Appendix 2.6C: Censoring Scenarios

Section 1. Description of Scenario Selection

I. Goal

Select a series of different locations, parameters, and censoring levels to evaluate the performance of multiple options for dealing with the pre-1999 data censoring in the CBP dataset. Tried to pick the scenarios so that they represent the common censoring cases we have in the data that might impact trend analysis.

II. Censoring features of our data sets that we want to evaluate:

- Censoring of data in the 1st half our record, and not the 2nd half, could impact long-term trend conclusions when we analyze 1985-present.
- Step-wise detection limits improvements as time went on may lead to spurious trend detection, or inability to identify a true trend.
- Censoring of individual constituents of a computed parameter result in interval censoring. Some interval censoring notes I've seen as I go:
    - TP is not much of a concern. >10% censoring only happens from 1985-1989 in 11% of the data sets. The worst case is CB5.2 surface at 21%
    - For TN, there is a lot of censoring in the first half of the record. Half of the data sets from 1985-1989 are censored >50%, and 1/3 of the data sets from 1990-1994 are censored >50%. This drops to almost nothing after 1994. However, looking at these plots, the censored and non-censored TN are all mixed together.
- Varied percent of data censored and how different methods might perform (see spreadsheets from Tetra Tech):
    - >70% censoring happens for
        - 1985-1989 PO4f, NO23F in VA (just a couple MD tribs), TN Potomac and VA tribs
        - 1990-1994: NO23F, TN VA tribs only, NH4F CB7.4N
    - 20 -70% happens for
        - All of the nutrient species and totals in 85-98
        - CHLA is only 1 sample, >10% happens a few more times for chla

III. Proposed Scenarios to generate censored data to test the methods we are incorporating into GAMs

Scenario 1) Constant detection limit in 1st half of record, with moderate amount of censoring. This scenario is to capture the chlorophyll-*a* censoring at 1ug/L that happened from 1985-1999. Only a few stations have >10% censoring at a constant 1 ug/L limit, but they are important ones for trends and water quality implications.

	Station	Layer-parameter	Time period 1: 1999-2006	Time period 2: 2007-2014	Reason picked
1a-b	CB3.3C	Bottom - chla	a. Censor lowest 10% of samples b. Censor lowest 25%	a. No censoring b. No censoring	Border-line increasing trend with censoring pre-99 (MD station)
1c-d	CB5.1	Bottom - chla	c. Censor lowest 10% d. Censor lowest 25%	a. No censoring b. No censoring	Significant increasing trend with censoring pre-99 (MD station)
1e-f	RET5.2	Surface - chla	e. Censor lowest 10% f. Censor lowest 25%	a. No censoring b. No censoring	Decreasing trend with censoring pre-99 (VA trib station)

Scenario 2) Very large to modest amount of data censored in first half of record, and little censoring after that. Both left and interval censoring conditions. This scenario is important for almost all the nutrients, and is a catch-all.
Note on LE3.1: roughly 2/3 of the censored data is left censored, and 1/3 is interval. So used that breakdown to create the scenarios based on LE3.1 (2g-i).

	Station	Layer-parameter	Time period 1: 1999-2006	Time period 2: 2007-2014	Reasons picked
2a-c	LE3.1	Surface - TN	a. Censor random 75% b. Censor random 50% c. Censor random 25%	No censoring No censoring No censoring	Interval censoring scenario, at a station with increasing trend and very high censoring pre-99 (Rapp, VA)
2d-f	TF4.2	Surface - TN	d. Censor random 75% e. Censor random 50% f. Censor random 25%	No censoring No censoring No censoring	Interval censoring scenario, at a station with no trend and very high censoring pre-99 (Pam, VA)
2g-i	LE3.1	Surface – NO23F	g. Censor a total of 75% (25% randomly selected to be interval, and 50% left censored) h. Censor total of 50% (17% interval censored, and 33% of the data left censored) i. Censor a total of 25% (8% interval censored, 17% left censored)	No censoring No censoring No censoring	Consistency with 2a-c above (Rapp, VA)

2j-l	EE3.0	Surface – NO23F	j. Censor lowest 75% k. Censor lowest 50% l. Censor lowest 25%	No censoring No censoring No censoring	MD station selected with similar censoring pattern pre-99.

Scenario 3) Step improvements in detection limits. For a few of the nutrient species, there are clear step improvements in detection limits, in some regions of the Bay. The most obvious parameters are NH4 and PO4, especially in VA mainstem. Test cases include data sets with increases and decreases because of different impacts this type of censoring could have.

.

	Station	Layer-parameter	Time period/cutoff	Reasons picked
3a	EE3.4	Bottom NH4F	1999-2000: 0.02 mg/L 2001-2003: 0.01 2004-2009: 0.004 2010-2014: none	VA eastern shore station with this pattern pre-99. Probably no trend.
3b	CB4.2C	Surface NH4F	1999-2000: 0.02 mg/L 2001-2003: 0.01 2004-2009: 0.004 2010-2014: none	Mainstem MD station with values that would be censored at these levels (although a lower percent than EE3.4), and probably a decreasing trend.
3c	CB7.4	Surface PO4F	1999-2000: 0.01 mg/L 2001-2005: 0.005 2006-2014: 0.0009	VA mainstem station that was with this pattern pre-1999. Probably no trend.
3d	CB4.1C	Surface PO4F	1999-2000: 0.01 mg/L 2001-2005: 0.005 2006-2014: 0.0009	MD mainstem station with values in this range. Possibly increasing trend to test impact.
3e	WT2.1	Surface PO4F	1999-2000: 0.01 mg/L 2001-2005: 0.005 2006-2014: 0.0009	MD tributary station with values in this range. Possibly decreasing trend to test impact.

Section 2. Results

Results are presented for each of the artificially-censored scenarios listed above. A summary table is provided first, and then graphical results are presented for each scenario. The graphics by panel are: uncensored data set (panel A), the uncensored GAM fit (panel B), the censored data set superimposing the data values that were censored (panel C), and five panels (D-H) depicting the censored GAM fit. In each case, the data are grey, the range of the censored data is shown with a grey line, the GAM model fit to the uncensored data is black, the individual runs from each Monte Carlo simulation are grey, the composite median of the Monte Carlo results is light blue, and the Expectation Maximization (referenced to as conditional expectation here also) result is a red dotted line.

Appendix 2.6C

Table 1. Summary of Results from Censoring Scenarios[a]					
Scenario	**Station**	**Layer**	**Parameter**	**Artificial censoring description**	**Notable results about EM method**
1a	CB3.3C	B	chla	Censor lowest 10% in first half of record	No change in any conclusions
1b	CB3.3C	B	chla	Censor lowest 25% in first half of record	No change in any conclusions
1c	CB5.1	B	chla	Censor lowest 10% in first half of record	No change in any conclusions
1d	CB5.1	B	chla	Censor lowest 25% in first half of record	No change in any conclusions
1e	RET5.2	S	chla	Censor lowest 10% in first half of record	No change in any conclusions
1f	RET5.2	S	chla	Censor lowest 25% in first half of record	No change in conclusions, except that the EM results suggest much more nonlinear pattern to the model than the Unc
2a	LE3.1	S	tn	Interval censoring of 75% in first half of record	No change in any conclusions
2b	LE3.1	S	tn	Interval censoring of 50% in first half of record	No change in any conclusions
2c	LE3.1	S	tn	Interval censoring of 25% in first half of record	No change in any conclusions
2d	TF4.2	S	tn	Interval censoring of 75% in first half of record	No change in any conclusions
2e	TF4.2	S	tn	Interval censoring of 50% in first half of record	No change in any conclusions
2f	TF4.2	S	tn	Interval censoring of 25% in first half of record	No change in any conclusions
2g	LE3.1	S	no23f	Combination left and interval censoring of 75% of first half of record	No change in conclusions at the p<0.05 level, but EM shows a higher % change and lower p-value than Unc
2h	LE3.1	S	no23f	Combination left and interval censoring of 50% of first half of record	No change in conclusions at the p<0.05 level, but EM shows a higher % change and lower p-value than Unc
2i	LE3.1	S	no23f	Combination left and interval censoring of 25% of first half of record	No change in any conclusions
2j	EE3.0	S	no23f	Censor lowest 75% in first half of record	Change in conclusion: EM results sugggest strong negative trend while Unc is no trend

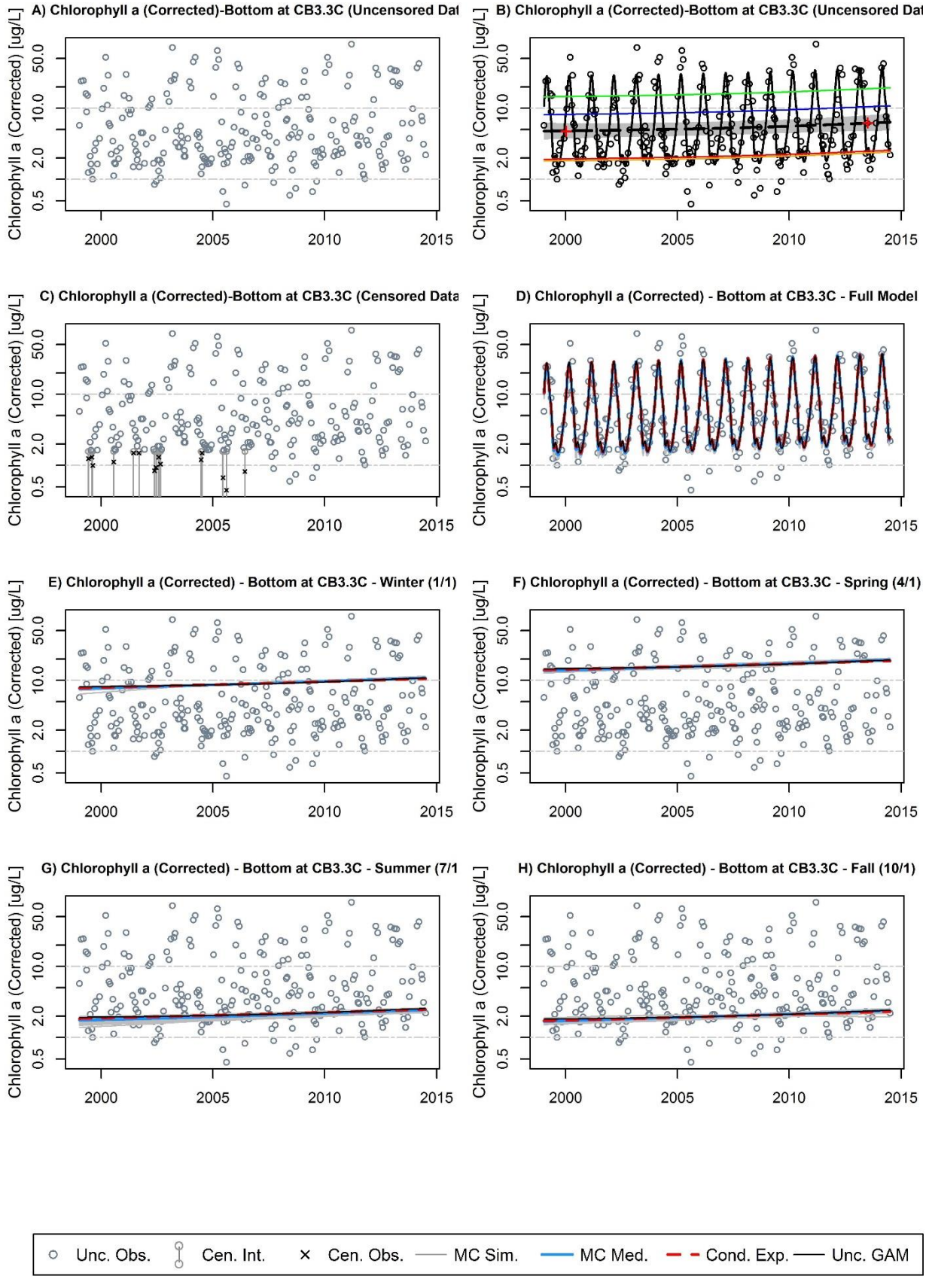
Appendix 2.6C

Table 1. Summary of Results from Censoring Scenarios[a]					
Scenario	**Station**	**Layer**	**Parameter**	**Artificial censoring description**	**Notable results about EM method**
2k	EE3.0	S	no23f	Censor lowest 50% in first half of record	Change in conclusion: EM results sugggest strong negative trend while Unc is no trend
2l	EE3.0	S	no23f	Censor lowest 25% in first half of record	No change in any conclusions
3a	EE3.4	B	nh4f	Step improvements in detection limits on no trend	No change in any conclusions
3b	CB4.2C	S	nh4f	Step improvements in detection limits on decr trend	No change in any conclusions
3c	CB7.4	S	po4f	Step improvements in detection limits on no trend	No change in any conclusions
3d	CB4.1C	S	po4f	Step improvements in detection limits on incr trend	Change in conclusions: EM direction and existance of long-term trend is different from Unc. This is due to the first 2 years being 100% censored.
3e	WT2.1	S	po4f	Step improvements in detection limits on decr trend	Change in conclusions: EM direction and existance of long-term trend is different from Unc. This is due to the first 2 years being 100% censored.

