Supplementary Materials-1

Using Real-Time Area VOC Measurements to Estimate Total Hydrocarbons Exposures to Workers involved in the *Deepwater* Horizon Oil Spill

Gurumurthy Ramachandran, Caroline P. Groth, Tran B. Huynh, Sudipto Banerjee, Patricia A. Stewart, Lawrence S. Engel, Richard K. Kwok, Dale P. Sandler and Mark Stenzel

1 Selecting the Regression relationship

1.1 Overview

Our goal in Ramachandran et al. (this issue) was to supplement full-shift Total Hydrocarbons (THC) personal measurements with volatile organic compounds VOC time-weighted averages (TWAs), developed from VOC area measurements. To do so, we matched the VOC measurements to daily THC estimates derived from using a VOC:THC relationship. This supplemental material explains the steps taken to determine that relationship using regression. The steps were the following:

1. Determine whether censored measurements should be included in the regression.
2. Determine what VOC:THC relationship(s) should be used as a predictive equation in the regression, i.e. what exposure determinants should be included.
3. Determine the number of THC full-shift measurements required for a daily THC estimate to include in the regression.

We conducted a regression analysis to predict THC exposure estimates from VOC measurements using daily observations of the mean of multiple THC measurements and the mean of corresponding VOC estimates.

1.2 Justification of Non-censored Regressions

1.2.1 Background

Since some of the THC measurements were censored, we needed to consider whether we should include these censored observations (i.e. values below the limit of detection (LOD)) in the regression analysis. Groth et al. (2017) previously had developed a Bayesian regression analysis method that accounted for censored data using personal measurements. In that study, they found that slopes and intercepts may be dramatically different in an analysis with high censoring in the response (86% in the response:benzene, 11% in the predictor:THC) where censored data are included compared to an analysis where the censored observations are omitted.

If all measurements were observed (i.e. no censoring), estimating means at the daily level would be straightforward and simple. Daily VOC estimates were all considered observed as they previously had been estimated at the instrument-hour level (i.e. all instrument-hour TWAs had an estimate, even if the average was below the LOD, in which case the average was replaced by a value of 0.05 ppm) (Groth et al., this issue). Some THC measurements, however, were censored. Development of a daily average required averaging the personal measurements in a day. Calculating such daily averages in the presence of censored data proved troublesome and complicated. While it was possible to use a substitution value like the LOD for each censored measurement, we did not consider this to be a good solution as substitution techniques take away the uncertainty in the values below the LOD, which could be important. Alternatively, developing a model that estimated each censored value, averaged the values, and developed the regression relationship would have required significant additional validation and testing. We wanted to use a validated method (such as that in Groth et al. 2017, 2018) to understand the impact of censored values on the regressions. The use of the methods in Groth et al. (2017, 2018) allowed us to investigate censoring at the personal measurement level assuming the relationships would be similar at the daily level (which should be true based on mathematical properties when assuming similar censoring levels of the daily measurements as are present in the personal measurements).

Therefore, we investigated whether regressions including censored information in our VOC:THC analysis generated different intercept, slope and correlation estimates than regressions that omitted censored information. If we found limited differences, we would use non-censored data only, as that approach greatly simplified the analysis.

1.2.2 Analysis of THC Personal Measurements and Matched VOC Time-Weighted Averages

Specifically, we identified cases where both a THC personal measurement and VOC hourly estimates were present on the same vessel on the same day and had at least 4 hours of overlap. Those corresponding VOC hourly averages were averaged to develop a VOC TWA estimate that corresponded to the specific hours of the day each THC personal measurement was collected on the same vessel-day. Each set of matched THC and VOC TWAs were datapoints for our regression analysis. Some of these THC measurements, however, were censored, and some overlapped within a day (i.e. typically within a day, THC samples on different workers were collected and began and ended at different times throughout the day).

Next, we performed weakly informative Bayesian regressions including and omitting censored information (censored datapoints) for several vessels with VOC measurements, both on the individual vessel level including the rig vessels, i.e. the *Discoverer Enterprise (Enterprise)* and *Development Driller III (DDIII)*, and the vessels piloting remotely operated vehicles (referred to as ROV vessels), i.e. the *Boa Deep C, Ocean Intervention III, Skandi Neptune,* and *Boa Sub C*, and for all vessels together (overall N=169; 12 censored measurements). These vessels were selected as they had some of the highest number of matching THC and VOC TWAs and covered the largest number of days in TP1a-3 (see text for definition of time periods). Rig vessel measurements were included in this analysis, but not used in the final analysis of ROV and marine vessels described here. Weakly informative normal priors were used for the regression coefficients and wide inverse gamma priors were used for the variance components to allow the data to drive the inference.

