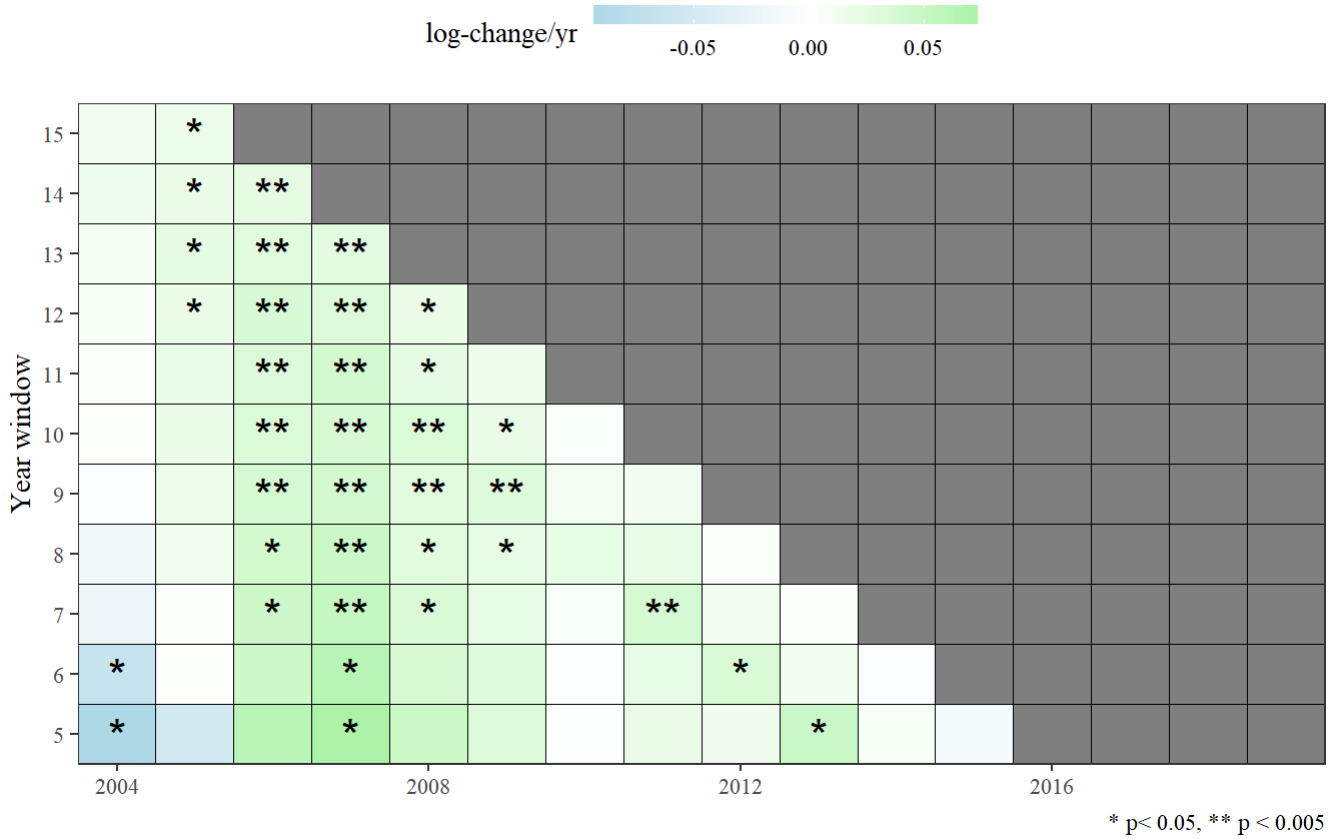


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope 0.05, log-slope 0.01 (0, 0.02), ns



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



Fourth Example

W4 TN: The results were Trend, Trend, No Trend for the SKT, Gam Avg, and Gam Raw, respectively.

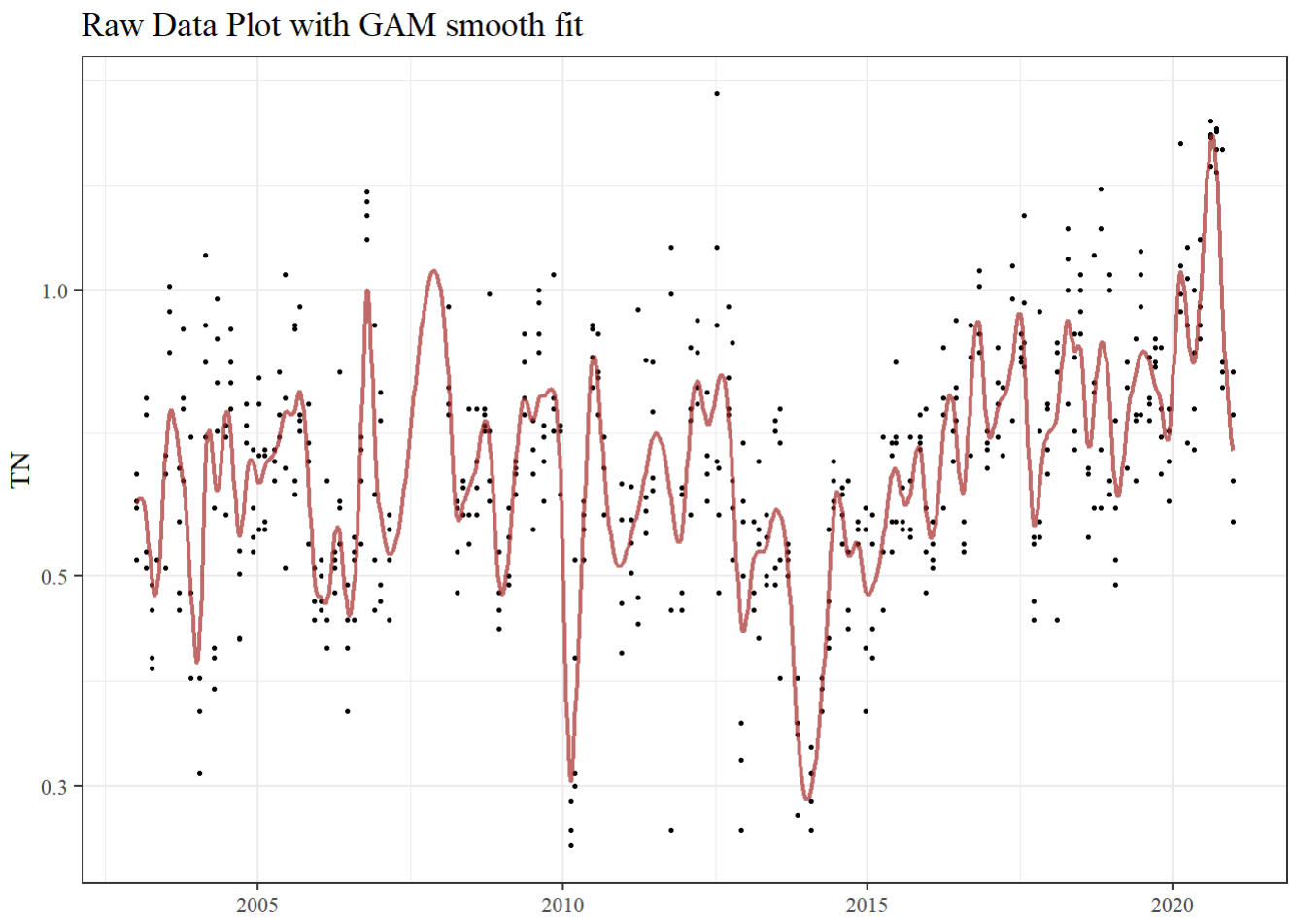
This site was not found to contain significant autocorrelation according to the SKT methodology.

The percent change plot suggests a significant difference between start and end points.

The regression plot with the box and whiskers below suggest a increasing slope.