[a] Abbreviations are: B=bottom layer, S = Surface layer; EM = expectation maximization; Unc = uncensored

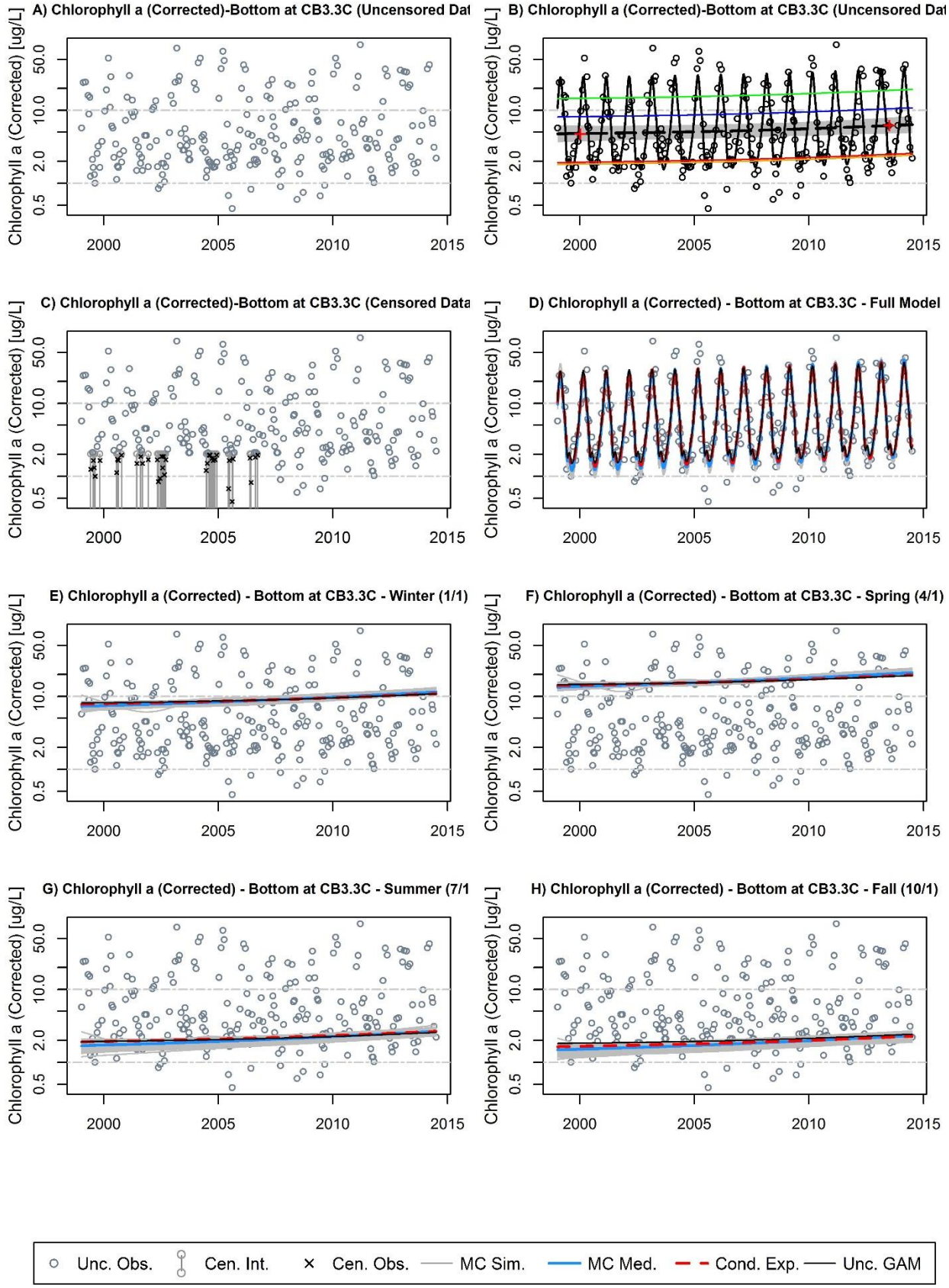
Appendix 2.6C

Scenario 1a



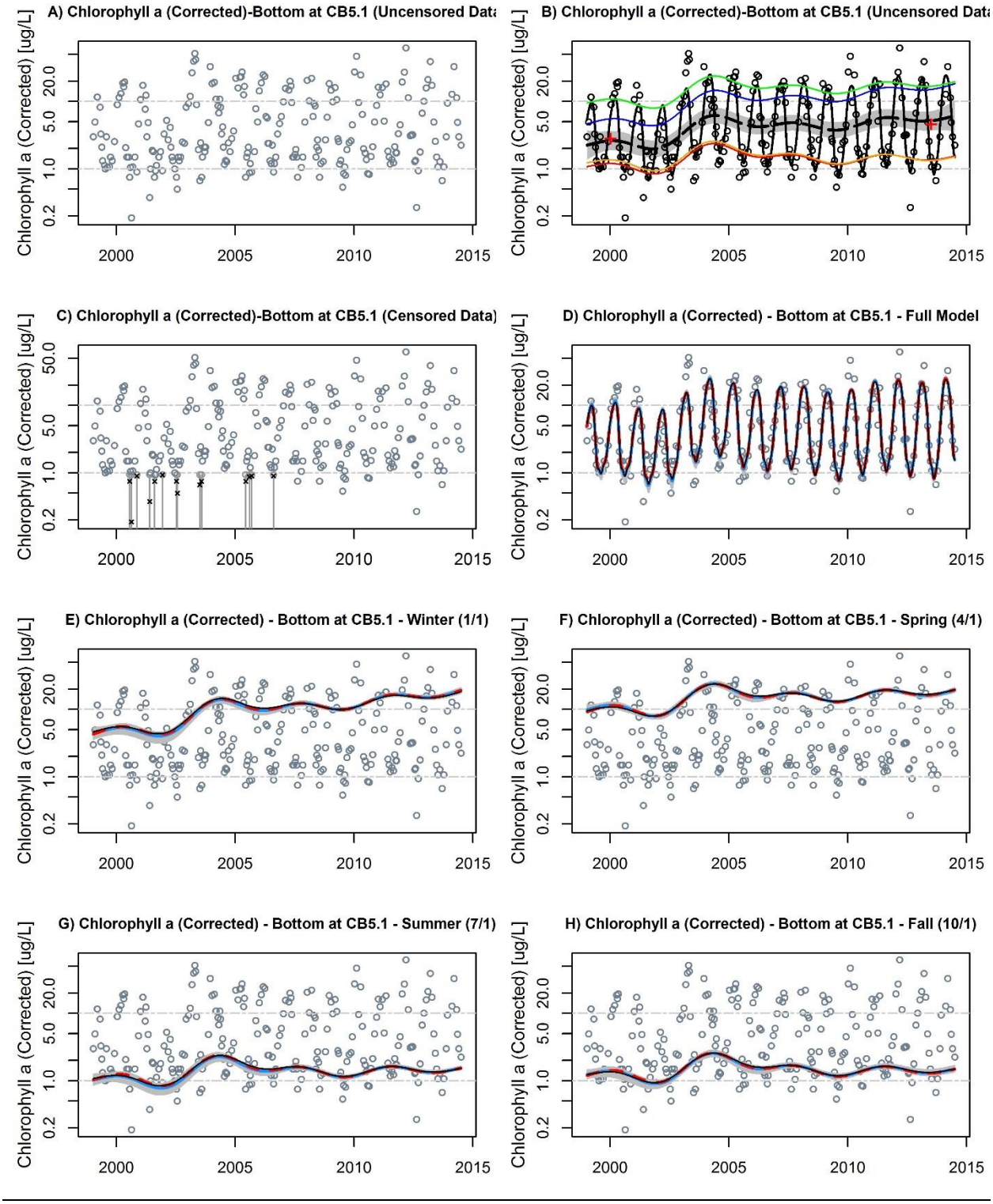
Appendix 2.6C

Scenario 1b



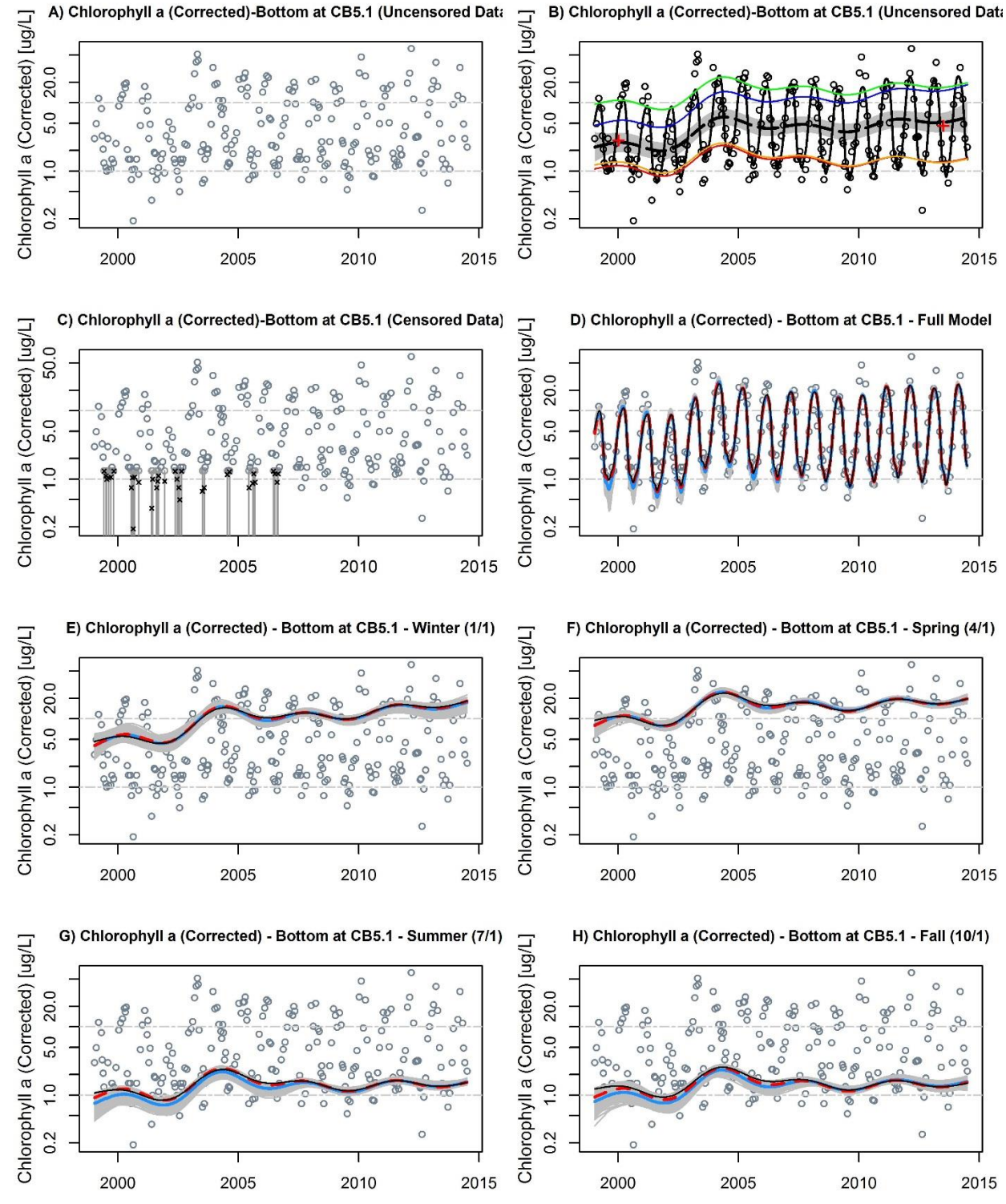
Appendix 2.6C

Scenario 1c



