Lake Rotorua: Trends in Water Quality (2001-2017)

JULY-2018

THOMAS STEPHENS (DAIRYNZ LTD), KEITH HAMILL (RIVER LAKE LTD) & CHRIS MCBRIDE (ENVIRONMENTAL RESEARCH INSTITUTE - UNIVERSITY OF WAIKATO)

CITATION:

Stephens T, Hamill KD, McBride C 2018. Lake Rotorua: Trends in Water quality (2001-2017). Technical report produced for Lake Rotorua Technical Advisory Group

Executive summary

This report has been prepared as part of the Science Review associated with Plan Change 10 of the Bay of Plenty Regional Council (BOPRC) Regional Water and Land Plan (RWLP). The report builds on previous analysis to and describe water quality trends in Lake Rotorua for the years 2001 to 2017.

The analysis focuses on the water quality variables of Secchi depth (SD), total phosphorus (TP), total nitrogen (TN) and chlorophyll-*a* (Chl-*a*) at two long term monitoring sites in Lake Rotorua. Water quality trends were analysed over two time-periods at multiple depths; the long term or full series (2001-2017) and the short term or more recent (2009-2017). Two different methods were used for determining trends:

- Season Kendall trend test with a Sen Slope Estimator (significant if the *p*-value was <0.05 and meaningful if the percent annual change was greater than 1%) and
- Directional confidence testing based on the Sen Slope Estimator (McBride et al. 2014). Confidence intervals were used to draw inferences about trend direction at 95% confidence, and classified as 'improving, 'indeterminant' or 'degrading'. A trend was 'indeterminant' if the confidence interval spanned zero i.e. that there was insufficient data to determine the trend direction.

The surface water time series was deconstructed using a statistical procedure to determine when changes in trend strength or direction occurred.

The trend analysis for the period 2001 to 2017 found a virtually certain improving trend for TN, TP, chlorophyll-*a* and Secchi depth. The trends were statistically significant and 'meaningful' at all water depths analysed - surface, mid-depth and bottom. Similarly, the directional confidence testing approach found improving" long-term trends for all four TLI attributes at all depths at both sites. The trend for TN and TP remain statistically significant and "virtually certain" after correcting for laboratory changes.

Seasonality components in TP, TN and Chl-*a* time-series recorded peak values in late summer / early autumn, while Secchi depth had a more complex seasonal pattern. Deconstruction of the trend found broadly similar patterns between the TLI components. Trend components of Chl-*a* time-series increased from 2001 to 2003 but decreased from 2006 to 2017. The period of greatest improvement in Chl-*a* concentrations was 2006 to 2012 and since 2016. SD appeared to decouple from Chl-*a* for a period from 2006 to 2009 (a decline followed by an improvement).

A non-parametric spearman rank correlation analysis of the surface water indicated that from 2001 to 2009 Chl-*a* was positively correlated with both TP and TN (with the correlation slightly stronger for TN), but since 2009 Chl-*a* was more strongly correlated with surface water TP. This was also supported by visual inspection, but notably, the analysis does not imply whether nutrients are causing and/or responding to change in Chl-*a*.

Trends were broadly similar between sites and between depths. However, at the deeper monitoring site, Site 5, improving trends in TN and TP were of consistently of greater magnitude with increasing depth. This could imply recent changes to water quality being greater in deeper nutrient reservoirs, including the potential for altered hypolimnetic and/or pore-water REDOX restricting on nutrient regeneration. This may point to alum having stronger effects on sedimentary nutrient regeneration than on flocculation of particulate nutrients in overlying waters, but this observation is tentative. More analysis is needed to investigate this possibility and it may be influenced by recent years data and monitoring depth.

Trends in TLI attributes since 2009 are largely of indeterminate direction, but there has been an “extremely likely” improving trend in Chl-*a* between 2009 -2017 (e.g., of >99% probability not due to change), albeit mostly occurring in the period 2009-2012.

Contents

1 Introduction

1.1 Background

1.1.1 Plan Change 10 Science Review

Bay of Plenty Regional Council (BOPRC) is in the process of approving Plan Change 10 to the Regional Water and Land Plan (RWLP). Plan Change 10 sets rules for Lake Rotorua nutrient management; it includes specific provisions to review the science supporting decisions associated with Lake Rotorua (i.e. LR M2 of the proposed plan change). Terms of Reference (ToR) have been prepared to establish the extent of the 2017 Science Review, including a task to:

"Review trends in-lake water quality attributes including N, P, chlorophyll-*a*, algal blooms, clarity and TLI. This will include updating the Joint UoW and DNZ report (Abell et al 2012)."

This report has been produced to complete this task by updating the lake trend analysis to cover recent years to 2017. Another report is also being prepared by BOPRC to assess the recent trends in water quality trends and loads from streams entering Lake Rotorua.

1.2 Objective

This report describes water quality trends in Lake Rotorua. The focus is on water quality variables associated with the Trophic Level Index (TLI), i.e. Secchi Depth (SD), total phosphorus (TP), total nitrogen (TN) and algal biomass (chlorophyll-*a*, Chl-*a*), and where appropriate, nitrate-nitrogen [NO3], ammoniacal-nitrogen [NH4], dissolved reactive phosphorus [DRP] and turbidity[1]. The analysis covers the two long-term monitoring sites with near-continuous monthly records, 'Site 2' and 'Site 5', for multiple depth series across the hydrological years 2001 to 2017.

This report builds upon earlier analysis by Abell et al. (2012), to determine for TLI attributes:

1. Long-term trends (2001-2017);
2. Variation in long-term trend between traditional trend testing (e.g., Abell et al., 2012) and the directional confidence testing approach (e.g., McBride et al., 2014);
3. Variation in long-term trend due to changes in analytical laboratory (between 2008-2010);
4. Variation in long-term trend between sites of water column sampling;
5. Variation in long-term trend by depth of water column sampling;
6. Variation between long-term (2001-2017) and more recent trends (2009-2017).

[1] Note that NO3 and NH4 are both expressed here, in terms of associated nitrogen (i.e., also described as NO3-N, NO3N in the literature).

1.3 Lake Rotorua

Lake Rotorua (Bay of Plenty) is a large (80.8 km^2), relatively shallow (~10 m mean depth) and polymictic lake that stratifies thermally and mixes numerous times per year in response to ambient weather conditions. Lake Rotorua is eutrophic. The water quality declined in the late 20th Century (Rutherford, 1984; PCE, 2006), but has improved in recent years due to multiple management interventions.

BOPRC sets water quality targets for the Rotorua Lakes in the Regional Water and Land Plan based on the Trophic Level Index (TLI). The target water quality for Lake Rotorua is a TLI of 4.2. The lake has been tracking close to this level in recent years, achieving a TLI ≤4.2 relatively frequently since 2012 (e.g., 2012, 2013, 2014, 2017; Figure 1.3). Prior to this, the lake's TLI was frequently considerably more elevated indicating more degraded water quality, historically.

Management interventions to improve the water quality of Lake Rotorua include: land disposal of the city's wastewater since 1991 (into the Puarenga Stream catchment), sewage reticulation of smaller communities, trial of nitrogen removal of water from Tikitere geothermal field (2011), alum dosing to lock phosphorus from Utuhina Stream (2006) and Puarenga Stream (2010), and regional rules to cap land-based inputs (Rule 11).

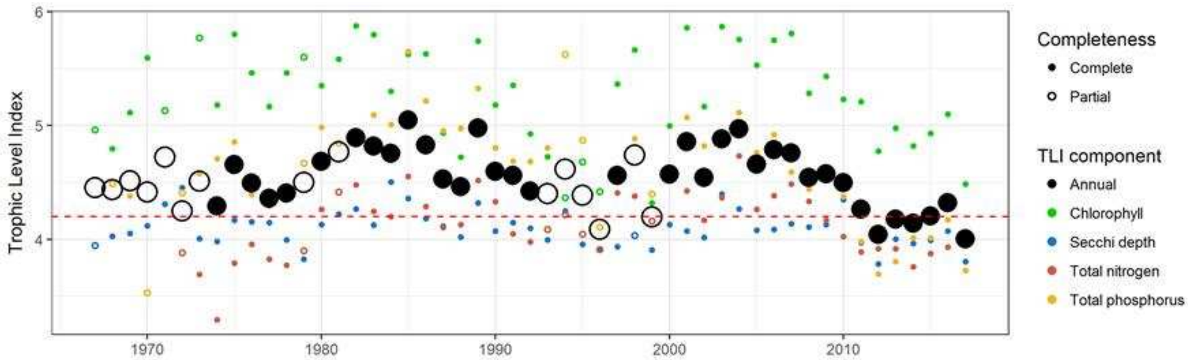


Figure 1.3 Long-term Trophic Level Index (TLI) measurements at mid-lake sites by hydrological year (previous July to present June). Each coloured circle is the mean of seasonal (quarterly) means, with the TLI equation applied (Burns et al., 1999). Solid dots denote years for which at least one measurement was available for all four seasons. Black circles are the overall annual TLI, where solid dots denote "complete" years in which all four seasons were sampled for all four component variables of the TLI. The lake TLI target is shown by the dashed red line[2].

[2] Davies-Colley et al. (2012:61) notes in reporting TLI, individual component observations should be converted into TLx scores before averaging into an annual TLx score for each of the four components (i.e., monthly TP converted into TLp, before converting these x12 TLp scores into an annual average TLp; repeat for TN, SD and Chl-a before combining into average annual TLI). Otherwise converting the annual average TP into a TLp and repeating this for TN, SD, and Chl-a, before averaging all four components, over-estimates annual TLI by ~10%. Figure 1.3 adopts the former approach.

1.4 Monitoring of Lake Rotorua

BOPRC collects lake water quality samples at two long term monitoring sites in Lake Rotorua. Lake Rotorua site 2 is south of Mokoia Island (-38.10017 / 176.28270) and Lake Rotorua site 5 is north of Mokoia Island. Water quality samples are collected monthly from three depths, if the lake is stratified the surface water samples are collected from the epilimnion, and the bottom water samples from both the hypolimnion and deep in the hypolimnion in general accordance with Burns et al. (2000). The top sample is collected as an integrated sample over the depth of the typical epilimnion and the bottom samples as discrete samples using a Van Dorn sampler.

Prior to July 2001 the depths sampled on any particular occasion was determined by the depth of the thermocline (as per Burns et al. 2000). Since July 2001, the samples were collected from consistent depths based on a typical thermocline; this simplified sample collection procedures. For Lake Rotorua Site 2 the depths are: integrated surface (0 – 6 m), discrete bottom/mid (15 m), and discrete bottom (18.5 – 24 m). For Lake Rotorua Site 5 the depths are: integrated surface (0 – 6 m), discrete bottom/mid (18 m), and discrete bottom (21 – 24 m) (Hamill and Scholes 2015).

1.5 Reporting by Abell et al. (2012)

Abell et al. (2012) investigated trends for the Trophic Level Index (TLI) and its four constituent variables over the interval of 2001-2012 in “surface” waters only; those from the surface to 6 m depth were collected by integrated-depth (tube) sampling.

Notable features of the Abell et al. (2012) report methodology include:

- No adjustment of nutrient time-series for laboratory changes and potential bias introduced by changing analytical procedures or equipment between 2008 and 2010;
- Use of the seasonal Kendall trend test (in TimeTrends v.3.31) and time-series deconstruction by loess (in the statistical package, R), to report on changes to TLI attributes whilst isolating for seasonal (monthly) effects on changes to time-series (i.e., to report on long-term, de-seasonalised trends);
- Characterisation of trend direction and magnitude, under the traditional approach of examining the statistical significance of a nil-hypothesis test on the median Sen slope estimate.

Abell et al. (2012) concluded that:

- TLI attributes improved from 2001-2012, at a rate that was statistically significant (different from zero; $p<0.001$);
- Improvement in TLI attributes were similar in direction, magnitude and statistical significance between Sites 2 and 5, suggesting basin-wide improvement from 2001-2012;
- Reduction (improvement) in TP concentrations were greater relative to those in TN concentrations, increasing the mass ratio of TN:TP in surface waters (by +6.1 to 6.3%/yr, 2001-2012);
- Improvement in TLI attributes ranged in order from TP (-8.2 to -9.0%/yr) > Chl-*a* (-6.7 to -7.0%/yr) > TN (-3.5 to -3.6%/yr) > SD (+2.6%/yr).

2 Methodology

2.1 Data sources

Water quality sample data were obtained from BoPRC's 'LakeWatch' database for Sites 2 and 5 at Lake Rotorua and for the period 2001-Jan to 2017-Dec.

2.2 Depths

In line with Abell et al. (2012), samples collected within 6 m of the surface were amalgamated into a "Surface" depth-series, those from 6.5-18 m into a "Mid" depth-series, those of 18.5-20 m into a "Deep" depth-series and those of >20 m into a "Bottom" depth-series. "Surface" samples were typically collected by a 6 m tube sample (and thus represent average water quality 0 – 6 m), whereas other samples were collected at discrete depths by Schindler-Patalas trap.

2.3 Time intervals

Where more than one observation was reported within a month, each was retained until trend analysis where median monthly observations were utilised.

In addition to depth, water quality data were also subset by time period to assess whether changes in trend have occurred since reporting by Abell et al. (2012). Periods assessed included the full dataset (June 2001 to July 2017), prior to and after the commercial lab change (July 2001 to November 2009; November 2009 to June 2017) and rolling 5-yearly intervals for the period November 2009 to July 2017 (e.g., November 2009 to June 2014, July 2010 to June 2015, July 2011 to June 2016 and July 2012 to June 2017).

2.4 Dataset quality control - anomalous values

Earlier reporting (Hamill and Scholes, 2016; Section 1.3) recommended exclusion of DRP observations at Site 2 ("Surface series) in April 2008 and TN at Site 2 ("Surface" series) in February 2010. However, neither observations appeared anomalous, so were retained.

However, values were excluded where observations of DRP>TP or NO_3+NH_4 > TN. DRP and TP observations were both excluded for: Site 2 on 20/09/2016 ("Mid" series); Site 5 on 21/11/2001 ("Mid" series), 17/12/2003 ("Mid" series), 19/09/2007 ("Mid" series) and 17/05/2005 ("Bottom" series). Exceedance of TN by NO3+NH4 occurred only at Site 2 on 22/02/2007 ("Mid" series) where NH_4 was excluded for being twice the concentration of TN but retaining TN and NO3N (as latter did not exceed former).

All observations reported as zero were excluded, affecting only DRP and NO3N series. It was unclear from the available dataset whether these samples were collected and analysed as below analytical detection limits or were simply not collected/analysed.

DRP observations were excluded for Site 2 on 16/01/2013 ("Mid" series), 16/11/2016 ("Tube" series), 16/11/2016 ("Bottom" series) and for Site 5 on 16/11/2016 ("Bottom" series). NO3N observations were excluded for Site 2 on 16/11/2016 ("Tube" series), 16/11/2016 ("Bottom" series), 18/01/2017 ("Mid" series) and for Site 5 on 18/01/2017 ("Tube" series), 18/01/2017 ("Mid" series) and 18/01/2017 ("Bottom" series).

2.5 Correction for laboratory change

BoPRC have used four laboratories for analysis of data reported here[3]:

- BoPRC internal laboratory service prior to 2001 until mid-February 2009;
- Hills commercial laboratory service from mid-February 2009 until October 2009;
- NIWA commercial laboratory service from November 2009 until June 2010;
- BoPRC internal (new) laboratory service from July 2010.

Prior reporting by Scholes (2013) and Hamill and Scholes (2016) demonstrated marked changes to reported TN, TP and DRP concentrations likely occurred due to change in laboratory. The latter identified a step-change in DRP from July 2008 and from November 2009 for TN and TP [4]. Despite the apparent oddity of the July 2008 step-change to DRP, we have applied the following corrections identified in those earlier reports to isolate the effects of those laboratory changes on nutrient time-series:

- DRP prior to July 2008 = 0.9826*DRP reported – 0.0081 (Scholes, 2013);
- TN from November 2009 = TN reported + 50.7 (Hamill and Scholes, 2016).
- TP from November 2009 = 0.829*TP reported (Hamill and Scholes, 2016)

Adjustments generated three new variables that were analysed for trends as per original variables, to ensure the effects of analytical changes on trend reporting are addressed: DRP_COR (corrected DRP), TN_COR (corrected TN) and TP_COR (corrected TP). The adjustments are made for the purpose of trend analysis and does not imply that the data prior to 2009 is more accurate.

2.6 Dataset quality control – data censoring

For trend analysis, it is common practice to censor or replace non-detected observations with values corresponding to one-half the detection limit (if left-censored; below detection) and by their value at detection plus one tenth (if right-censored; above detection) (Larned et al. 2015).

Left detection limits were only recorded by metadata for TP and DRP time-series. The latter shifted from 4 mg DRP/m^3 to 1 mg DRP/m^3 after November 2009, and from 8 mg TP/m^3 to 4 mg TP/ m^3 from February 2009 to 1 mg TP/m^3 from November 2009. However, the latter did not appear to correspond to reported minima in DRP which varied well below 8 mg DRP/m^3 prior to November 2009 (to 1 mg DRP/m^3). In the absence of knowing whether these observations were accurate (with inaccurate detection limits) or an inaccurate entry of data (with accurate detection limits), and that if the latter, these likely represented random output, detection limits were instead consistently set for DRP to 1 mg DRP/m3 throughout the time-series.

[3] Dates provided appear to be indicative, and the precise boundary of the lab change might vary for indicator. We have attempted to identify the most likely boundary for each laboratory. Notably, any effects of a laboratory-induced step-change have been minimised by splitting the time-series as per recommendations in Scholes (2013) and Hamill and Scholes (2016). Likewise, all time-series analysed extend widely beyond or before the latter date, with the trend analysis methodology relatively insensitive to a handful of "outliers" that these either late or early step-changes could introduce.

[4] The timing of step-change in DRP is not consistent with reported laboratory change.

The rationale for this approach was that randomised output should not then markedly affect the overall direction of pairwise differences between observations, the basis of determining the direction of a trend (i.e., trend direction should continue to be correctly identified from observations above detection limit or otherwise fail to be identified with confidence if the latter affected observations predominate). However, information about the magnitude of trend then becomes potentially misleading. Given similarly low observations of NO_3 and NH_4 as per DRP, and in the absence of knowledge of the corresponding detection limits, equivalent detection limits were applied to the latter (e.g. 1 mg/m3).

The proportion of NO_3, NH_4 and DRP observations that failed to exceed either 4 mg/m3 prior to November 2009 and 1 mg/m3 after November 2009, are presented for depth-series at Sites 2 and 5 in Table 3.6.1. NH_4 appears least affected whilst between 17.6- 42.7% of NO_3 and DRP series would be affected should the latter left detection limits actually apply.

TN detection limits were presumed to be equivalent to TP whilst turbidity and SD detection limits were set to 0.1 NTU and 0.1 m respectively, on basis of expert opinion.

No variable warranted right-censoring whereas even with less conservative detection limits adopted as above, NO_3, NH_4 and DRP were left-censored in all site by depth by time-interval combinations.

Table 3.6.1: Proportion of reported NO_3, NH_4, DRP observations that failed to exceed 4 mg/m3 prior to November 2009 and 1 mg/m3 thereafter.

Site	**Depth**	**Series *n* including NA (Year length)**	**NO_3**	**NH_4**	**DRP**
2	Surface	189 (2001 to 2017)	35.4%	7.7%	32.2%
2	Mid	237 (2001 to 2017)	39.7%	10.4%	27.4%
2	Deep	85 (2003 to 2017)	42.7%	7.6%	32.9%
2	Bottom	34 (2005 to 2017)	17.6%	2.9%	21.9%
5	Surface	184 (2001 to 2017)	41.5%	12.2%	36.5%
5	Mid	191 (2001 to 2017)	40.5%	8.1%	32.4%
5	Deep	14 (2002 to 2016)	21.4%	7.1%	35.7%
5	Bottom	156 (2003 to 2017)	36.0%	6.7%	31.5%

2.7 Exclusion of Series

It is good practice in trend analysis to exclude time-series that offer insufficient temporal span or frequency of detection (e.g., Helsel and Hirsch, 1992). Both Larned et al. (2015) and Snelder (2018) recommended excluding time-series for lakes where <15% of samples were censored, <80% of years and/or <80% of seasons lacked observations[5].

Applying the seasonal and inter-annual coverage rules (>80% observations distributed across both) resulted in the exclusion of all Site 2 "Deep" time-series (all indicators) as well as Site 5 "Deep" time-series for NO_3 (2003-2009), NH_4 (2003-2017; 2003-2009) and DRP (2003-2017; 2003-2009) (Table 3.7.1).

Applying the censoring rule (<15% censored), excluded NO_3 and DRP time-series for all site and depth combinations, with the exception of NO_3 at both Site 2 "Surface" and "Mid" depths for 2011-2016 and 2012-2017 time-intervals (e.g., NO_3 is included for analysis of 5-yearly rolling trends at Site 2, in both depth-series, but otherwise is not robust for interpretation of trends therein) (Table 3.7.2).

NH_4 time-series generally possessed fewer censored observations across all depth by site combinations of time-interval. At Site 2, only "Mid" depths over the 2001-2017 and 2001-2009 intervals failed to adhere to the rule for <15% censoring of NH_4. At Site 5, only "Surface" depths over the 2001-2017 and 2001-2009 intervals failed the same rule. Hence at both sites, more recent NH_4 observations for intervals since 2009 are suitable for trend analysis, on basis of censoring and coverage of seasons and years.

Note: output for NO_3, NH_4 and DRP are included in Appendix 1 regardless of exclusion from the main report results and discussion, given the potential that direction information is valid nonetheless.

[5] Note LAWA (www.LAWA.org.nz) use a modified set of rules to exclude time series from analysis if >90% of data within any one year are not reported or that >30% of reported data is censored. The latter was deemed too permissive for analysis here given the already permissive approach taken in censoring (e.g., low detection limits applied).

Table 3.7.1: Proportion of monthly observations for which a raw or censored observation was reported, by site and depth, for hydrological year interval 2001 to 2017. Rule applied requires ≥80% annual and inter-annual coverage. Values greater than 100% is due to more than one observation in some months.

Site	**Depth**	**SD**	**TP**	**TN**	**Chl-*a***	**NO_3**	**NH_4**	**DRP**	**Turbidity**
2	Surface	104%	100%	99%	98%	96%	96%	96%	99%
2	Mid	NA	122%	12%	NA	118%	119%	119%	121%
5	Surface	100%	97%	95%	96%	93%	95%	94%	97%
5	Mid	NA	97%	98%	NA	96%	97%	95%	99%
5	Deep*	NA	93%	91%	NA	89%	90%	89%	93%

*Deep time-series extend from 2003 to 2001.

Table 3.7.2: Proportion of left-censored (< detection limit) monthly observations by site and depth for Sites 2 and 5 across valid depth intervals, for hydrological year interval 2001 to 2017. Rule applied requires <15% for inclusion in trend analysis; no observations were right censored.

Site	Depth	SD	TP	TN	Chl-*a*	NO_3	NH_4	DRP	Turbidity
2	Surface	0%	0%	0%	0%	17%	2%	13%	0%
2	Mid	NA	0%	0%	NA	22%	3%	11%	0%
5	Surface	0%	0%	0%	0%	13%	4%	12%	0%
5	Mid	NA	0%	0%	NA	21%	2%	11%	0%
5	Deep*	NA	0%	0%	NA	21%	1%	12%	0%

*Deep time-series extend from 2003 to 2001.

The outcome of seasonal/inter-annual coverage and censoring rules is recorded by Table 3.7.3. All TN, TP, Chl-*a*, SD and Turbidity time-series across station, depth and time-interval combinations are suitable for trend analysis. NH_4 time-series were generally also suitable for trend analysis, except at four of the 35 site by depth by time-interval combinations (e.g., for each indicator there are five site by depth combinations for each of seven time-intervals). Notably, 2001-2017 and 2001-2009 series had different starting dates depending on depth but otherwise all other series retained consistent starting and finish dates for hydrological year intervals; “Mid” series began in July 2001, Site 5 “Bottom” in July 2003 and “Surface” series in October 2001.

Table 3.7.3: Combinations of site by depth and time-intervals, with corresponding water quality indicators selected for trend analysis after censoring and coverage rules were applied.

Time interval		Site 2		Site 5		
	Hydrological year	Surface	Mid	Surface	Mid	Bottom
Full	2001*-2017	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity	SD, TP, TN, Chl-*a*, Turbidity	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4
Pre NIWA-lab change	2001*-2009*'**	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity	SD, TP, TN, Chl-*a*, Turbidity	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4
Post NIWA-lab change	2009**-2017	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4
Rolling 5-years	2009**-2014	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4
	2010-2015	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4
	2011-2016	SD, TP, TN, Chl-*a*, Turbidity, NH_4, NO_3	SD, TP, TN, Chl-*a*, Turbidity, NH_4, NO_3	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4
	2012-2017	SD, TP, TN, Chl-*a*, Turbidity, NH_4, NO_3	SD, TP, TN, Chl-*a*, Turbidity, NH_4, NO_3	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4	SD, TP, TN, Chl-*a*, Turbidity, NH_4

*Denotes Site 2 and 5 "Surface" series begin October 2001 whereas all other depth series begin July 2001 with exception of Site 5 "Bottom" that begins July 2003; ** denotes up to or including November 2009.

2.8 Trend analysis

All trend analysis was implemented within the statistical software TimeTrends (Jowett, 2018), utilising the recently updated seasonal Kendall trend test routine to permit directional confidence testing as per recommendations in McBride et al (2014).

2.8.1 Trend Direction

Trend analysis approaches have advanced since Abell et al (2012), avoiding earlier emphasis on the term "significant" in favour of referring to the magnitude and equivalence of trend direction. These changes have been implemented following recommendations in McBride et al (2014), first nationally by Larned et al (2015) and more recently, by Snelder et al (2018) and Young et al (2018).

McBride et al (2014) identified that trends over time almost always exist in ecological data due to stochastic events and dynamic equilibrium (e.g., for climate-driven lake indicators, due to changing seasonality in response to El Nino Southern Oscillation over 4 to 5-year intervals or to the Pacific Decadal Oscillation over 10-12 year intervals). They proposed a strength-of-evidence procedure, that interprets trend tests as direction detectors and performs them simultaneously with equivalence tests.