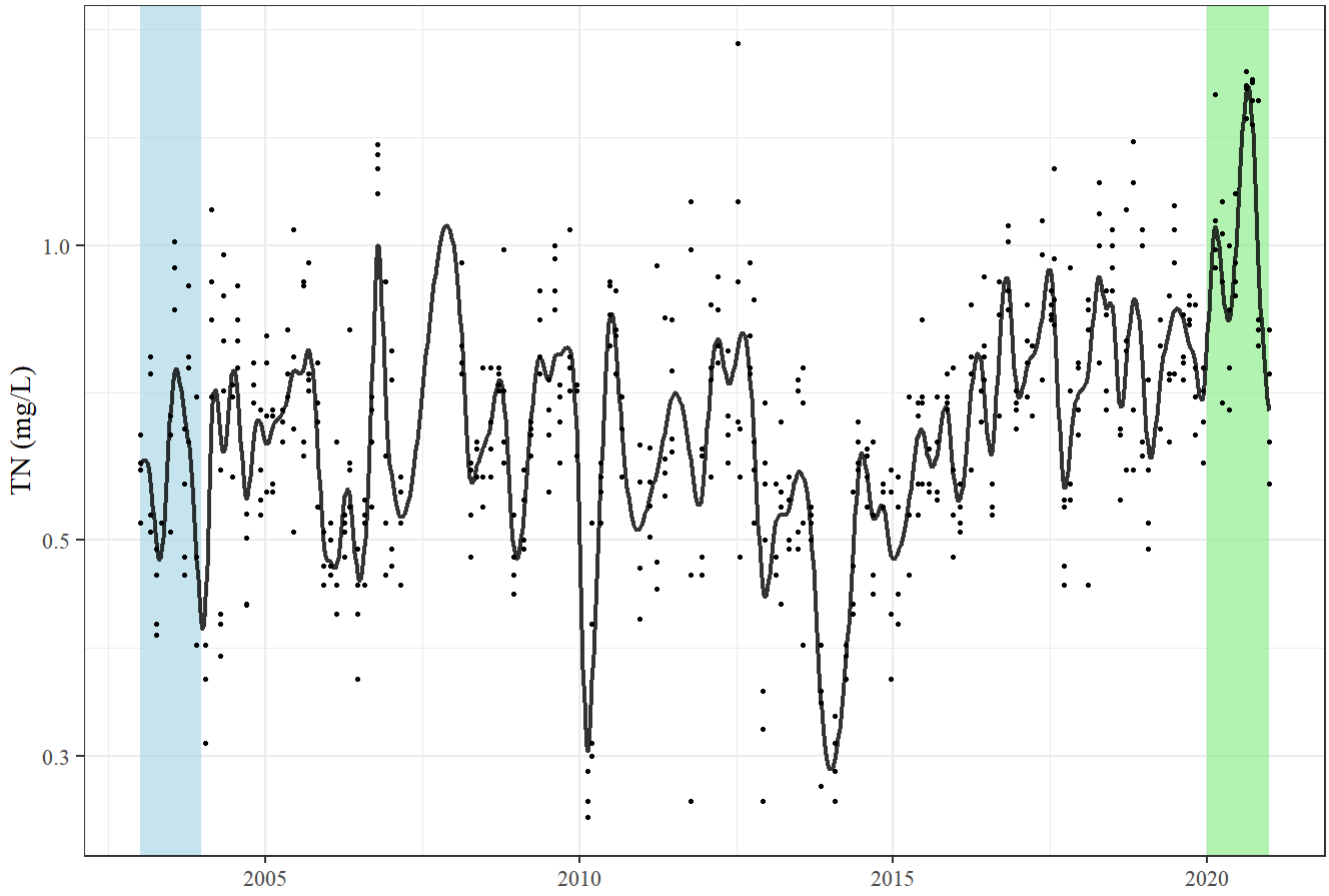
The GAM averaged (and SKT) result was significant but the GAM Raw was not