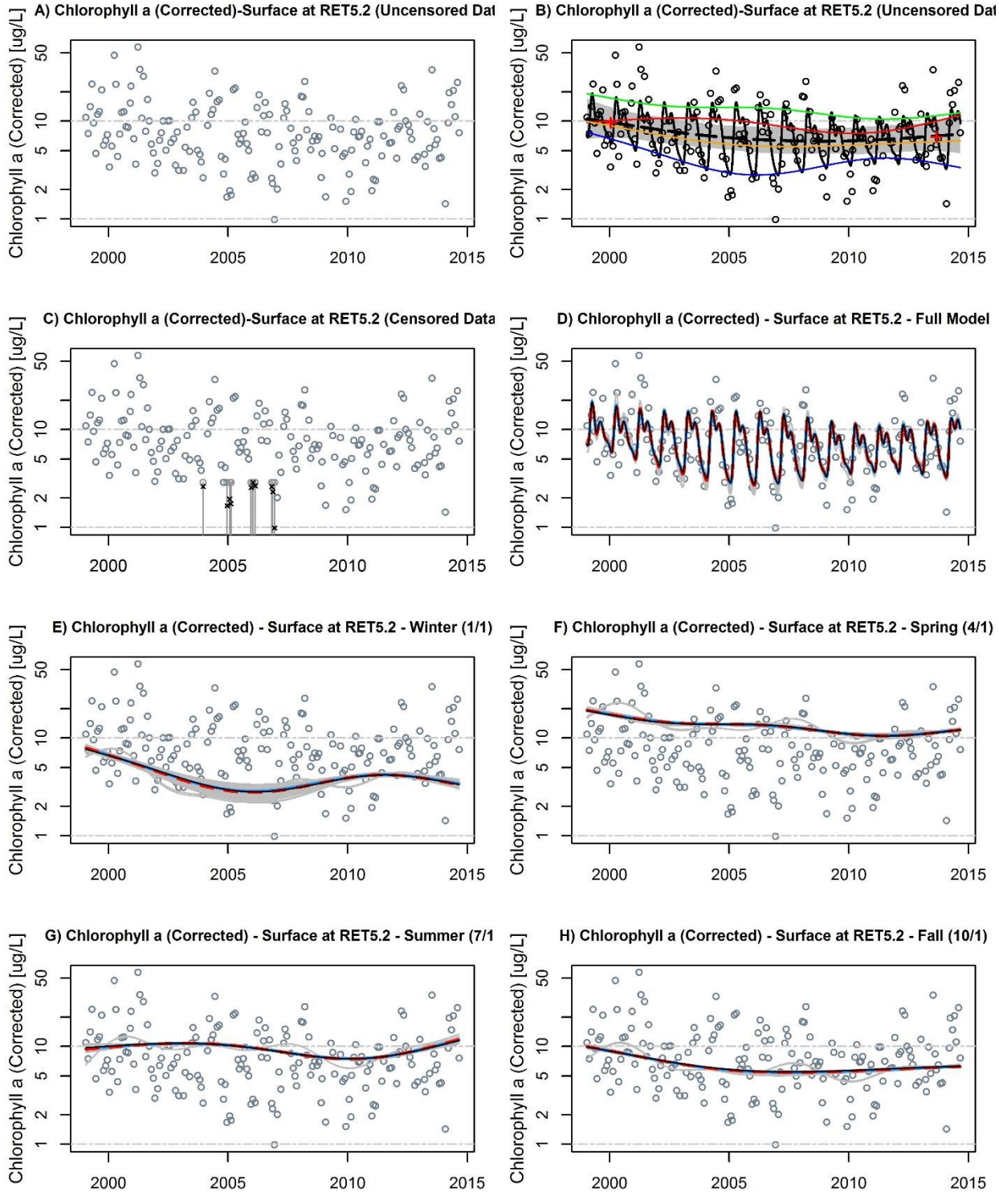
Appendix 2.6C

Scenario 1d



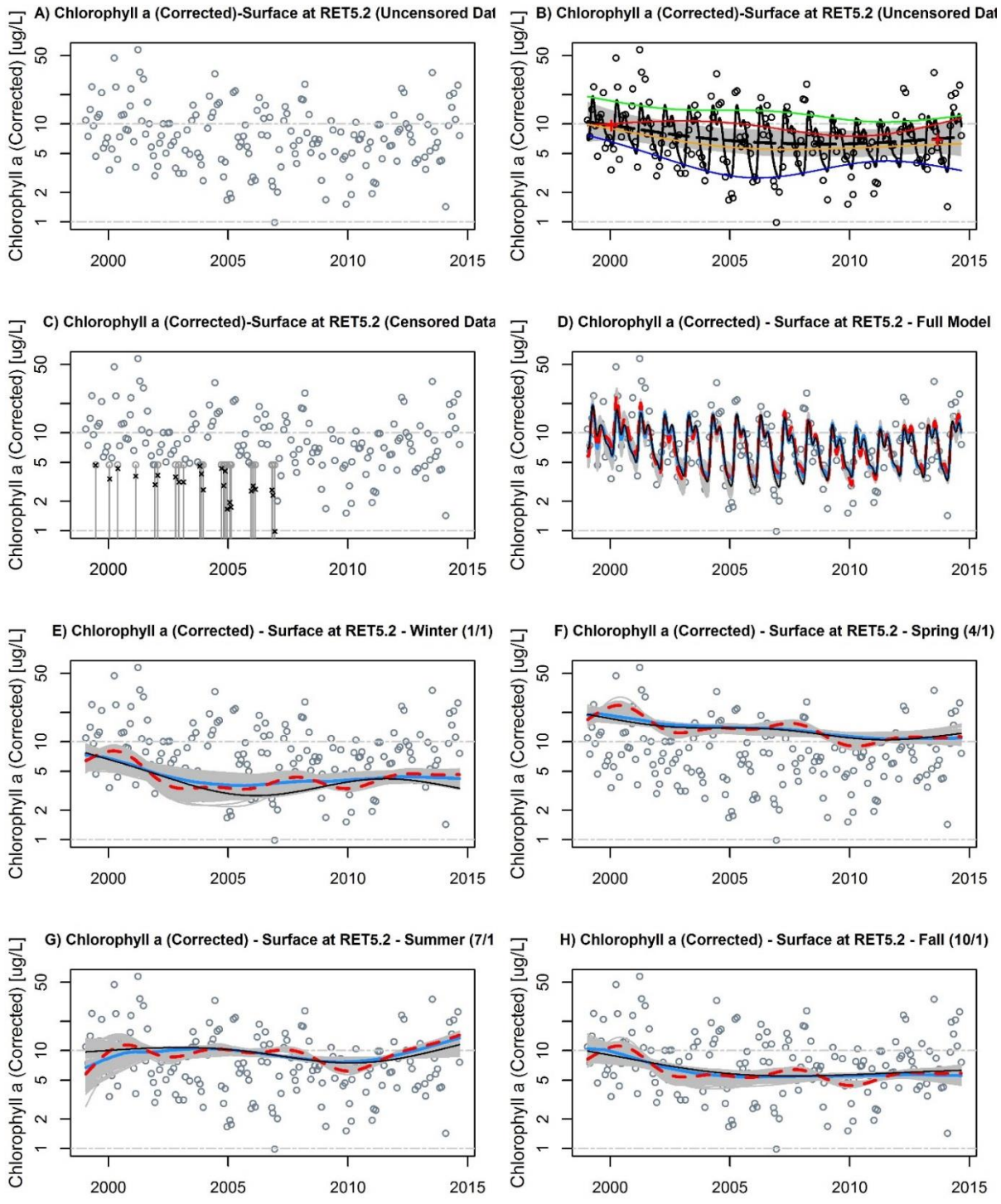
Appendix 2.6C

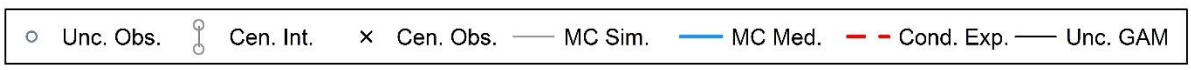
Scenario 1e



Appendix 2.6C

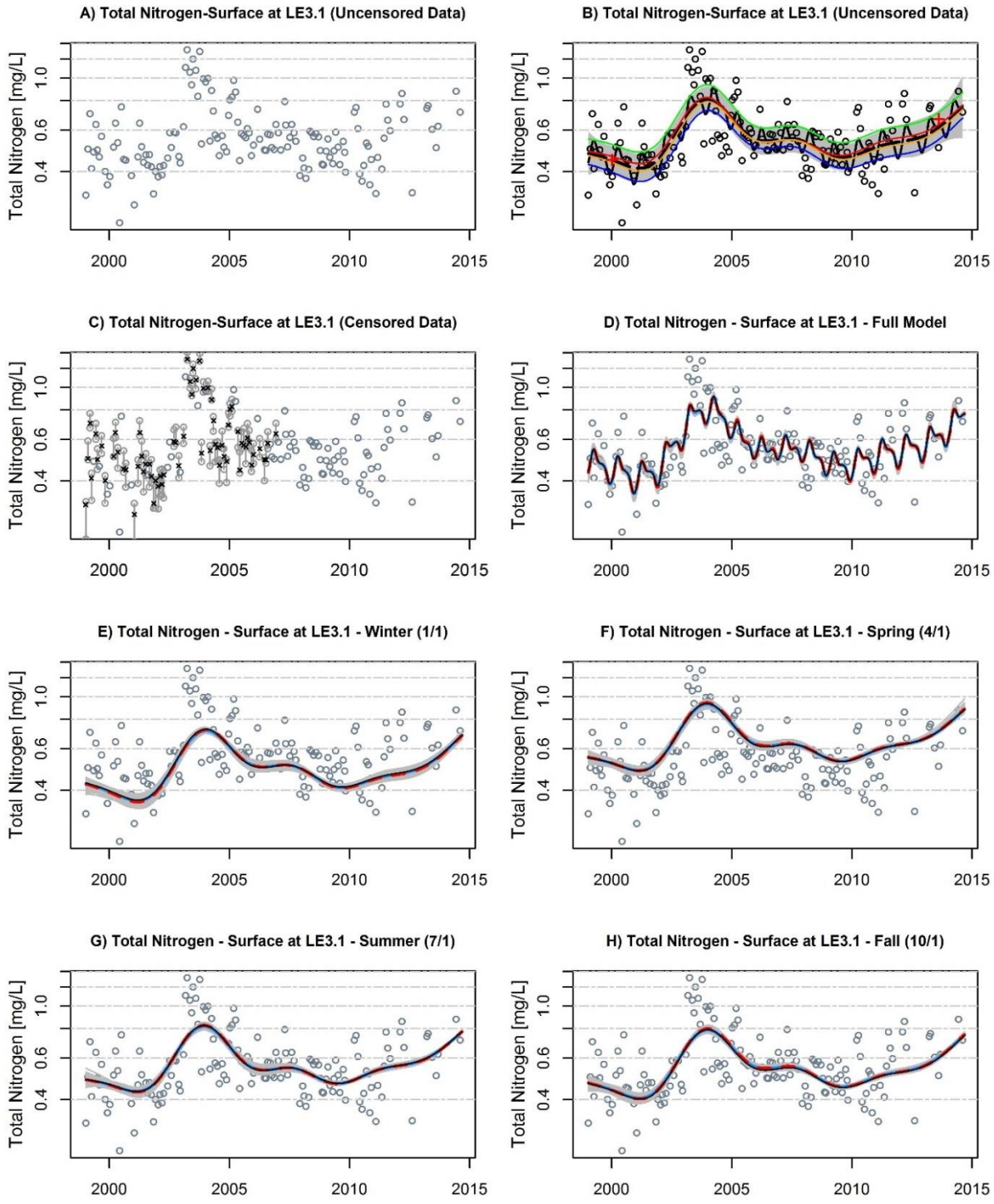
Scenario 1f





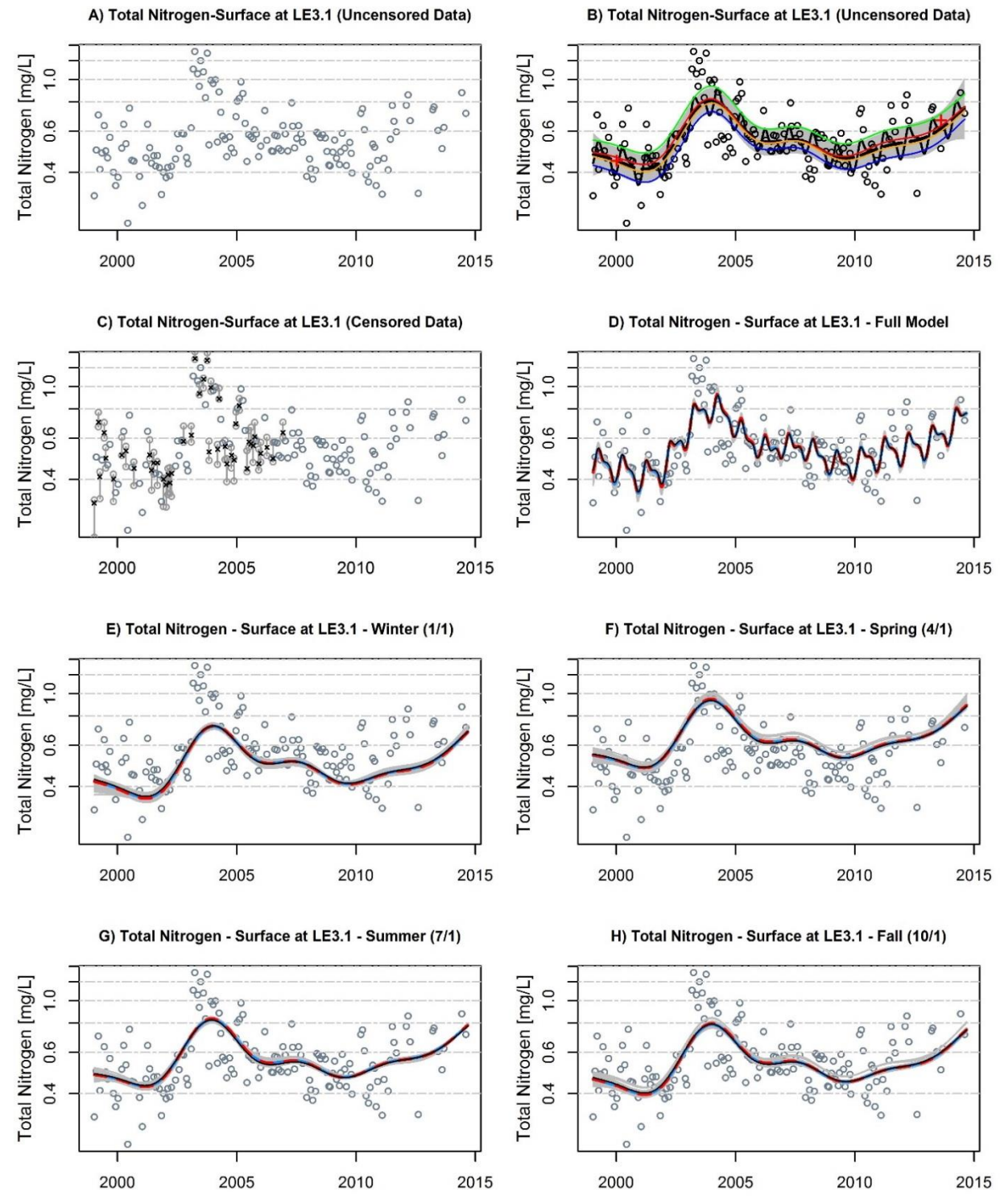
Appendix 2.6C