For each regression, we obtained the median and 95% credible interval (CI) for the intercept, slope, and correlation coefficient. We were interested in identifying if the 95% CIs overlapped between analyses. Lack of overlap would suggest different inferences between the results; although, no statistical difference can be inferred directly since both the non-censored and the non-censored + censored databases contained the same non-censored data. Results of this analysis are included in Tables S1 and S2.

Table S1: All-Data (censored and non-censored) regressions using full-shift THC personal measurements and VOC TWAs.

		Intercept		Slope		Correlation	
Scenario	N	Median	95% CI	Median	95% CI	Median	95% CI
Overall (all Ships)	169	1.48	(0.62, 2.30)	0.76	(0.66, 0.87)	0.74	(0.67, 0.80)
Enterprise	52	3.78	(2.86, 4.68)	0.50	(0.39, 0.62)	0.80	(0.67, 0.88)
DDIII	26	6.28	(3.63, 8.84)	-0.01	(-0.42, 0.42)	-0.01	(-0.38, 0.39)
Boa Deep C	8	9.89	(9.10, 10.73)	-0.07	(-0.17, 0.02)	-0.54	(-0.89, 0.18)
Ocean Intervention III	17	-2.55	(-10.87, 6.69)	1.19	(0.18, 2.10)	0.59	(0.09, 0.84)
Skandi Neptune	49	-0.18	(-1.81, 1.37)	0.98	(0.79, 1.19)	0.84	(0.73, 0.91)
Boa Sub C	17	-0.22	(-3.19, 2.31)	0.75	(0.36, 1.21)	0.73	(0.39, 0.89)

THC=Total Hydrocarbons; VOC=Volatile Organic Compounds; N=number of measurements; CI=95%ile credible intervals; TWA=time-weighted average for the full work shift.

Table S2: Non-censored data regressions using full-shift THC personal measurements and VOC TWAs

		Intercept		Slope		Correlation	
Scenario	N	Median	95% CI	Median	95% CI	Median	95% CI
Overall (all Ships)	157	1.42	(0.37, 2.43)	0.77	(0.64, 0.90)	0.71	(0.62, 0.78)
Enterprise	48	3.26	(2.13, 4.35)	0.56	(0.43, 0.70)	0.80	(0.66, 0.88)
DDIII	23	8.02	(4.10, 11.84)	-0.27	(-0.86, 0.33)	-0.21	(-0.56, 0.25)
Boa Deep C	8	9.89	(9.10, 10.73)	-0.07	(-0.17, 0.02)	-0.54	(-0.89, 0.18)
Ocean Intervention III	17	-2.55	(-10.87, 6.69)	1.19	(0.18, 2.10)	0.59	(0.09, 0.84)
Skandi Neptune	49	-0.18	(-1.81, 1.37)	0.98	(0.79, 1.19)	0.84	(0.73, 0.91)
Boa Sub C	12	0.03	(-6.40, 6.61)	0.72	(-0.21,1.64)	0.48	(-0.13, 0.82)

THC=Total Hydrocarbons; VOC=Volatile Organic Compounds; N=number of measurements; CI=95%ile credible intervals; TWA=time-weighted average for the full work shift.

Comparing the results of these two analyses, we note that there is little difference between the slope coefficients reported in the all-data case vs in the non-censored data only case. This result is likely related to the fact that most observations were observed, rather than censored so the datasets did not differ greatly. All of the credible intervals overlapped. In some cases the regression was not significant (as shown on the *DDIII* and *BoaDeep C*), but nevertheless, we reached the same conclusion of a non-significant relationship in both analyses (shown by the interval not crossing 0). While censoring in this example was limited (7%), THC was not highly censored in the final dataset used to form the regression equations either on the ROV vessels (19%) or on the marine vessels (MVs) (12%). Thus, the slightly higher censoring levels were unlikely to substantially influence the findings above. These results suggested that including censored data may not be necessary in the VOC:THC evaluation as the relationships remained consistent when censored data were omitted.

1.2.3 Conclusion

Based on these results, we used only non-censored data in the regressions because there were not major or significant differences in the relationships when censored data were included. In addition, this solution was desired as it greatly simplified the methods going forward for regressing daily averages.