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
## reducing knots for cont_year spline from 144
```



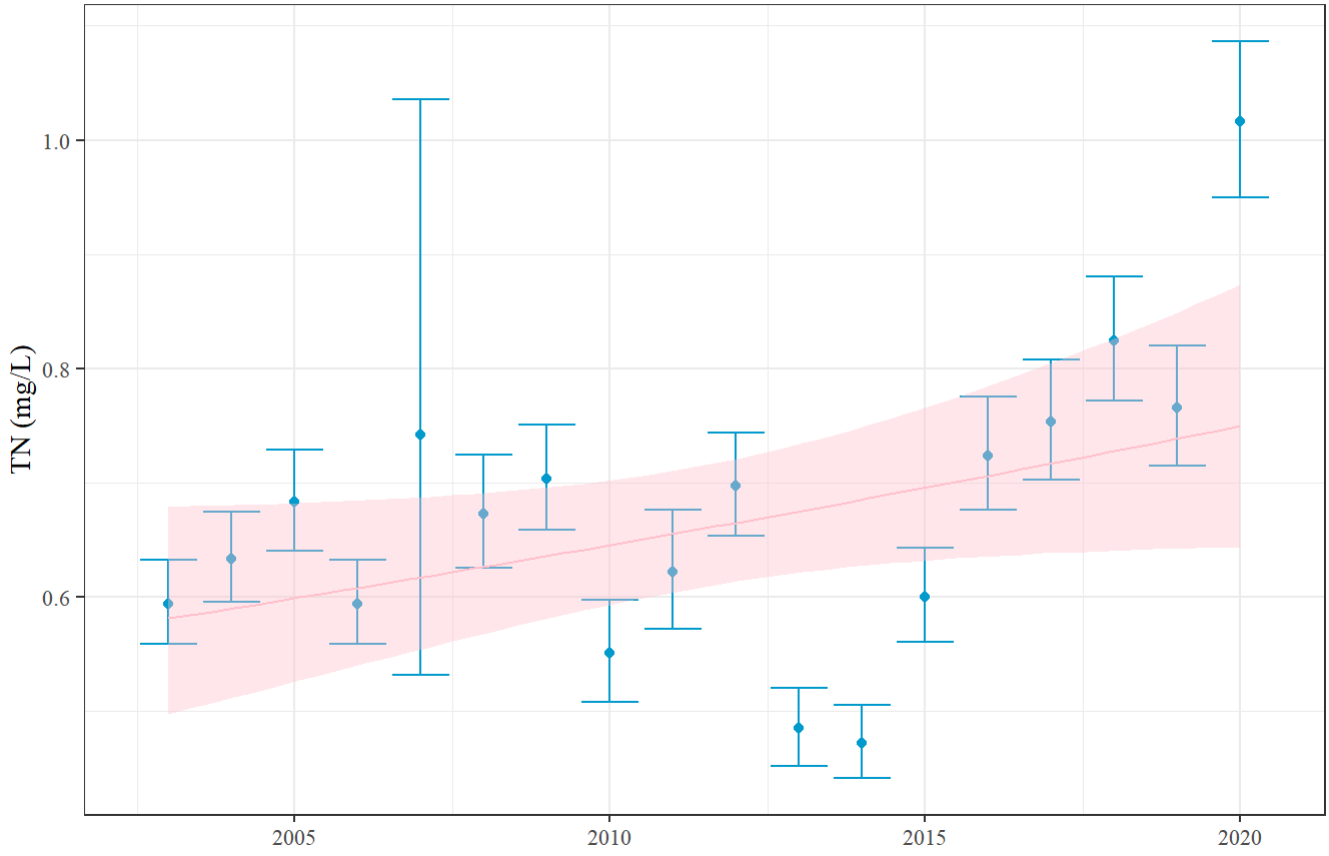

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 0.59, Test: 1.02, Change: 70.97%, p < 0.001

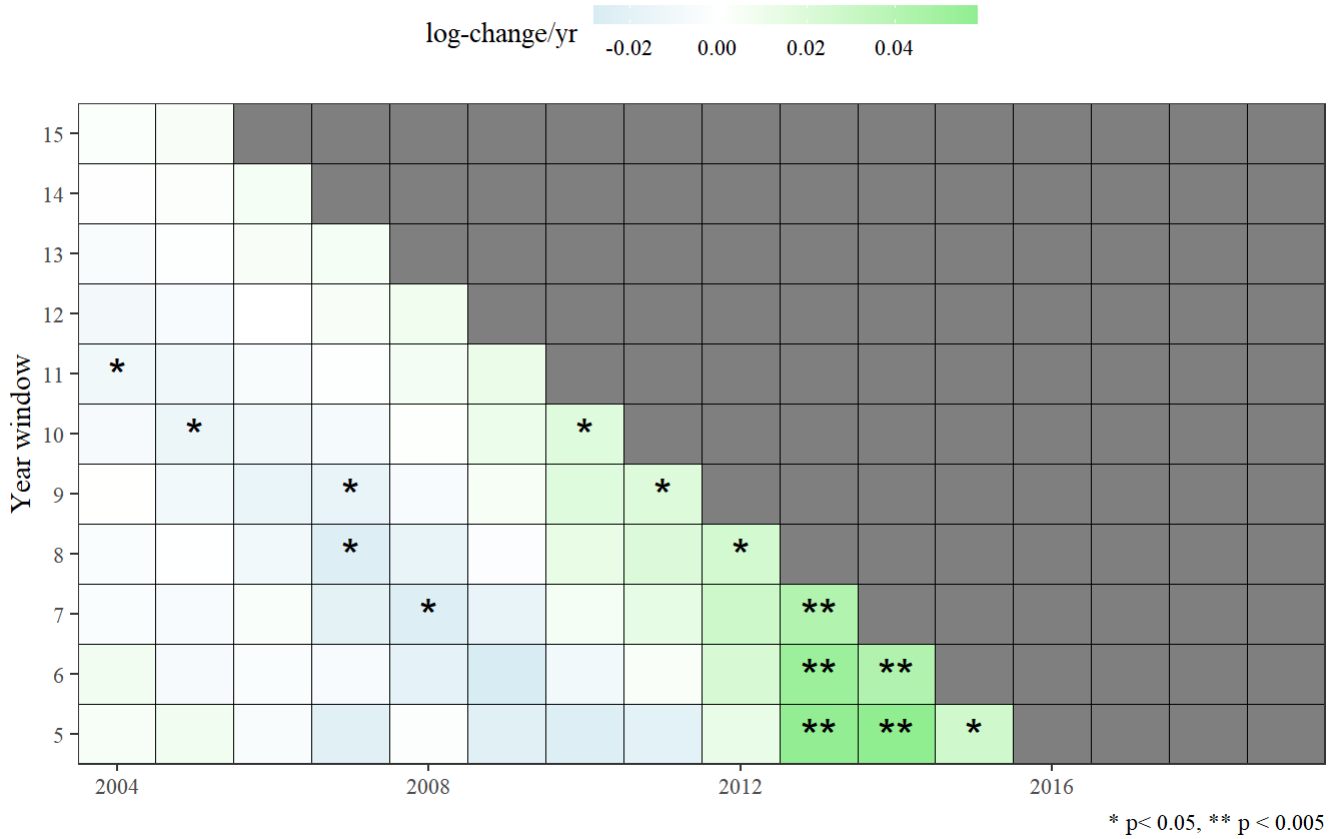


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope 0.01, log-slope 0.01 (0, 0.01), ns



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



Fifth Example

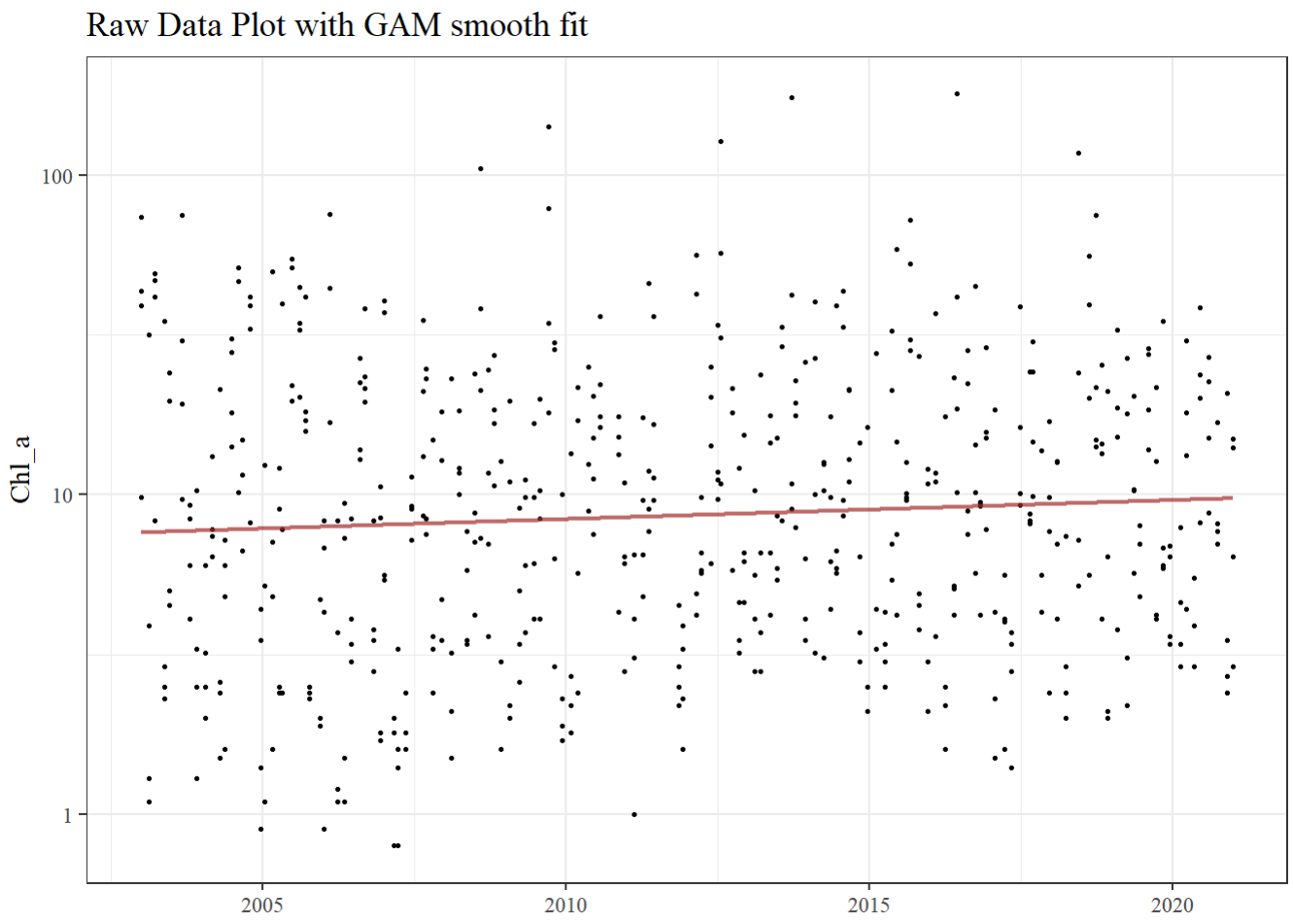
W5 Chl_a: In this example the exact opposite occurs: No Trend, No Trend, Trend.

The raw GAM fit suggest no seasonality - looks like a bad fit overall

The percent change plot suggests NO significant difference between start and end points.

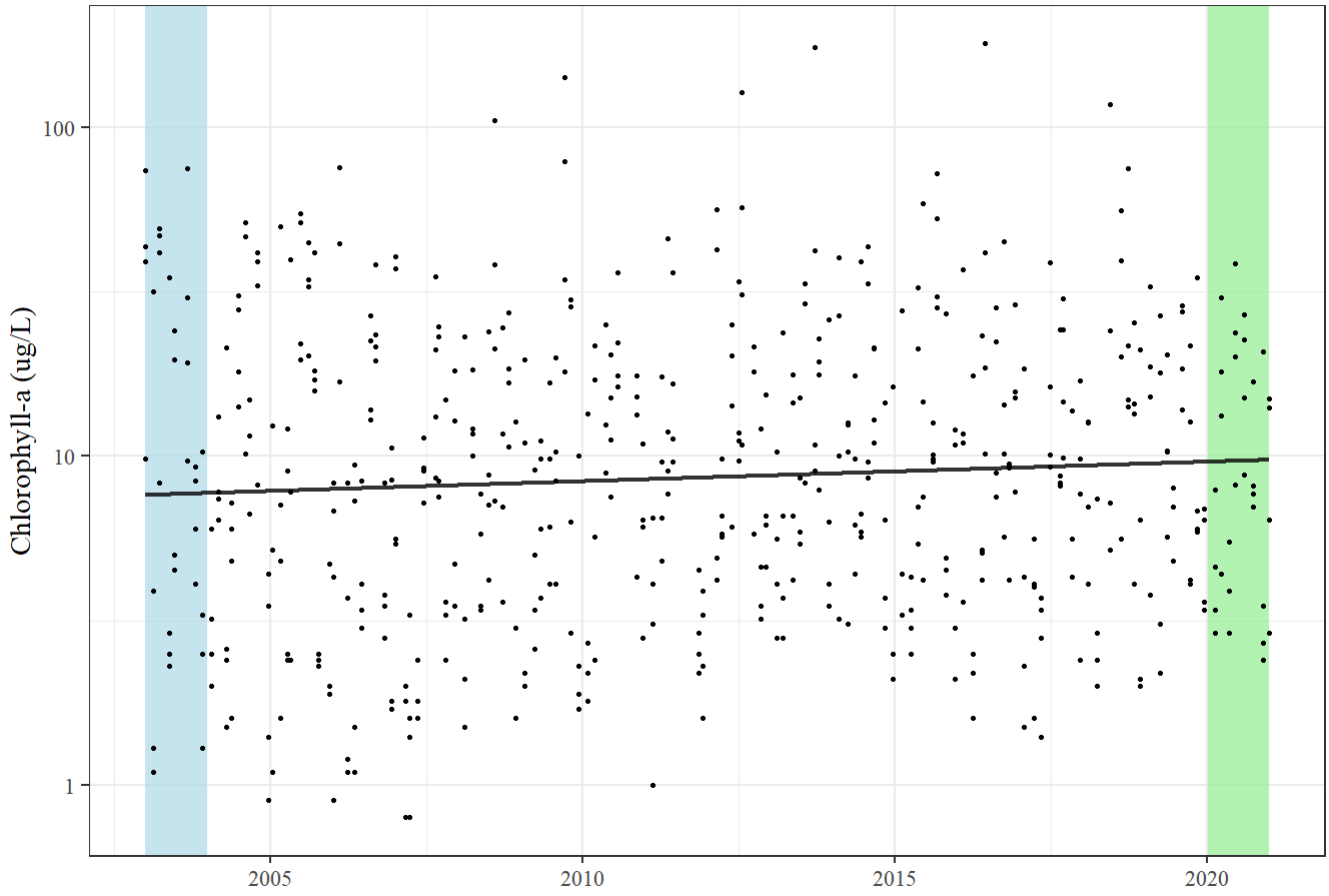
However, the regression plot with the box and whiskers below suggest a increasing slope.

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
```