Scenario 2a



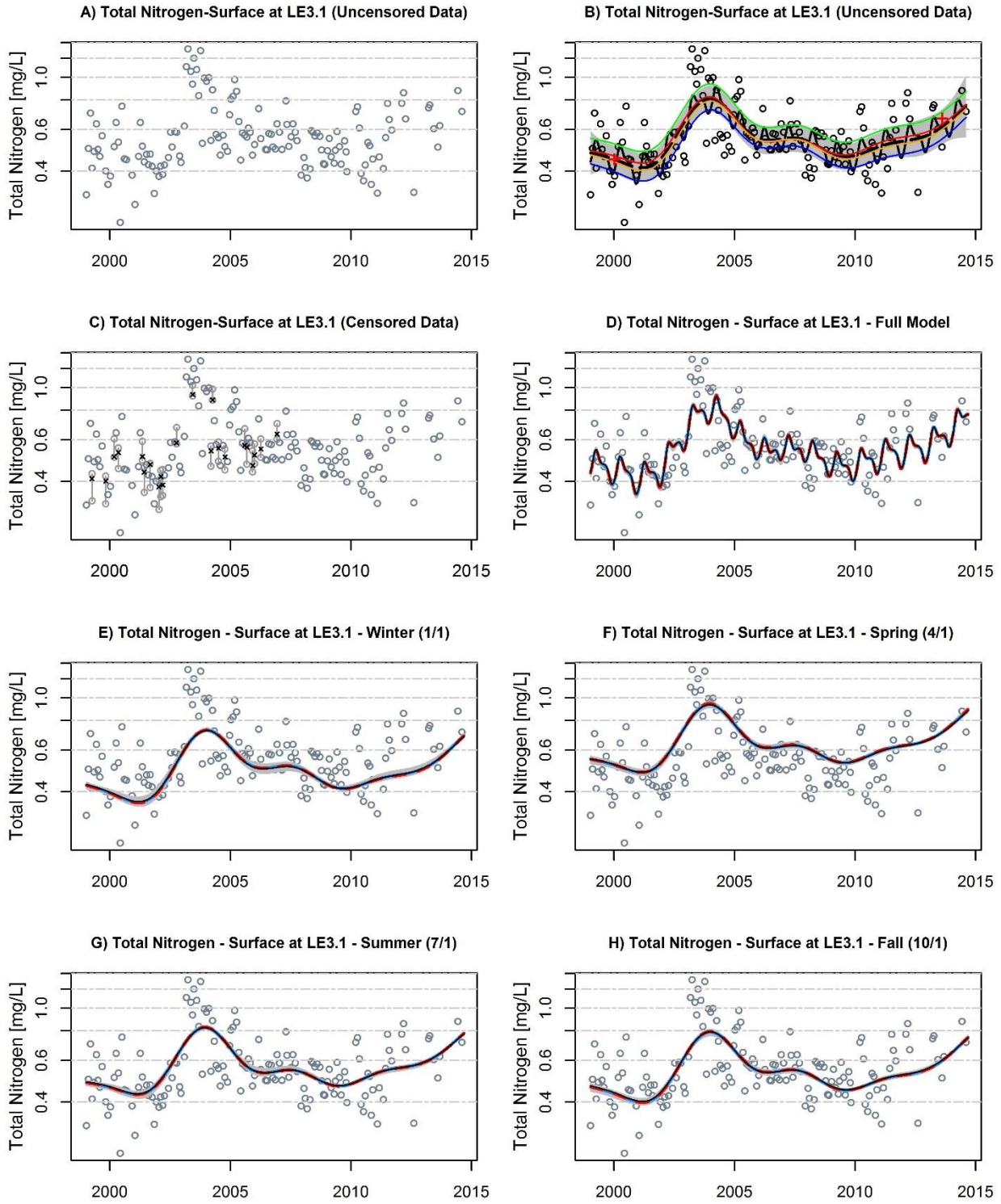
Appendix 2.6C

Scenario 2b



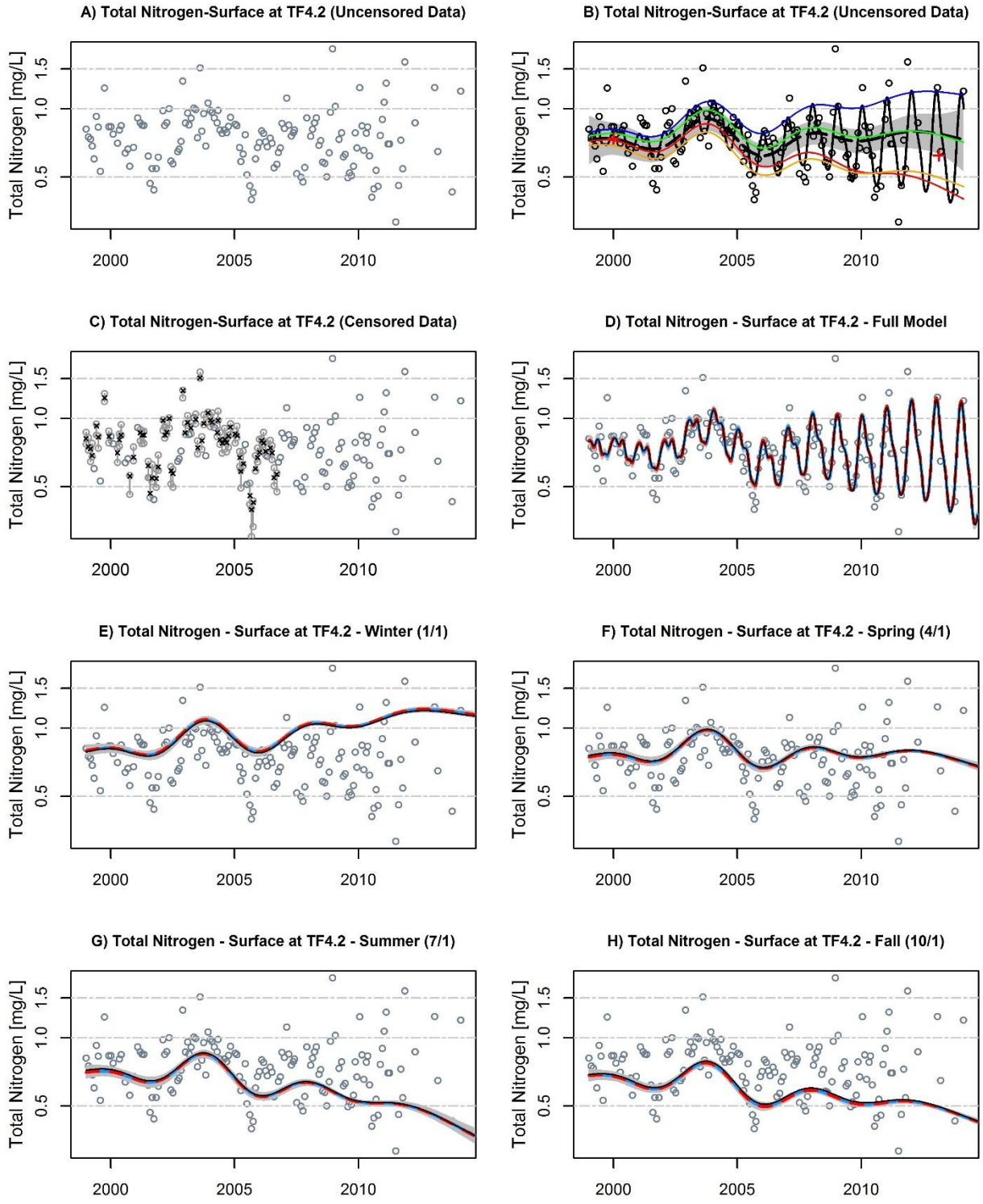
Appendix 2.6C

Scenario 2c



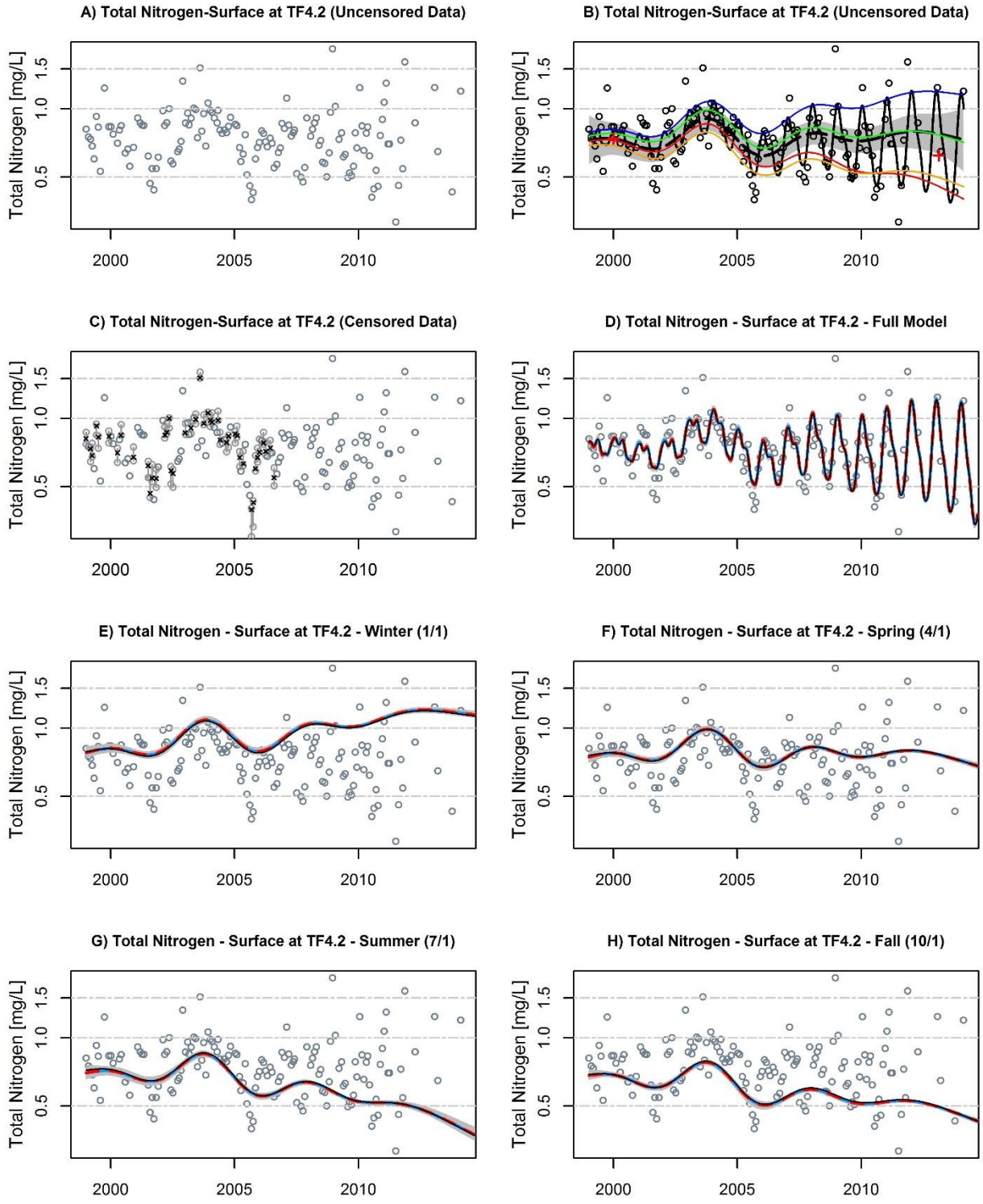
Appendix 2.6C

Scenario 2d



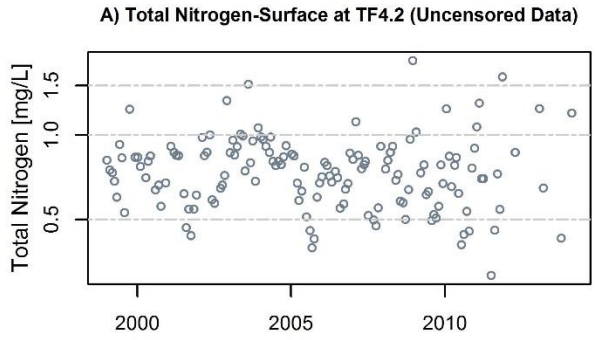
Appendix 2.6C

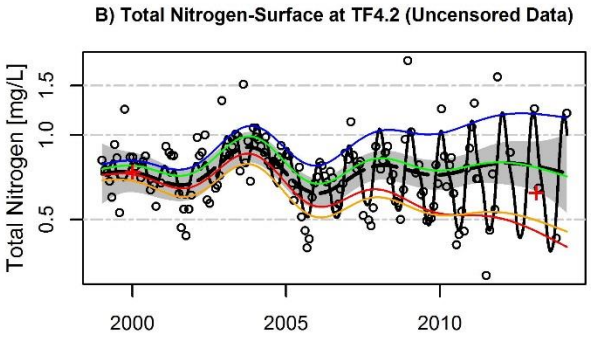