1.3 ROV Vessels and MV Relationships

1.3.1 Background

In the next step, we considered what determinants should be used to evaluate the VOC:THC relationship for the ROV vessels and the marine vessels (MVs) The two rig vessels, the *Enterprise* and *DDIII*, had relatively complete coverage of THC personal measurements over the duration of exposure. In addition, it was unclear how to assign VOC area measurements to individual jobs on the rigs (in contrast to the ROV vessels and MVs that estimated exposure on the vessel level, not the job level. Therefore, we chose not to supplement the THC information with VOC data on these vessels, but rather, chose to use them for validation). Possible determinants were the type of vessel (ROV vessel or MV), the vessel function, and the time period (1a,1b, 2, and 3, see Ramachandran et al., (this issue) for details) for which there were measurements.

To develop daily averages for the regression analyses, we used the THC measurements and their corresponding VOC TWAs. As indicated in section 2, we first identified the days for which there was a THC measurement and a VOC TWA (derived from multiple 1-hr estimates). We then excluded all pairs of measurements in which either the THC measurement or the VOC TWA was <4 hr in duration to be consistent with the criteria we used when developing other THC estimates (Huynh et al., this issue a, b, c) or there was <4 hr overlap. We also eliminated censored THC samples. Following this we averaged all THC measurements on each vessel-day with at least three THC measurements and then averaged the corresponding VOC TWAs in the same vessel-day to develop pairs of daily averages of THC and VOC. These THC-VOC daily average pairs were used to form the regression relationships described below (see sections 3-4).

1.3.2 ROV Results

Matching Dataset Descriptions

We developed daily averages on all 13 ROV vessels with THC data individually and on all ROV vessels combined. The following table provides the number of measurements present at each of the steps described in section 3.1 above. Two ROV vessels (*Helix Express* and *Iron Horse*) did not have any days on which both THC and VOC were available and two more vessels (*Casey Chouest* and *HOS Achiever*) did not have the required 4-hr overlap. Thus a total of 9 ROV vessels had 723 THC/VOC measurements that were available for development of daily averages. From these 723 measurements, we developed 160 vessel-day averages with at least 3 THC personal measurements per day.

Table S3: Counts of the THC measurements and daily averages of the THC:VOC pairs by ROV vessel. The totals (left-most column) reflect the sum by row across the columns. There were 160 daily averages with at least 3 THC measurements per vessel-day, some of which were on the same day but on different vessels.

		Boa Deep C	*Boa Sub C*	*Casey Chouset*	*Chouest Holiday*	*Helix Express*	*HOS Achiever*	*Iron Horse*	*Normand Fortress*	*Ocean Intervention I*	*Ocean Intervention III*	*Olympic Challenger*	*REM Forza*	*Skandi Neptune*	**Total**
Total THC Measurements	THC measurements	135	239	20	50	13	62	18	77	49	364	71	137	343	1578
	Matches on the same vessel day with VOC hourly averages	131	125	20	5	0	7	0	69	43	350	62	80	285	1177
	$\geq$ 4 hours overlap of VOC to THC on the same vessel-day	125	116	20	5	-	5	-	68	39	350	62	80	284	1154
	Non-censored	96	73	15	2	-	5	-	65	17	338	33	65	223	932
	Daily averages (N$\geq$ 3)	60	58	6	0	-	0	-	60	7	312	11	30	179	723
Daily Averages	All time periods	14	15	2	0	-	0	-	11	2	67	2	9	38	160
	Time period 1a (N$\geq$ 3)	0	1	0	-	-	-	-	0	0	4	0	0	4	9
	Time period 1b (N$\geq$ 3)	11	2	2	-	-	-	-	10	2	50	2	9	34	122
	Time period 2 (N$\geq$ 3)	3	12	0	-	-	-	-	1	0	10	0	0	0	26
	Time period 3 (N$\geq$ 3)	0	0	0	-	-	-	-		0	3	0	0	0	3
	Time periods 1a-1b (N$\geq$ 3)	11	3	2	-	-	-	-	10	2	54	2	9	38	131
	Time periods 2-3 (N$\geq$ 3)	3	12	0	-	-	-	-	1	0	13	0	0	0	29

THC=Total hydrocarbons; VOC=volatile organic chemicals; ROV vessel: vessel operating a remotely operated vehicle

We combined TP1a and 1b because they reflected the period of oil release and TP1a had a small number of observations (N=9). Similarly, we combined TP2 and 3, the period when the oil release was controlled (N=26 and N=3 respectively). We notice that for many vessels, we are left with less than 5 daily observations across all time periods. Specifically, only the *Boa Deep C*, *Boa Sub C*, *Normand Fortress*, *Ocean Intervention III*, *REM Forza*, and *Skandi Neptune* had more than 5 daily averages available. Individual regressions were run for these vessels but all vessel-day averages were used for the regressions of all ROV vessels combined.