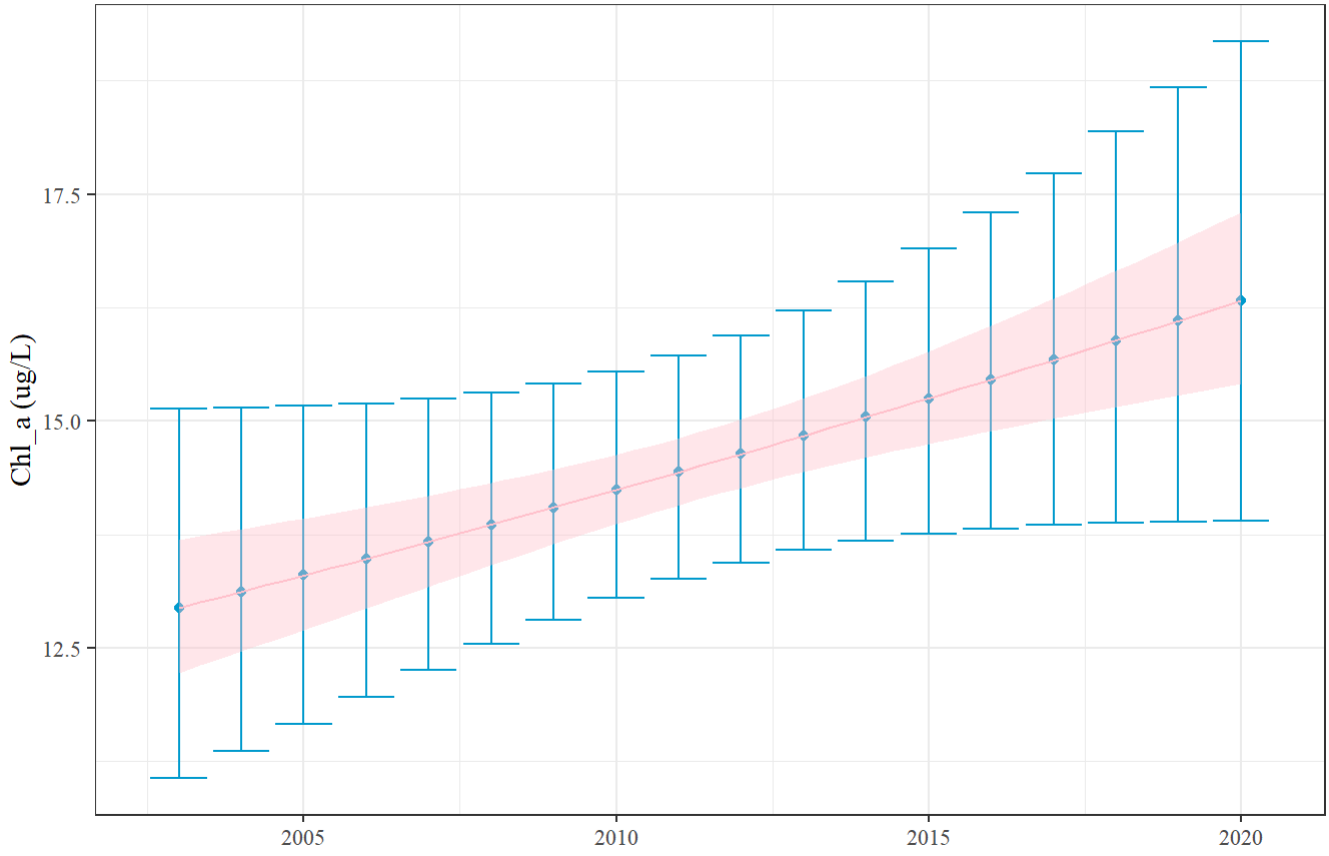

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 12.94, Test: 16.33, Change: 26.18%, ns

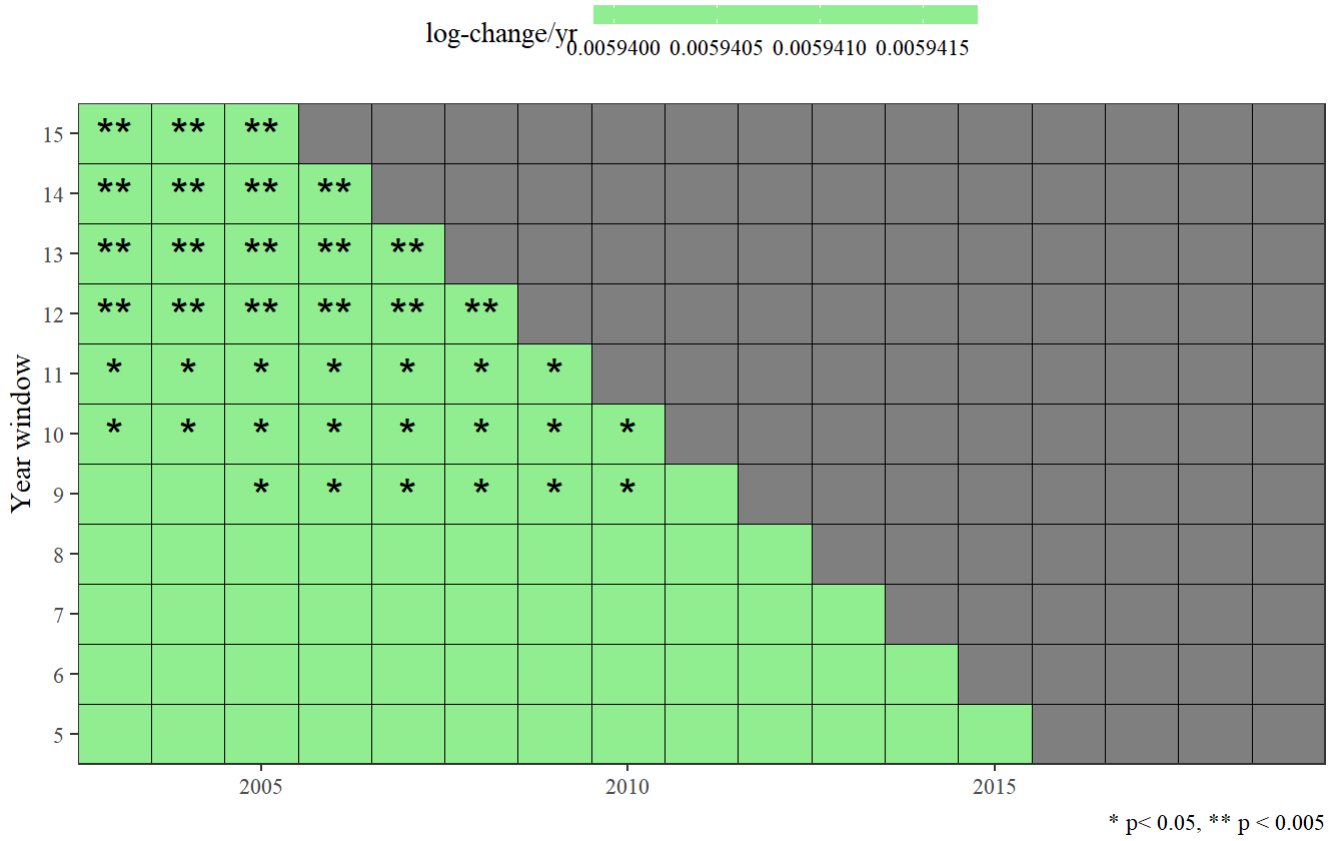


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope 0.2, log-slope 0.01 (0, 0.01), $p < 0.001$