Scenario 2e

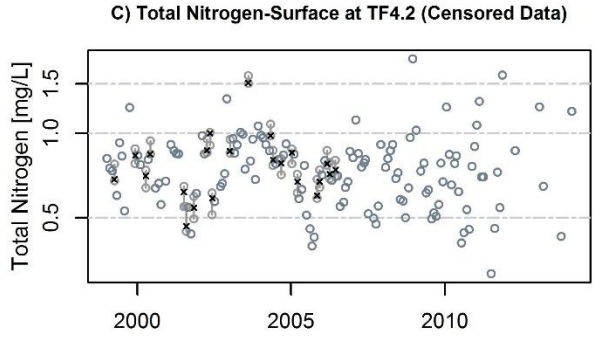


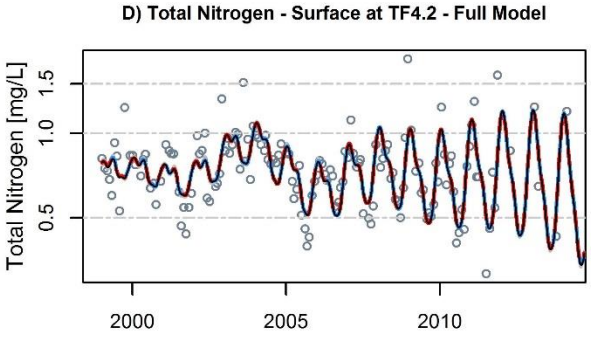
Appendix 2.6C

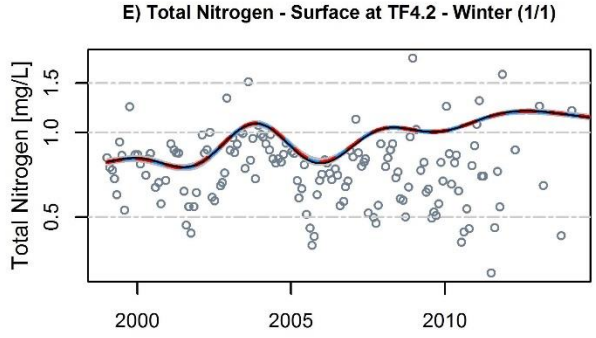
Scenario 2f

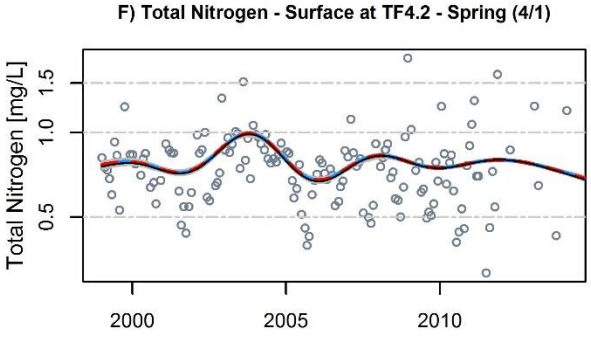


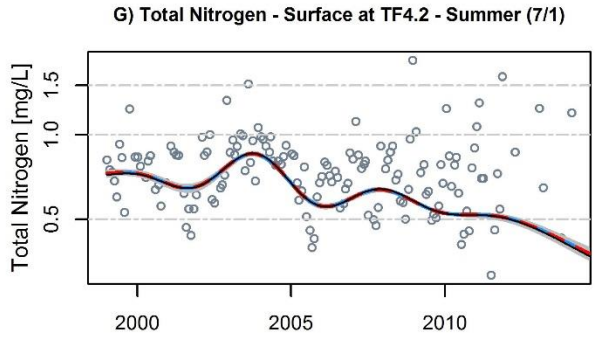


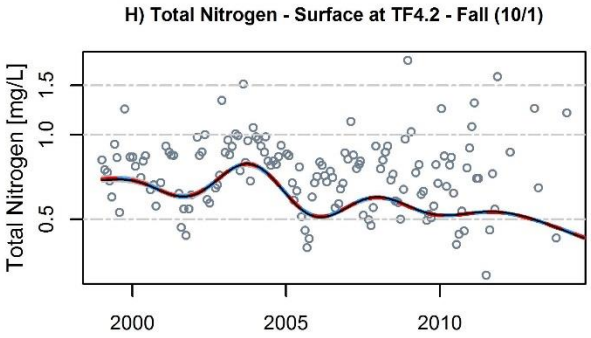