Daily Average Regressions: Overall Regressions for All ROV Vessels

First we performed a series of simple linear regressions for each ROV vessel and for all ROV vessels combined. We performed a regression including all time periods combined and one for each of our two grouped time periods (TP1a-1b and TP2-3). We grouped ROV vessels by function, but results were similar to the all ROV vessels regression and are not shown.

These regressions allowed us to identify if a linear relationship was present and provided an R-squared, correlation, slope and intercept estimates. Results of these regressions are shown in Table S4.

Table S4: Regressions on the THC:VOC matched vessel-days for various ROV vessels and time periods. The mean, median, and 95% CI of the posterior samples are reported for the intercept, slope, and correlation. We also report the number of daily averages per regression (N) and the median posterior estimate of the R-squared. Regressions were not performed if less than 5 observations were available.

	Time		Intercept			Slope			Correlation			R-Squared
ROV(s)	Period(s)	N	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI	Median
All	All	160	-0.81	-0.80	(-1.84, 0.19)	1.01	1.01	(0.88, 1.15)	0.78	0.78	(0.71, 0.83)	0.61
All	1a-1b	131	-0.41	-0.40	(-1.54, 0.68)	0.98	0.98	(0.83, 1.13)	0.78	0.78	(0.70, 0.84)	0.61
All	2-3	29	5.78	5.81	(3.76, 7.72)	-0.15	-0.16	(-0.46, 0.17)	-0.19	-0.20	(-0.52, 0.21)	0.04
Boa Deep C	All	14	-0.87	-0.85	(-3.81, 2.04)	1.05	1.05	(0.66, 1.45)	0.84	0.86	(0.60, 0.95)	0.74
Boa Sub C	All	15	3.15	3.13	(0.22, 6.26)	0.29	0.29	(-0.21, 0.76)	0.32	0.34	(-0.23, 0.73)	0.12
Casey Chouset	All	2										
Chouest Holiday	All	0										
Helix Express	All	0										
HOS Achiever	All	0										
Iron Horse	All	0										
Normand Fortress	All	11	-0.78	-0.70	(-6.54, 4.84)	0.76	0.75	(-0.01, 1.55)	0.53	0.58	(-0.01, 0.87)	0.33
Ocean Intervention I	All	2										
Ocean Intervention III	All	67	-1.72	-1.70	(-3.33, -0.19)	1.13	1.13	(0.93, 1.35)	0.82	0.82	(0.72, 0.89)	0.68
Olympic Challenger	All	2										
REM Forza	All	9	-0.56	-0.55	(-2.84, 1.68)	0.93	0.93	(0.62, 1.25)	0.90	0.92	(0.68, 0.98)	0.85
Skandi Neptune	All	38	2.69	2.70	(1.17, 4.15)	0.65	0.65	(0.47, 0.84)	0.77	0.78	(0.60, 0.88)	0.60
Boa Deep C	1a-1b	11	-0.43	-0.42	(-4.34, 3.46)	1.01	1.00	(0.51, 1.51)	0.79	0.83	(0.44, 0.95)	0.68
Boa Sub C	1a-1b	3										
Normand Fortress	1a-1b	10	-2.08	-1.98	(-7.69, 3.19)	0.96	0.95	(0.23, 1.74)	0.67	0.71	(0.17, 0.92)	0.50
Ocean Intervention III	1a-1b	54	-0.78	-0.76	(-2.41, 0.77)	1.04	1.03	(0.84, 1.25)	0.83	0.83	(0.72, 0.90)	0.69
REM Forza	1a-1b	9	-0.56	-0.55	(-2.84, 1.68)	0.93	0.93	(0.62, 1.25)	0.90	0.92	(0.68, 0.98)	0.85
Skandi Neptune	1a-1b	38	2.69	2.70	(1.17, 4.15)	0.65	0.65	(0.47, 0.84)	0.77	0.78	(0.60, 0.88)	0.60
Boa Sub C	2-3	12	5.72	5.74	(1.68, 9.87)	-0.16	-0.17	(-0.87, 0.52)	-0.14	-0.17	(-0.67, 0.45)	0.06
Ocean Intervention III	2-3	13	5.82	5.86	(0.75, 11.08)	-0.14	-0.15	(-0.94, 0.63)	-0.11	-0.12	(-0.62, 0.47)	0.06