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



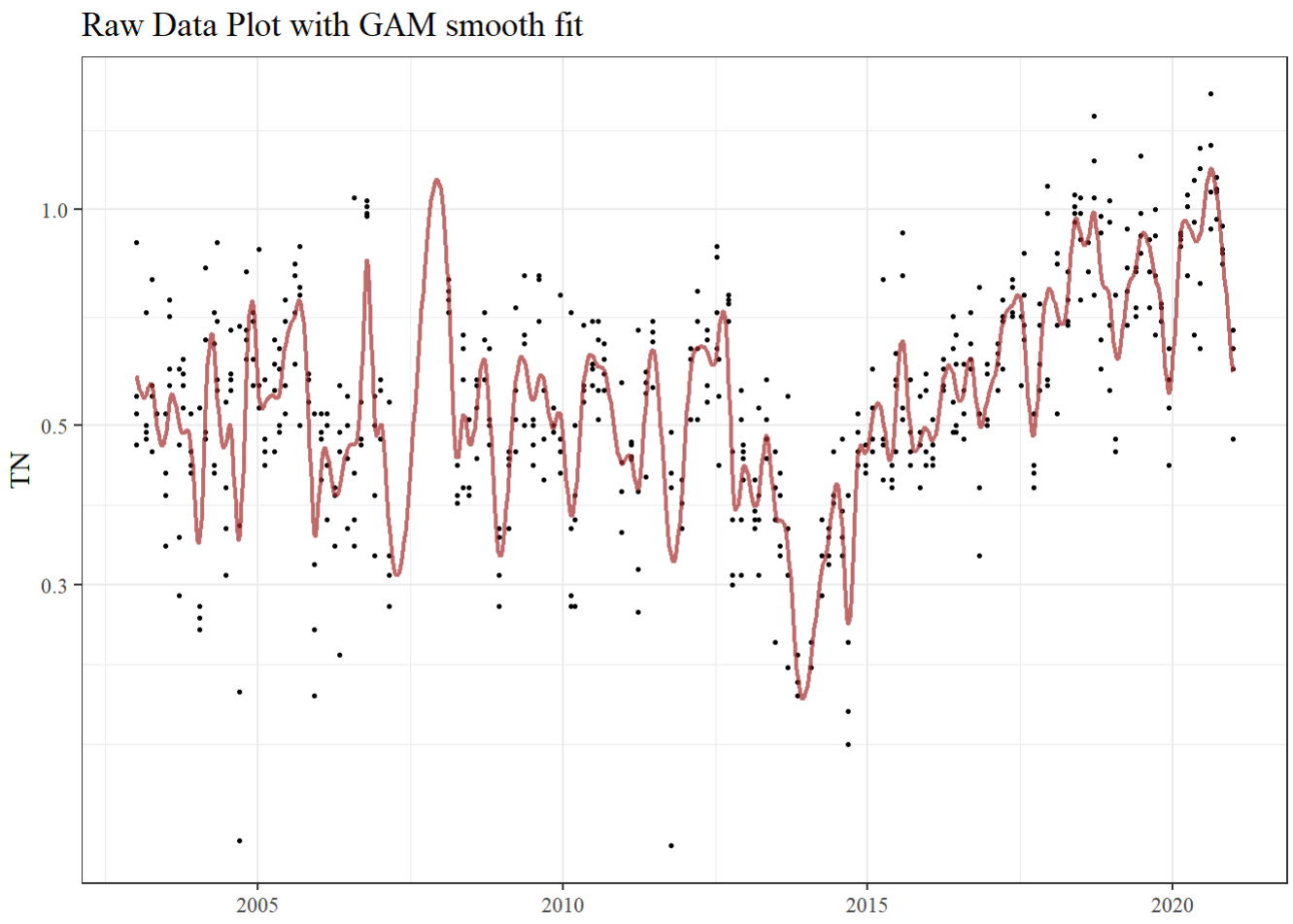
Sixth Example

W6 TN: The Results of this example suggest: Trend, Trend, No Trend for SKT, GAM avg, and GAM raw, respectively

The percent change plot suggests a significant difference between start and end points.

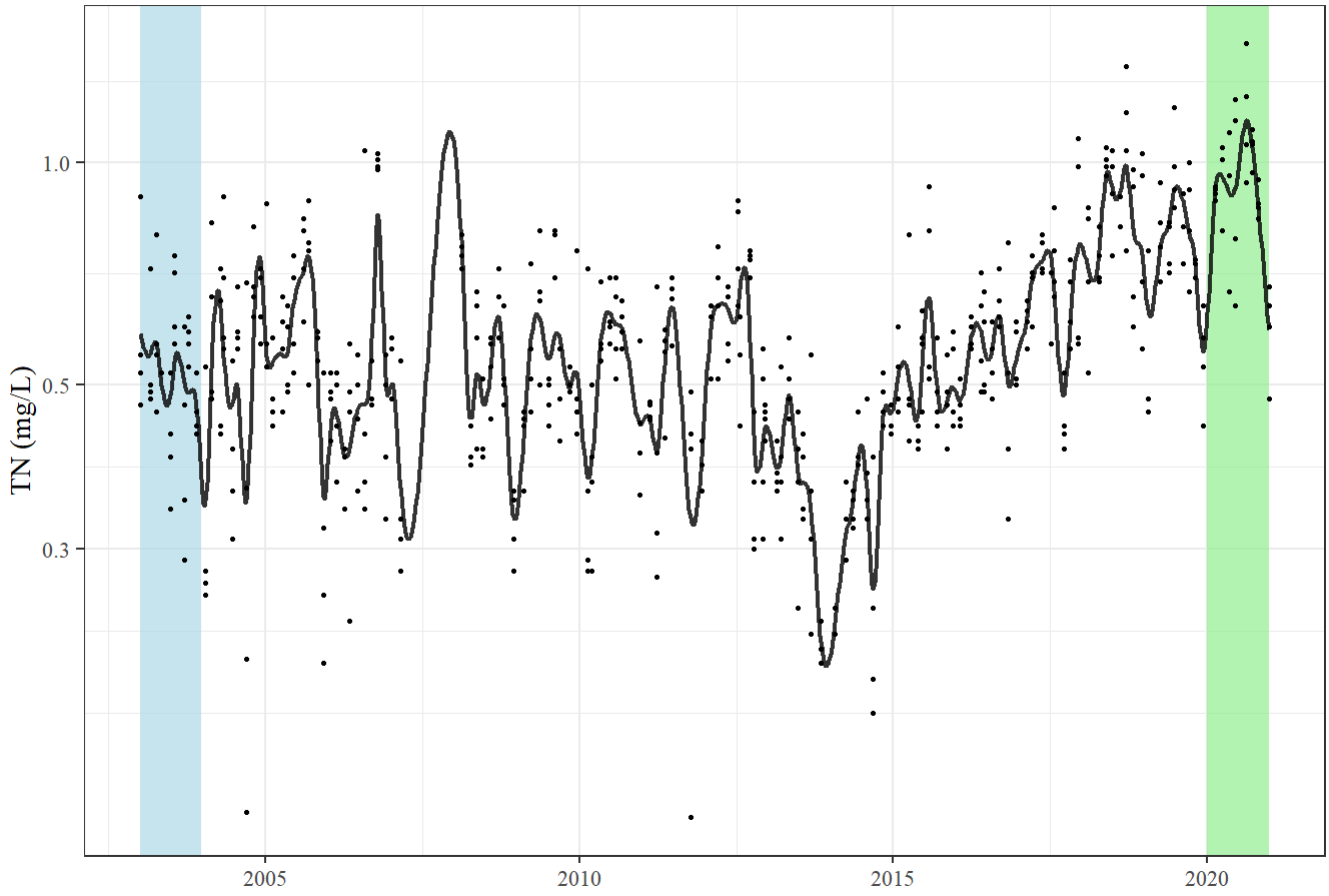
The regression plot with the box and whiskers below suggest a increasing slope.