In contrast, traditional approaches to trend analysis typically relied on testing a form of null hypothesis that the difference between populations is exactly zero (i.e. a 'nill hypothesis'). That hypothesis is only rejected if the test's 'P value' is less than a chosen significance level (generally where $p<0.05$)[6], in which case the result is called 'statistically significant'. This approach can suffer the perverse outcome of reporting statistically significant trends of zero magnitude (meaningless but significant) or misclassifying a trend as "stable" or "not changing" ($p\geq0.5$) where insufficient samples are available to confidently reveal a trend slope[7] (meaningful but insignificant). Notably, neither criticism affected earlier reporting in Abell et al. (2012) due to the high rate and significance of trends for TN, TP, Chl-*a* and SD (rates of change of ± 2.6 to 9.0 %/yr; p = 0.001 to <0.05).

Recent trend reporting in New Zealand uses the direction of trend to be established with varying levels of confidence by a new direction-testing procedure (see McBride et al., 2014). This uses information about slopes between paired observations derived from the traditional non-parametric seasonal Kendall trend test (i.e., the output of, the Seasonal Sen Slope Estimate (SSSE)). The SSSE determines the median of all possible within-season slopes in the dataset (e.g., all possible slopes for paired January by January observations across a time-series; forward comparing later to earlier observations). The latter enables seasonal effects to be isolated from observations, classed into twelve seasons for Lake Rotorua time-series given their monthly resolution.

To report the results of directional confidence tests, we have used two approaches in line with two recent trend reporting approaches for New Zealand water quality. The first approach utilised by the Land and Water Aotearoa network classifies the direction into "improving", "degrading" or "indeterminate data to establish trend direction" class, on basis of whether the 90% confidence interval

[6] Where the chances of incorrectly identifying a difference from zero is less than 1 in 20; $p<0.05$.

[7] For further discussion of problems associated with earlier significance testing of nil hypotheses in trend testing, see Appendix 1 in Larned et al (2015)

about median Sen slope includes zero. The outcome is that a trend direction is determined with 95% confidence or assigned “indeterminate” status rather than simply deemed insignificant (this is the approach used by Larned et al. 2015, Young et al., 2018). The second reporting approach allows a more nuanced understanding of confidence in whether the data shows a trend (Table 3.8.1). This is the approach used for the Intergovernmental Panel on Climate Change (IPCC) and has been incorporated into regional state of the environment reporting in New Zealand by Snelder (2018).

Table 3.8.1: Level of confidence categories used to convey the likelihood that water quality was improving (modified from Snelder 2018).

LAWA and MfE categorical level of confidence	IPCC categorical level of confidence of improvement	Probability (%)
Improving	Virtually certain	99-100
	Extremely likely	95-99
Indeterminate	Very likely	90-95
	Likely	67-90
	About as likely as not	33-67
	Unlikely	10-33
	Very unlikely	5-10
Degrading	Extremely unlikely	1-5
	Exceptionally unlikely	0-1

Here, we apply both improved direction-testing approaches alongside the more traditional nil-hypothesis to be both relevant with recent changes but consistent to earlier trend reporting for Lake Rotorua water quality (e.g. Abell et al. 2012).

2.8.2 Meaningful Trends

For the purpose of reporting a meaningful rate of change we have used a threshold of 1%/yr of percent annual change (PAC). This is consistent with past approaches used for reporting water quality trends in Lake Rotorua (e.g. Abell et al 2012, Hamill and Scholes 2015), and nationally (e.g. Ballantine et al. 2001). This implicitly assumes that trends are unidirectional and might exclude some important reversals in trend directions.

One recent national report (Larned et al. 2015) used an alternative approach based on a nominal rate of change equivalent to lake water quality degrading from National Objective Framework (NOF) Band “A” (near-reference) through to Band “D” (nationally unacceptable) over a 10-year interval. These thresholds and rates are coarse and clearly an unacceptable for the Rotorua Lakes for which there is a mandate to improve lake water quality and targets have been set that are more stringent than the D band. The long-term target for the lake is to achieve a TLI of 4.2 or better (rather than remain within a eutrophic class or NOF banding).

Determining what is a meaningful trend is specific to the purpose of the analysis and the management intention. We recommend the Technical Advisory Group (TAG) consider setting a "meaningful" rate threshold for future reporting and or management purposes.

2.9 Time-series Deconstruction

Time-series deconstruction is not a hypothesis-testing routine but is complementary to statistical approaches like the seasonal Kendall trend test. It is more diagnostic and breaks a time-series into three components. In the deconstruction process, a time-series is split into the seasonal component (typical inter-monthly variation), a long-term or trend component and an unaccounted proportion of variance, typically based on Locally Weighted Scatterplot Smoothing (LOESS) (e.g., Cleveland et al., 1990).

Combined with information on the median rate of change over a time-series (median seasonal Sen slope), determining dates for the likely emergence, acceleration, deceleration and/or return of a trend to equivalence (zero), can help better understand its causes. Hence, deconstruction is applied here to deliver on the objective of describing recent changes in lake water quality, albeit limited to the surface (tube-integrated) time-series.

All time-series deconstruction was implemented by the STL package in R (R Core Team, 2018) for the lengthiest interval (2001 to 2017) at all "surface" and "mid" depths. STL requires continuous series with observations throughout all periods and consistent frequency – here, at monthly resolution. To accommodate this, all time-series were examined for repeated within-period observations, which were then resolved into a monthly median as necessary. Where monthly observations were absent, a simple approach of infilling with the average of the immediate prior and following monthly observation was adopted (Table 3.9.1). Given the need for continuous, frequent observations and otherwise relatively high rates of censoring for dissolved nutrient species, time-series deconstruction was limited to SD, TP, TN and Chl-*a* (e.g., TLI attributes).

Table 3.9.1: Proportion of monthly observations replaced in time-series for STL deconstruction (*n*) at Sites 2 and 5 over interval July 2001 to June 2017 for TLI attributes (July 2001 to June 2017 for SD at both sites; October 2001 to June 2017 at Site 2 for TP, TN, Chl-*a*; October 2001 to May 2017 for Site 5 TP, TN, Chl-*a*).

Site	Depth-series	SD	TP	TN	Chl-*a*
2	Surface	4% (8)	1% (2)	2% (4)	2% (5)
2	Mid	NA	4% (10)	5% (10)	NA
5	Surface	3% (6)	3% (5)	4% (8)	3% (6)
5	Mid	NA	3% (6)	2% (4)	NA

Deconstructed trend components were then extracted for each site at "surface" and "mid" depths, before non-parametric paired correlation (Spearman rank) to determine degree of association between long-term changes TN or TP and corrected TN or TP, as well as to Chl-*a* (i.e., pairing indicator trend scores within equivalent depth, to demonstrate degree of association between indicators following correction for laboratory change).

3 Results & Discussion

3.1 Long-term trends (2001-2017)

Results for seasonal Kendall trend tests on SD, Chl-*a*, TP and TN concentrations are presented separately by site and depth for the hydrological year interval 2001 to 2017 in Appendix 1 and summarised by Table 4.1.1.

Table 4.1.1: Seasonal Kendall trend test output for TLI attributes over the interval 2001 to 2017, reported using the approach of Abell et al. (2012). SD and Chl-*a* not monitored at "mid" or "bottom" depths. Statistical significance of nill hypothesis (p), Seasonal Kendal Slope Estimate (SKSE) and relative SKSE (RSKE) where latter is the SKSE standardised to median series concentration to yield relative change per annum in attribute. **Note**: confidence in trends is described later but all four attributes experienced "highly confident" improvements under new trend reporting approaches.

Site	Depth	Seasonal Kendall trend test											
		SD			Chl-*a*			TP (uncorrected)			TN (uncorrected)		
		p	SKSE	RSKE (%/yr)	p	SKSE	RSKE (%/yr)	p	SKSE	RSKE (%/yr)	P	SKSE	RSKE (%/yr)
2	Surface	<0.01	0.06	2.2	<0.01	-0.94	-6.4	<0.01	-1.72	-7.2	<0.01	-11.7	-3.3
	Mid	NA			NA			<0.01	-1.80	-7.2	<0.01	-10.9	-3.2
5	Surface	<0.01	0.06	2.3	<0.01	-0.97	-6.9	<0.01	-1.67	-6.9	<0.01	-11.2	-3.2
	Mid	NA			NA			<0.01	-1.95	-7.8	<0.01	-14.5	-4.1
	Bottom	NA			NA			<0.01	-1.87	-8.1	<0.01	-17.8	-5.0

Under earlier trend reporting approaches (e.g., Abell et al., 2012), all attributes experienced "significant and meaningful" rates of change for improvement, over the long-term, at all depths tested (p<0.05; 2001-2017). Rates of long-term improvement (RSKE) across the four TLI attributes in "surface" waters varied from -7.2 to +2.3%/yr, with values similar to those reported previously for the period 2004-2013 in Abell et al. (2012) in the order: SD (+2.2 to 2.3%/yr) < TN (-3.2 to -3.3%/yr) < Chl-*a* (-6.4% to -6.9%/yr) < TP (-6.4 to -6.9%/yr). Within surface waters (0-6 m), reductions in algal biomass (Chl-*a*) at both sites were nearly double those of TN concentration but near-equivalent to those of TP concentration – also equivalent to findings in Abell et al. (2012). Long-term rates of improvement whilst "meaningful" are less than those in Abell et al. (2012); SD and TN improving at slightly lesser rate but Chl-*a* and TP improving by ~0.5-1%/yr less in "surface" waters (0-6 m depth), than over the earlier interval of 2004-2013. Combined however, long-term trends in TLI attributes share similar characteristics with those reported by Abell et al. (2012) (e.g., of "significant and meaningful" improvement in each).

Under the directional confidence testing approach (McBride et al. 2014), all four TLI attributes showed "improving" long-term trends at all three depths tested, across both sites (e.g., 95% confidence intervals of SKSE do not include zero for period 2001 to 2017 – Figure 4.1.1). Likewise, under the more refined approach to reporting adopted by Snelder (2018), all site by depth combinations of TLI attribute time-series experienced a "virtually certain" improvement in the long-term (>99% probability – Figure 4.1.2).

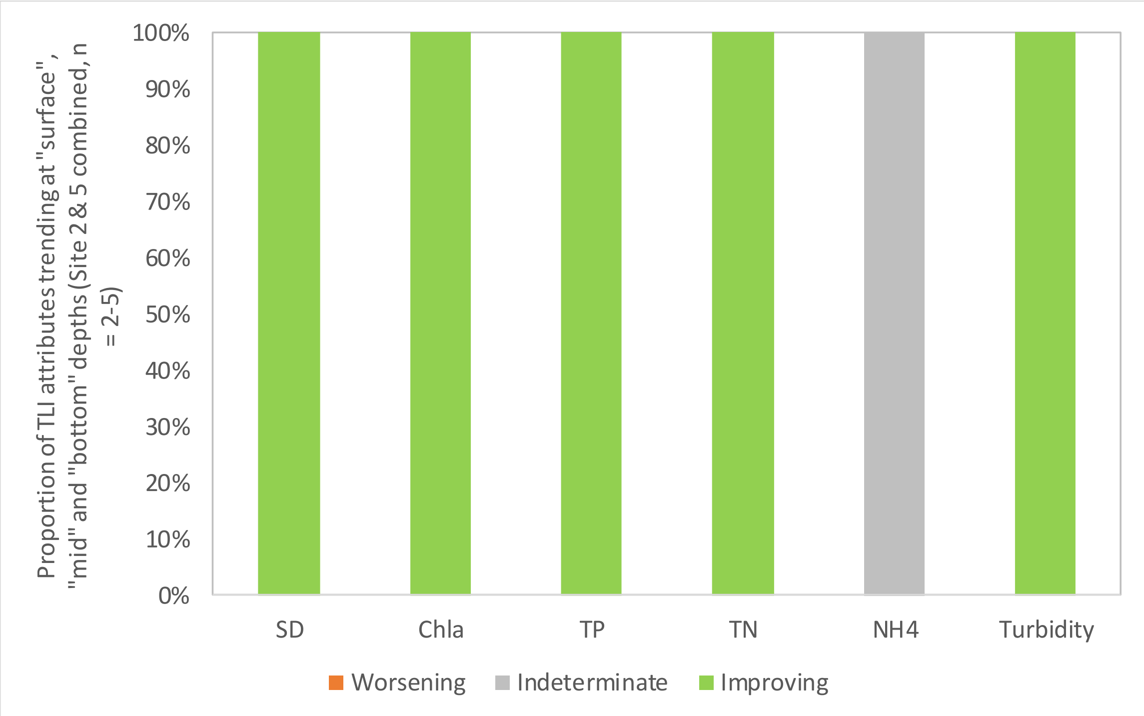


Figure 4.1.1: Trend direction reported under directional confidence testing approach recommended by McBride et al. (2014), for TLI attributes across all depth and site combinations. **Note**: *n* = 2, 2, 5, 5, 5 and 5 for SD, Chl-*a*, TP, TN, NH_4 and turbidity respectively.

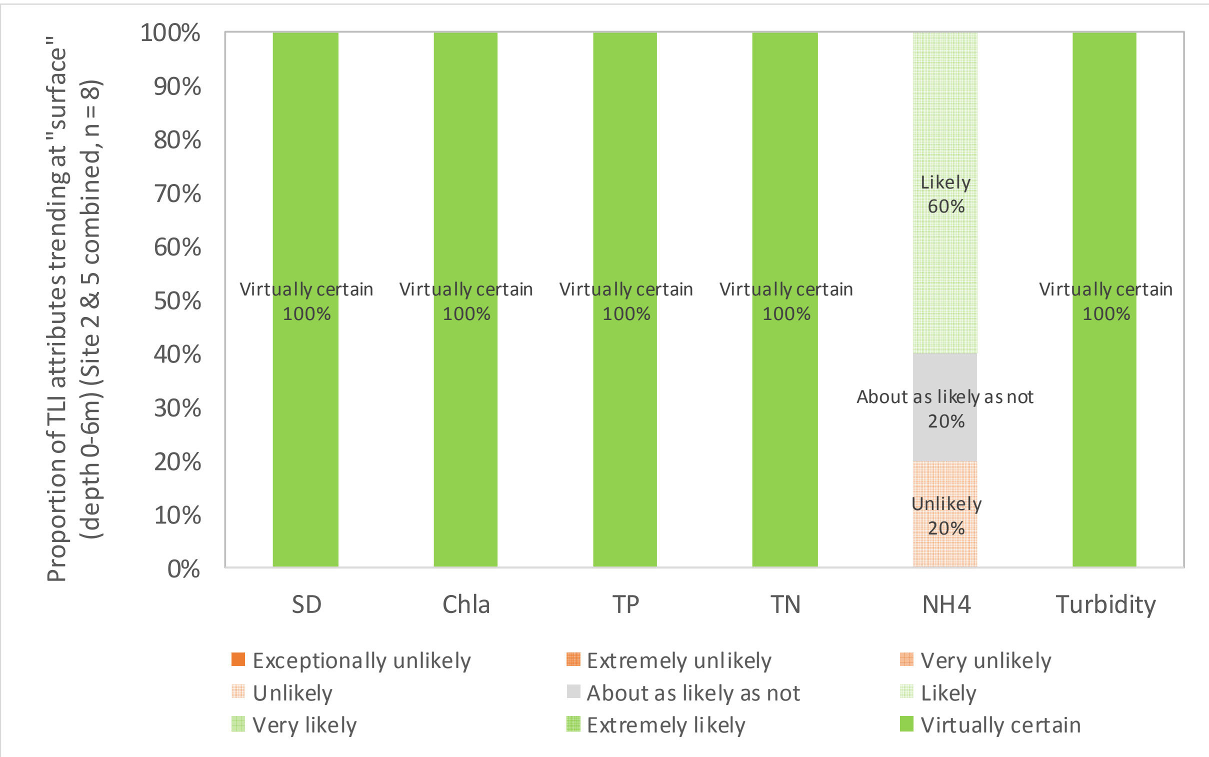
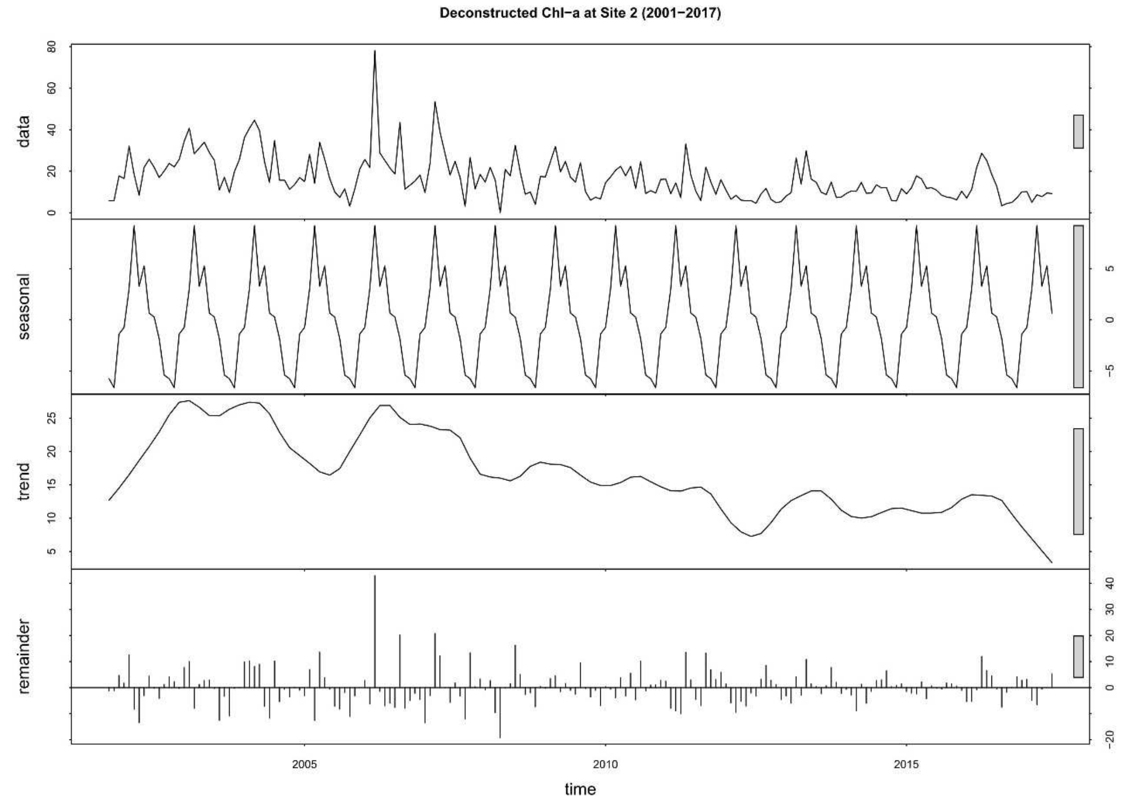


Figure 4.1.2: Likelihood of improvement under confidence interval reporting of median Sen slope direction as adopted by Snelder (2018), for TLI attributes across all depth and site combinations. All TLI attributes experienced a virtually certain likelihood of improvement (>99% probability). **Note**: *n* = 2, 2, 5, 5, 5 and 5 for SD, Chl-*a*, TP, TN, NH_4 and turbidity, respectively.



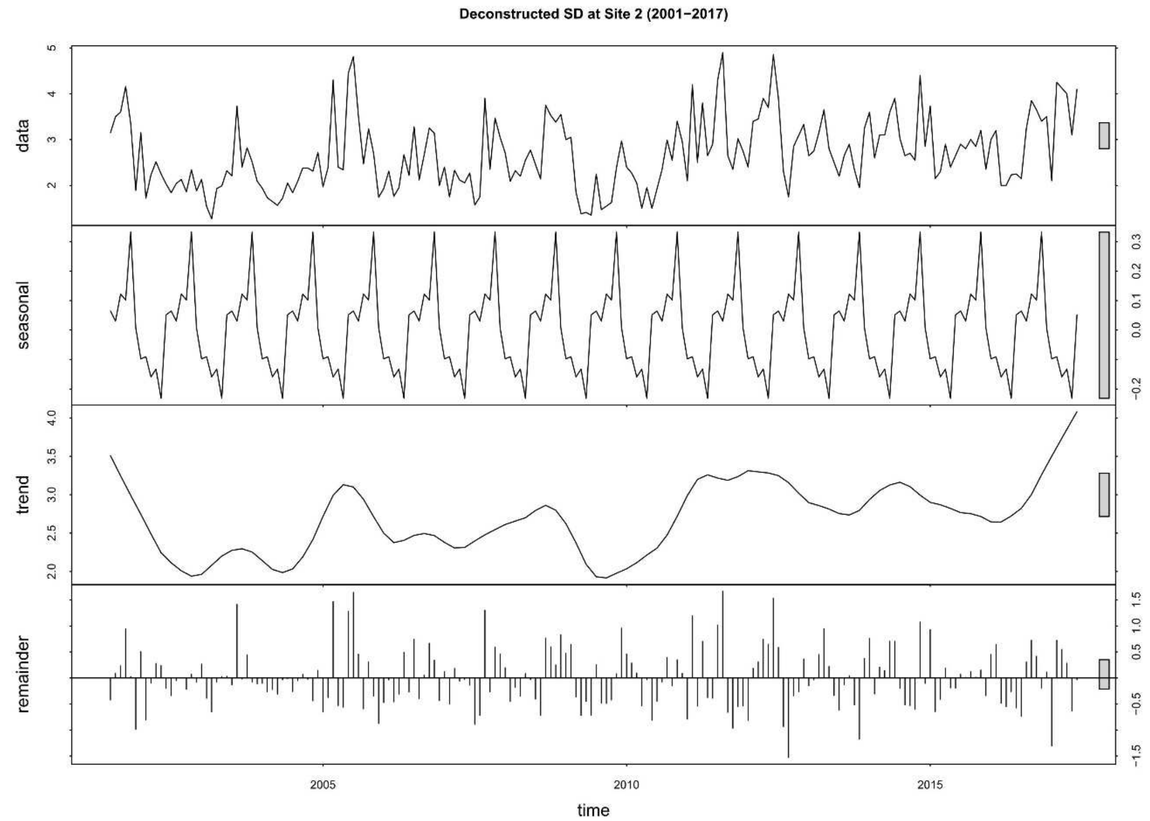
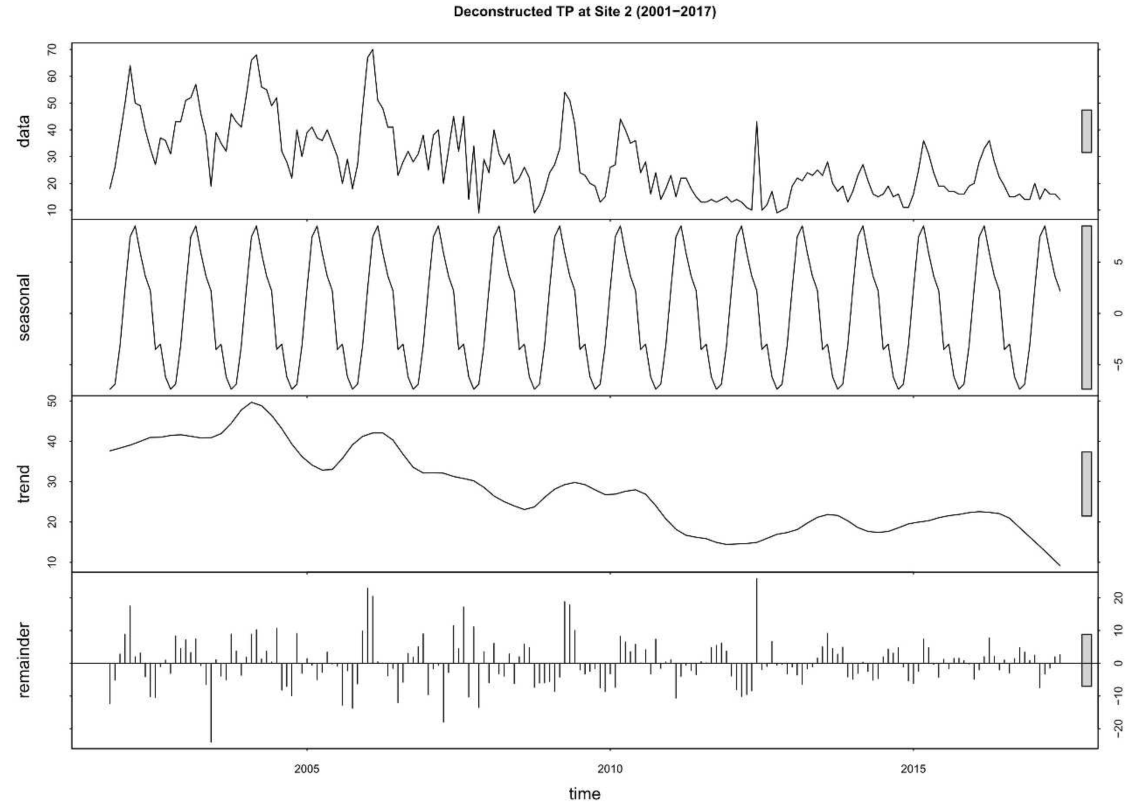


Figure 4.1.3a: Long-term (2001-2017) trend deconstruction output for Site 2 at "surface" depth for Chl-*a* and SD. See Appendix 2 for individual long-term time-series deconstruction of TN and TP at "mid" depths.