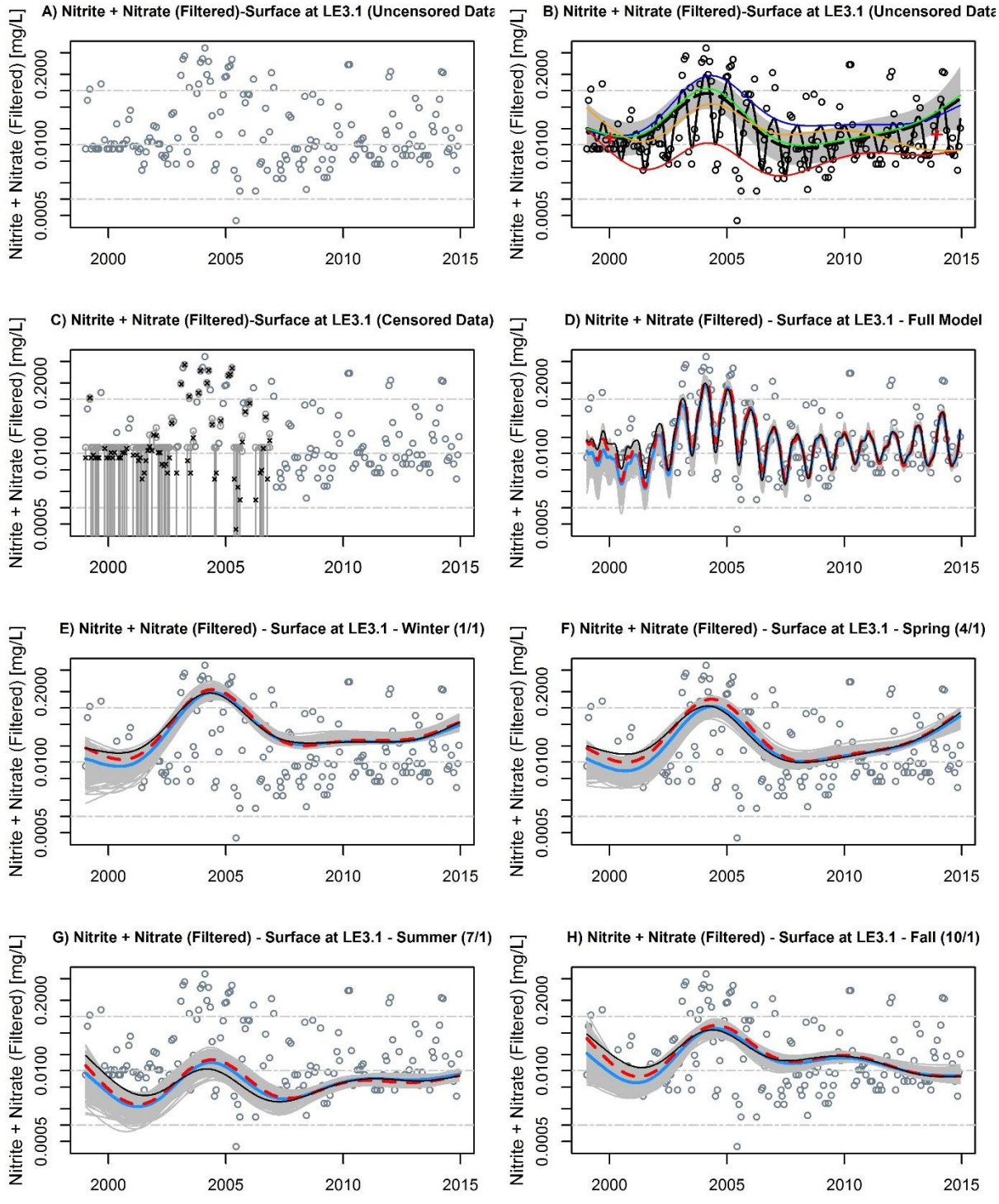


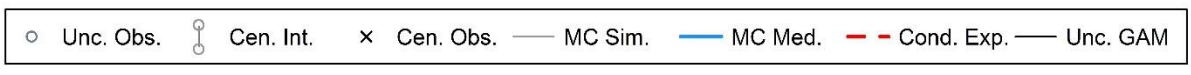




Appendix 2.6C

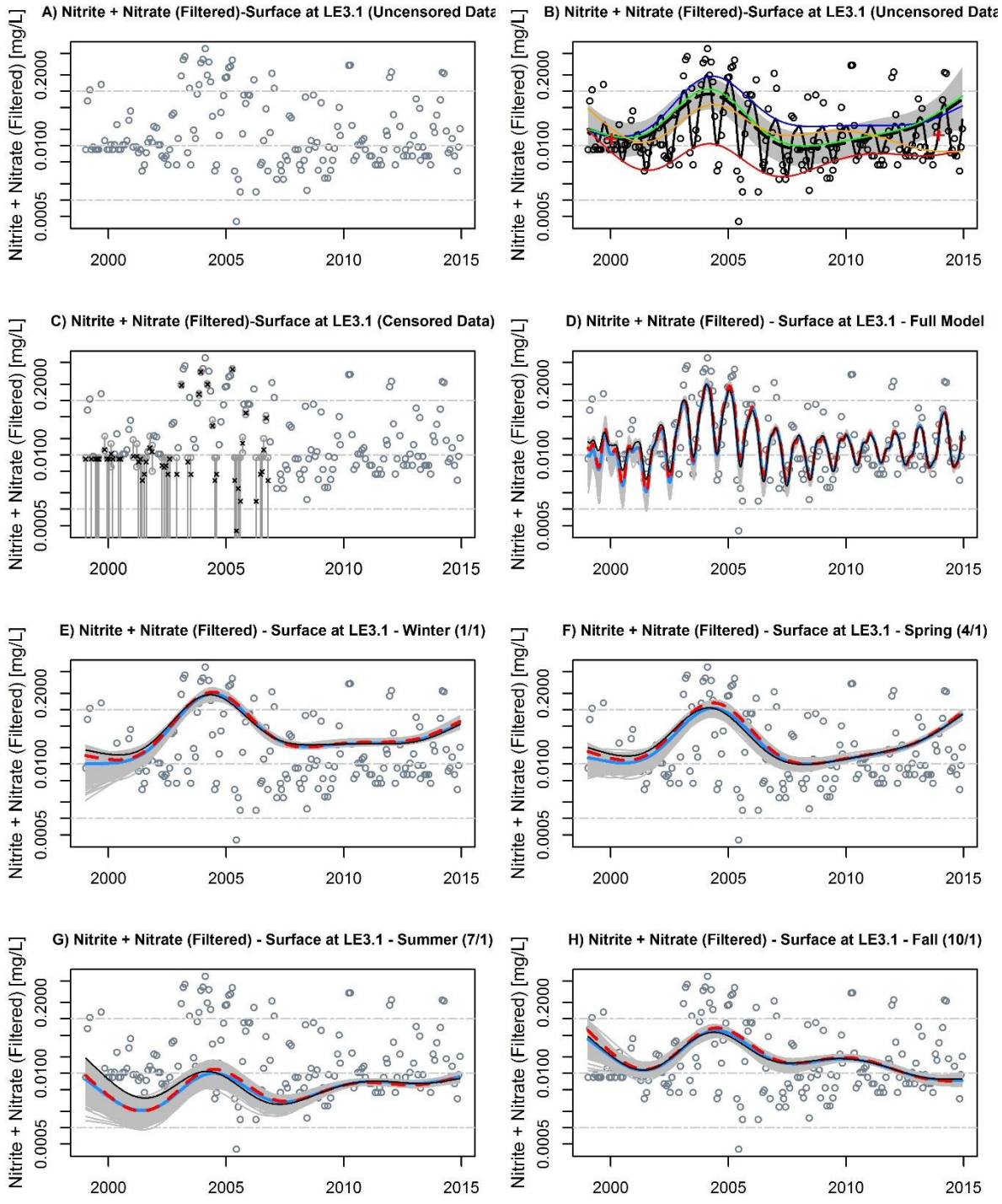
Scenario 2g





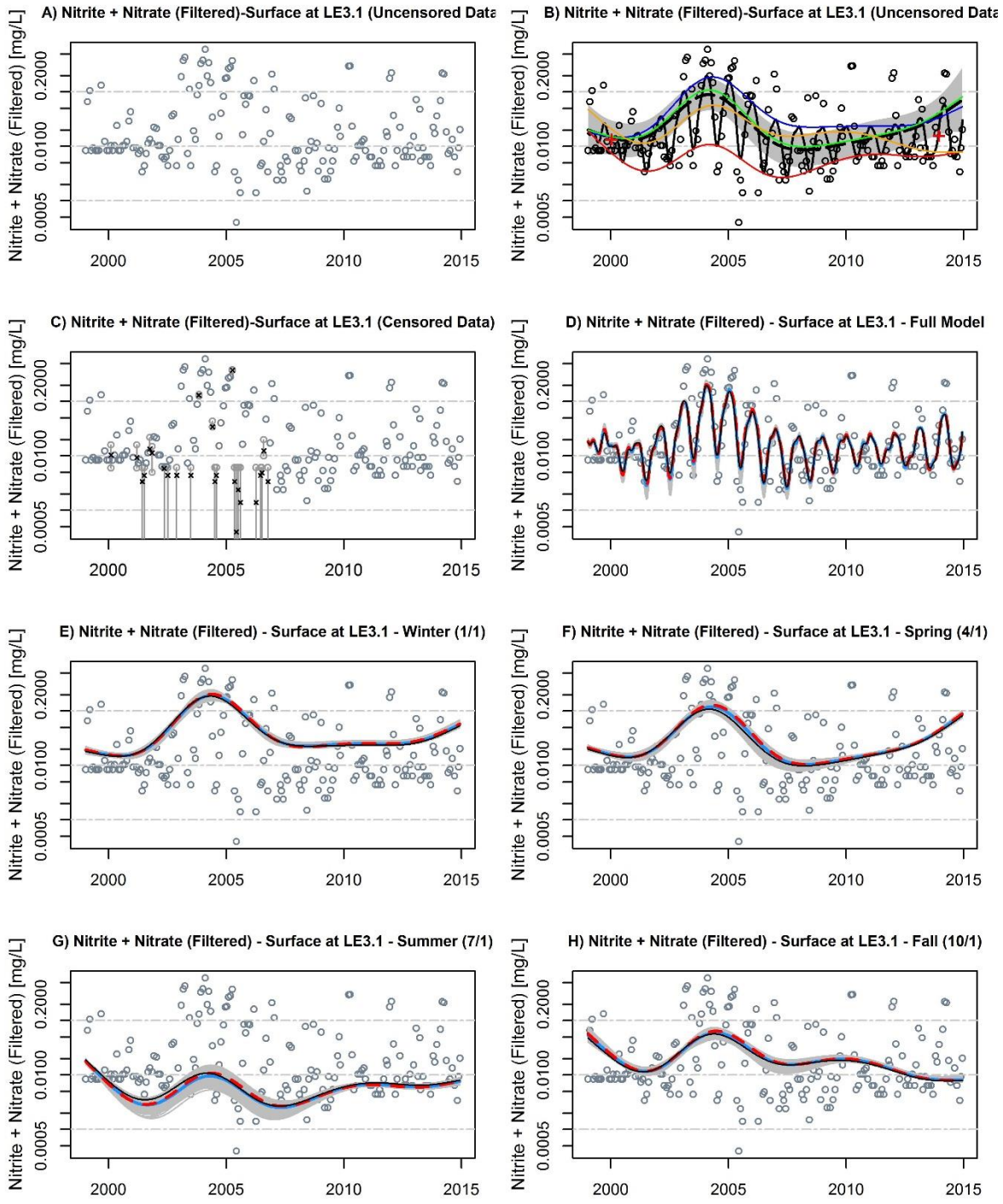
Appendix 2.6C

Scenario 2h



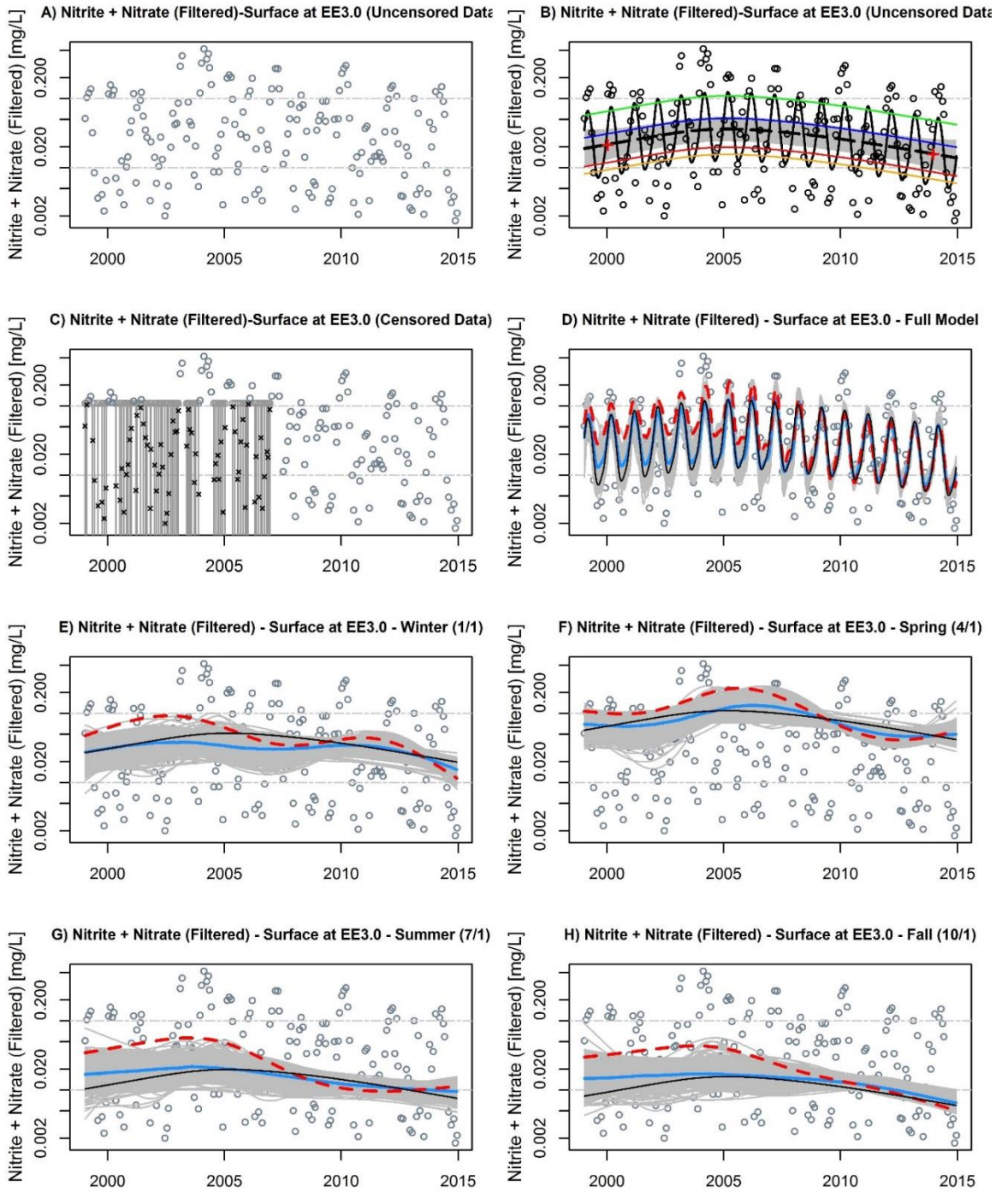
Appendix 2.6C

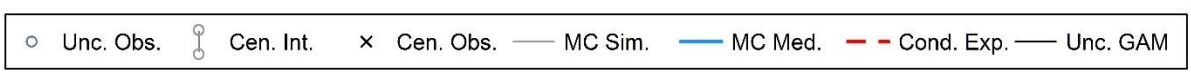