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
## reducing knots for cont_year spline from 144
```



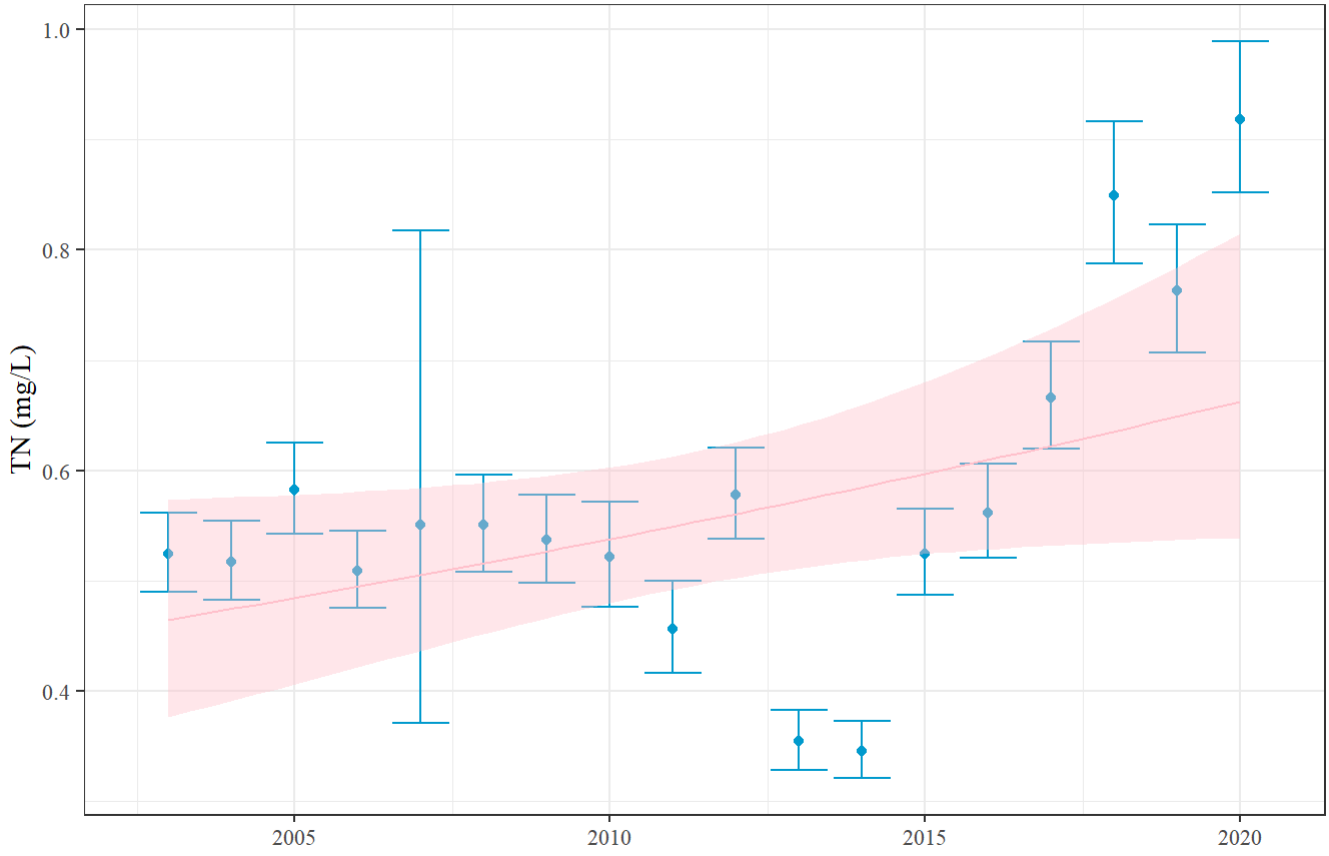

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 0.52, Test: 0.92, Change: 74.83%, p < 0.001

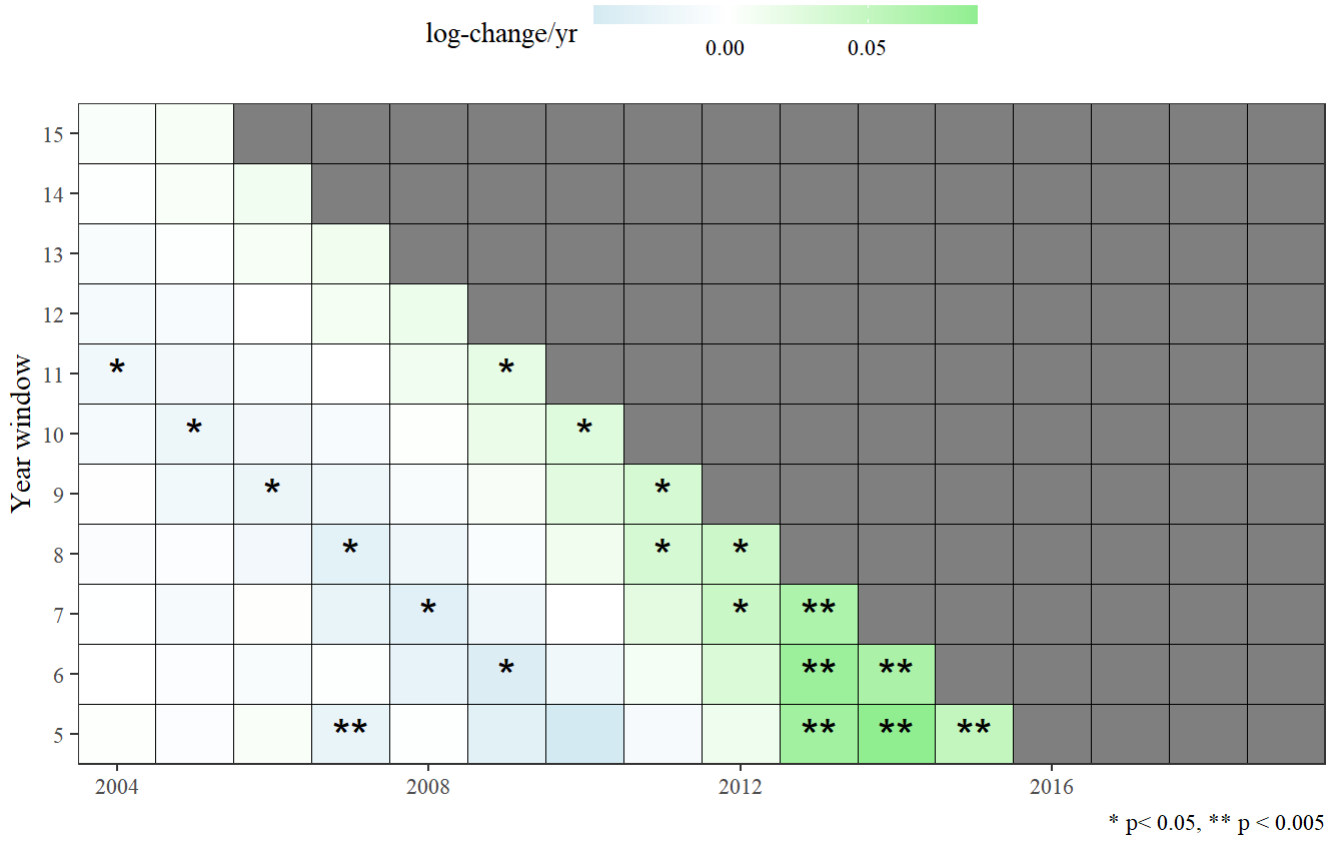


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope 0.01, log-slope 0.01 (0, 0.02), ns



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



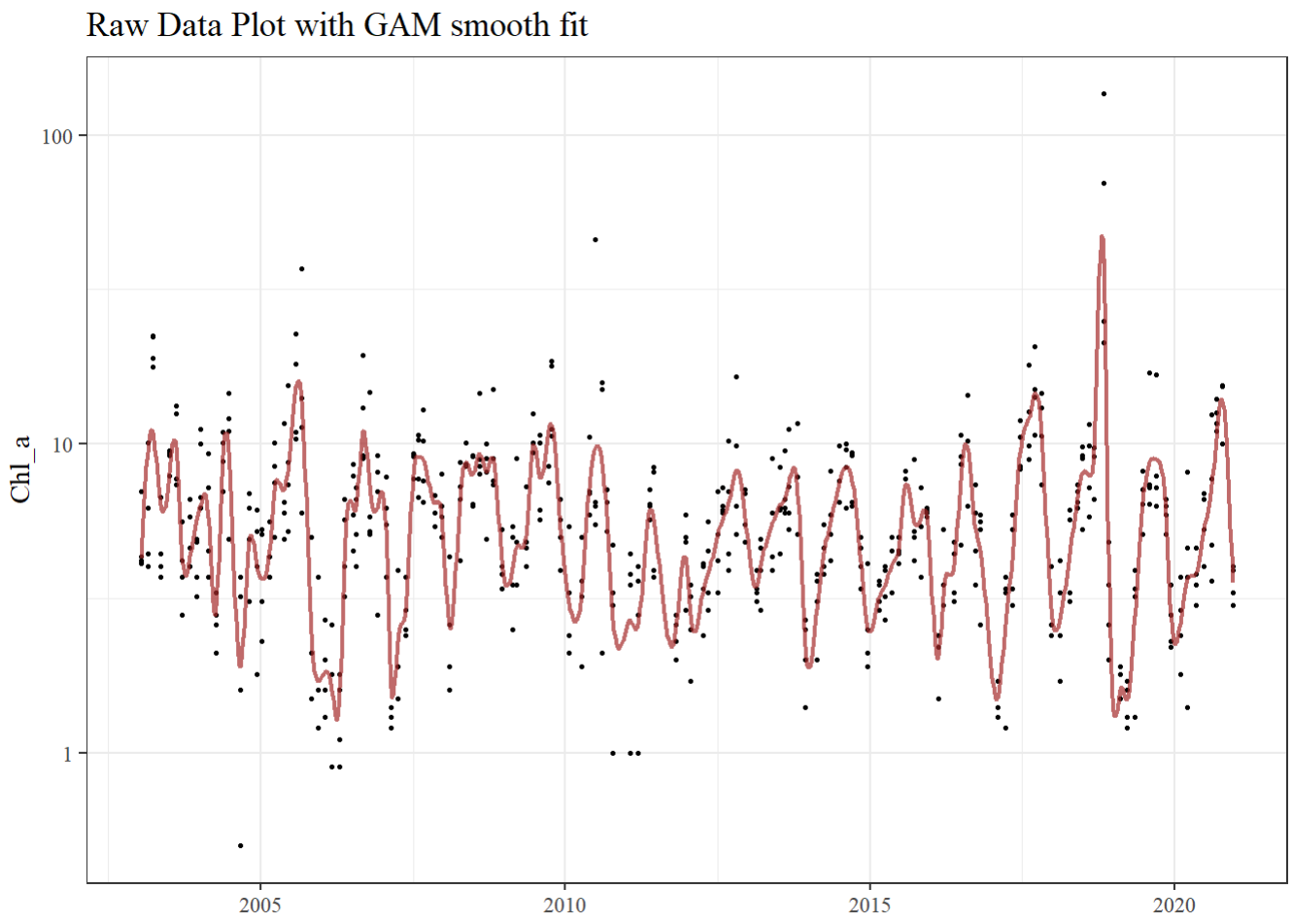
Seventh Example

W7 Chla Result: Trend, No Trend, No Trend

The percent change plot suggests a significant difference between start and end points.

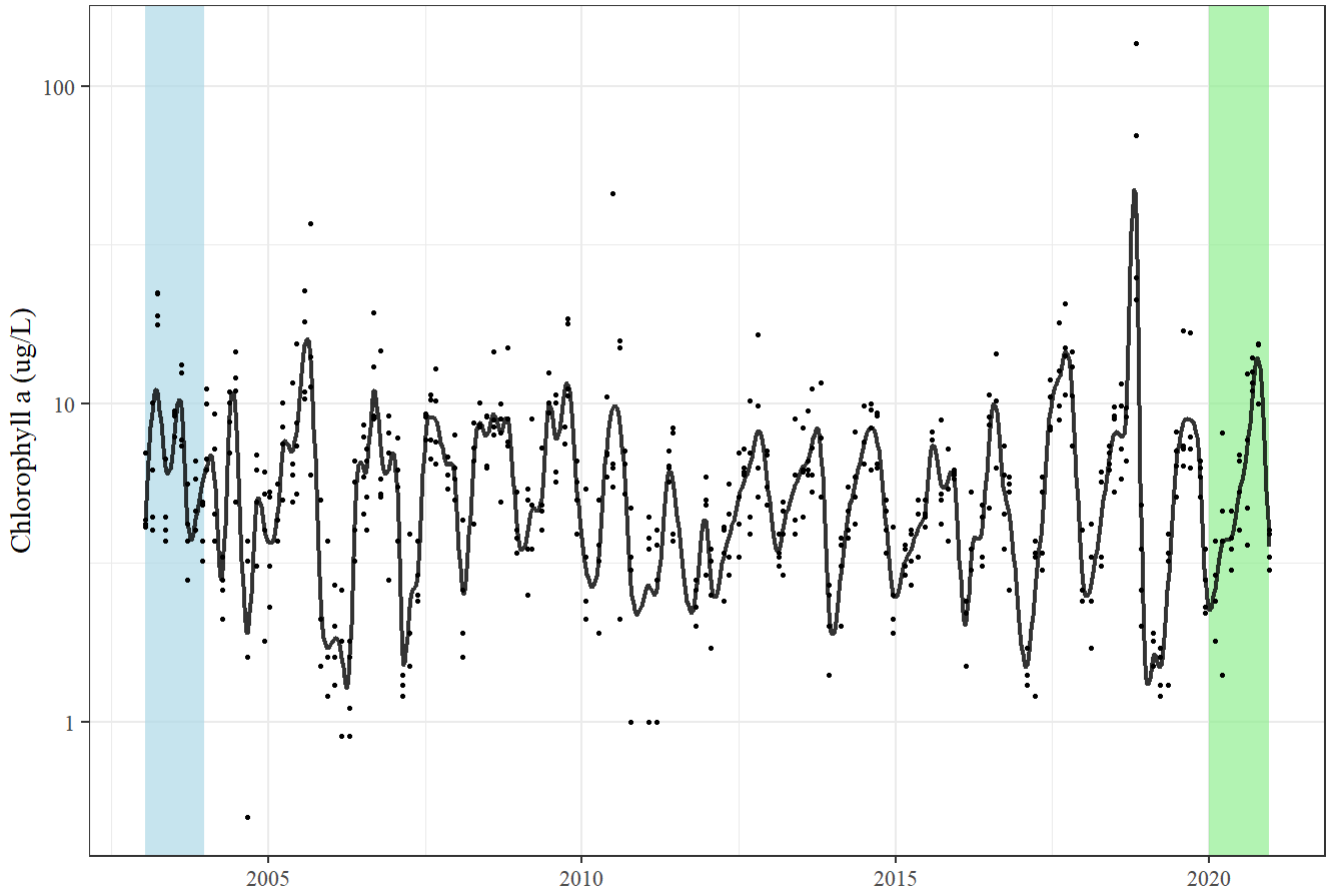
The regression plot with the box and whiskers below suggest a slight decreasing slope.

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
```