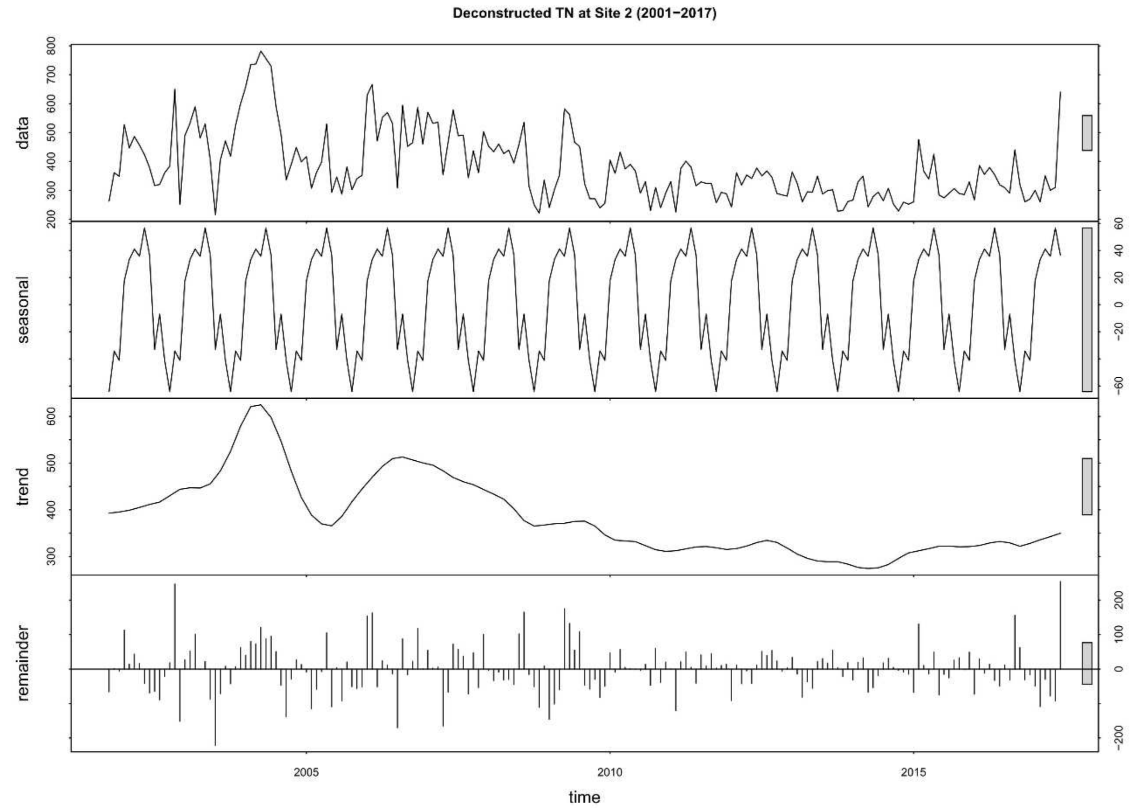
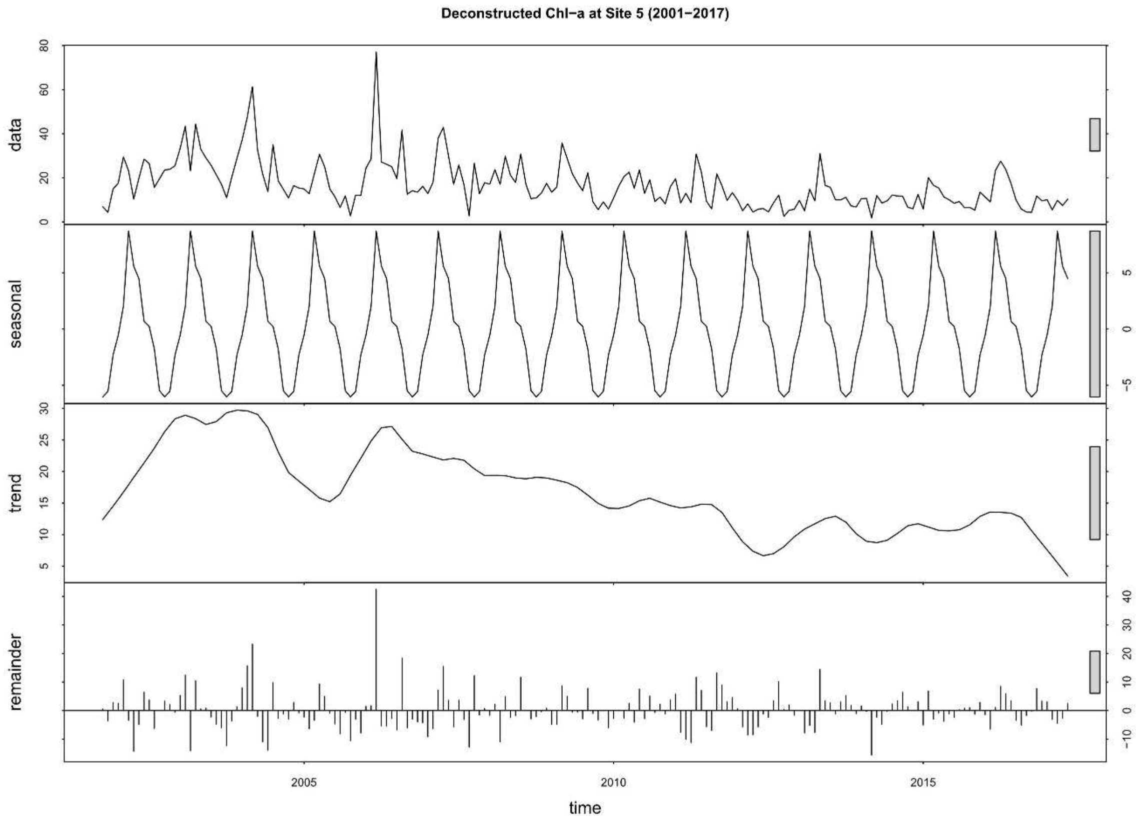


Figure 4.1.3b: Long-term (2001-2017) trend deconstruction output for Site 2 at "surface" depth for TP and TN. See Appendix 2 for individual long-term time-series deconstruction of TN and TP at "mid" depths.



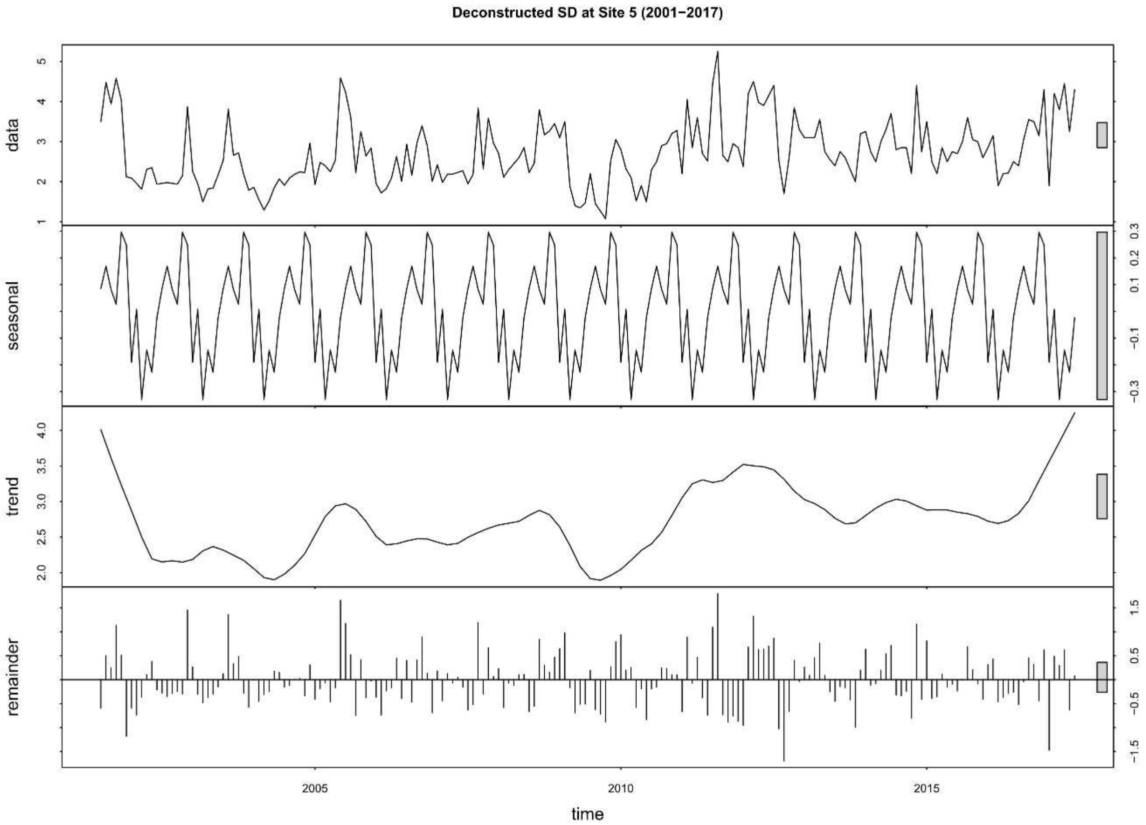
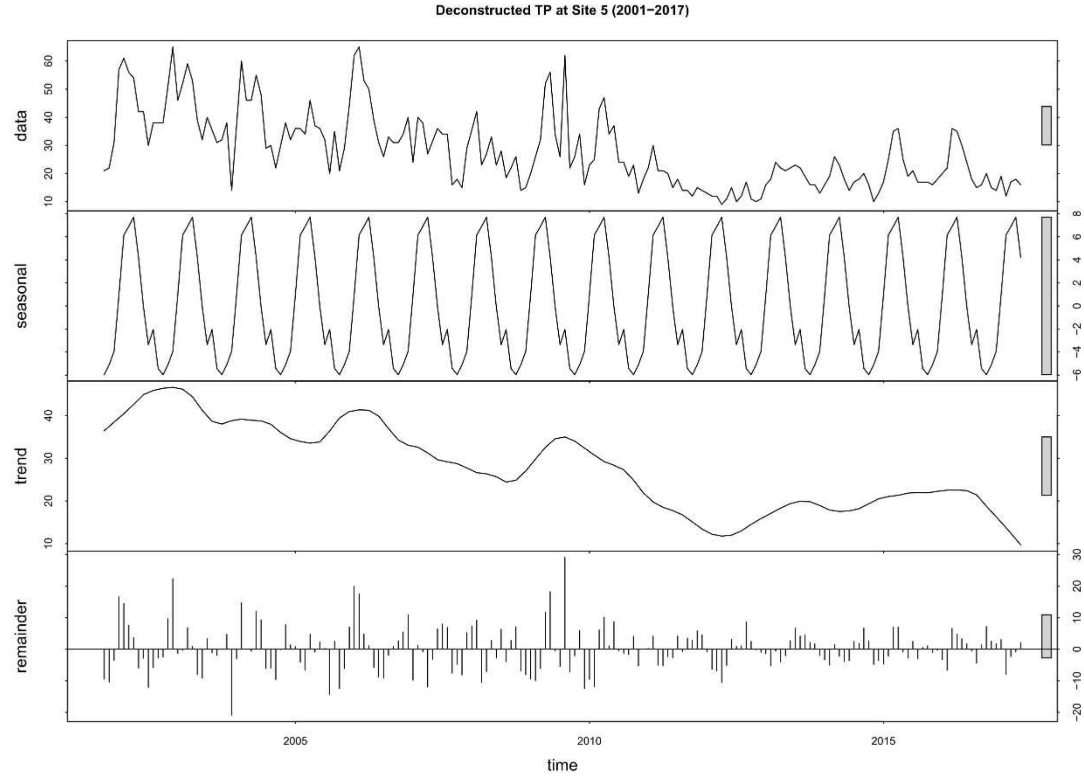


Figure 4.1.4a: Long-term (2001-2017) trend deconstruction output for Site 5 at "surface" depth for Chl-*a* and SD. See Appendix 2 for individual long-term time-series deconstruction of TN and TP at "mid" depths.



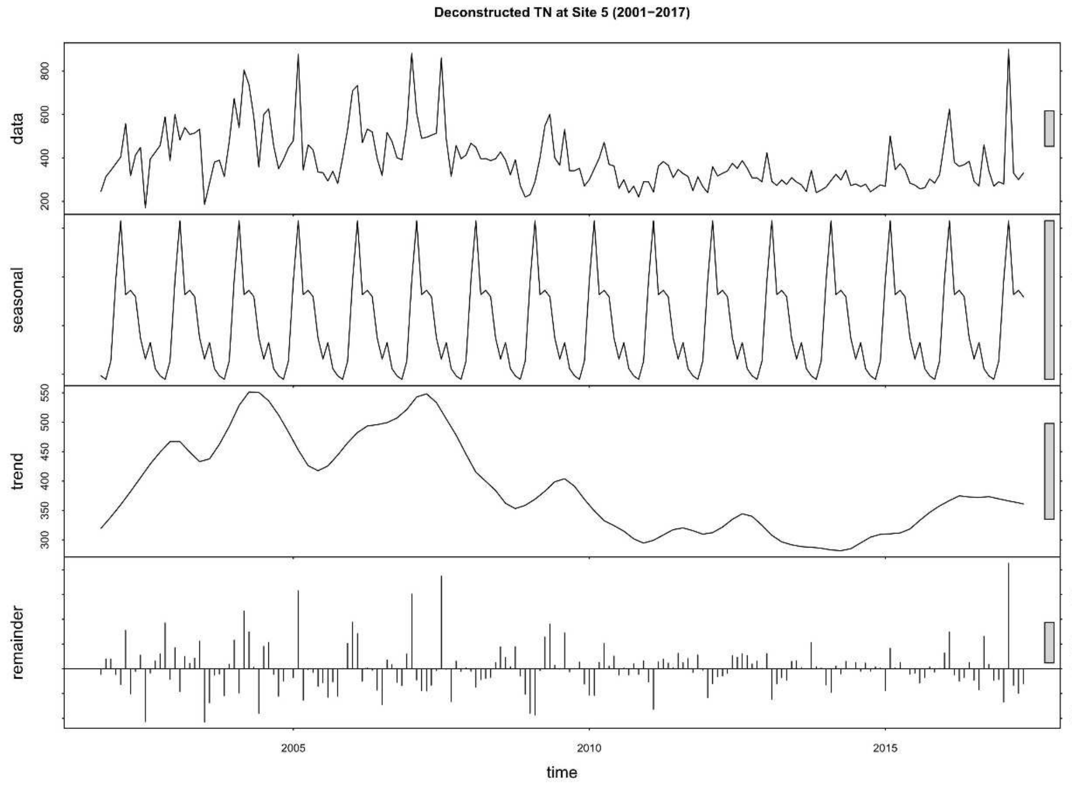


Figure 4.1.4b: Long-term (2001-2017) trend deconstruction output for Site 5 at "surface" depth for TP and TN. See Appendix 2 for individual long-term time-series deconstruction of TN and TP at "mid" depths.

All “virtually certain” improvements in TLI attributes within Lake Rotorua had a rate of change ≥1%/yr. Regarding dissolved nutrient species, long-term trends for DRP and NO_3 are unlikely to be robust due to excessive censoring and/or insufficient temporal coverage across seasons or years. However, long-term (2001-2017) trends in NH_4 at all site by depth combinations tested were “insignificant” (under the earlier trend reporting approach) and otherwise possessed “insufficient data to infer trend direction”' (under the directional confidence approach). By contrast, long-term (2001-2017) trends in turbidity represented “significant and meaningful” ($p<0.05$; ≥1%/yr) and “virtually certain” improvements at all site by depth combinations, supporting equivalent findings of widespread, long-term improvements to SD in Lake Rotorua. The rates of turbidity improvement were considerably greater than those in SD; reductions of -4 to -4.5%/yr at Site 2 and -3.4 to -4.4%/yr at Site 5 (for all depth-series; 2001-2017).

Long-term trend deconstructions are presented in Figures 4.1.3 to 4.1.4. Patterns of seasonal and long-term trend time-series were equivalent for each TLI attribute across sites from 2001 to 2017. For instance, amongst surface waters seasonal maxima in Chl-*a* occurred in late summer to early autumn (February-April), antiphased to seasonality in SD which reached maxima from winter to early summer (August-December). Seasonality in TP and TN amongst surface waters was near identical in timing to that of Chl-*a* (peaking February-April), likely reflecting the predominance of organic particulate P and N dominating latter nutrient pools in the upper 6 m (e.g., Burger et al., 2008; Smith et al., 2016). Notably, seasonal components of SD time-series at both sites varied considerably more than nutrient or algal time-series demonstrating more complex seasonal variation (i.e., responses to more than simply changing algal biomass, although responding negatively to peaks in algal biomass over late summer to early autumn).

Long-term (2001-2017) trend components to SD time-series are similar to those reported in Abell et al. (2012), and alike between stations. Trend series at both sites suggest marked reductions from ~4m SD in 2001 to ~2m by 2003, followed by modest improvement to ~3m by 2005. Thereafter, trend components in SD declined to ~2m in 2009 before reversing to greater clarity since 2009. Marked rapid improvement has occurred most recently from 2016, approaching the earlier trend components in SD equivalent to that earlier in 2001 (rising by >1m to ~4m SD).

Trend components within Chl-*a* time-series appear broadly inverse to those of SD but with a notable period from 2009 to 2011 where the two effectively decouple. For instance, at both sites Chl-*a* exhibits a near-monotonic pattern of improvement in trend components from 2006 until 2017, accelerating rapidly from 2016 (trend component Chl-*a* reducing from ~13 mg/m^3 to <5 mg/m^3 by 2017). However, marked degradation of SD from late 2008 through 2009 and recovery from 2010, is absent in Chl-*a* trend components. From this, it appears changes to SD from 2008 to 2012 were not strongly influenced by or in turn, influenced algal production at Sites 2 and 5. Whereas, outside of that interval, the two appear more strongly coupled (i.e., respond coherently to variation in the other whether one or both are drivers). The overall pattern of change in the trend components of Chl-*a* time-series at both sites are similar in magnitude, timing and rate of change (e.g., similar profiles over time). A finding supported by the near-equivalent rates and direction of change reported from seasonal Kendall trend testing.

Trend components of algal biomass demonstrated marked increases at both sites from 2001 to 2003 (from <15 to ~26 mg/m^3), albeit with marked “random” (unexplained) variation imposed over this. Little consistent change occurred to trend components of Chl-a from 2003 until 2006. Thereafter trend components of Chl-*a* decreased fairly consistently until June 2017 (record-end). Notably, the period of greatest reduction in Chl-*a* trend component was from 2006 to 2008, marking a near 10 mg/m^3 decrease over three years in trend component. Thereafter, a more modest but continued decrease in Chl-*a* trend component occurred until 2016 (declining by ~5 mg/m^3) before further rapid decrease until 2017.

Trend components within TP and TN time-series appear similar to those of Chl-*a*, though with distinct differences in the recent past (Figures 4.1.3 to 4.1.4). For instance, changes in Chl-*a* trend component prior to 2009 appear more similar to those of TN than TP; TP remained consistently elevated rather than rising from 2001 to 2003 as per TN and Chl-*a*. However, from 2009 onwards although trend components in TN decrease markedly until 2011 (from ~525 to 300 mg/m3), a further decrease does not occur until present and instead a trend for increasing TN appears from 2014 (from ~300 to 375 mg/m3, respectively). By contrast, the TP trend component time-series displays only a modest increase from 2001 to 2003 (from ~35 to 47 mg/m3) a broadly consistent decrease in TP trend component occurs from 2003 until 2017 (from ~47 to <10 mg/m3, respectively). The latter appears interrupted by a cyclical but short-lived (~1 year) increase in trend component at 4-yearly period with a rapid decrease in TN trend component from 2016 (by ~10 mg/m3 until 2017) coinciding with that cyclicity. The latter is a speculative finding and further analysis is needed to determine that a ~4-year periodic process is capable of marked variation in TP concentration, but this would align well with ENSO periodicity that has been demonstrated to alter water chemistry in other New Zealand waterbodies (e.g., Mosley, 2000; Scarsbrook et al., 2003; Reid and Ogden, 2006). Combined, this could suggest tighter coupling of trend components in TN and Chl-*a* for the earlier part of the long-term record (e.g., 2001-2009) and weaker thereafter, with the reverse evident between TP and Chl-*a*. The most recent period demonstrates this most clearly, with marked rapid improvement in SD, Chl-*a* and TP since 2016 that is not evident in TN, whose trend component increased in concentration.

3.2 Laboratory correction, site and depth: Long-term trend effects (2001-2017)

Given the omission of trend reporting on DRP, for lack of sufficient uncensored and temporally distributed samples, the effects of laboratory corrections are limited to long-term trends in TN and TP. Long-term trend findings are unchanged whether using corrected or uncorrected time-series of either. For instance, under the now outdated trend reporting approach of Abell et al. (2012), both corrected and uncorrected TN and TP concentrations across all depth by site combinations represented “significant and meaningful” improvements. That is unexpected for TN given the laboratory correction is a continual offset (addition) of 50.7 mg/m3 on samples since November 2009 (increasing all pairwise comparisons to later years thereby creating an apparent worsening on the uncorrected TN time-series), but less so in TP which was scaled (reduced) by a factor of 0.829. Clearly, despite the laboratory correction being introduced, long-term trends were sufficiently marked to otherwise remain of a meaningful rate and statistically significant from zero. Under the directional confidence testing approach, both corrected and uncorrected long-term trends for improvement in TN and TP at all depth by site combinations are “virtually certain” (e.g., 95% confidence intervals do not include zero).

Table 4.2.1: Seasonal Kendall trend test output for uncorrected and corrected TN and TP attributes over the interval 2001 to 2017, reported under the approach of Abell et al. (2012). **Note**: all trends reported below are "virtually certain" improvements under the directional confidence approach (e.g., 95% confidence intervals about median Sen slope do not include zero).

Site	Depth	Seasonal Kendall trend test											
		TP			TP corrected			TN			TN corrected		
		p	SKSE	RSKE (%/yr)	p	SKSE	RSKE (%/yr)	p	SKSE	RSKE (%/yr)	P	SKSE	RSKE (%/yr)
2	Surface	<0.01	-1.72	-7.2	<0.01	-1.98	-8.8	<0.01	-11.7	-3.3	<0.01	-7.26	-1.9
	Mid	<0.01	-1.80	-7.2	<0.01	-2.03	-8.8	<0.01	-10.9	-3.2	<0.01	-6.57	-1.8
5	Surface	<0.01	-1.67	-6.9	<0.01	-1.96	-8.9	<0.01	-11.2	-3.2	<0.01	-6.30	-1.6
	Mid	<0.01	-1.95	-7.8	<0.01	-2.16	-9.0	<0.01	-14.5	-4.1	<0.01	-9.63	-2.5
	Bottom	<0.01	-1.87	-8.1	<0.01	-2.10	-10.1	<0.01	-17.8	-5.0	<0.01	-12.2	-3.1

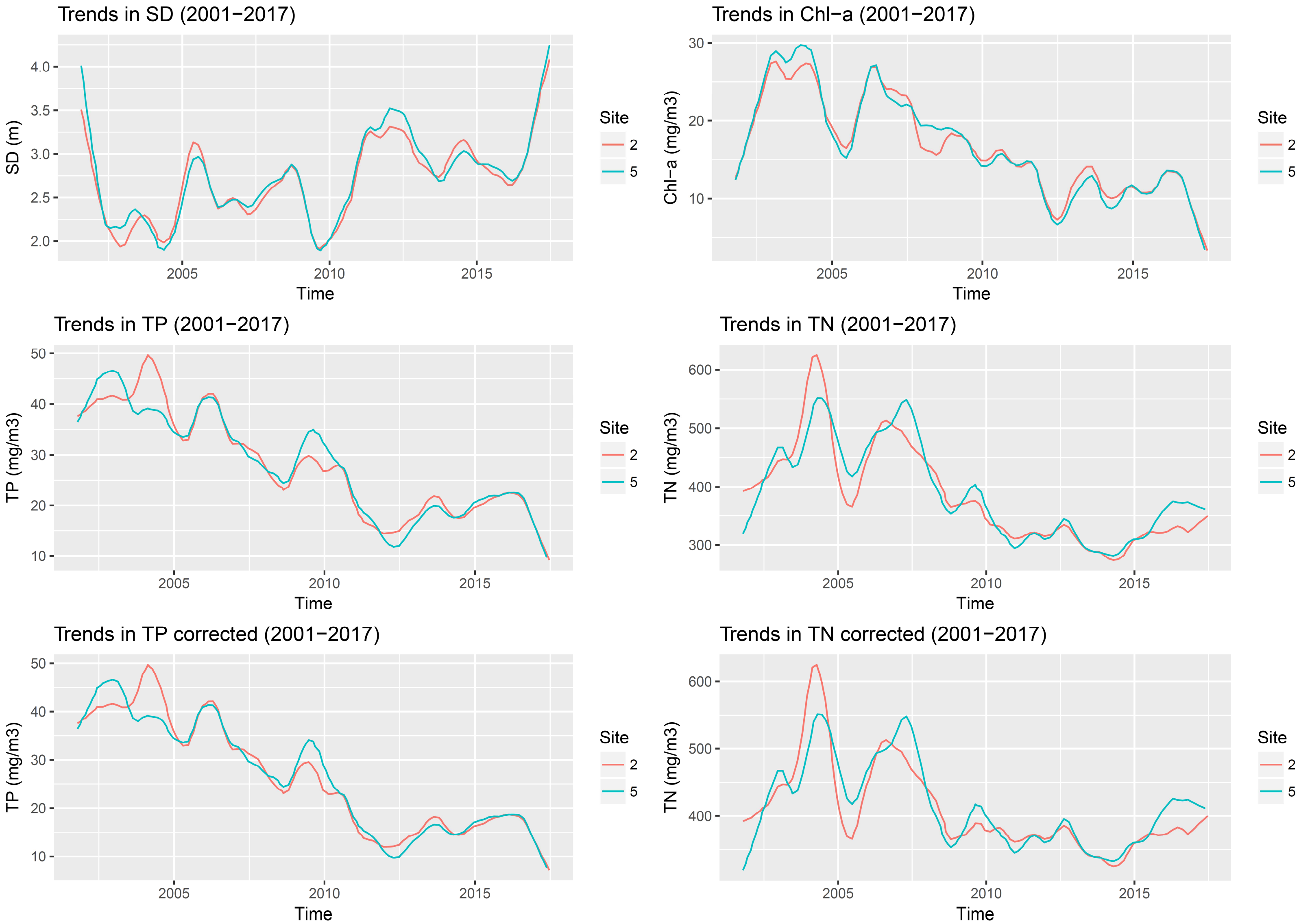


Figure 4.2.1: Long-term (2001-2007) trend deconstruction output for Sites 2 and 5 combined at "surface" depth for Chl-*a*, SD, TP and TN.

Figure 4.2.2: Long-term (2001-2007) trend deconstruction output for Site 5 by "surface" and "mid" depths for Chl-*a*, SD, TP and TN

Figure 4.2.3: Long-term (2001-2007) trend deconstruction output for Site 5 by “surface” and “mid” depths for Chl-*a*, SD, TP and TN.