THC=Total hydrocarbons; VOC=volatile organic chemicals; ROV vessel: vessel operating a remotely operated vehicle; 95% CI=95% credible interval

Conclusion

The regression for all ROV vessels in TP1a-1b was credibly notable (credible interval does not include 0) at the 0.05 level for the slope and correlation coefficient and was associated with a high R-squared (0.61). Meanwhile, the slope and correlation coefficient for all ROV vessels in TP2-3 were statistically different from those of TP1a-1b (as the 95% CIs did not overlap for these independent groups), but was not statistically different from zero. See section 5 below for possible reasons for this difference. The R-squared of the individual vessel regressions ranged from 0.12 to 0.85. The vessel with the poorest relationship (*Boa Sub C*) had a R-squared of 0.12 (compared to the ROV vessel regression R-squared of 0.61). This vessel, however, contributed only 3 observations in TP1a-1b, the period of higher

R-squares, and 12 observations in TP2-3, the period of lower R-squares. The *Boa Sub C* had a R-squared of 0.06 for the latter time period, which was similar to the overall R-squared for TP2-3 (0.04).

1.3.3 MV Results

Matching Dataset Descriptions

We performed the same exclusion criteria to the THC measurements on the MVs as we did to the THC measurements on the ROV vessels (section 3.1.1). Many more MVs were excluded throughout this process. Two MVs had no THC measurements (*IOS Pipeliner* and *West Sirius*). Five more MVs were excluded because they did not have matching THC:VOC days (*Adriatic, Seacor Reliant, Seacor Vanguard, HOS Strongline,* and *War Admiral*). All remaining MVs had at least a 4-hr overlap and were non-censored. Finally, we included only vessel-days for which there were at least 3 THC:VOC pairs for the vessel-day. Seven more MVs were excluded for this reason (*Blue Dolphin, Overseas Cascade, Helix Producer, Loch Rannoch, Odyssea Diamond, Seacor Pride,* and *Seacor Venture*). As a result, 5 MVs remained, contributing only 75 THC measurements with which to develop daily averages. The following table provides the number of THC measurements present at each step in the development of the daily exposure metric. It also provides the number of daily averages available.

Table S5: Counts of THC measurements and daily averages of the THC:VOC pairs by MVs. The totals (left-most column) reflect the sums by row across the columns. There were 19 daily averages with at least 3 THC measurements per vessel-day in TP1a-1b and only 3 in TP2-3.

		Adriatic	*Blue Dolphin*	*Overseas Cascade*	*Helix Producer*	*IOS Pipeliner*	*Loch Rannoch*	*Massachusetts*	*Monica Ann*	*Normand Commander*	*Odyssea Diamond*	*Seacor Pride*	*Seacor Reliant*	*Seacor Vanguard*	*Seacor Venture*	*Stim Star III*	*HOS Strongline*	*Tyler Stephen*	*War Admiral*	*West Sirius*	**Total**
Total THC Samples	THC measurements	3	20	3	20	0	31	19	69	91	35	33	17	7	60	36	22	111	23	0	**600**
	Matches on the same vessel-day with VOC hourly averages	0	2	3	7	-	25	19	34	73	17	2	0	0	24	32	0	93	0	-	**331**
	$\geq$ 4 hours overlap of VOC to THC on the same vessel-day	-	2	3	7	-	4	19	34	71	17	2	-	-	23	31	-	92	-	-	**305**
	Non-censored	-	2	2	1	-	3	18	31	64	16	2	-	-	22	23	-	84	-	-	**268**
	Daily averages (N$\geq$ 3)	-	0	0	0	-	0	3	3	9	0	0	-	-	0	16	-	44	-	-	**75**
	All time periods	-	0	0	0	-	0	1	1	3	0	0	-	-	0	5	-	12	-	-	**22**
	Time periods 1a-1b (N$\geq$ 3)	-	-	-	-	-	-	1	1	2	-	-	-	-	-	3	-	12	-	-	**19**
	Time periods 2-3 (N$\geq$ 3)	-	-	-	-	-	-	-	-	1	-	-	-	-	-	2	-	-	-	-	**3**

THC=Total hydrocarbons; VOC=volatile organic chemicals; MV=marine vessel; TP=time period.