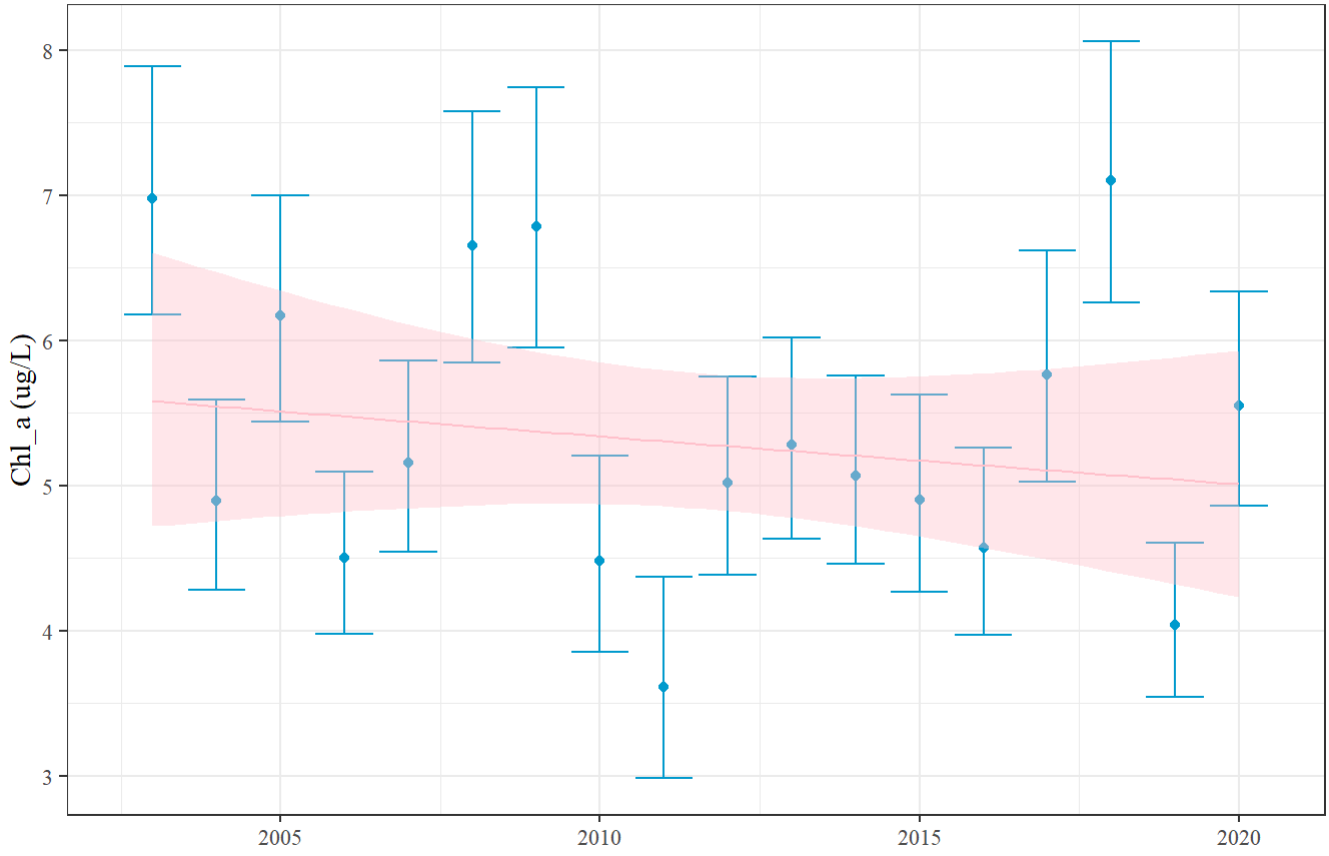

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 6.99, Test: 5.54, Change: -20.67%, p < 0.05

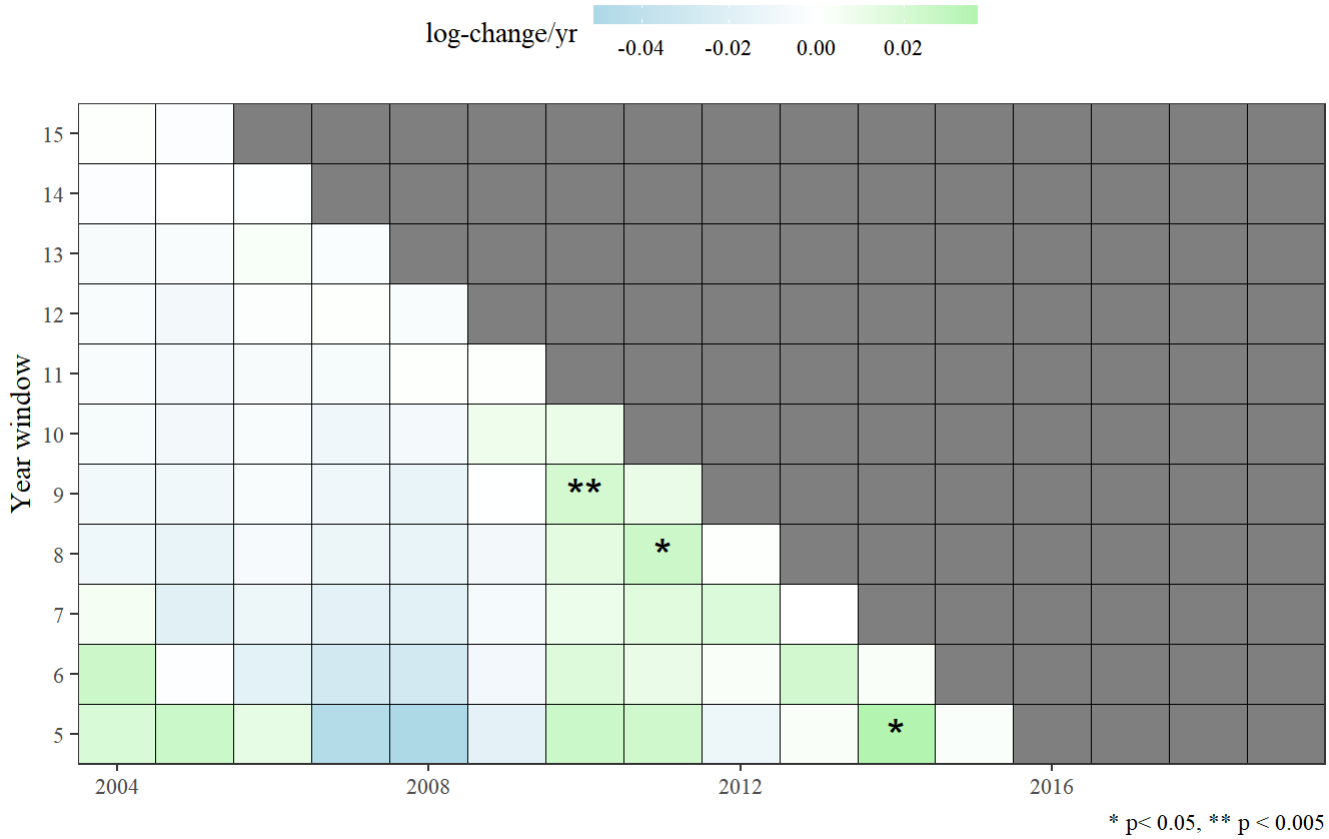


Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope -0.03, log-slope 0 (-0.01, 0), ns