From Table 4.2.1, it is clear that the laboratory correction to TP has consistently increased the rate of change (RSKE) by ~1-2%/yr over the long-term time-series (2001-2017). The opposite is true of trends in TN whose rate is reduced following the laboratory correction, also by ~1-2%/yr. Earlier reported trends in Abell et al. (2012) did not adjust for the effects of the four laboratory changes between 2008 and 2010, and although doubt exists about the correction applied here (owing to a lack of dual laboratory testing of identical samples), it is likely that earlier reporting in Abell et al. (2012) and above (Section 4.1) under-estimated improving trends in TP and over-estimated improving trends in TN. Whilst, this makes no impact on the qualitative reporting of long-term trends in TP and TN being for "significant and meaningful" or "virtually certain" improvement to both, it does have a notable effect on the order of improvement by RSKE amongst TLI attributes in "surface" waters. Notably, the corrected order of long-term improvement in TLI attributes across both sites in "surface" waters differs from Section 4.1 (and earlier reporting in Abell et al., 2012). Corrected "surface" TN concentrations improved least in the long-term between sites, in the order: corrected TN (-1.6 to -1.9%/yr) < SD (+2.2 to 2.3%/yr) < Chl-*a* (-6.4% to -6.9%/yr) < TP (-8.8 to -8.9%/yr) (2001-2017).

The overall pattern of long-term deconstructed trend components in corrected and uncorrected TN and TP is similar across the two depth-series investigated (e.g., 0-6 m "surface" and 12.5-18 m "mid" depth; Figure 4.2.2). Both corrected and uncorrected trend components exhibit similar patterns of change, described for uncorrected TN and TP in Section 4.1. This was demonstrated by strong, positive and significant (r_s= 0.96-0.99; p<0.001) correlations between trend components, paired (by time) within each depth by site combination (Table 4.2.2; Figure 4.2.1) – **note**: this might alternatively indicate a high degree of temporal autocorrelation which is common to time-series that have not undergone deconstruction but it should be noted that this is the correlation of deconstructed trend components (i.e., filtered to remove for seasonal effects likely to increase temporal autocorrelation). Trend components may yet remain autocorrelated even after the effects of seasonal processes are removed from a time-series, where they share the same long-term driver, which would then explain strong positive association. Regardless, it is clear from the strong positive association reported for corrected and uncorrected trend components, that the laboratory correction of TN and TP has not altered inferences about the timing and pattern of changes in both noted above in Section 4.1.

Table 4.2.2: Spearman rank correlation coefficients reported for corrected and uncorrected TN and TP trend components, by site and depth, paired by time for the long-term (2001-2017) time-series. **Note**: all correlations were significant to p<0.001.

Site	Depth	Spearman's correlation coefficient (r_s; all p<0.001)	
		TP~TP Corrected	TN~TN Corrected
2	Surface	0.997	0.957
	Mid	0.999	0.985
5	Surface	0.997	0.954
	Mid	0.998	0.996

Table 4.2.3 presents findings from further examination of the effects of laboratory corrections on trend components of TN and TP time-series. This examines the association between uncorrected and corrected TN or TP to Chl-*a* within "surface" waters, for both sites and paired by time. In this instance, the strength of association is expected to be relatively high due to temporal autocorrelation as TN and TP can both drive an effect in Chl-*a* and vice-versa, changes in Chl-*a* can reduce or increase TN and

TP availability. This temporal autocorrelation would be expected to be lessened in the trend component having been de-seasonalised, but regardless, should operate consistently over time on those trends if TN and TP are both equally driving or responding to changes in Chl-*a* (i.e., will result in trend components of TN~Chl-*a* and TP~Chl-*a* sharing similar direction and degree of association when paired by time). Unusually therefore, the results of non-parametric correlations demonstrate equivalent and strength of association between temporally-paired trend components of TP~Chl-*a* after 2009 than prior to (whether corrected or uncorrected), in "surface" waters at both sites (TP~Chl-*a* r_s = +0.54 to +0.79 [2001-2009] vs +0.54 to +0.65 [2009-2017]). Whereas, the direction of association between temporally-paired trend components in TN~Chl-*a* is reversed and weakened after 2009 (TN~Chl-*a* r_s = +0.70 to +0.81 [2001-2009] vs -0.09 to -0.31 [2009-2017]).

Table 4.2.3: Spearman rank correlation coefficients (r_s) reported for uncorrected and corrected TN~Chl-*a* and TP~Chl-*a* at "surface" depth, paired by date for the time interval 2001-2009, 2009-2017 and 2001-2017. **Note**: former time-intervals were split at November 2009 in line with the timing of laboratory corrections.

Site	**Correction to nutrient time-series**	**Correlation coefficients (r_s) for "surface" water trend components**					
		TN ~ Chl-*a*			TP ~ Chl-*a*		
		2001-2009	2009-2017	2001-2017	2001-2009	2009-2017	2001-2017
2	Uncorrected	0.809	-0.309	0.837	0.771	0.538	0.912
2	Corrected	0.766	-0.307	0.733	0.788	0.539	0.915
5	Uncorrected	0.736	-0.089	0.803	0.537	0.650	0.870
5	Corrected	0.698	-0.088	0.686	0.596	0.646	0.883

Further research is needed to delve into the causes for the apparent disassociation between "surface: trend components in TN and Chl-*a* from 2009, which might in turn correspond to some artefact of the TN-correction. However, the correlation output supports visual interpretation of trend components in Section 4.1 that noted the association between TN and Chl-*a* was visibly weaker from 2009 onwards at both sites for uncorrected time-series (see Figures 4.1.2 and 4.1.3). Whether temporal autocorrelation has artificially inflated the magnitude of correlations, the lack of an equivalent direction of correlation between trend components of Chl-*a* and TN or Chl-*a* and TP after 2009, raises an important question of whether both total nutrients are directly driving or responding to changing Chl-*a* more recently (i.e., whether or not temporal autocorrelation has operated to strengthen reported correlations above).

As to the effects of depth altering trend inferences, variation in trend amongst TN and TP was evident only for Site 5 and limited to the magnitude rather than nil-hypothesis significance or confidence of direction (Table 4.2.1). That is, at all depths tested and over the long-term (2001-2017) we can be "virtually certain" of improvement in TN and TP or alternatively, that both TN and TP experienced "significant and meaningful" improvement. However, at site 5 improving trends in corrected and uncorrected TN and TP were all of consistently of greater magnitude with increasing depth. "Bottom" time-series reported reductions in TP of -1.2%/yr greater rate than "surface" time-series at Site 5 with an equivalently greater rate in TN of -1.5 to -1.8%/yr (corrected and uncorrected, respectively). Hence, either the correction procedure has introduced an artefact otherwise absent at Site 2, or the behaviour of Site 5 differs from that of Site 2 by depth for possessing greater improvement at depth than in surface water total nutrient availability. Further monitoring and analysis is needed particularly at Site 2 “bottom” depths to determine this with greater rigour, but if the case, this implies that recent changes to water quality (and their cause[s]) seem better able to operate on limiting deeper nutrient reservoirs. If as has been implied by Smith et al. (2016) and Stephens et al. (2017), alum is the key driver of water quality improvement in Lake Rotorua over the long-term (since 2006), this finding would point to the effects of alum on deeper and possibly hypolimnetic nutrient reservoirs being greater than within surface waters.

To further investigate differences in trend by depth, deconstructed trend components for TN and TP were compared within site between “surface” and “mid” depths (Figures 4.2.2 and 4.2.3). Overlaying the deconstructed trend components for each of TN and TP by site, revealed a similar pattern of change in trend over time between “surface” and “mid” depths, in line with earlier findings in Table 4.2.1. that the rates of change, outcome of nil-significance and trend directional confidence outcomes were similar. Despite the RSKE increasing with depth over the period 2001 to 2017 at Site 5, but otherwise appearing not to at Site 2, the similarity in trend components suggests the timing of improving TN and TP concentrations within Lake Rotorua were similar throughout the water column.

3.3 Recent trends (2009-2017)

Shorter-term, more recent trends for “surface” waters across both sites are presented in Table 4.3.1 and also in Figures 4.3.1 and 4.3.2 which adopt reporting conventions from LAWA and MfE (Snelder, 2018), respectively (see Appendix 1 for seasonal Kendall trend test output at “mid” and “bottom” depths over equivalent intervals).

Trends in TLI attributes since 2009 appear largely of indeterminate direction, a possible consequence of insufficient data for analysis and/or trends of lesser magnitude occurring more recently – interpretation of long-term trend components in Section 4.1 and comparison to Figure 4.3.2 suggests the latter is more reasonable, albeit with recent improvement to SD, Chl-*a* and TP since 2016 (see Section 4.1 discussion of trend component). For instance, of the 40 possible pairings of TLI attributes (SD, Chl-*a*, TN, TP) across both sites and the five time intervals tested (spanning 2009-2017), 12 experienced a “meaningful and significant” trend within “surface” waters. All such 12 “meaningful and significant” trends were also “extremely likely” or “virtually certain” trends (95-99% and >99% confident, respectively), with 8 for improvement and 4 for worsening under the directional confidence testing approach of Larned et al (2015). Amongst recent time-intervals tested, 95% confident improving trends were limited to earlier years (largely from 2009-2014) whilst worsening trends from later years (only 2011-2016) (Figure 4.3.2).

We can only speculate as to what might have caused this interannual variability. Alum dosing began in 2006 to the Utuhina Stream and substantially increased in late 2010 with alum dosing to the Puarenga, this may have contributed to low TP in 2012 and 2013(e.g., Ling, 2016). Around the same time there have been climatic variability with a period of high rainfall 2011 and low rainfall during 2013 to 2015, this may have modified nutrient inputs or lake stratification (Smith et al. 2016). Also the period of 2015 and 2016 appeared to be associated with a greater influence of geothermal activity in Sulphur Bay that could contribute both TP and aluminium.

Overall, lesser emphasis should be placed in more recent trend intervals given both the reduced number of samples compared to the long-term (2001-2017) but also, trend deconstruction suggesting most marked changes in TLI attribute trends occurred from 2006 to 2009 and more recently after 2016 meaning considerably lesser change to trend components occurred from 2009 to 2016 (e.g. Figures 4.1.2 and 4.1.3). However, earlier inferences about the effects of laboratory corrections are supported by recent trend intervals; corrected and uncorrected TN and TP seasonal Kendall trend test outputs are alike in terms of “significant and meaningful” or “extremely likely” status, whilst corrected TN (TP) rates of change are generally lesser (greater) than uncorrected equivalents. Hence, long-term and recent trend inferences are equivalent under older or more recent reporting approaches (i.e., reporting approach does not affect objective findings).

Table 4.3.1: Seasonal Kendall trend test output for "surface" water quality indicators over intervals since 2009 for both sites. Confidence of direction is classified as "indeterminate" if the 95% confidence interval about median Sen slope estimate includes zero, and otherwise "improving" or "worsening" depending on latter direction. Trend reporting under the outdated approach is included, with "significant and meaningful" trends underlined. Trend output for corrected TN and TP is presented in brackets beneath uncorrected TN and TP output.

Date	**Seasonal Kendall trend test**	**Site 2**				**Site 5**			
		SD	Chl-*a*	TN (Corrected)	TP (Corrected)	SD	Chl-*a*	TN (Corrected)	TP (Corrected)
2009-2017	p	0.204	0.022	0.333 (0.333)	0.588 (0.588)	0.167	0.114	0.54 (0.54)	0.661 (0.661)
	SKSE	0.067	-0.641	-3.629 (-3.629)	-0.071 (-0.059)	0.081	-0.54	-1.66 (-1.66)	0 (0)
	RSKE	2.3	-6.3	-1.2 (-1.0)	-0.4 (-0.4)	2.8	-5.3	-0.6 (-0.5)	0 (0)
	Direction (LAWA)	Indeterminate	Improving	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate
	Likelihood of improvement	Likely	Extremely likely	Likely	Likely	Very likely	Very likely	Likely	About as likely as not
2009-2014	p	0.111	0.055	0.004 (0.022)	0.102 (0.074)	0.054	0.038	0.004 (0.017)	0.007 (0.007)
	SKSE	0.155	-1.254	-11.867 (-11.35)	-1.011 (-1.296)	0.121	-1.53	-14.785 (-12.77)	-1.875 (-1.712)
	RSKE	5.6	-12.1	-3.8 (-3.1)	-5.6 (8.7)	4.4	-15.3	-4.8 (-3.6)	-10.4 (-11.5)

	Direction (LAWA)	Indeterminate	Improving	Improving	Indeterminate	Improving	Improving	Improving	Improving
	Likelihood of improvement	Extremely likely	Extremely likely	Virtually certain	Very likely	Extremely likely	Extremely likely	Virtually certain	Virtually certain
2010-2015	p	0.137	1	0.014 (0.014)	0.467 (0.467)	0.2	1	0.372 (0.372)	0.164 (0.164)
	SKSE	-0.07	-0.084	-9.709 (-9.709)	0.585 (0.485)	-0.1	-0.017	-3.889 (-3.889)	0.989 (0.82)
	RSKE	-2.5	-0.8	-3.2 (-2.8)	3.4 (3.4)	-3.5	-0.1	-1.3 (-1.1)	5.8 (5.8)
	Direction (LAWA)	Indeterminate	Indeterminat e	Improving	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate
	Likelihood of improvement	Very unlikely	About as likely as not	Virtually certain	Unlikely	Very unlikely	About as likely as not	Likely	Very unlikely
2011-2016	p	0.042	0.156	0.717 (0.717)	0.001 (0.001)	0.033	0.111	0.767 (0.767)	0 (0)
	SKSE	-0.2	0.709	-4.601 (-4.601)	2.007 (1.664)	-0.187	0.967	1.835 (1.835)	2 (1.658)
	RSKE	-7.0	6.8	-1.5 (-1.3)	11.8 (11.8)	-6.6	9.9	0.6 (0.5)	11.8 (11.8)
	Direction (LAWA)	Worsening	Indeterminat e	Indeterminate	Worsening	Worsening	Indeterminate	Indeterminate	Worsening

	Likelihood of improvement	Extremely unlikely	Very unlikely	Likely	Exceptionally unlikely	Extremely unlikely	Extremely unlikely	About as likely as not	Exceptionally unlikely
2012-2017	p	0.505	0.525	0.346 (0.346)	0.664 (0.664)	0.469	0.941	0.255 (0.255)	0.085 (0.085)
	SKSE	0.075	-0.249	2.495 (2.495)	0 (0)	0.113	0.051	3.803 (3.803)	0.867 (0.719)
	RSKE	2.6	-2.5	0.8 (0.7)	0 (0)	4.0	0.5	1.3 (1.1)	4.8 (4.8)
	Direction (LAWA)	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate	Indeterminate
	Likelihood of improvement	Likely	Likely	Unlikely	About as likely as not	Likely	About as likely as not	Unlikely	Extremely unlikely

Figure 4.3.1: Trend direction reported under directional confidence interval approach recommended by McBride et al. (2014) for SD, Chl-*a*, TN and TP for both sites at "surface" depth, across rolling 5-year intervals since 2009. **Note**: *n* = 8 per time interval. For specific trend output refer to Table 4.3.1.

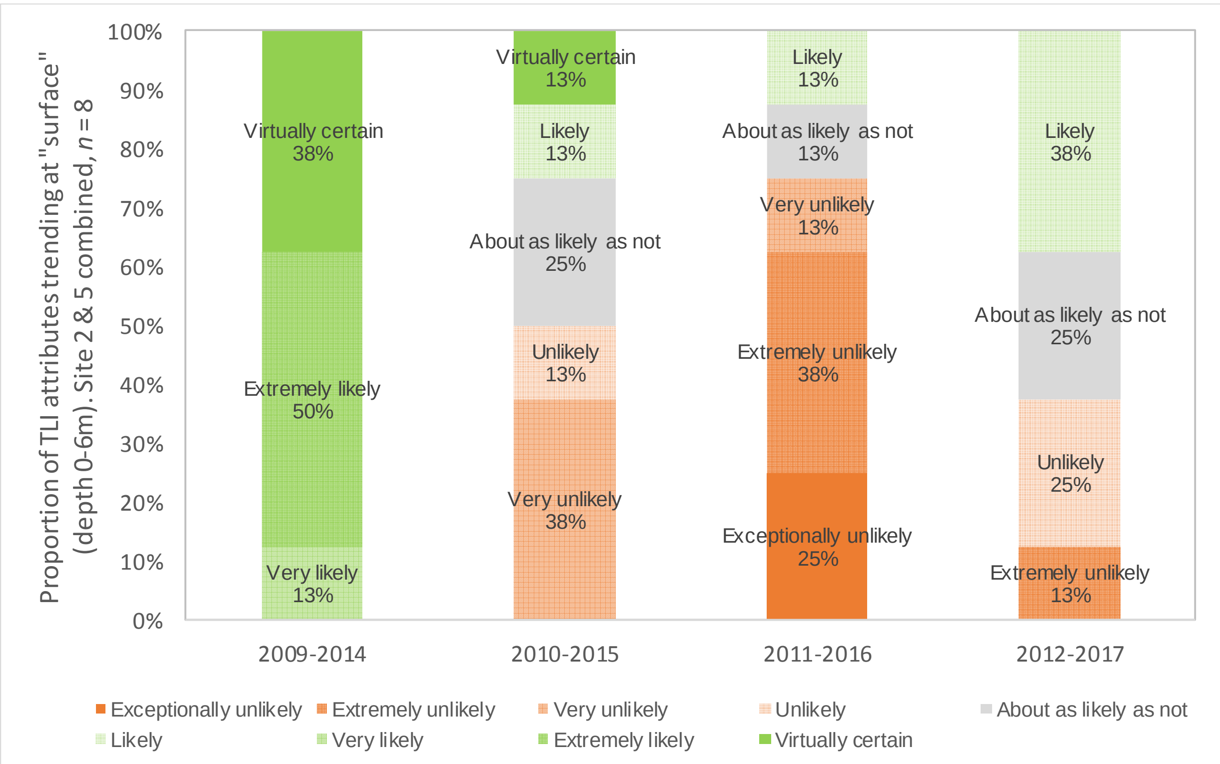


Figure 4.3.2: Probability of improving trend reported under confidence interval approach recommended by Snelder (2018) for SD, Chl-*a*, TN and TP for both sites at "surface" depth, across rolling 5-year intervals since 2009. **Note**: *n* = 8 per time interval. For specific trend output refer to Table 4.3.1.

4 Summary

This report addresses long-term and recent trends in water quality indicators over the period of July 2001 until June 2017, for both Sites 2 and 5, at three differing depths within Lake Rotorua. Censoring (for non-detects) and uneven distribution of observations over years and seasons, has generally limited robust statistical inferences to Trophic Level Index (TLI) attributes rather than dissolved nutrient species. **Note**: information on changes to ammoniacal-nitrogen (NH_4), nitrate-nitrogen (NO_3), dissolved reactive phosphorus (DRP) and turbidity is presented in Appendices 1 and 2. Long-term (2001-2017) and more recent (2012-2017) seasonal Kendall trend test output is summarised for each of the TLI attributes in Tables 5.1 and 5.2.

Table 5.1: Seasonal Kendall trend test output for “surface” water quality indicators from 2001-2017 at Site 2 and 5 within Lake Rotorua. Confidence of direction is classified as “indeterminate” if the 95% confidence interval about median Sen slope estimate includes zero, and otherwise “improving” or “worsening” depending on latter direction. Trend reporting under the outdated nil-hypothesis approach is included, with “significant and meaningful” trends underlined. Trend output for corrected TN and TP is presented in brackets beneath uncorrected TN and TP output.

Date	**Seasonal Kendall trend test**	**Site 2**				**Site 5**			
		SD	Chl-*a*	TN adjusted	TP adjusted	SD	Chl-*a*	TN adjusted	TP adjusted
2001-2017	p	<0.01	<0.01	<0.01 (<0.01)	<0.01 (<0.01)	<0.01	<0.01	<0.01 (<0.01)	<0.01 (<0.01)
	SKSE	0.057	-0.939	-11.678 (-7.262)	-1.721 (-1.978)	0.06	-0.968	-11.238 (-6.304)	-1.665 (-1.958)
	RSKE (%/yr)	2.2	-6.4	-3.3 (-1.9)	-7.2 (-8.8)	2.3	-6.9	-3.2 (-1.6)	-6.9 (-8.9)
	Direction (LAWA)	All trends were “improving” (95% confidence intervals of SKSE do not include zero)							
	Likelihood of improvement	All trends were “virtually certain” (>99% probability)							

Table 5.2: Seasonal Kendall trend test output for "surface" water quality indicators from 2012-2017 at Site 2 and 5 within Lake Rotorua. Confidence of direction is classified as "indeterminate" if the 95% confidence interval about median Sen slope estimate includes zero, and otherwise "improving" or "worsening" depending on latter direction. Trend reporting under the outdated nil-hypothesis approach is included, with "significant and meaningful" trends underlined. Trend output for corrected TN and TP is presented in brackets beneath uncorrected TN and TP output.

Date	**Seasonal Kendall trend test**	**Site 2**				**Site 5**			
		SD	Chl-*a*	TN adjusted	TP adjusted	SD	Chl-*a*	TN adjusted	TP adjusted
2012-2017	P	0.505	0.525	0.346 (0.346)	0.664 (0.664)	0.469	0.941	0.255 (0.255)	0.085 (0.085)
	SKSE	0.075	-0.249	2.495 (2.495)	0 (0)	0.113	0.051	3.803 (3.803)	0.867 (0.719)
	RSKE (%/yr)	2.6	-2.5	0.8 (0.7)	0 (0)	4.0	0.5	1.3 (1.1)	4.8 (4.8)
	Direction (LAWA)	Indeterminate (95% confidence intervals of SKSE include zero)							
	Likelihood of improvement	Likely (82%)	Likely (75%)	Unlikely (18%)	About as likely as not (50%)	Likely (83%)	About as likely as not (43%)	Unlikely (11%)	Extremely unlikely (3%)

Key findings on water quality trends in Lake Rotorua, are listed by objective below.

Objectives 1 & 2 – Long-term trends in water quality indicators (2001-2017) & variation in long-term trends between traditional trend tests and directional confidence testing approaches:

- All four TLI attributes at both sites (2,5) and over three depth combinations ("surface", "mid", "bottom"), experienced "significant and meaningful" or "virtually certain" trends for improvement over the long-term (2001-2017) using directional confidence testing reporting.

Note: seasonal Kendall trend output for each depth by site combination is available in Tables 4.1.1 and 4.2.1.

Objective 3 – Variation in long-term trend due to changes in analytical laboratory (between 2008-2010):

- Uncorrected median long-term rates of improvement in "surface" waters (≤ 6m depth) were greatest in TP and least in SD, in line with earlier reporting by Abell et al. (2012). Trends varied in decreasing order from: TP (-6.9 to -7.1%/yr) >Chl-*a* (-6.4 to -6.9%/yr) >TN (-3.2 to -3.3%/yr) >SD (+2.2 to +2.3%/yr). Across all three depths and both sites, uncorrected TP reduced by a rate of -6.9 to -8.1%/yr, nearly double the rate of reduction in uncorrected TN of -3.2 to -5.0%/yr (2001-2017).

- The Abell et al. (2012) report did not correct for bias in TN and TP introduced by laboratory changes after November 2009. In the long-term (2001-2017) across all three depths and both sites, corrected TP reduced by a greater rate (-8.8 to -10.1%/yr) and corrected TN reduced by a lesser rate (-1.6 to -3.1%/yr) (2001-2017). The corrected order of median long-term rate of change in TLI attributes within "surface" waters differs from those reported in Abell et al. (2012) as TP (-8.8 to -8.9%/yr) >Chl-*a* (-6.4 to -6.9%/yr) >SD (+2.2 to +2.3%/yr) >TN (-1.6 to -1.9%/yr). Hence, after correcting for laboratory bias, TN concentrations improved least over the long-term, at a rate of one quarter that of improvements to Chl-*a*, whilst TP concentrations improved by nearly five times the rate of TN.

Objectives 4 & 5 – Variation in long-term trend between sites and depth of water column sampling:

- Long-term (2001-2017) "surface" water trends within each TLI attribute time-series varied little between Sites 2 and 5 (e.g., equivalent classification of trends as "significant and meaningful" or "virtually certain"; similar rates of change; similar pattern of changes in trend component between sites at equivalent depth). Hence, changes to TN, TP, SD and Chl-*a* individually, appeared consistent across the lake.
- Long-term (2001-2017) trends within each TN or TP time-series were similar from "surface" to "mid" depths (e.g., equivalent classification of trends as "significant and meaningful" or "virtually certain"; similar rates of change; similar pattern of change in trend component).
- Long-term (2001-2017) trends in "bottom" waters (~20 m) were only investigated at Site 5. Although trends in TN and TP concentrations were classified alike, as "significant and meaningful" or "virtually certain" improvements, rates of change were consistently greater at depth. "Bottom" waters experienced a long-term rate of reduction in TP and TN concentration approximately 1.2%/yr and 1.5-1.8%/yr greater than "surface" waters, respectively (see Table 4.2.1 – for instance, "surface" declines in TP at Site 5 were of -6.9%/yr [uncorrected] and -8.9%/yr [corrected] compared to equivalent rates of -8.1%/yr and -10.1%/yr in "bottom" water).

Objective 6 – Variation between long-term and more recent trends

- Deconstruction of long-term (2001-2017) "surface" water time-series, indicated marked worsening in the trend components of all four TLI attributes occurred between 2001 and 2003. Marked improvement to all four TLI trend components occurred from 2006 to 2009. Thereafter, trend components of Chl-*a*, SD and TP time-series continued to generally improve, notably accelerating from 2016. Whereas, trend components of TN time-series ceased to improve after 2014, worsening thereafter (i.e., trend components of TN time-series rise in concentration after 2014).
- Over the long-term (2001-2017), trend components in corrected or uncorrected TN and TP are strongly correlated to those of Chl-*a* (r_s = +0.69 to +0.92). However, from 2009-2017 trend components of TN (corrected and uncorrected) time-series weakened and reversed to become negatively associated with Chl-*a* (r_s =-0.09 to -0.31), whilst those of TP (corrected and uncorrected) remained strongly, positively associated with Chl-*a* (r_s = +0.54 to +0.79).
- More recent (2009-2017) seasonal Kendall trend test output supports above findings of lesser rates and more divergent trends in TLI attributes.

