TETRA TECH

Appendix 2.6B Summary of Censoring

To:	Jennifer Keisman
Cc:	Rebecca Murphy, Elgin Perry
From:	Jon Harcum, Alex DeWire
Date:	January 27, 2016
Subject:	CBP GAM Technical Support: Summary of Censoring (Nutrient Species Data)

This communication fulfills the deliverable requirement to provide an evaluation of the 1985-2015 data record at each long-term tidal monitoring station to compute the frequency of censored data values for surface and bottom for seven nutrient species: NH4F, NO23F, PN, PO4F, PP, TDN, and TDP. An excel spreadsheet, CBP_WQ_dataCensoringSummary_species.xlsx, contains the detailed results. This is a follow-up to the January 14, 2016 evaluation of data for TN, TP, chlorophyll *a*, DO, and Secchi depth in the "January 14 Summary of Censoring.docx" memo and corresponding Excel table "CBP_WQ_dataCensoringSummary.xlsx". Where appropriate, some statistics and figures are carried from the previous memo.

Data Preparation. Two data sets were provided by Rebecca Murphy. These data sets included water quality observations for NH4F, NO23F, PN, PO4F, PP, TDN, and TDP and were separated into the different files by station type (main bay vs. tributary). These data were combined into one dataset (n=843,409). Observations with a qualifier of "<" were treated as censored data. Twenty records with a qualifier of ">" for NO23F and TDN were treated as uncensored for this analysis; 6,443 records with a qualifier of "G" for all parameters except TDN were treated as uncensored for this analysis. Non-primary trend sites were excluded (i.e., sites not included in CBPstats::tidalStations). Observations of the format "*B" in the method column were excluded. These values represent censored observations where ½ the detection limit was substituted. Observations of the format "*A" in the method column were maintained as they indicate results where the detection limit was substituted. The data set was further reduced to one observation per station, parameter, layer, and date/time. When there is more than one observation for a given station, parameter, layer, and date/time, then the record that is censored (i.e., "<") was maintained. If more than one uncensored sample or more than one censored observation was present for a given station, parameter, layer, and date/time sample, then the results were averaged. The final number of observations used for further analysis was 772,067.

Tetra Tech Inc.
Tel 864-650-5815 tetratech.com

Results. Overall, the highest levels of censoring occurs in the 1985-1989 time period (see Table 1). For the nutrient species, censoring for NH4F, NO23F, and PO4F have the highest levels of censoring with 19.7 percent, 27.9 percent, and 27.4 percent respectively during the 1985-1989 time span. The percentage censoring is reduced during more recent time periods. (Table 1 also includes the statistics for the core variables from the January 14, 2016 memorandum.)

Table 1. Overall Number of Observations and Percent Censoring by Parameter and Date Range.

PARM	PARAMETER NAME[1]	Number of Observations				% Censored Data			
		1985-1989	*1990-1994*	*1995-1998*	*1999-2015*	*1985-1989*	*1990-1994*	*1995-1998*	*1999-2015*
NH4F	AMMONIUM NITROGEN AS N (FILTERED)	19,394	23,103	15,458	58,332	19.7	19.3	10.5	8.4
NO23F	NITRITE+NITRATE NITROGEN AS N (FILTERED)	19,441	22,653	15,653	58,234	27.9	19.6	6.3	7.43
PN	PARTICULATE NITROGEN	5,266	11,830	11,205	58,480	3.6	0.9	0.4	0.69
PO4F	ORTHOPHOSPHATE PHOSPHORUS AS P (FILTERED)	19,702	23,078	15,677	58,375	27.4	14.9	8.1	5.4
PP	PARTICULATE PHOSPHORUS	19,084	21,941	15,490	58,545	11.9	6.3	3.9	0.0
TDN	TOTAL DISSOLVED NITROGEN	13,298	17,935	15,138	58,372	12.9	5.0	1.3	0.0
TDP	TOTAL DISSOLVED PHOSPHORUS	19,835	22,577	15,598	58,373	14.6	6.7	4.1	0.6
CHLA[2]	ACTIVE CHLOROPHYLL-A	18,683	20,541	12,634	50,611	1.5	1.6	0.6	0.3
DO[2]	DISSOLVED OXYGEN	21,100	23,545	15,810	56,747			0.1	0.02
SECCHI[2]	SECCHI DEPTH	10,759	11,585	7,853	28,518				0.05
TN[2]	TOTAL NITROGEN	19,788	22,552	15,065	54,194	29.3	17.4	1.5	0.9
TP[2]	TOTAL PHOSPHORUS	21,550	23,477	15,688	54,528	3.8	1.7	0.4	0.4

[1] From Water Quality Database: Database Design and Data Dictionary (January 2004).
[2] Data for these parameters did not include data from 2015.