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```



Eighth Example

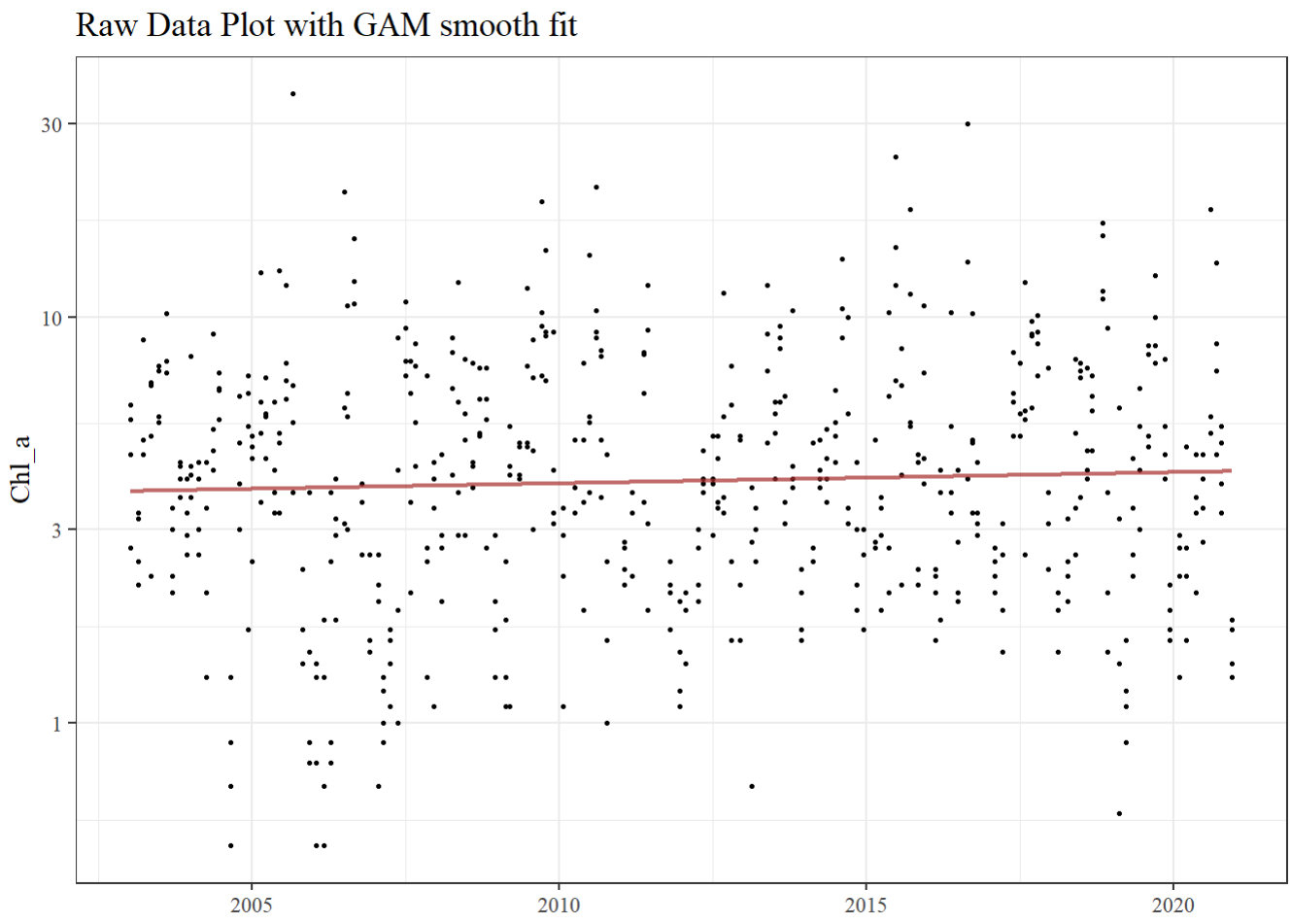
W8 Chla Result: No Trend, No Trend, Trend

Looks like a poor GAM fit overall, similar to example 5

The percent change plot suggests NO significant difference between start and end point.

The regression plot with the box and whiskers below suggest a increasing slope.

```
## reducing knots for cont_year spline from 216
## reducing knots for cont_year spline from 198
## reducing knots for cont_year spline from 180
## reducing knots for cont_year spline from 162
```

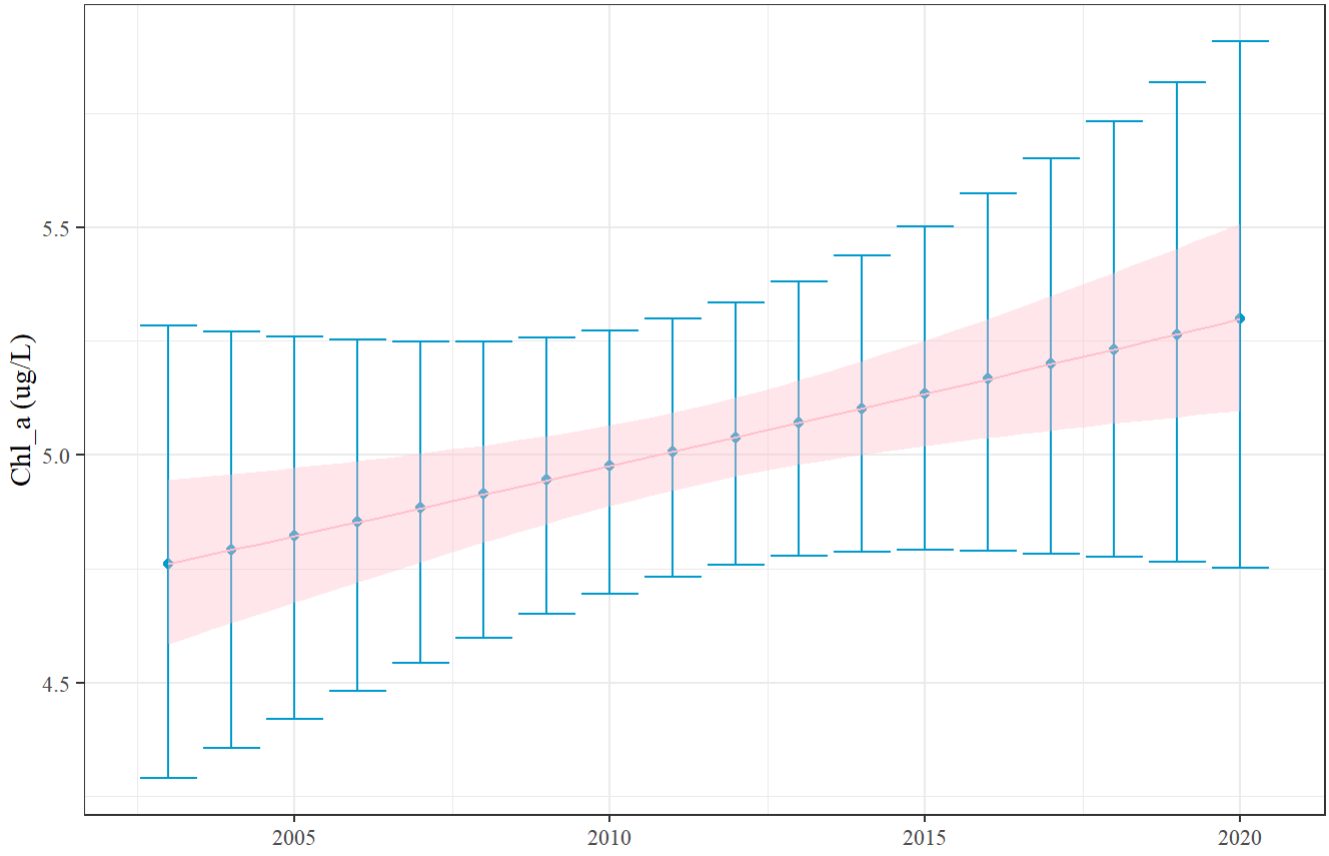



```
## Warning: Transformation introduced infinite values in continuous y-axis
```

Base: 4.76, Test: 5.3, Change: 11.27%, ns

Est. mean with 95% confidence intervals: Jan 1-Dec 30

Trend from 2003 to 2020: approximate slope 0.03, log-slope 0 (0, 0), p < 0.01



```
## Warning: Removed 99 rows containing missing values (`geom_text()`).
```