- Directional confidence testing found “improving” trends limited largely to the interval 2009-2014, “worsening” to the interval 2011-2016 and with all trends “indeterminate” during the most recent interval of 2012-2017.
- Reporting directional confidence testing on the more nuanced scale of Snelder (2018) also showed a pattern of weakening of long-term improvement more recently. For instance, all TLI attribute trends from 2009-2014 were of “very likely”, “extremely likely” or “virtually certain” improvement (probability of improvement ≥90%). By 2011-2016, three quarters of TLI attributes reported a “very unlikely” or worse likelihood of improvement (probability of improvement <10%). Most recently from 2012-2017, equivalent proportions of TLI attributes experienced “likely” improvement (probability of improvement ≥67%) as “unlikely” or “extremely unlikely” improvement (probability of improvement <33%).

5 Conclusion

There has been a highly confident trend of improving water quality in Lake Rotorua from July 2001 to July 2017 (2017 (i.e., “virtually certain”, >99% probability of improvement). This has occurred for all TLI variables (TN, TP, Chl-*a* and Secchi depth), at both monitoring sites and at multiple sampling depths. Much of the improvements occurred in the period 2006 to 2012.

Trend components in Chl-*a* time series have historically been strongly correlated with both TN and TP (period 2001 to 2009), but since 2009 trend components in Chl-*a* were only strongly correlated with surface water TP and not with TN.

6 Recommendations for further research

- Determine trends in dissolved oxygen availability at bottom depths, both throughout the entire record length and for periods of stratification, from which to determine if alum-associated changes to internal nutrient release.
- Determine trends in nutrient ratios amongst “surface” waters and assess trend components for association to alum dosing (e.g., to alum mass dosed, to alum concentration instream at point of application).
- Determine a complementary and appropriate method to analyse for trends in overall TLI score, whether at annual or monthly time-step, that otherwise does not prevent trends in each component cancelling each other when combined.
- Revise the definition of “meaningful” to be context-specific to the targets set for water quality outcomes in Lake Rotorua and adopt one of the two rather than both trend directional methods currently in use for reporting in New Zealand.

7 References

Abell, J., Stephens, T., Hamilton, D., McBride, C. and Scarsbrook, M. 2010. *Analysis of Lake Rotorua water quality trends: 2001-2012*. Environmental Research Institute Report: 010. Prepared in response to an Environment Court mediation on the Bay of Plenty Regional Policy Statement of 21 November 2012.

Ballantine, D., Booker, D., Unwin, M., and Snelder, T. 2010. Analysis of national river water quality data for the period 1998–2007. NIWA Client Report: CHC2010-038 for Ministry for the Environment (Project MFE10502): sec. 2.4.3.

Burger, D., Hamilton, D.P., and Pilditch, C.A. 2008. Modelling the relative importance of internal and external nutrient loads on water column nutrient concentrations and phytoplankton biomass in a shallow polymictic lake. *Ecological Modelling*, **211**: 411-423.

Cleveland, R., Cleveland, W., McRae, J., and Terpenning, I. 1990. STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics*, **6**: 3-73.

Hamill, K. D., and Scholes, P. 2015. *Rotorua Te Arawa Lakes State and Trends to 2014/15*. Prepared for Bay of Plenty Regional Council.

Helsel, D.R., and Hirsch, R.M. 1992. Statistical Methods in Water Resources. USGS Techniques of Water Investigations Book 4, Chapter A3, 510 p.26.

Jowett, I. 2018.TimeTrends. v.6.30 2017. Trends and Equivalent Analysis. Funded by FRST Envirolink and sponsorship from Northland Regional Council, Hawke's Bay Regional Council, Southland Regional Council and Auckland Council.

Larned, S., Snelder, T., Unwin, M., McBride, G., Verburg, P., and McMillan, H. 2015. Analysis of water quality in New Zealand lakes and rivers. NIWA client report: CHC2015-033. Prepared for Ministry for the Environment.

Ling, N. 2016. Puarenga Stream alum dosing effects 2015. ERI client report No. 82. Prepared for Bay of Plenty Regional Council.

Mosley, P. Regional differences in the effects of El Nino and La Nina on low flows and floods. *Hydrological Sciences Journal*, **45**: 249-267.

McBride, G.B., Cole, R., Westbrooke, I., and Jowett, I.G. 2014. Assessing environmentally significant effects: A better strength-of-evidence than a single P value? *Environmental Monitoring and Assessment*, **186**: 2729–2740.

Parliamentary Commissioner for the Environment (PCE). 2006. Restoring the Rotorua lakes: The ultimate endurance challenge. Wellington, New Zealand. p. 50.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reid, M., and Ogden, R. 2006. Trend, variability or extreme event? The importance of long-term perspectives in river ecology. *River Research and Applications*, **22**: 167-177.

Rutherford, J. C. 1984. Trends in Lake Rotorua water quality. New Zealand Journal of Marine and Freshwater Research. 18: 355-365.

Scarsbrook, M., McBride, C., McBride G., and Bryers, G. 2003. Effects of climate variability on rivers: Consequences for long term water quality analysis. *Journal of the North American Water Resources Association*, **108**: 1435-1447.

Scholes, P. 2013. Trends and state of nutrients in Lake Rotorua streams 2013. Bay of Plenty Regional Council Environmental Publication 2013/08.

Smith, V., Wood, S., McBride, C., Atalah, J., Hamilton, D., and Abell, J. 2016. Phosphorus and nitrogen loading restraints are essential for successful eutrophication control of Lake Rotorua, New Zealand. Inland Waters, 6: 273-283.

Snelder, T. 2018. Assessment of recent reductions in E.coli and sediment in rivers of the Manawatū-Whanganui region. LWP Client Report for Horizons Regional Council and the Ministry for the Environment.

Stephens, T., Hamilton D., and Bruere, A. 2017. Summary of expert evidence of Dr. Thomas Stephens for DairyNZ Limited and Fonterra. Updated to include outcome of water quality expert caucusing, 3rd April 2017, Millennium Hotel Rotorua between Dr. Thomas Stephens, Prof. David Hamilton and Mr. Andy Bruere. https://www.boprc.govt.nz/media/610363/summary-of-evidence-to-read-out-pc-10-stephens-post-caucus-final.pdf

Young, R. et al. 2018. LAWA National River Water Quality 10-year trends. Published online: https://www.lawa.org.nz/explore-data/river-quality/#/tb-national. Last visited: 25/05/2018.

Appendix 1 – Summary output for seasonal Kendall trend test by TimeTrends

Lake Rotorua Site 2

Depth							Summary information			Surface			2001-2017
Site 2	Sample size	P	Median Sen slope (annual)	PAC	90% confidence limits for slope	Trend direction	Sufficient DL (>=15%)	Sufficient coverage (>80% annual)	Conclusive	Important	P(lwp)	Confidence of improvement	LAWA Direction
SD	184	0	0.057	2.221	0.039 to 0.075	increasing	Yes	Yes	Yes	Yes	1	Virtually certain	increasing
TP	184	0	-1.721	-7.171	-2.005 to -1.505	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN	182	0	-11.678	-3.319	-14.372 to -9.251	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
Chla	181	0	-0.939	-6.435	-1.116 to -0.706	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
NO3	176	0	0.25	8.335	0.165 to 0.392	increase	No	Yes	Yes	Yes	0	Exceptionally unlikely	increasing
NH4	177	0.792	0	0	-0.251 to 0.226	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
DRP	176	0	-0.285	-14.256	-0.374 to -0.200	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	182	0	-0.104	-4.541	-0.128 to -0.089	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TP_COR	184	0	-1.978	-8.838	-2.256 to -1.687	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	182	0	-7.262	-1.908	-9.709 to -5.017	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
DRP_COR	176	0	-0.278	-13.893	-0.365 to -0.200	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							Summary information			Mid			2001-2017
Site 2	Sample size	P	Median Sen slope (annual)	Percent annual change	90% confidence limits for slope	Trend direction	Sufficient DL (>=15%)	Sufficient coverage (>80% annual)	Conclusive	Important	P(lwp)	Confidence of improvement	LAWA Direction
TP	184	0	-1.803	-7.211	-2.080 to -1.538	decreasing	Yes	Yes	Yes	No	1	Virtually certain	decreasing
TN	181	0	-10.943	-3.218	-13.780 to -8.696	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	180	0	0.188	6.255	0.083 to 0.30	increasing	No	Yes	Yes	Yes	0	Exceptionally unlikely	increasing
NH4	180	0.214	-0.332	-2.215	-0.739 to 0	decreasing	No	Yes	No	Yes	0.893	Likely	Insufficient data to infer trend direction
DRP	179	0	-0.335	-16.732	-0.461 to -0.251	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	183	0	-0.084	-4.007	-0.100 to -0.060	decreasing	Yes	Yes	No	Yes	1	Virtually certain	decreasing
TP_COR	184	0	-2.033	-8.84	-2.366 to -1.811	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	181	0	-6.567	-1.761	-9.417 to -4.247	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
DRP_COR	179	0	-0.334	-16.724	-0.458 to -0.250	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							**Summary information**			Surface			2001-2009
Site 2	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
SD	95	0.161	0.025	1.077	-0.010 to 0.091	increasing	Yes	Yes	No	Yes	0.927	Very likely	Insufficient data to infer trend direction
TP	93	**0**	-2.091	-5.809	-3.008 to -1.499	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN	92	0.43	-2.994	-0.676	-17.154 to 6.238	decreasing	Yes	Yes	No	No	0.787	Likely	Insufficient data to infer trend direction
Chla	90	**0.018**	-1.198	-6.473	-1.663 to -0.263	decreasing	Yes	Yes	Yes	Yes	0.993	Virtually certain	decreasing
NO3	86	0.09	0.031	1.667	0.000 to 0.249	increasing	No	Yes	No	Yes	0	Exceptionally unlikely	Insufficient data to infer trend direction
NH4	86	0.866	0	0	-0.869 to 1.207	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
DRP	86	**0**	-0.794	-19.842	-1.687 to -0.393	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	91	0.566	-0.05	-1.718	-0.101 to 0.099	decreasing	Yes	Yes	No	Yes	0.717	Likely	Insufficient data to infer trend direction
TP_COR	93	**0**	-2.091	-5.809	-3.008 to -1.499	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	92	0.43	-2.994	-0.676	-17.154 to 6.238	decreasing	Yes	Yes	No	No	0.787	Likely	Insufficient data to infer trend direction
DRP_COR	86	**0**	-0.78	-19.883	-1.657 to -0.386	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							**Summary information**			**Mid**			**2001-2009**
Site 2	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
TP	95	**0**	-2.2	-6.112	-3.513 to -1.379	decreasing	Yes	Yes	Yes	No	1	Virtually certain	decreasing
TN	92	0.68	-2.538	-0.597	-10.092 to 6.935	decreasing	Yes	Yes	No	No	0.663	About as likely as not	Insufficient data to infer trend direction
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	92	0.907	0	0	-0.083 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
NH4	91	0.465	-0.316	-1.975	-1.331 to 0.536	decreasing	No	Yes	No	Yes	0.745	Likely	Insufficient data to infer trend direction
DRP	91	**0**	-0.82	-16.4	-1.495 to -0.500	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	94	0.36	-0.05	-1.886	-0.125 to 0.025	decreasing	Yes	Yes	No	Yes	0.787	Likely	Insufficient data to infer trend direction
TP_COR	95	**0**	-2.2	-6.112	-3.513 to -1.379	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	92	0.68	-2.538	-0.597	-10.092 to 6.935	decreasing	Yes	Yes	No	No	0.663	About as likely as not	Insufficient data to infer trend direction
DRP_COR	91	**0**	-0.806	-16.427	-1.469 to -0.492	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							Summary information				Surface			2009-2017
Site 2	Sample size	P	Median Sen slope (annual)	Percent annual change	90% confidence limits for slope	Trend direction	Sufficient DL (>=15%)	Sufficient coverage (>80% annual)	Conclusive	Important	P(lwp)	Confidence of improvement	LAWA Direction
SD	89	0.204	0.067	2.341	-0.014 to 0.100	increasing	Yes	Yes	No	Yes	0.864	Likely	Insufficient data to infer trend direction
TP	91	0.588	-0.071	-0.42	-0.690 to 0.332	decreasing	Yes	Yes	No	No	0.706	Likely	Insufficient data to infer trend direction
TN	90	0.333	-3.629	-1.178	-7.101 to 1.071	decreasing	Yes	Yes	No	Yes	0.842	Likely	Insufficient data to infer trend direction
Chla	91	**0.022**	-0.641	-6.282	-1.022 to -0.198	decreasing	Yes	Yes	Yes	Yes	0.989	Extremely likely	decreasing
NO3	90	0.937	0	0	-0.400 to 0.499	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
NH4	91	0.728	0	0	-1.003 to 0.587	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
DRP	90	0.801	0	0	0.000 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
Turb	91	**0.036**	-0.071	-3.73	-0.123 to 0.000	decreasing	Yes	Yes	No	Yes	0.975	Extremely likely	Insufficient data to infer trend direction
TP_COR	91	0.588	-0.059	-0.42	-0.572 to 0.275	decreasing	Yes	Yes	No	No	0.706	Likely	Insufficient data to infer trend direction
TN_COR	90	0.333	-3.629	-1.012	-7.101 to 1.071	decreasing	Yes	Yes	No	Yes	0.842	Likely	Insufficient data to infer trend direction
DRP_COR	90	0.801	0	0	0.000 to 0.000	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction

Depth							Summary information				Mid			2009-2017
Site 2	Sample size	P	Median Sen slope (annual)	Percent annual change	90% confidence limits for slope	Trend direction	Sufficient DL (>=15%)	Sufficient coverage (>80% annual)	Conclusive	Important	P(lwp)	Confidence of improvement	LAWA Direction
TP	89	0.502	-0.333	-1.85	-0.865 to 0.497	decreasing	Yes	Yes	No	No	0.736	Likely	Insufficient data to infer trend direction
TN	89	0.157	-3.33	-1.129	-7.498 to 0.603	decreasing	Yes	Yes	No	Yes	0.92	Very likely	Insufficient data to infer trend direction
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	88	0.354	-0.354	-7.08	-1.337 to 0.167	decreasing	No	Yes	No	Yes	0.823	Likely	Insufficient data to infer trend direction
NH4	89	0.476	-0.334	-2.787	-1.999 to 0.500	decreasing	Yes	Yes	No	Yes	0.762	Likely	Insufficient data to infer trend direction
DRP	88	0.282	0	0	-0.086 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
Turb	89	**0.004**	-0.06	-3.157	-0.100 to -0.032	decreasing	Yes	Yes	Yes	Yes	0.998	Virtually certain	decreasing
TP_COR	89	0.502	-0.276	-1.85	-0.717 to 0.412	decreasing	Yes	Yes	No	Yes	0.736	Likely	Insufficient data to infer trend direction
TN_COR	89	0.157	-3.33	-0.963	-7.498 to 0.603	decreasing	Yes	Yes	No	No	0.92	Very likely	Insufficient data to infer trend direction
DRP_COR	88	0.282	0	0	-0.086 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction

Lake Rotorua Site 5

Depth							Summary information			Surface			2001-2017
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
SD	183	**0**	0.06	2.308	0.044 to 0.074	increasing	Yes	Yes	Yes	Yes	1	Virtually certain	increasing
TP	179	**0**	-1.665	-6.937	-1.913 to -1.417	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN	176	**0**	-11.238	-3.193	-13.336 to -8.183	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
Chla	178	**0**	-0.968	-6.912	-1.178 to -0.800	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
NO3	172	**0**	0.286	14.295	0.167 to 0.501	increasing	No	Yes	Yes	Yes	0	Exceptionally unlikely	increasing
NH4	177	0.662	0.083	1.042	-0.167 to 0.400	increasing	No	Yes	No	Yes	0.321	Unlikely	Insufficient data to infer trend direction
DRP	175	**0**	-0.3	-15.01	-0.370 to -0.214	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	179	**0**	-0.1	-4.353	-0.130 to -0.071	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TP_COR	179	**0**	-1.958	-8.899	-2.192 to -1.702	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	176	**0**	-6.304	-1.611	-8.408 to -4.218	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
DRP_COR	175	**0**	-0.289	-14.459	-0.361 to -0.208	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							Summary information			Mid			2001-2017
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
TP	182	**0**	-1.951	-7.803	-2.199 to -1.563	decreasing	Yes	Yes	Yes	No	1	Virtually certain	decreasing
TN	184	**0**	-14.484	-4.097	-17.601 to -12.183	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	181	**0.002**	0.12	4.007	0 to 0.250	increasing	No	Yes	No	Yes	0.004	Exceptionally unlikely	Insufficient data to infer trend direction
NH4	182	0.536	-0.125	-0.759	-0.524 to 0.200	decreasing	Yes	Yes	No	No	0.705	Likely	Insufficient data to infer trend direction
DRP	179	**0**	-0.301	-15.041	-0.416 to -0.214	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	186	**0**	-0.075	-3.404	-0.100 to -0.050	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TP_COR	182	**0**	-2.156	-8.977	-2.488 to -1.856	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	184	**0**	-9.632	-2.506	-12.368 to -7.053	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
DRP_COR	179	**0**	-0.3	-14.994	-0.406 to -0.208	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							**Summary information**			**Surface**			**2001-2009**
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
SD	94	0.178	0.035	1.549	-0.016 to 0.070	increasing	Yes	Yes	No	Yes	0.923	Very likely	Insufficient data to infer trend direction
TP	90	**0**	-1.986	-5.841	-3.029 to -1.122	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN	89	0.722	1.921	0.465	-9.412 to 11.501	increasing	Yes	Yes	No	No	0.337	About as likely as not	Insufficient data to infer trend direction
Chla	89	0.12	-0.728	-3.895	-1.333 to 0.030	decreasing	Yes	Yes	No	Yes	0.939	Very likely	Insufficient data to infer trend direction
NO3	85	0.06	0.063	6.251	0.000 to 0.249	increasing	No	Yes	No	Yes	0.378	About as likely as not	Insufficient data to infer trend direction
NH4	88	0.904	0	0	-0.500 to 1.226	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
DRP	88	**0**	-0.901	-30.041	-1.300 to -0.502	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	90	0.14	-0.1	-3.345	-0.200 to 0.004	decreasing	Yes	Yes	No	Yes	0.931	Very likely	Insufficient data to infer trend direction
TP_COR	90	**0**	-1.986	-5.841	-3.029 to -1.122	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	89	0.722	1.921	0.465	-9.412 to 11.501	increasing	Yes	Yes	No	No	0.337	About as likely as not	Insufficient data to infer trend direction
DRP_COR	88	**0**	-0.886	-30.124	-1.277 to -0.493	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing

Depth							**Summary information**			**Mid**			**2001-2009**
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**PAC**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
TP	91	**0.001**	-1.803	-4.872	-2.481 to -0.995	decreasing	Yes	Yes	Yes	No	1	Virtually certain	decreasing
TN	93	1	-0.188	-0.043	-10.246 to 9.076	decreasing	Yes	Yes	No	No	0.504	About as likely as not	Insufficient data to infer trend direction
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	91	0.407	0	0	0.000 to 0.213	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
NH4	91	0.731	-0.167	-1.115	-0.754 to 0.599	decreasing	Yes	Yes	No	Yes	0.627	About as likely as not	Insufficient data to infer trend direction
DRP	88	**0**	-0.669	-13.385	-1.003 to -0.214	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing
Turb	95	0.664	0.02	0.766	-0.076 to 0.1	increasing	Yes	Yes	No	No	0.279	Unlikely	Insufficient data to infer trend direction
TP_COR	91	**0.001**	-1.803	-4.872	-2.481 to -0.995	decreasing	Yes	Yes	Yes	Yes	1	Virtually certain	decreasing
TN_COR	93	1	-0.188	-0.043	-10.246 to 9.076	decreasing	Yes	Yes	No	No	0.504	About as likely as not	Insufficient data to infer trend direction
DRP_COR	88	**0**	-0.658	-13.407	-0.986 to -0.211	decreasing	No	Yes	Yes	Yes	1	Virtually certain	decreasing

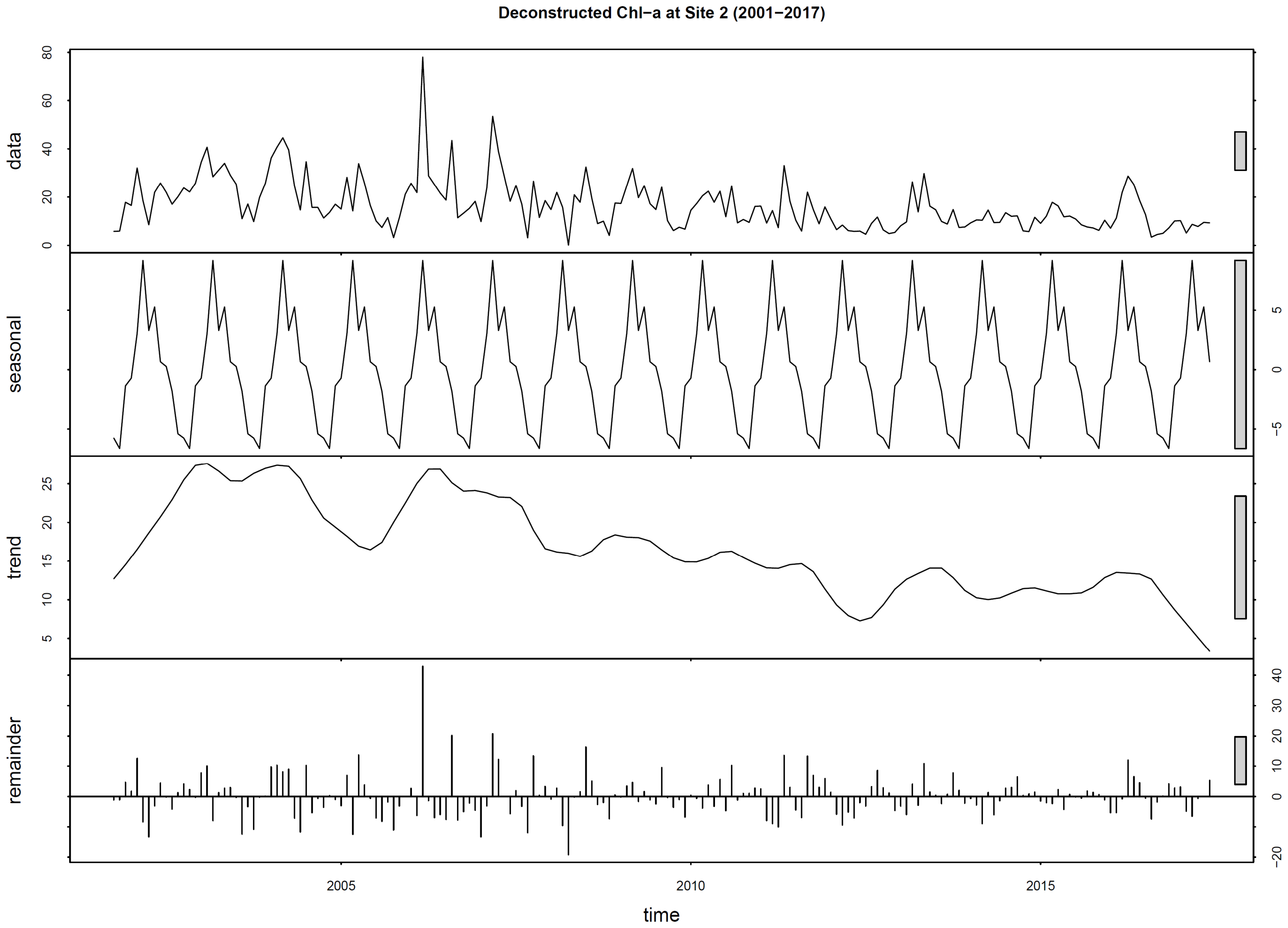
Depth							**Summary information**			**Bottom**			**2003-2009**
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
TP	66	**0.004**	-2.449	-6.71	-4.249 to -1.003	decreasing	Yes	Yes	Yes	No	0.998	Virtually certain	decreasing
TN	64	**0.003**	-25.095	-5.429	-38.795 to -14.286	decreasing	Yes	Yes	Yes	Yes	0.999	Virtually certain	decreasing
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	62	0.536	0	0	0.000 to 0.165	uncertain (median slope zero)	No	No	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
NH4	62	0.85	0	0	-1.988 to 1.999	uncertain (median slope zero)	Yes	No	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
DRP	61	0.183	-0.25	-5.009	-0.990 to 0.000	decreasing	No	No	No	Yes	0.868	Likely	Insufficient data to infer trend direction
Turb	66	0.347	-0.1	-3.83	-0.250 to 0.075	decreasing	Yes	No	No	Yes	0.827	Likely	Insufficient data to infer trend direction
TP_COR	66	**0.004**	-2.449	-6.71	-4.249 to -1.003	decreasing	Yes	Yes	Yes	Yes	0.998	Virtually certain	decreasing
TN_COR	64	**0.003**	-25.095	-5.429	-38.795 to -14.286	decreasing	Yes	Yes	Yes	Yes	0.999	Virtually certain	decreasing
DRP_COR	61	0.183	-0.246	-5.018	-0.973 to 0.000	decreasing	No	No	No	Yes	0.868	Likely	Insufficient data to infer trend direction