Figure 1 displays cumulative density function of percent censoring at the station/layer level by date range and parameter for the nutrient species. Figure 2 displays similar cumulative density functions for the core variables from the January 14, 2016 memorandum. The general observations related to which parameters having the highest levels of censoring carries forward to Figures 1 and 2. For example, about 40 percent of the station/layers have censoring exceeding 20 percent in 1985-1989 and 1990-1994 for NH4F (Figure 1, top left panel). Table 2 tabulates the percent of censoring exceeded by 50 percent and 25 percent of the station/layers by date range. For example, 50 percent of the station/layers with NH4F data from 1985-1989 have more than 14.6 percent or more censoring; 25 percent have more than 37.7 percent or more censoring.

Figure 3 displays a scatterplot of the maximum censoring limit as a function of the minimum censoring limit for each site/layer combination by parameter and time period. In some cases the maximum censoring limit is twice or greater than the minimum censoring limit.

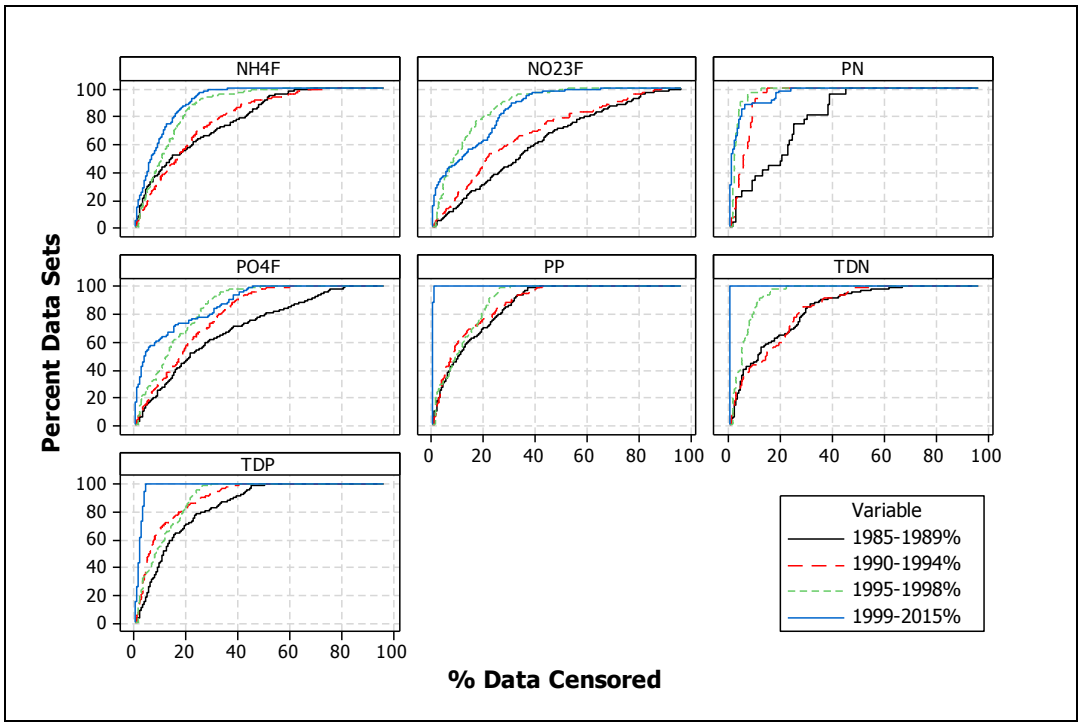


Figure 1. Cumulative Density Function of Percent Censoring for Nutrient Species.