Scenario 2i



Appendix 2.6C

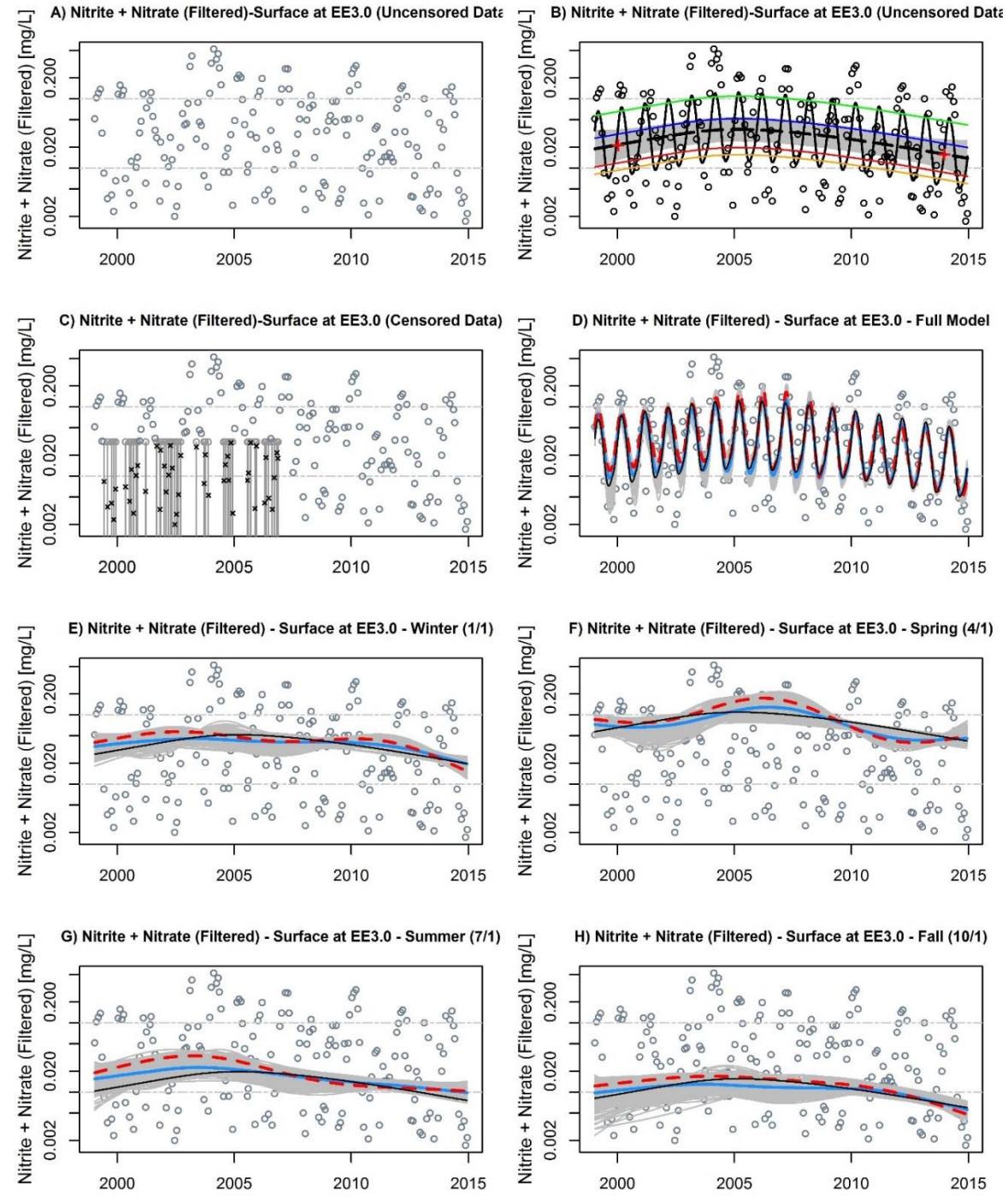
Scenario 2j

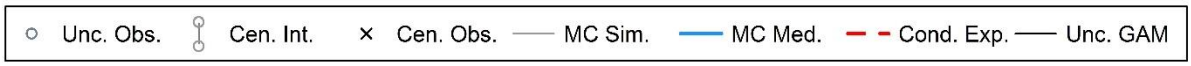




Appendix 2.6C

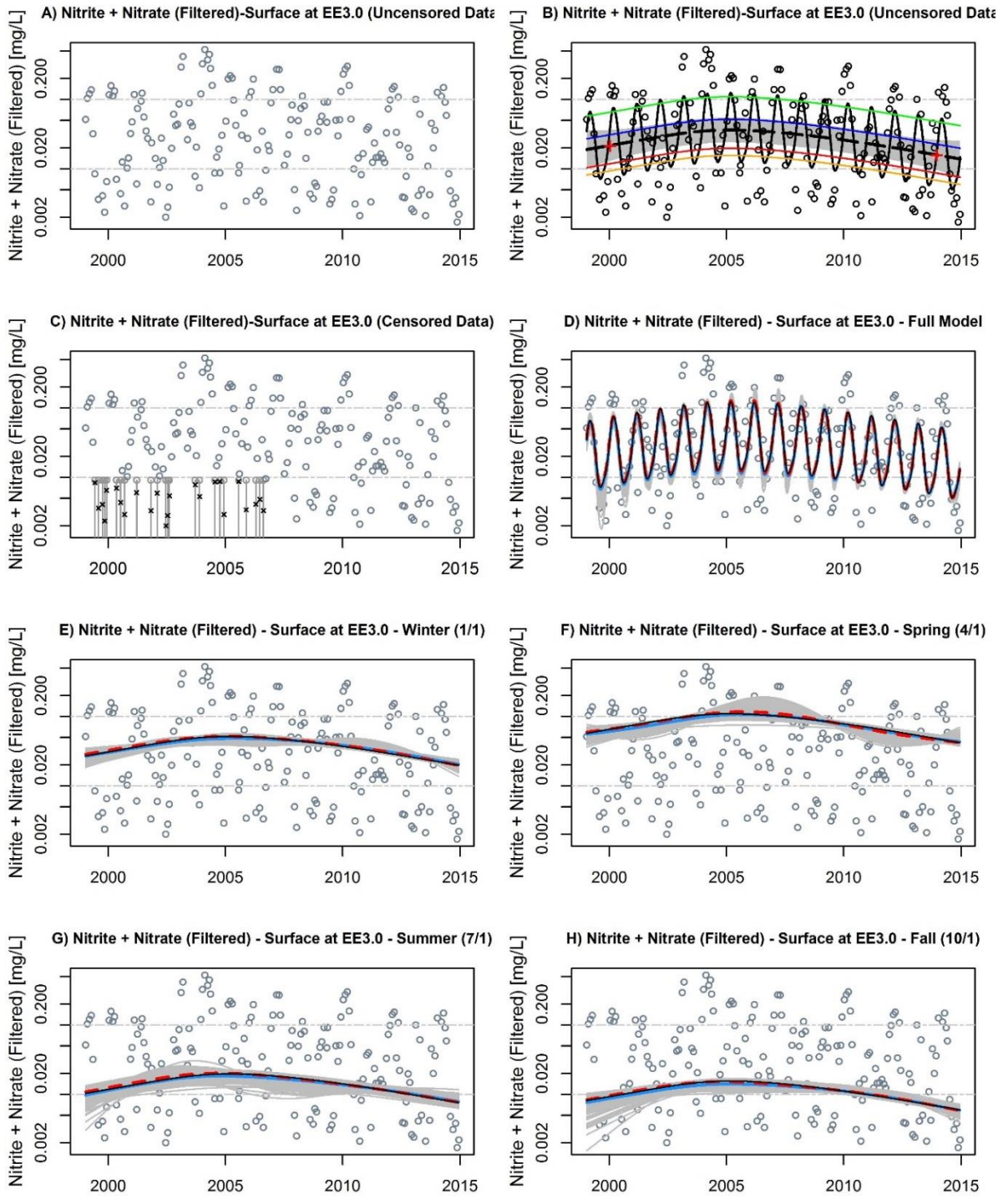
Scenario 2k

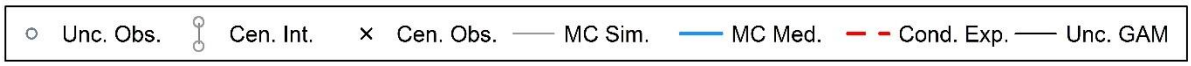




Appendix 2.6C

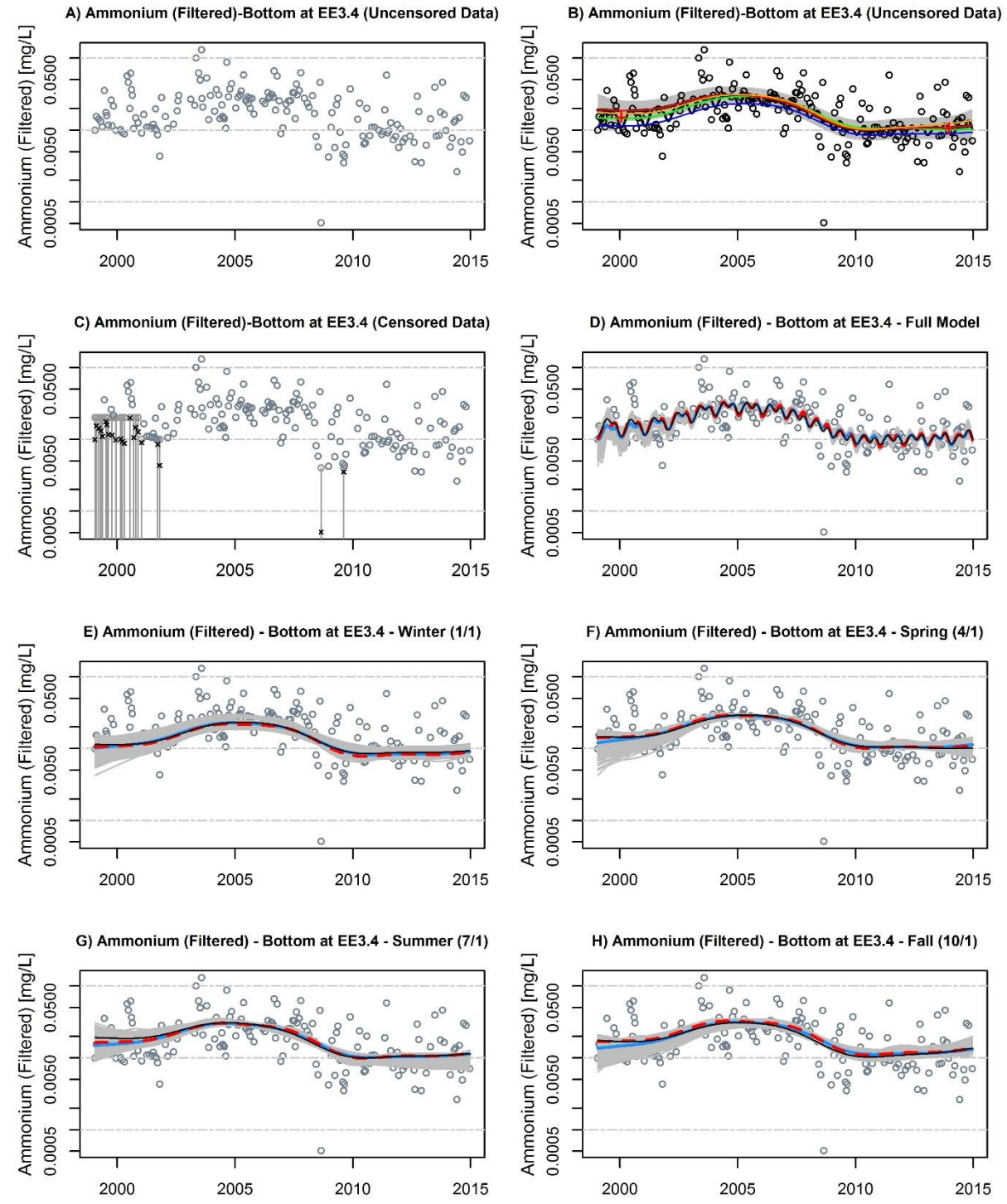
Scenario 2l