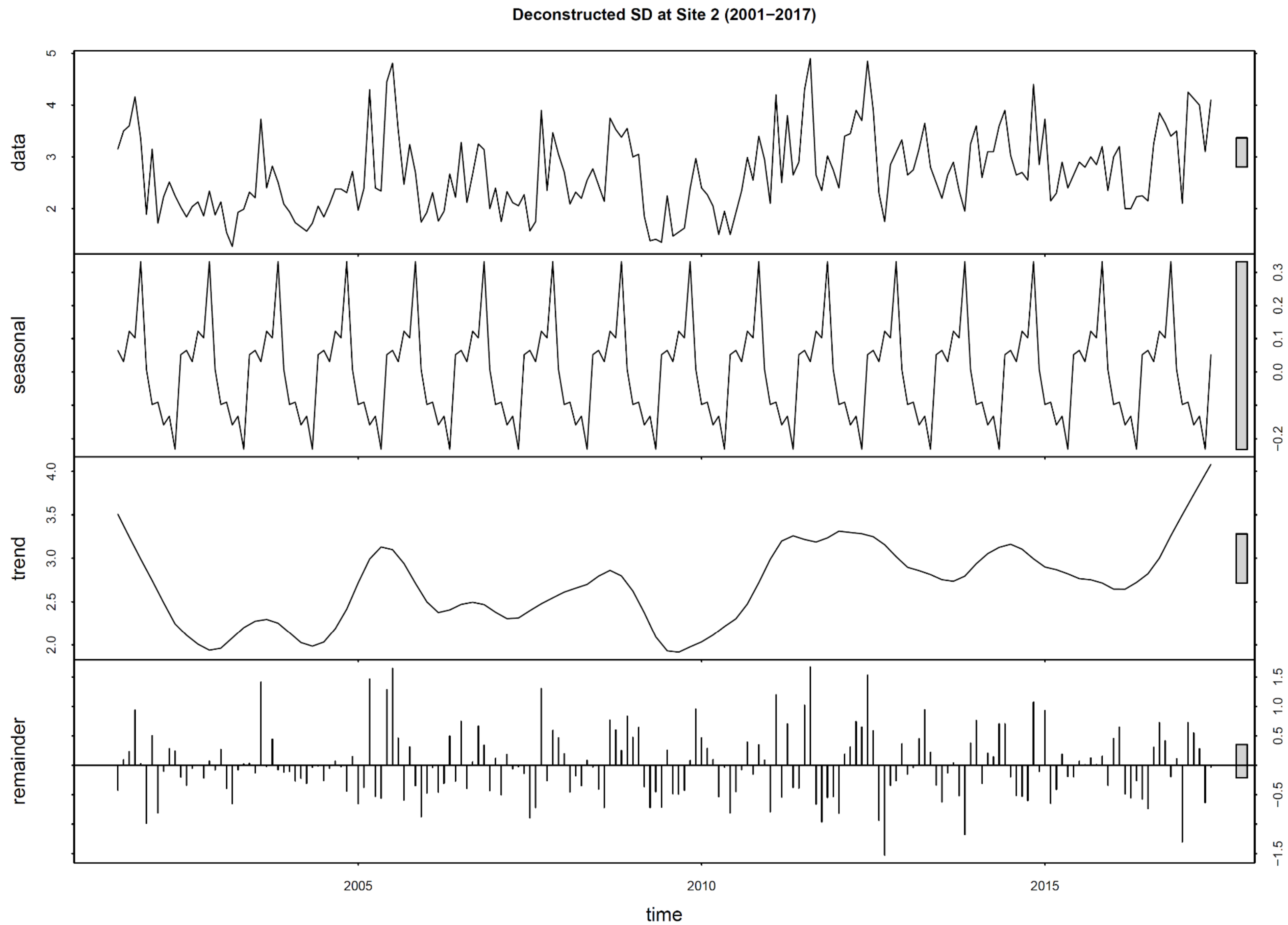
Depth							Summary information			Surface			2009-2017
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
SD	89	0.167	0.081	2.843	-0.012 to 0.120	increasing	Yes	Yes	No	Yes	0.934	Very likely	Insufficient data to infer trend direction
TP	89	0.661	0	0	-0.667 to 0.497	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
TN	87	0.54	-1.666	-0.55	-5.436 to 1.888	decreasing	Yes	Yes	No	No	0.688	Likely	Insufficient data to infer trend direction
Chla	89	0.114	-0.535	-5.349	-1.014 to 0.016	decreasing	Yes	Yes	No	Yes	0.943	Very likely	Insufficient data to infer trend direction
NO3	87	0.232	0.167	2.38	0.000 to 0.946	increasing	No	Yes	No	Yes	0.178	Unlikely	Insufficient data to infer trend direction
NH4	89	0.812	0.184	2.045	-1.113 to 1.003	increasing	Yes	Yes	No	Yes	0.369	About as likely as not	Insufficient data to infer trend direction
DRP	87	0.858	0	0	0.000 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
Turb	89	0.338	-0.025	-1.251	-0.064 to 0.019	decreasing	Yes	Yes	No	Yes	0.843	Likely	Insufficient data to infer trend direction
TP_COR	89	0.661	0	0	-0.553 to 0.412	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
TN_COR	87	0.54	-1.666	-0.471	-5.436 to 1.888	decreasing	Yes	Yes	No	No	0.688	Likely	Insufficient data to infer trend direction
DRP_COR	87	0.858	0	0	0.000 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction

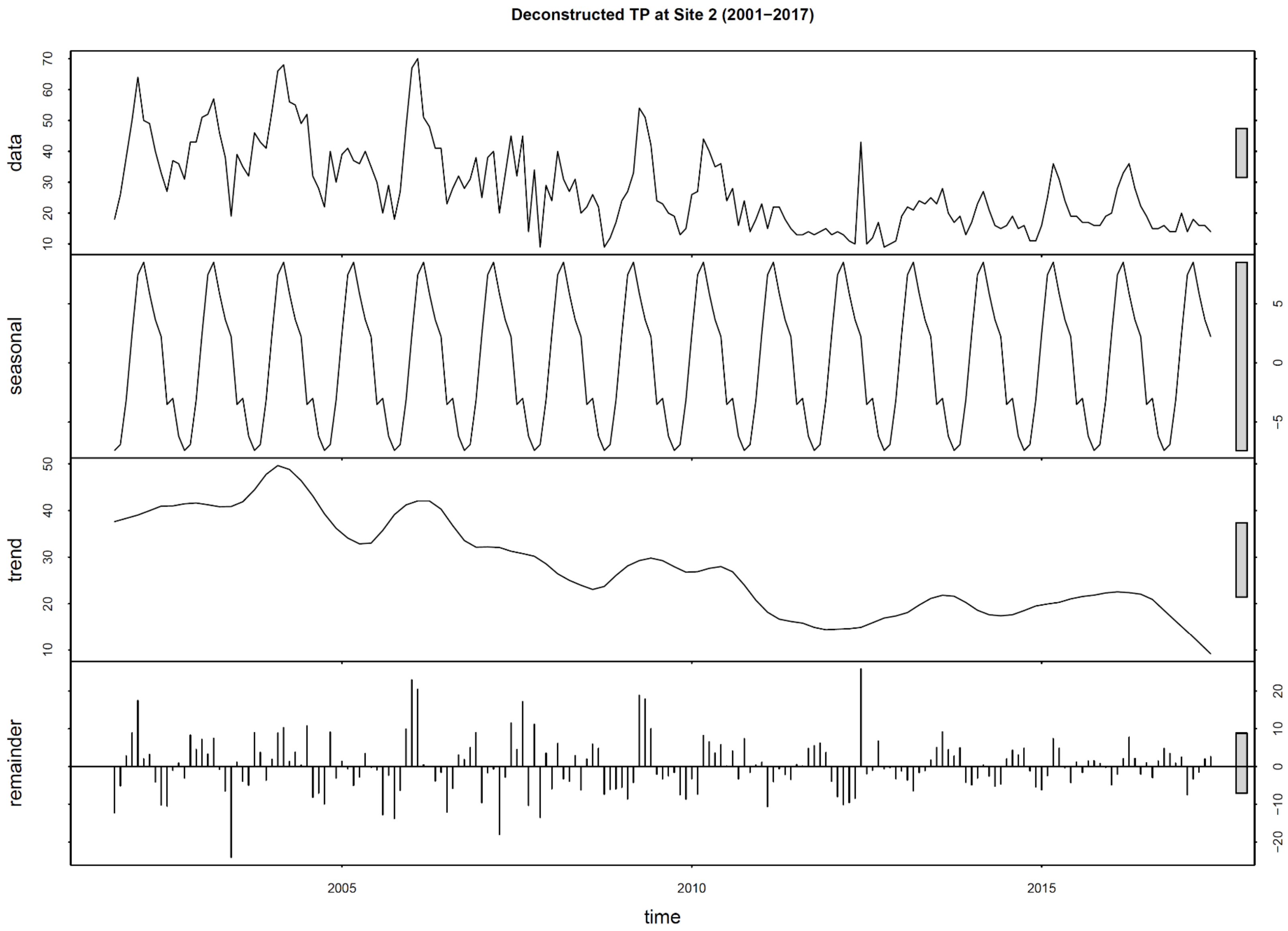
Depth							Summary information			Mid			2009-2017
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
TP	91	0.785	0	0	-0.666 to 0.413	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
TN	91	**0.005**	-5.297	-1.796	-10.689 to -2.482	decreasing	Yes	Yes	Yes	Yes	0.996	Virtually certain	decreasing
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	90	0.305	-0.201	-3.089	-1.080 to 0.000	decreasing	No	Yes	No	Yes	0.83	Likely	Insufficient data to infer trend direction
NH4	91	0.192	-1.004	-5.907	-3.331 to 0.315	decreasing	Yes	Yes	No	Yes	0.9	Very likely	Insufficient data to infer trend direction
DRP	91	0.446	0	0	-0.112 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
Turb	91	**0.037**	-0.055	-2.909	-0.133 to 0.000	decreasing	Yes	Yes	No	Yes	0.98	Extremely likely	Insufficient data to infer trend direction
TP_COR	91	0.785	0	0	-0.552 to 0.342	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
TN_COR	91	**0.005**	-5.297	-1.532	-10.689 to -2.482	decreasing	Yes	Yes	Yes	Yes	0.996	Virtually certain	decreasing
DRP_COR	91	0.446	0	0	-0.112 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction

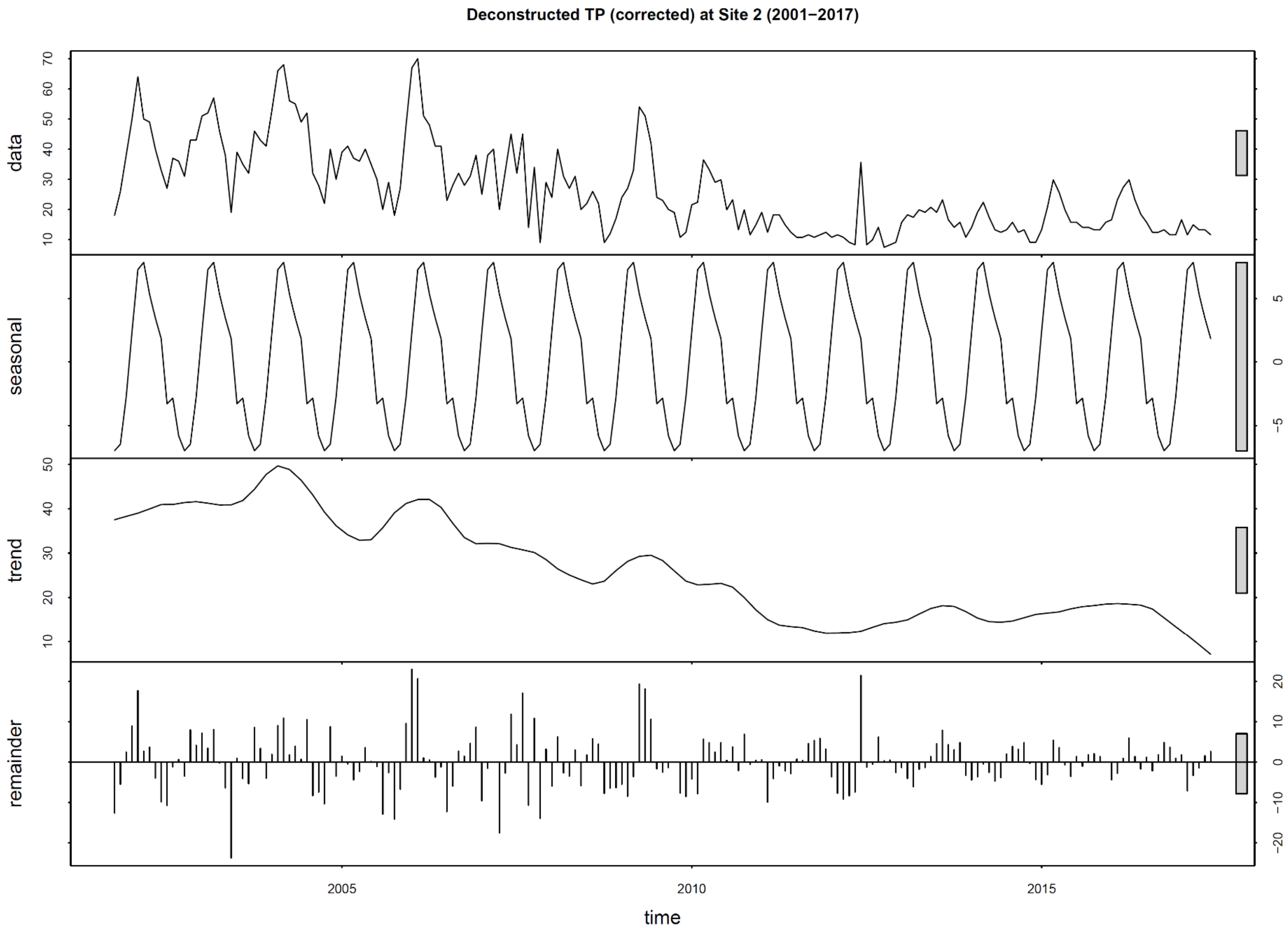
Depth							Summary information			Bottom			2009-2017
Site 5	**Sample size**	**P**	**Median Sen slope (annual)**	**Percent annual change**	**90% confidence limits for slope**	**Trend direction**	**Sufficient DL (>=15%)**	**Sufficient coverage (>80% annual)**	**Conclusive**	**Important**	**P(lwp)**	**Confidence of improvement**	**LAWA Direction**
TP	85	0.525	0.225	1.25	-0.349 to 0.603	increasing	Yes	Yes	No	Yes	0.208	Unlikely	Insufficient data to infer trend direction
TN	85	0.293	-3.794	-1.224	-10.731 to 1.668	decreasing	Yes	Yes	No	Yes	0.866	Likely	Insufficient data to infer trend direction
Chla	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NO3	84	0.695	0	0	-1.082 to 0.313	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
NH4	85	0.332	-0.992	-4.509	-3.985 to 0.693	decreasing	Yes	Yes	No	Yes	0.834	Likely	Insufficient data to infer trend direction
DRP	85	1	0	0	0.000 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
Turb	85	0.932	0	0	-0.100 to 0.050	uncertain (median slope zero)	Yes	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction
TP_COR	85	0.525	0.187	1.25	-0.290 to 0.500	increasing	Yes	Yes	No	Yes	0.208	Unlikely	Insufficient data to infer trend direction
TN_COR	85	0.293	-3.794	-1.052	-10.731 to 1.668	decreasing	Yes	Yes	No	Yes	0.866	Likely	Insufficient data to infer trend direction
DRP_COR	85	1	0	0	0.000 to 0.000	uncertain (median slope zero)	No	Yes	No	No	0.5	About as likely as not	Insufficient data to infer trend direction

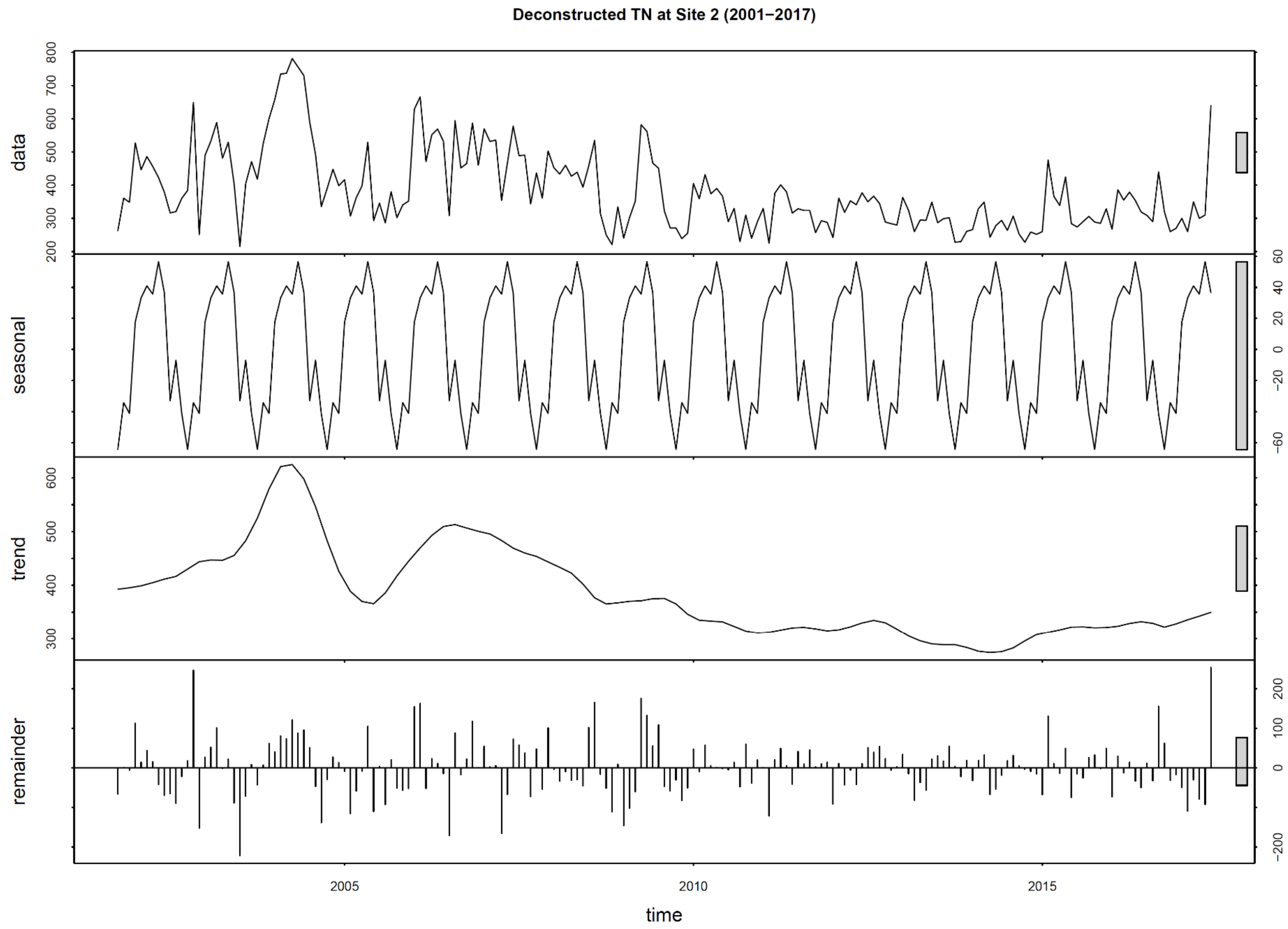
Appendix 2 – Long-term (2001-2017) time-series deconstruction for Sites 2 and 5

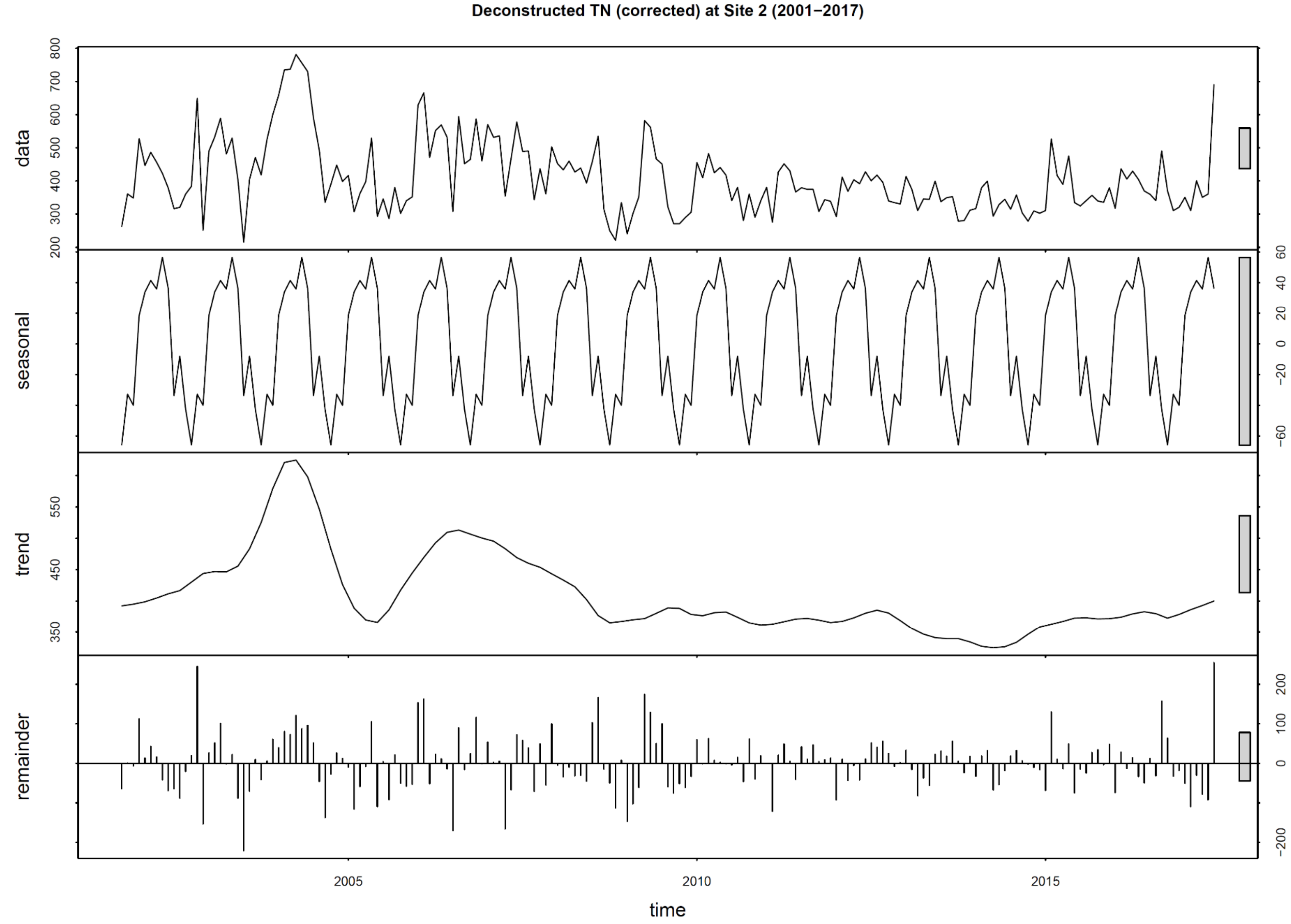


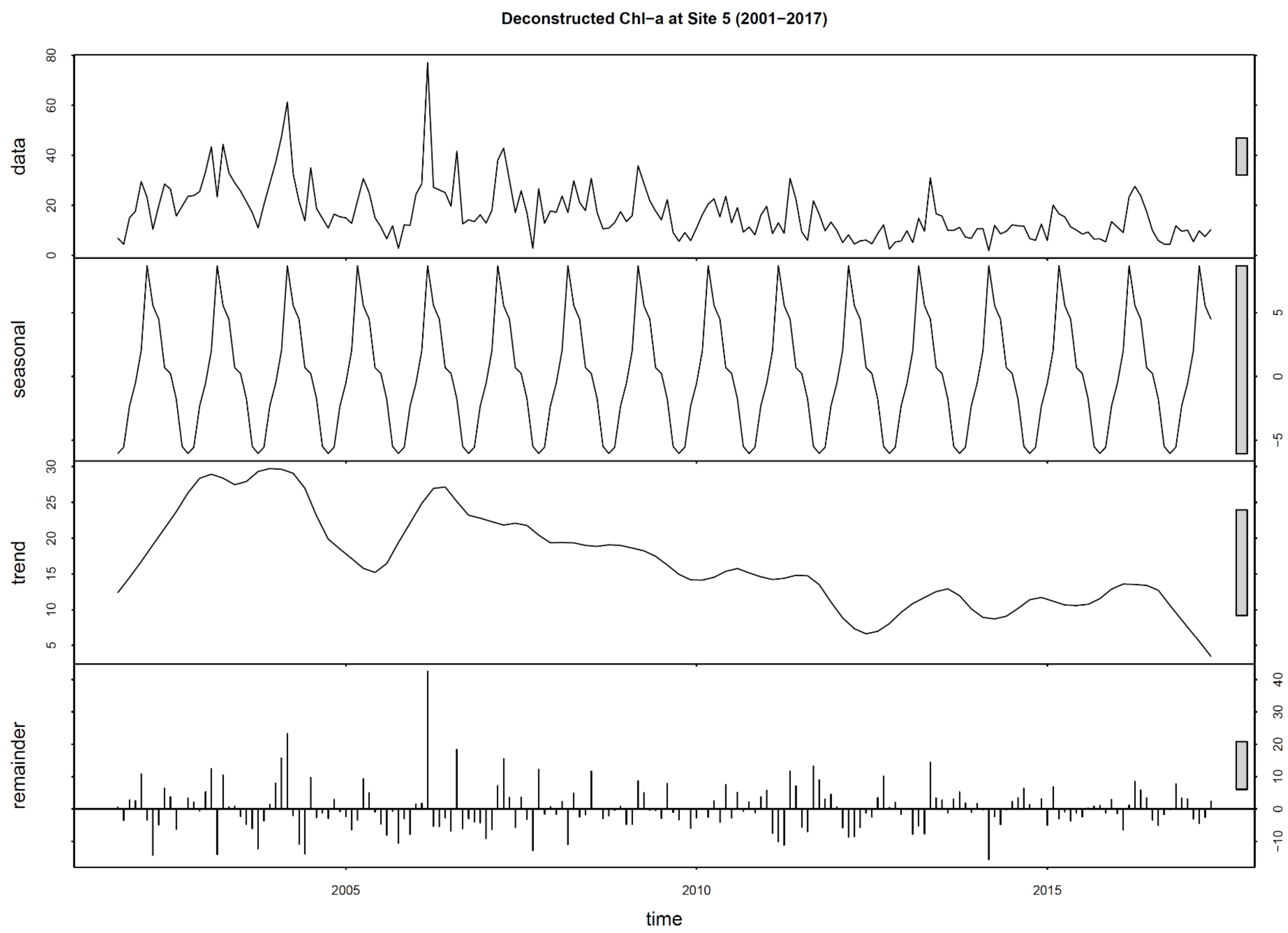

Deconstructed SD at Site 2 (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015
2
3
4
5
−0.2
0.0
0.1
0.2
0.3
2.0
2.5
3.0
3.5
4.0
−1.5
−0.5
0.5
1.0
1.5