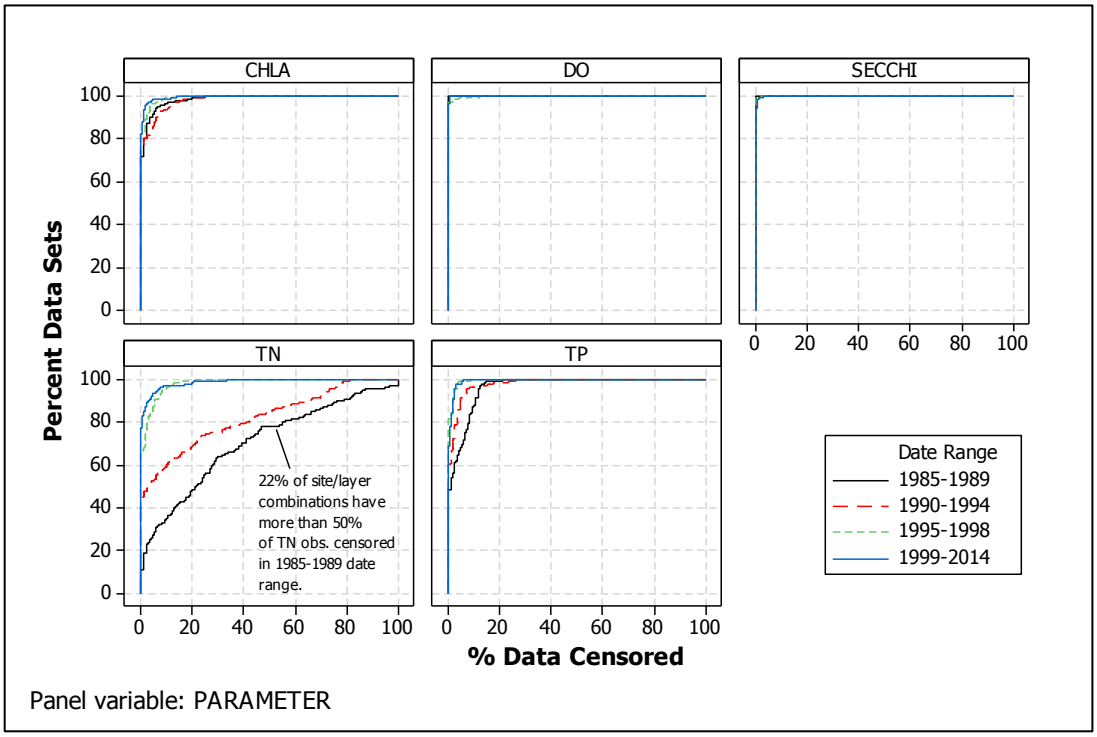
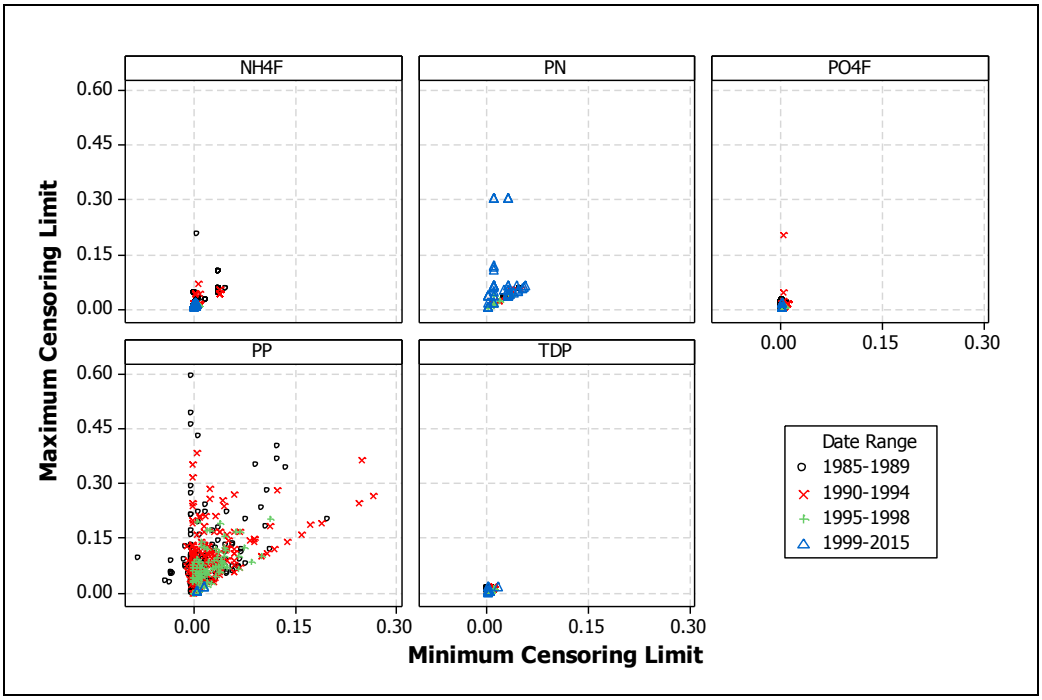


Table 2. Level of Percent Censoring Exceeded in 50 Percent and 25 Percent of Station/Layers by Parameter and Date Range.