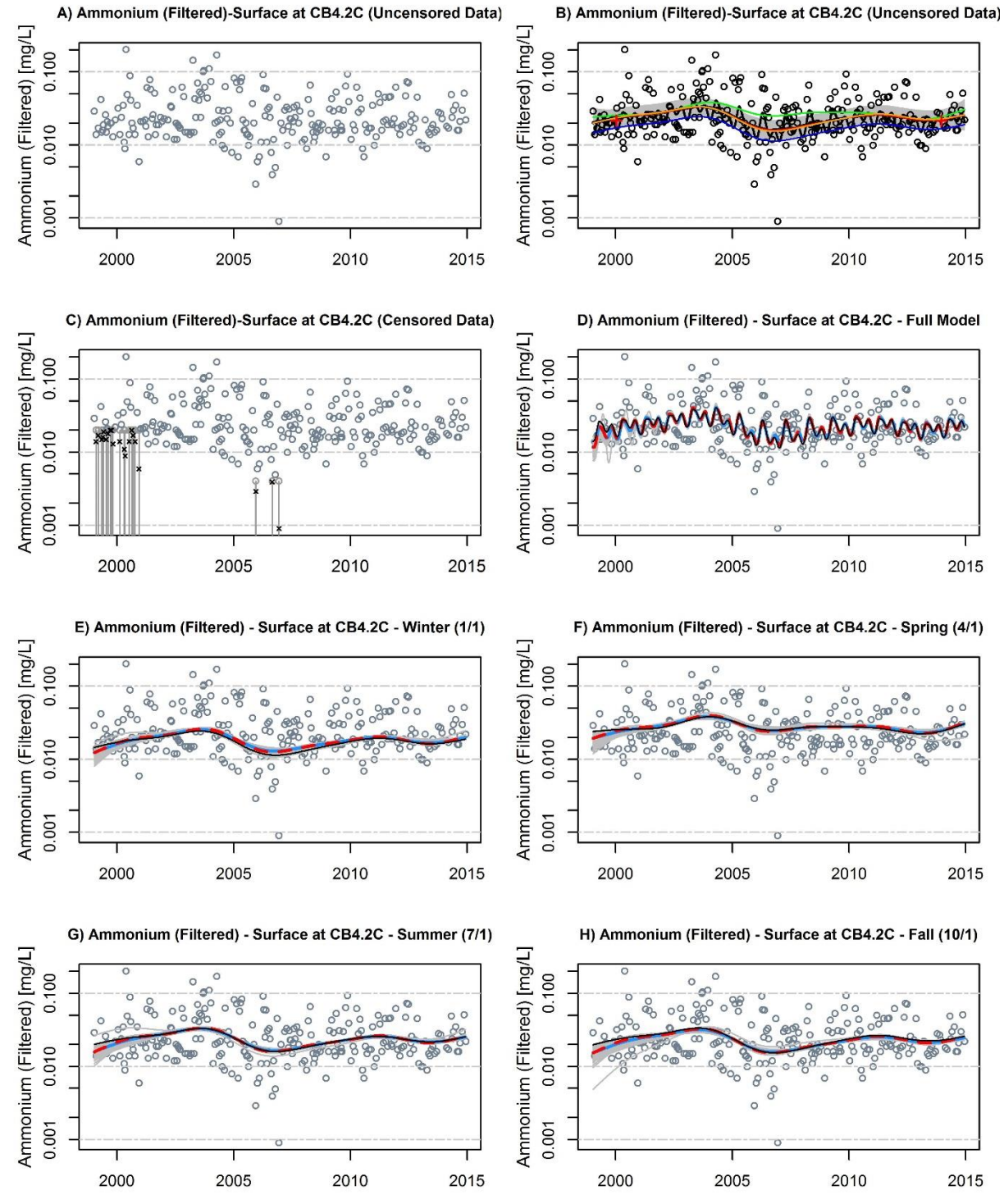
Appendix 2.6C

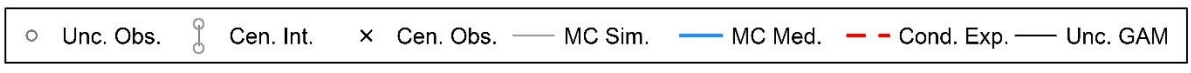
Scenario 3a



Appendix 2.6C

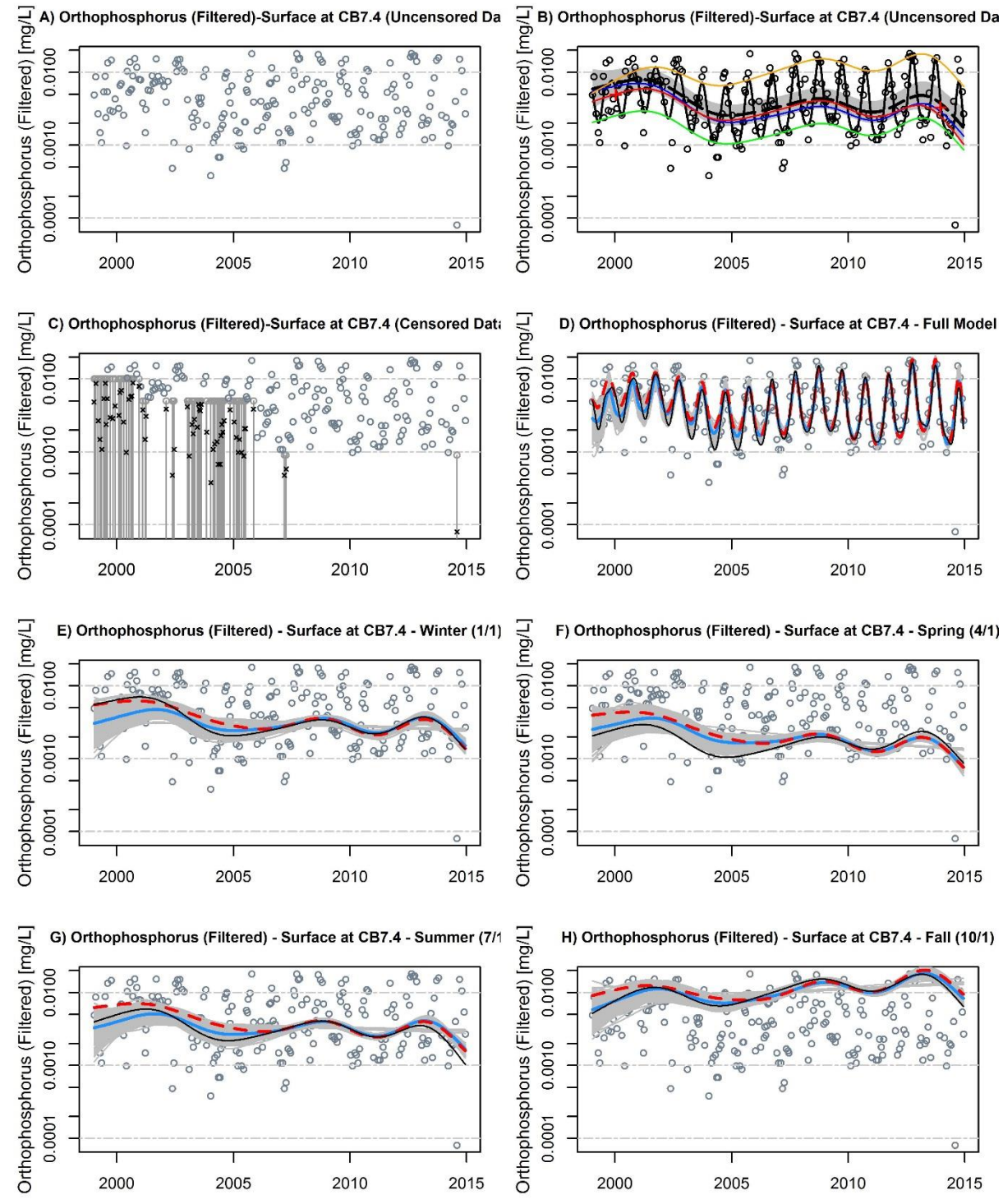
Scenario 3b





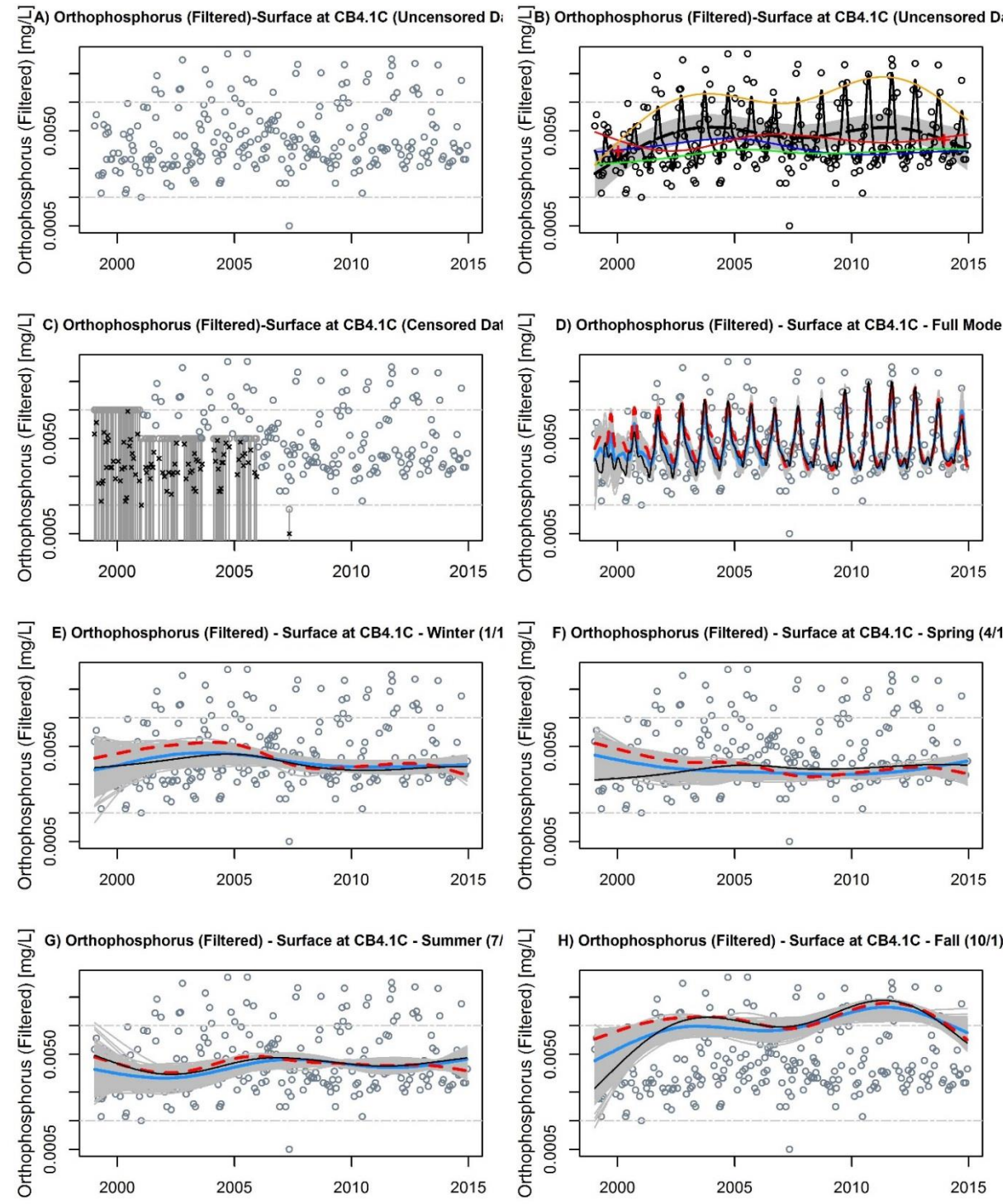
Appendix 2.6C

Scenario 3c



Appendix 2.6C

Scenario 3d



Appendix 2.6C

Scenario 3e