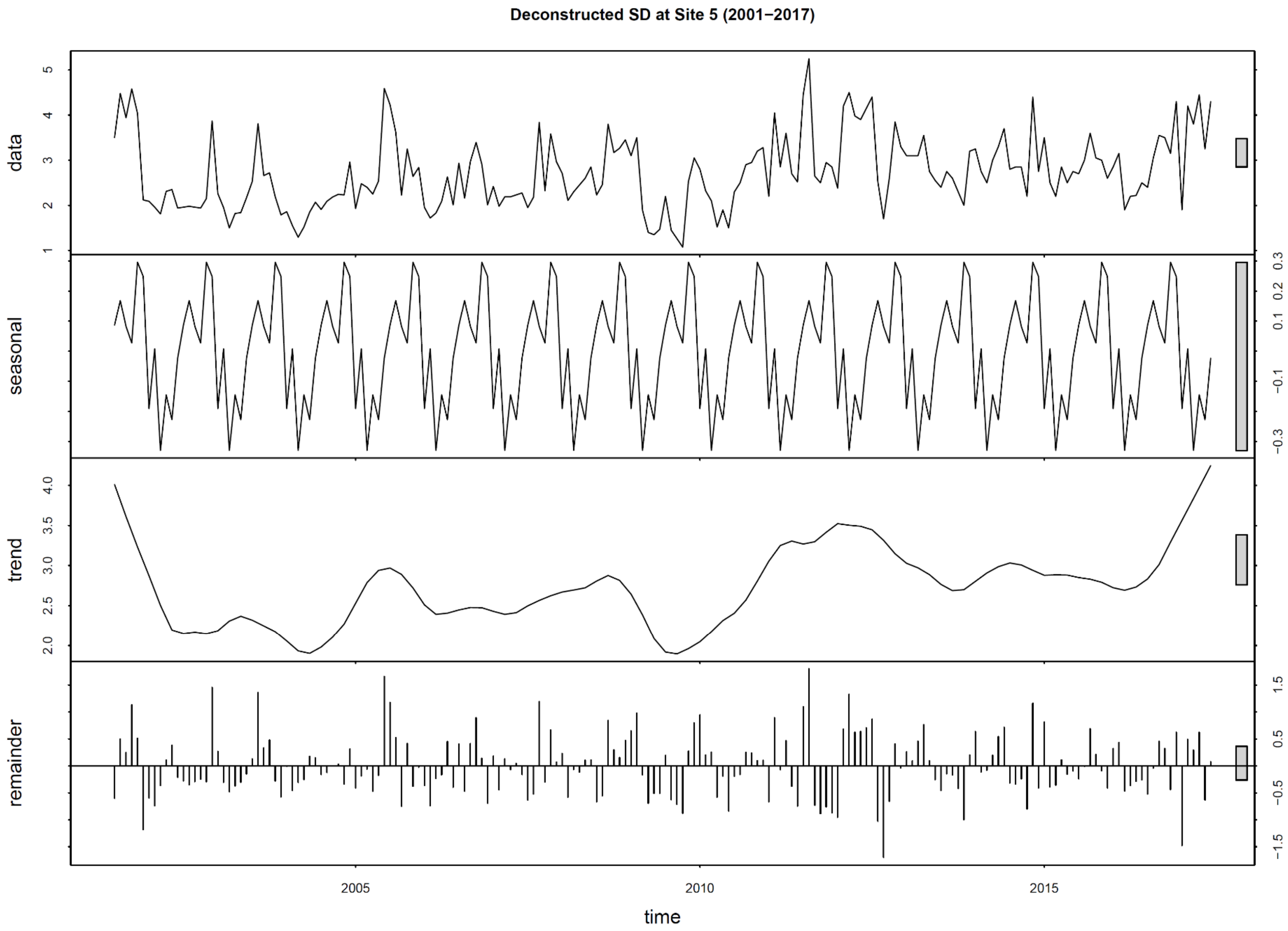

Deconstructed TP at Site 2 (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015
10
20
30
40
50
60
70
5
0
−5
−10
−20

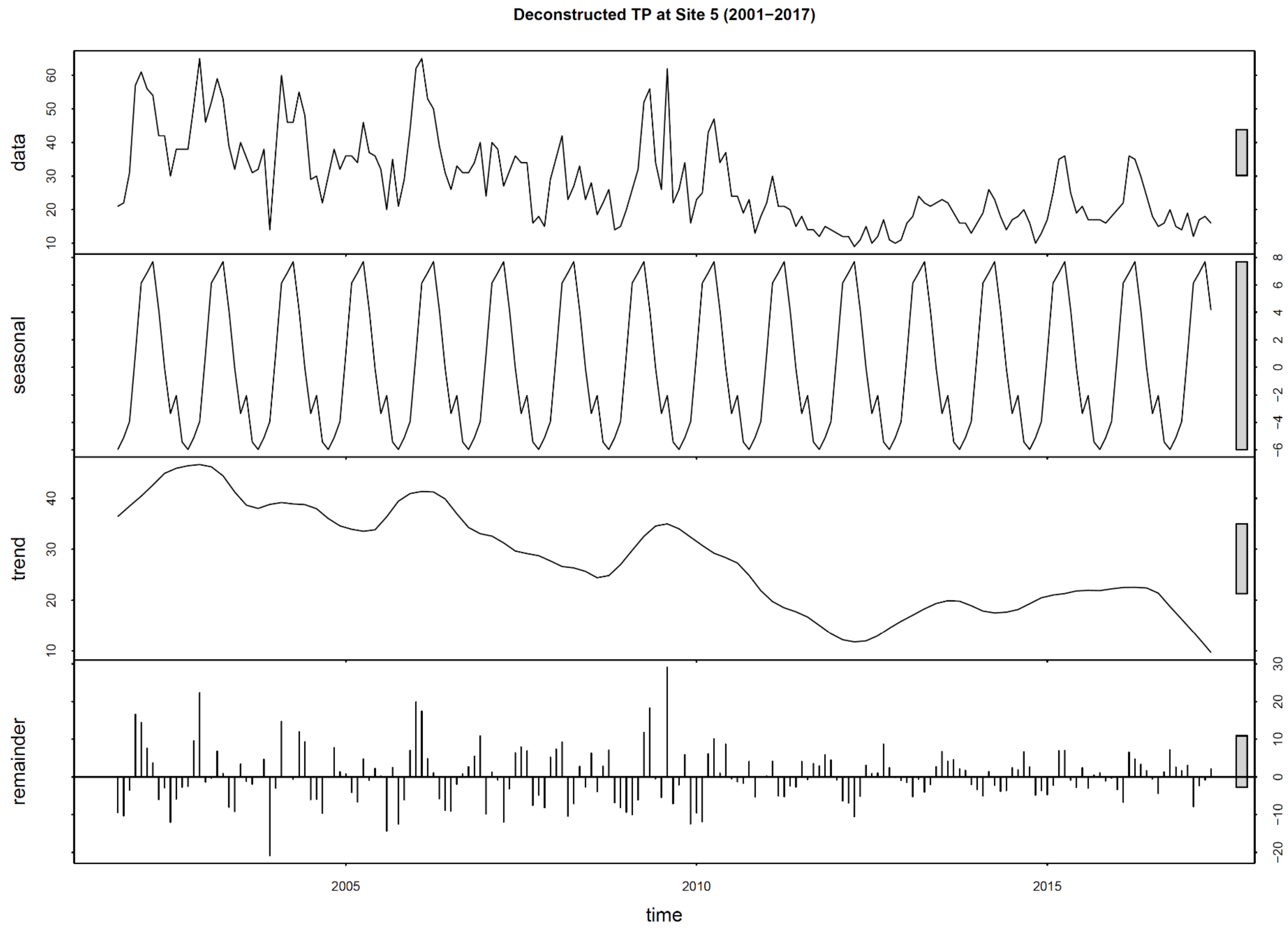

Deconstructed TP (corrected) at Site 2 (2001–2017)
data
seasonal
trend
remainder
time
10
20
30
40
50
60
70
5
0
-5
10
20
30
40
50
20
10
0
-10
-20
2005
2010
2015

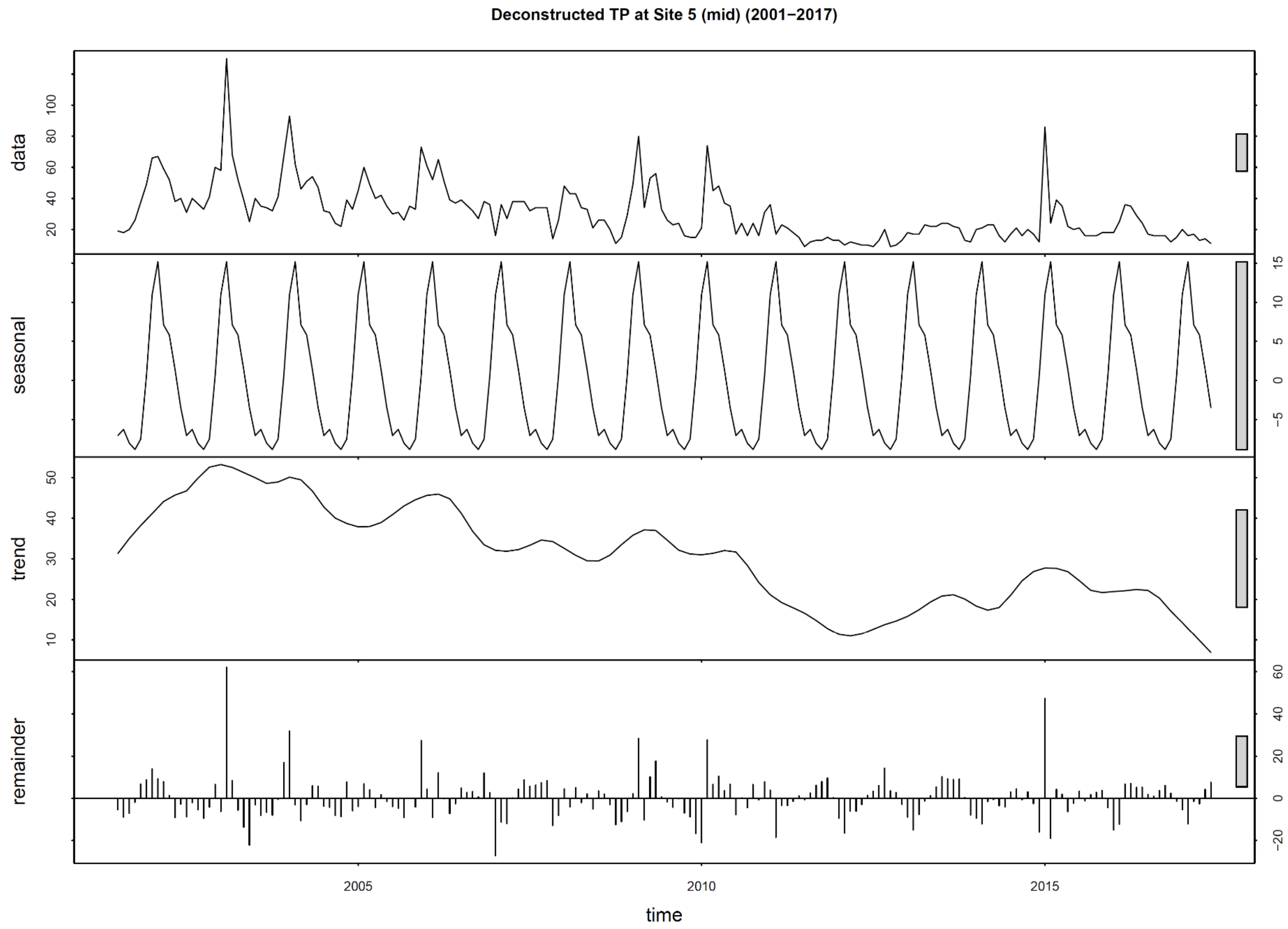

Deconstructed TN at Site 2 (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015

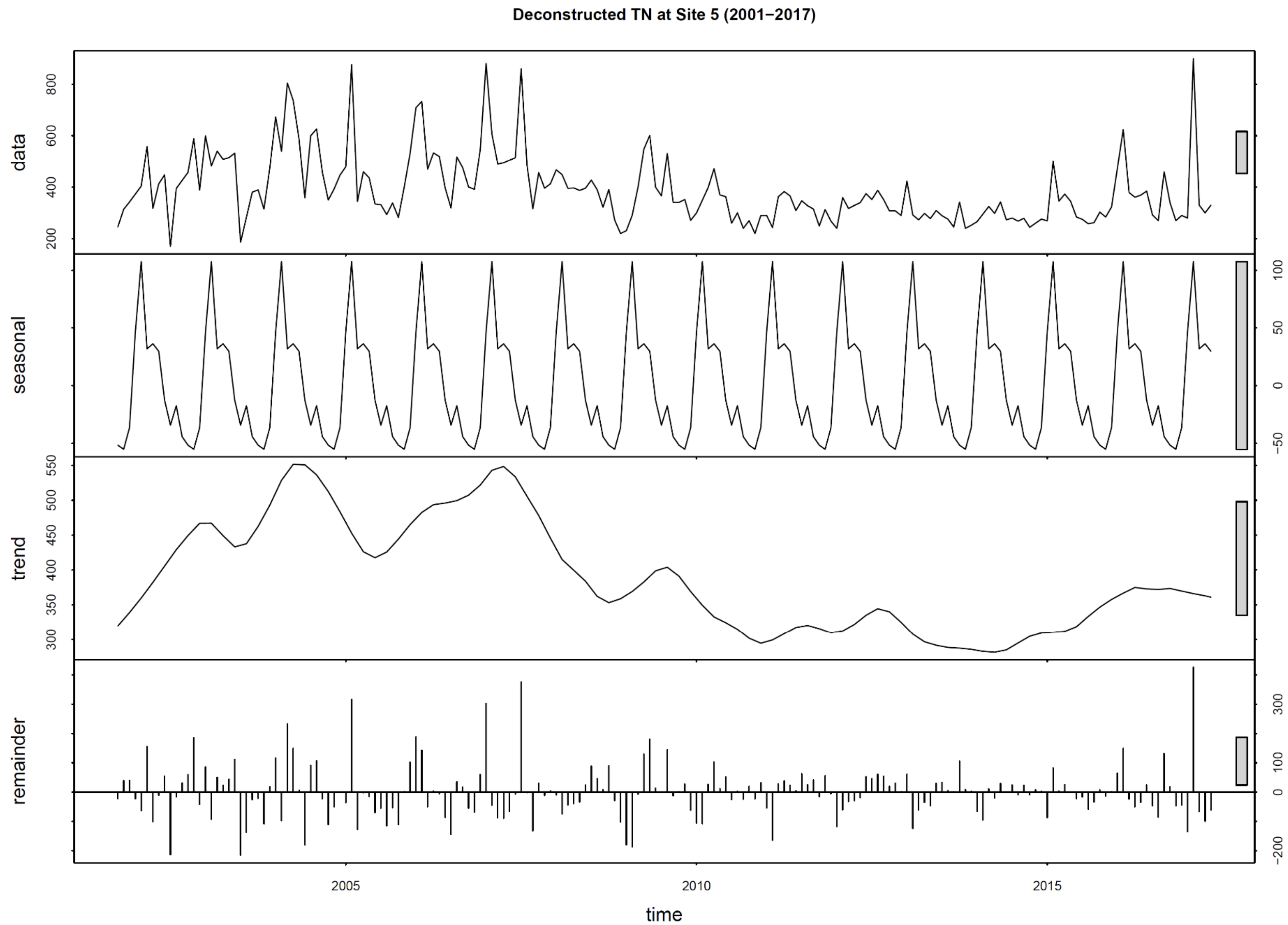

Deconstructed TN (corrected) at Site 2 (2001–2017)
data
seasonal
trend
remainder
time
200
300
400
500
600
700
800
60
40
20
0
−20
−60
350
450
550
200
100
0
−200
2005
2010
2015


Deconstructed Chl−a at Site 5 (2001−2017)
data
seasonal
trend
remainder
time
0
20
40
60
80
5
0
−5
5
10
15
20
25
30
40
30
20
10
0
−10
2005
2010
2015

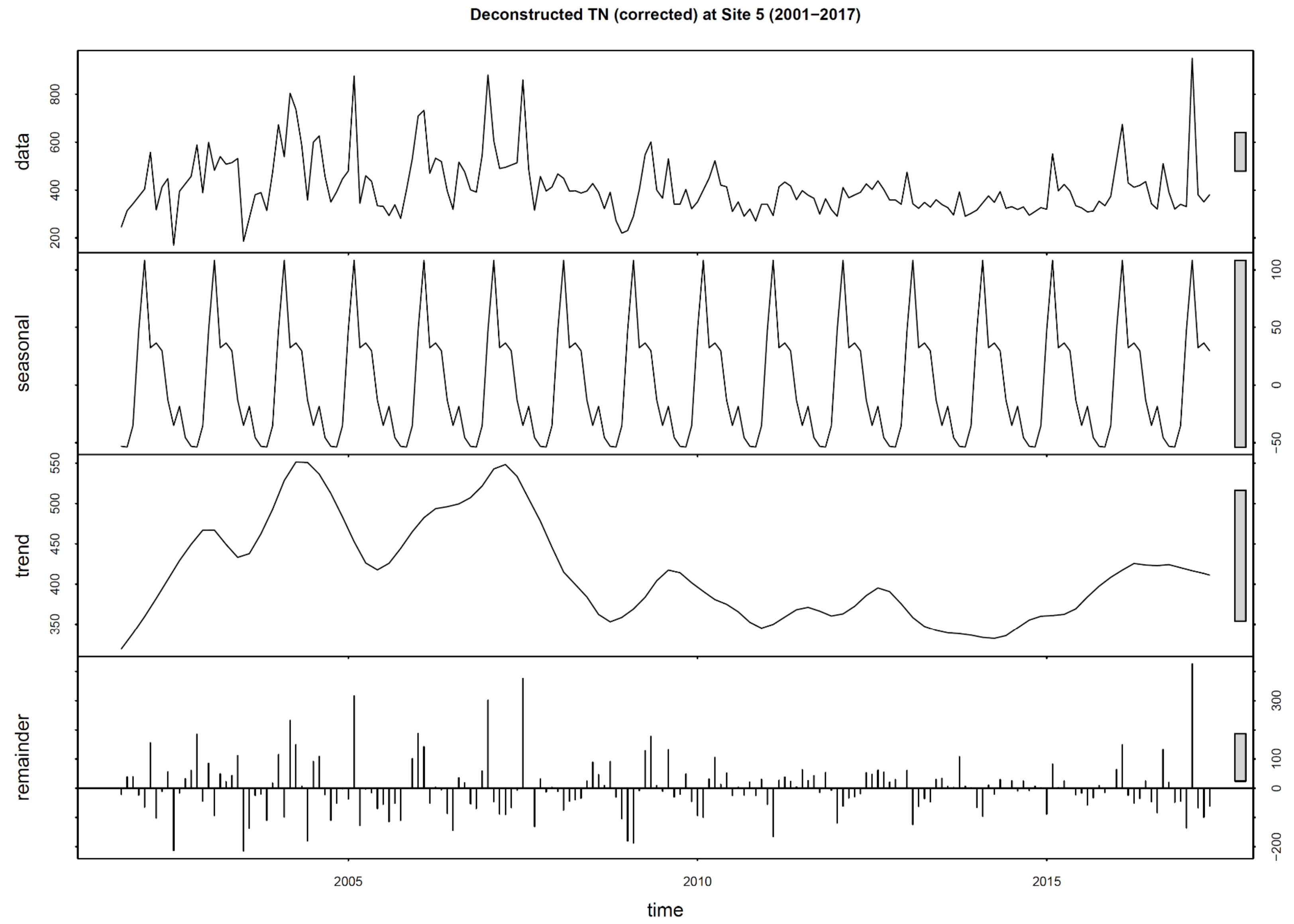

Deconstructed SD at Site 5 (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015
1
2
3
4
5
0.3
0.2
0.1
−0.1
−0.3
2.0
2.5
3.0
3.5
4.0
1.5
0.5
−0.5
−1.5

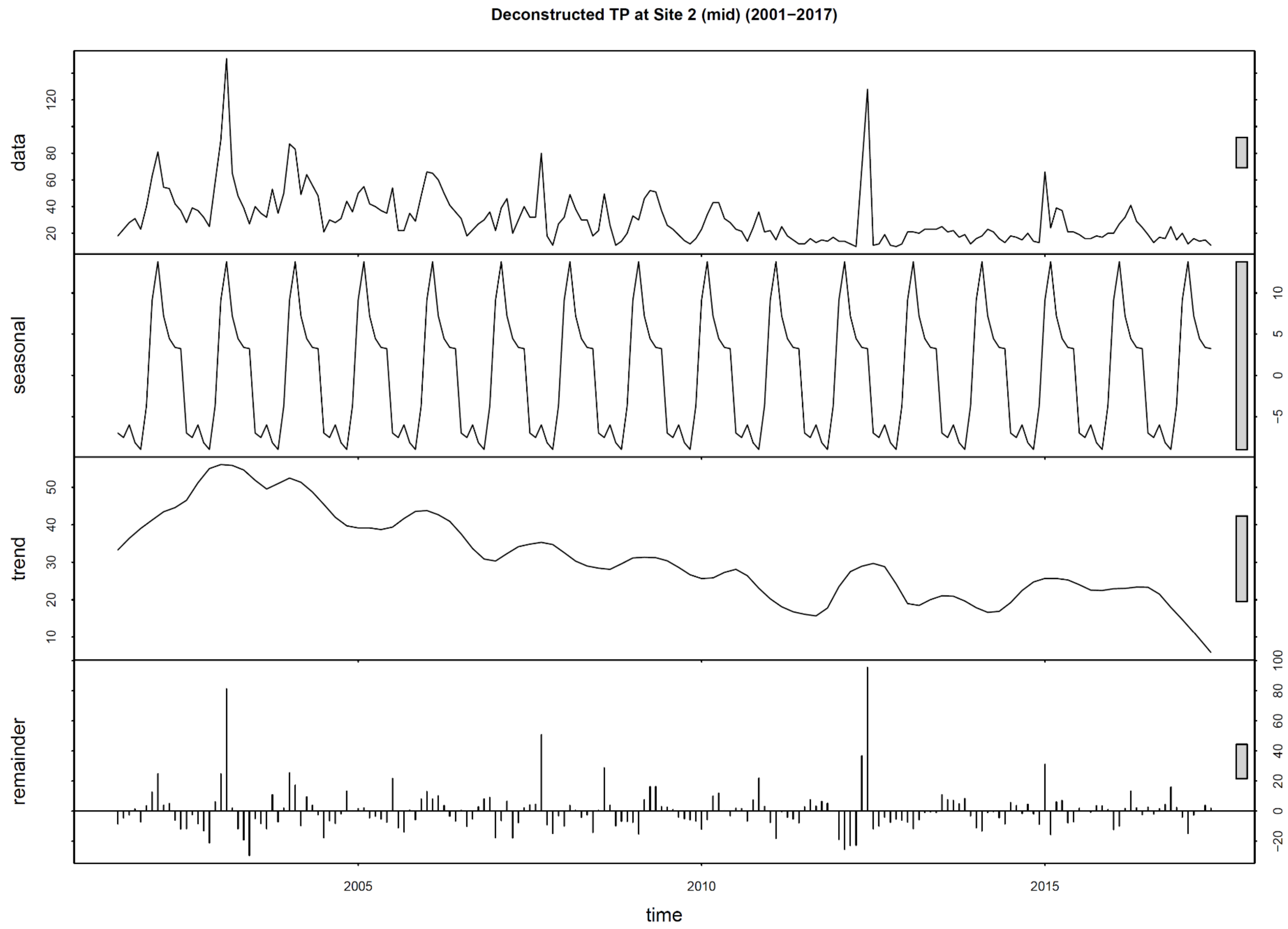

Deconstructed TP at Site 5 (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015

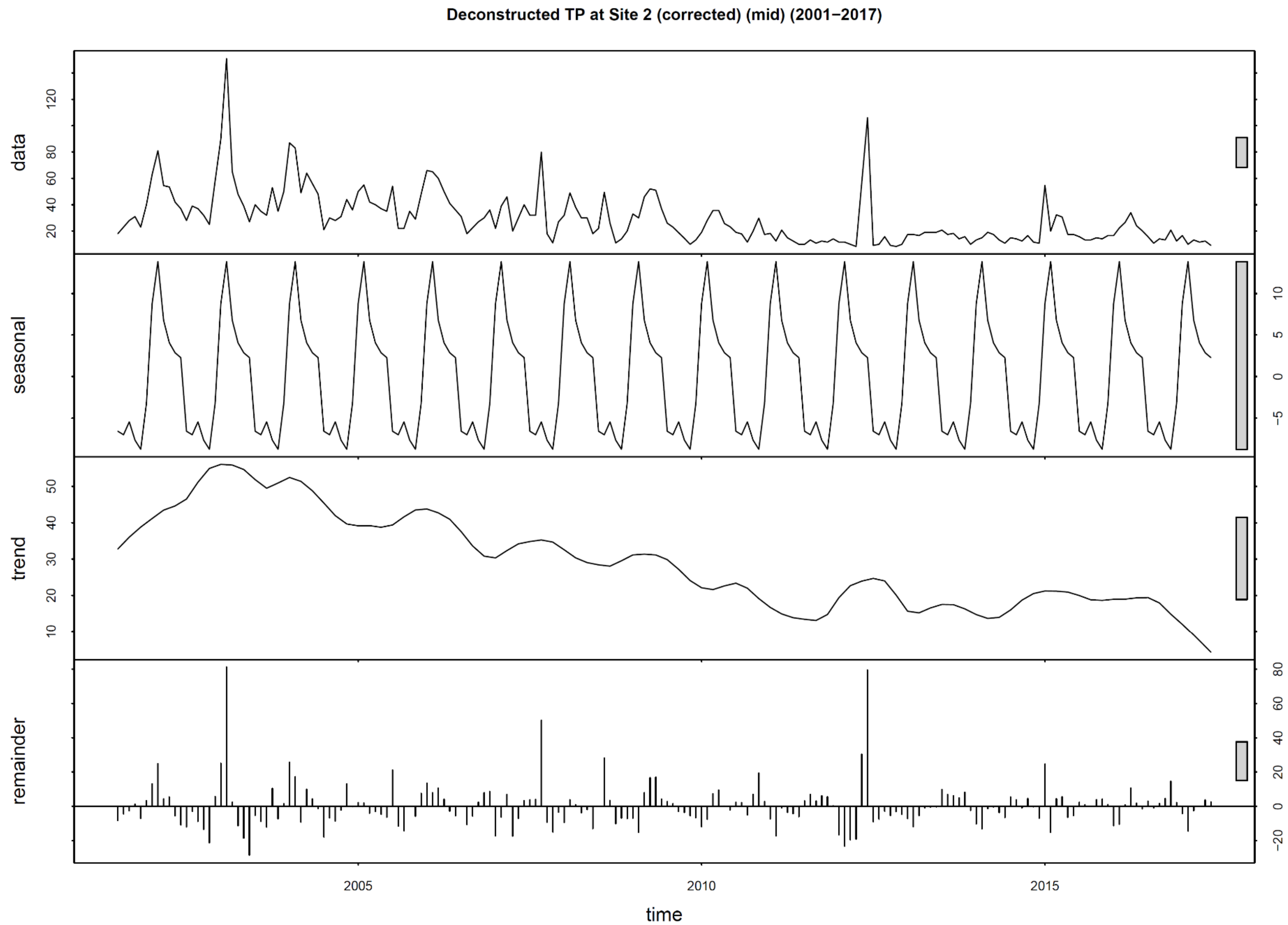

Deconstructed TP at Site 5 (mid) (2001–2017)
data
seasonal
trend
remainder
time
2005
2010
2015