	1985-1989%		1990-1994%		1995-1998%		1999-2015%	
	50% of Station/ Layers	25% of Station/ Layers	50% of Station/ Layers	25% of Station/ Layers	50% of Station/ Layers	25% of Station/ Layers	50% of Station/ Layers	25% of Station/ Layers
NH4F	14.6	37.7	16.9	29.4	10.5	17.4	6.5	13.7
NO23F	33.7	54.7	21.4	44.2	8.5	17.0	11.6	25.0
PN	20.5	29.0	5.7	8.8	2.2	3.6	1.1	3.9
PO4F	21.7	44.8	17.8	30.7	12.5	23.3	3.8	21.2
PP	10.5	23.1	7.7	19.3	9.6	18.0	0.5	0.6
TDN	11.5	27.1	14.4	25.0	4.9	8.2	0.5	*
TDP	11.2	23.1	5.9	14.0	8.1	17.5	2.2	3.1
CHLA[1]	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0
DO[1]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SECCHI[1]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TN[1]	21.3	45.4	3.3	27.8	0.0	1.8	0.0	0.0
TP[1]	1.0	6.5	0.0	2.1	0.0	0.0	0.0	0.6

[1] Data for these parameters did not include data from 2015.



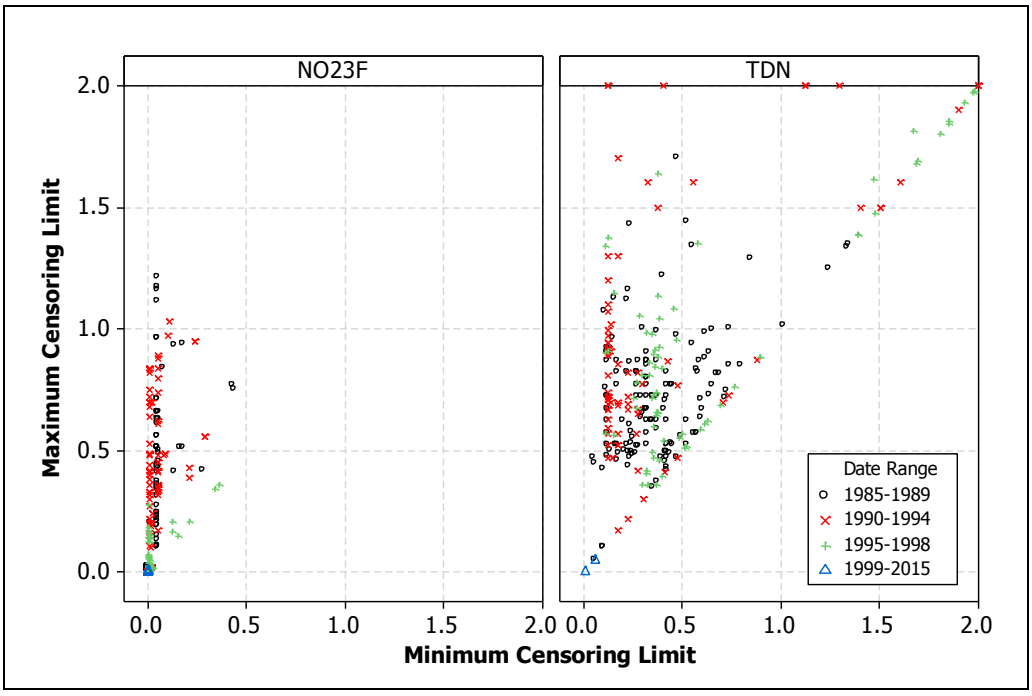


Figure 3. Scatterplot of Maximum Censoring Limit as a Function of Minimum Censoring Limit by Parameter and Date Range.