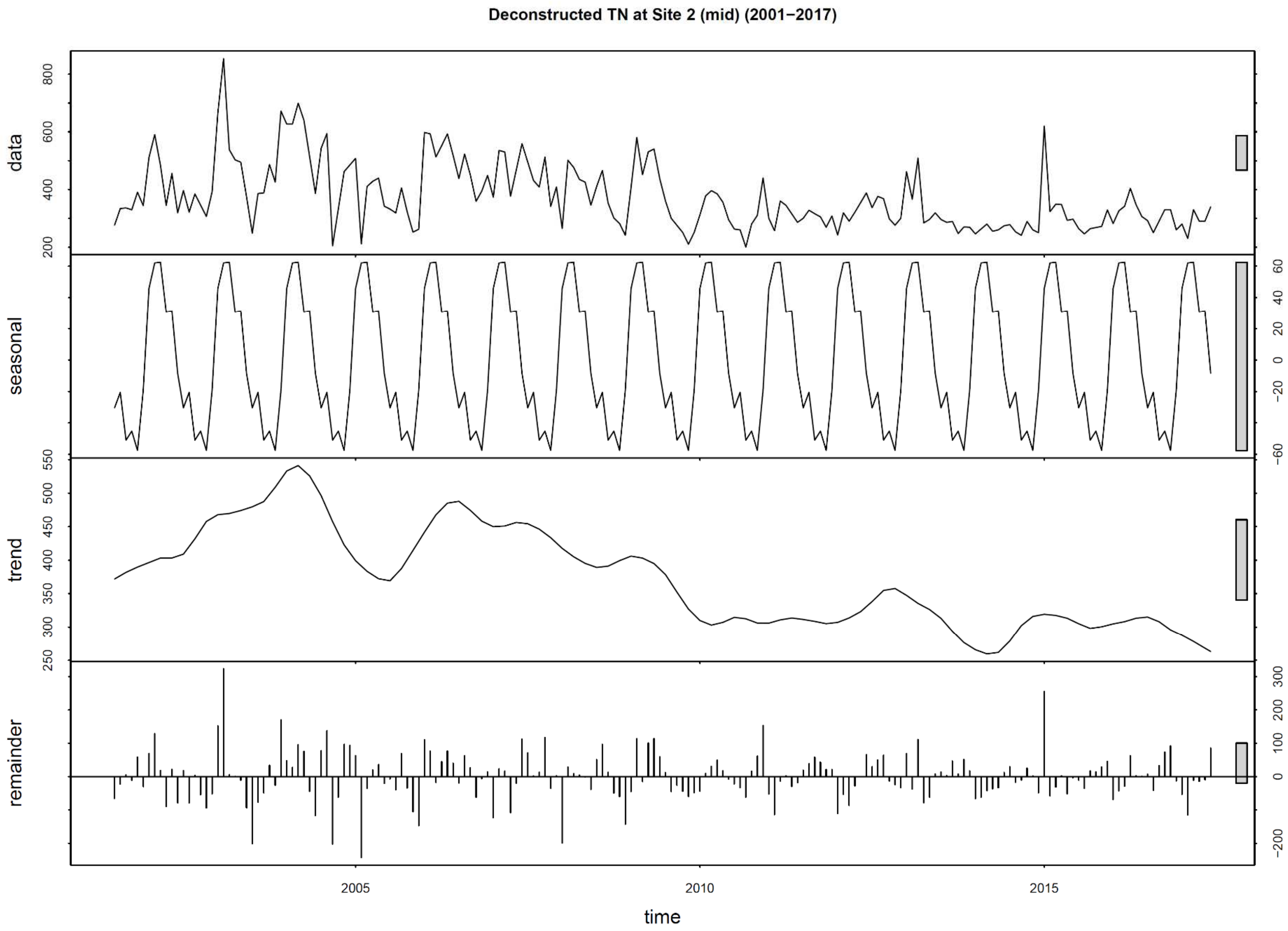
Deconstructed TN at Site 5 (2001−2017)
data
200
400
600
800
seasonal
100
50
0
−50
trend
300
350
400
450
500
550
remainder
300
100
0
−200
2005
2010
2015
time

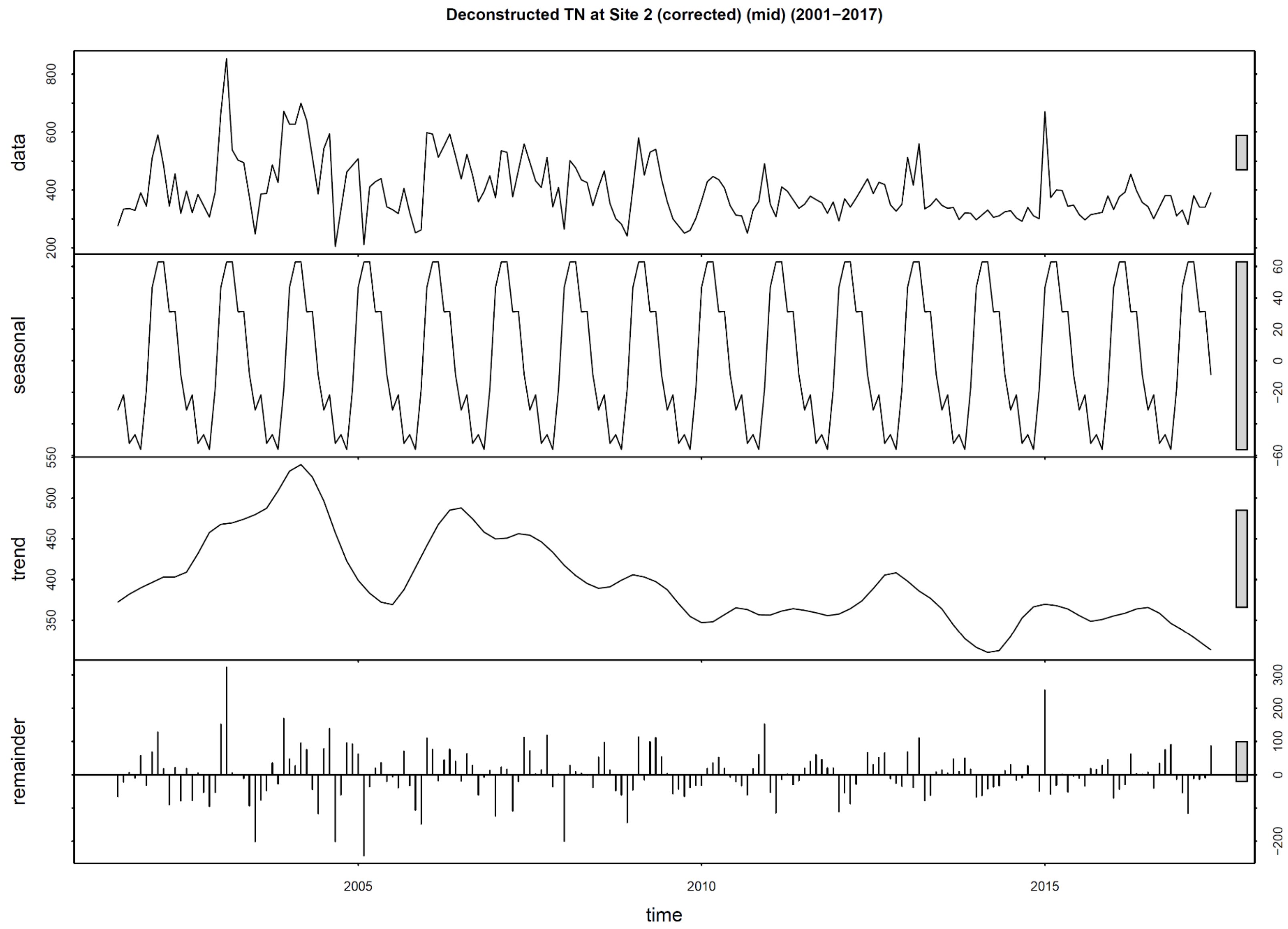

Deconstructed TN (corrected) at Site 5 (2001−2017)
data
seasonal
trend
remainder
time
200
400
600
800
−50
0
50
100
350
400
450
500
550
−200
0
100
300
2005
2010
2015

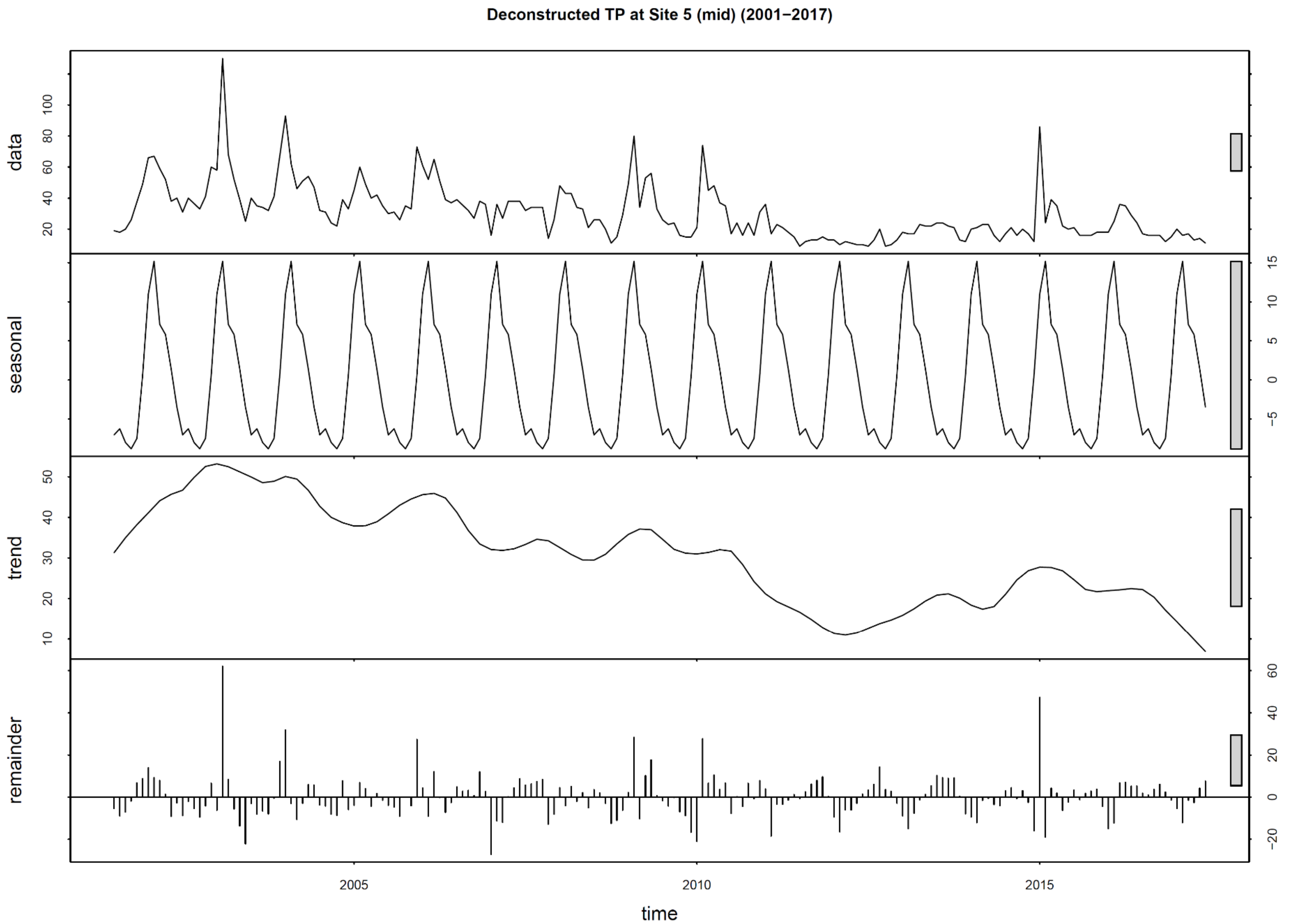

Deconstructed TP at Site 2 (mid) (2001–2017)
data
120
80
60
40
20
seasonal
10
5
0
−5
trend
50
40
30
20
10
remainder
100
80
60
40
20
0
−20
2005
2010
2015
time

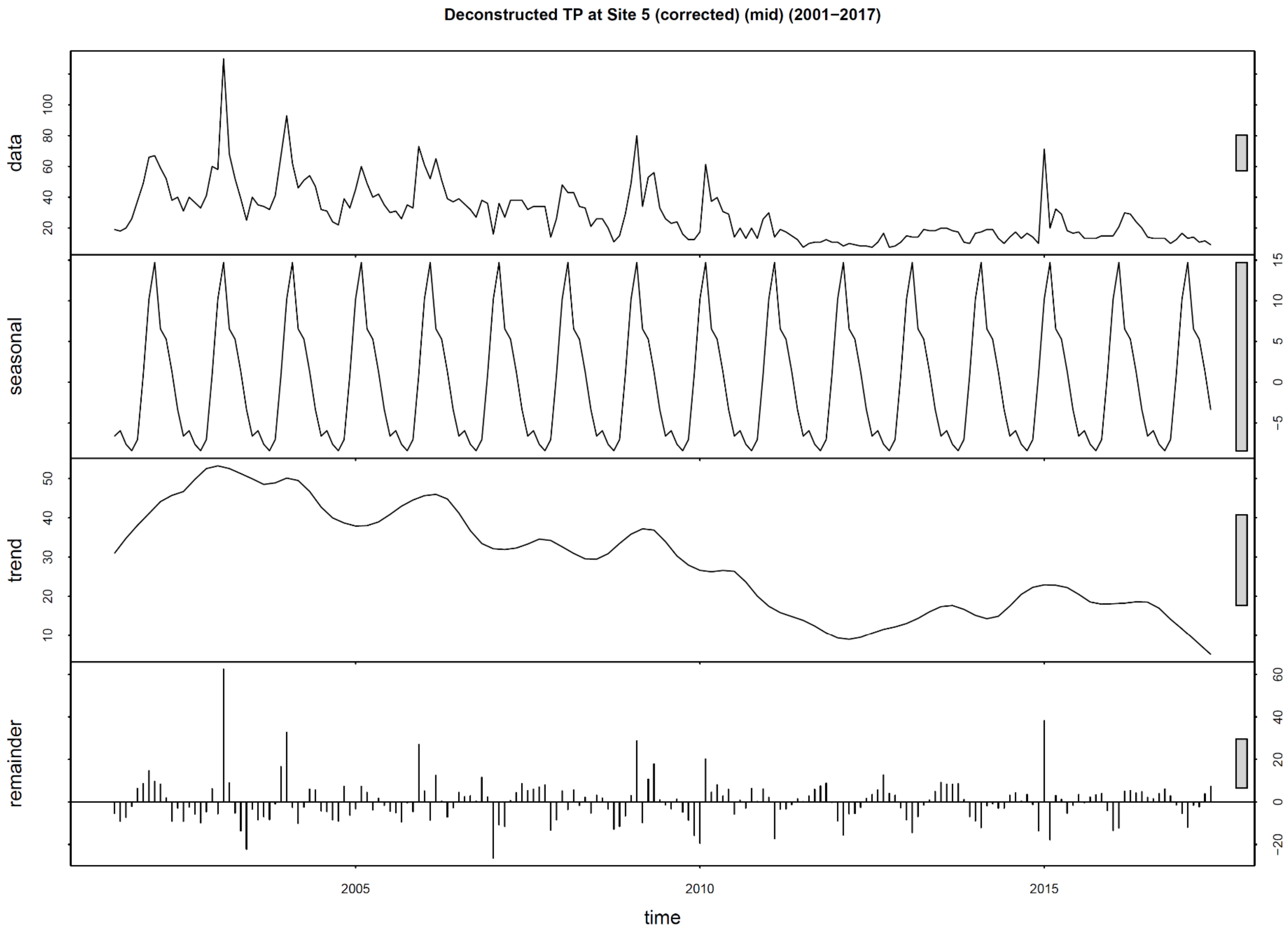


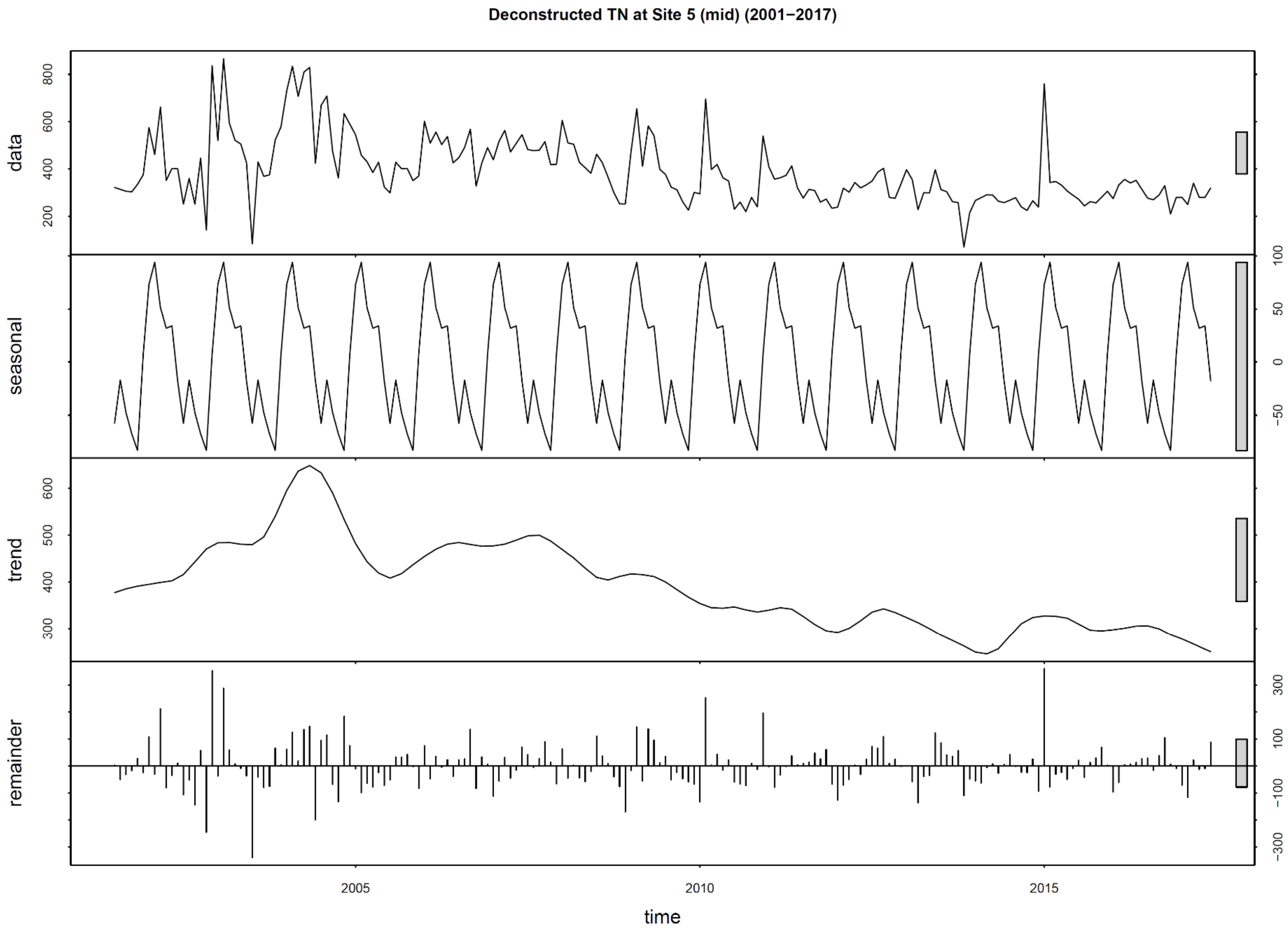
Deconstructed TP at Site 2 (corrected) (mid) (2001–2017)
data
seasonal
trend
remainder
time
2005
2010
2015

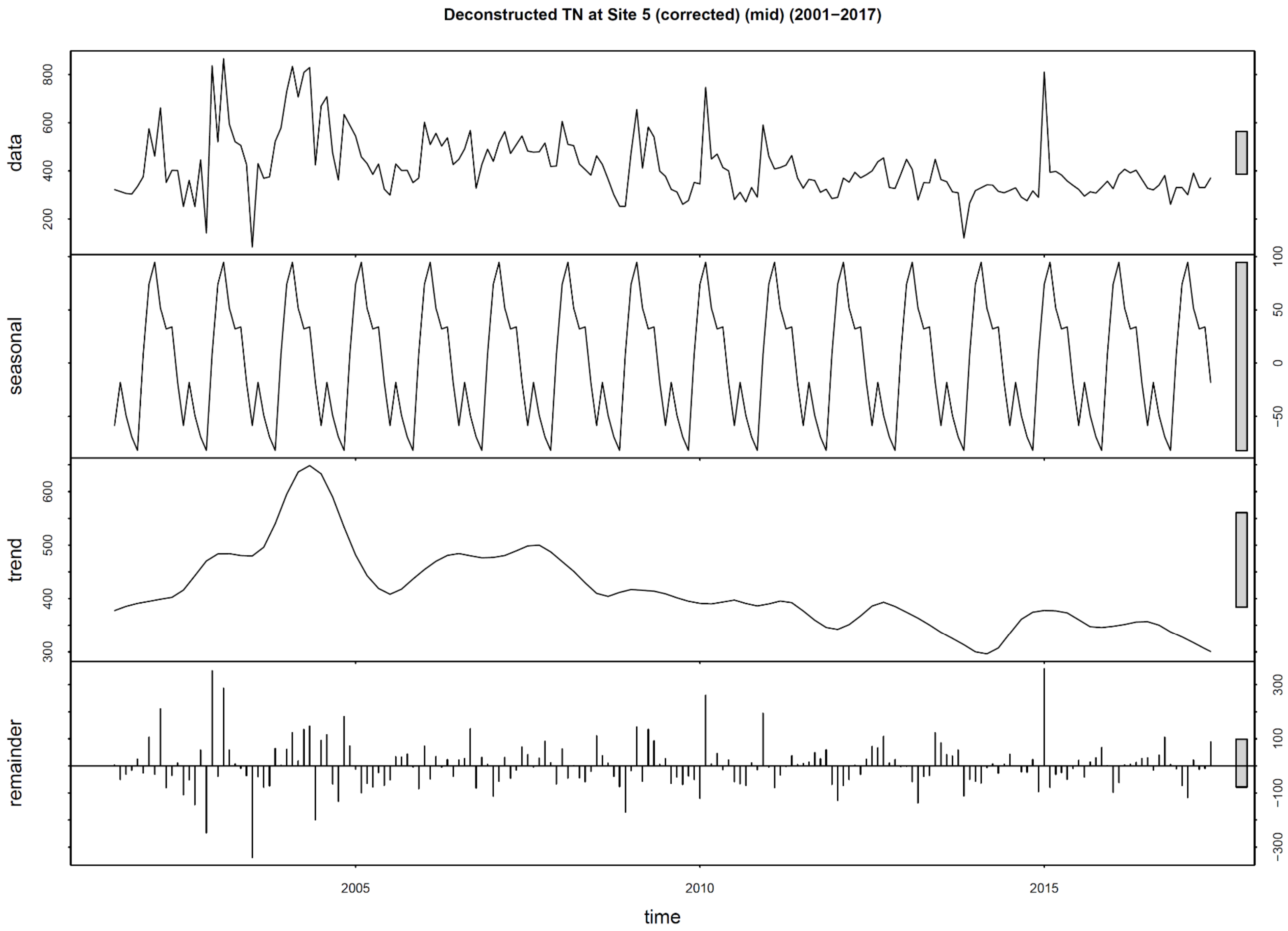

Deconstructed TN at Site 2 (mid) (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015


Deconstructed TN at Site 2 (corrected) (mid) (2001−2017)
data
seasonal
trend
remainder
time
200
400
600
800
−60
−20
0
20
40
60
350
400
450
500
550
−200
0
100
200
300
2005
2010
2015


Deconstructed TP at Site 5 (mid) (2001–2017)
data
seasonal
trend
remainder
20
40
60
80
100
15
10
5
0
−5
10
20
30
40
50
60
40
20
0
−20
2005
2010
2015
time


Deconstructed TP at Site 5 (corrected) (mid) (2001–2017)
data
seasonal
trend
remainder
time
2005
2010
2015
20
40
60
80
100
15
10
5
0
−5
10
20
30
40
50
60
40
20
0
−20


Deconstructed TN at Site 5 (mid) (2001−2017)
data
seasonal
trend
remainder
200
400
600
800
100
50
0
−50
300
400
500
600
300
100
−100
−300
2005
2010
2015
time


Deconstructed TN at Site 5 (corrected) (mid) (2001−2017)
data
seasonal
trend
remainder
time
2005
2010
2015
200
400
600
800
100
50
0
−50
300
400
500
600
300
100
−100
−300