Conclusion
This low number of daily averages in TP1a-1b (N=19) and TP2-3 (N=3) likely resulted in unstable and possibly unrepresentative conclusions. We therefore investigated the possibility of increasing the number of daily matches by decreasing the number of measurements required for a day (Section 1.4).

1.4 Investigation of the Number of Measurements Per Day: ROV Vessel vs. MVs Regressions

1.4.1 Background

In this section, we evaluated regressions on the THC:VOC daily matches by vessel type and time period based on different numbers of observations (N) required per day to achieve sufficient sample sizes for stable regression results.

Due to the small number of days with $N \geq 3$ observation pairs per day on the MVs, we considered reducing the number of observations required to be included in a daily average by comparing regression results requiring at least 1, at least 2, and at least 3 observations per vessel-day. These regressions should not be compared statistically because the datasets are not the same and not independent, but if we see a higher R-squared statistic with sufficient sample size for a particular regression, intuitively, that finding may provide justification for choosing that regression and the corresponding requirement for a number of observations per day.

For each N per day, we report the overall relationship (all ROVs and MVs combined and separately for all time periods) and relationships by time period by vessel type. For each N per day, our table provides intercept, slope, and correlation posterior samples (mean, median, and 95% CI). We also report the median posterior estimate of R-squared.

1.4.2 Comparing Results: Number of THC:VOC Matches (N) Per Day

Regressions by N Per Day

Table S6: Comparison of N THC:VOC matches per day. The mean posterior estimate, median posterior estimate, and 95% credible interval (CI) are reported for the intercept, slope, and correlation parameters. The median posterior estimate of the R-squared is also provided.

N Per Day	Vessel Type	Time Periods	N	Intercept			Slope			Correlation			R-Squared
				Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI	
1	Marine and ROV	All	449	1.01	1.02	(0.29, 1.71)	0.71	0.71	(0.61, 0.82)	0.57	0.57	(0.51, 0.64)	0.33
1	ROV	All	306	0.52	0.53	(-0.26, 1.27)	0.82	0.81	(0.71, 0.93)	0.67	0.67	(0.60, 0.73)	0.45
1	Marine	All	143	4.50	4.50	(3.05, 5.88)	0.12	0.12	(-0.09, 0.34)	0.10	0.10	(-0.07, 0.28)	0.01
1	Marine and ROV	1a-1b	311	0.70	0.71	(-0.15, 1.52)	0.79	0.79	(0.68, 0.91)	0.63	0.63	(0.56, 0.70)	0.40
1	Marine and ROV	2-3	138	5.81	5.81	(4.74, 6.83)	-0.14	-0.14	(-0.30, 0.03)	-0.15	-0.15	(-0.32, 0.03)	0.02
1	ROV	1a-1b	231	0.10	0.11	(-0.81, 0.98)	0.89	0.89	(0.77, 1.01)	0.71	0.71	(0.64, 0.77)	0.51
1	Marine	1a-1b	80	5.11	5.13	(3.10, 7.02)	0.09	0.08	(-0.19, 0.38)	0.07	0.07	(-0.16, 0.31)	0.01
1	ROV	2-3	75	6.66	6.67	(5.24, 8.01)	-0.27	-0.27	(-0.48, -0.03)	-0.28	-0.29	(-0.49, -0.04)	0.08
1	Marine	2-3	63	4.25	4.27	(2.51, 5.89)	0.08	0.08	(-0.16, 0.35)	0.09	0.09	(-0.17, 0.35)	0.01
2	Marine and ROV	All	317	0.47	0.48	(-0.38, 1.27)	0.80	0.80	(0.69, 0.92)	0.63	0.63	(0.56, 0.69)	0.40
2	ROV	All	223	-0.10	-0.09	(-1.04, 0.80)	0.90	0.90	(0.78, 1.03)	0.71	0.71	(0.63, 0.77)	0.50
2	Marine	All	94	4.47	4.49	(2.67, 6.17)	0.15	0.15	(-0.10, 0.42)	0.13	0.13	(-0.08, 0.34)	0.02
2	Marine and ROV	1a-1b	143	0.56	0.57	(-0.43, 1.51)	0.81	0.81	(0.69, 0.95)	0.65	0.65	(0.56, 0.72)	0.42
2	Marine and ROV	2-3	79	5.17	5.19	(3.65, 6.62)	-0.04	-0.04	(-0.26, 0.20)	-0.04	-0.04	(-0.27, 0.21)	0.01
2	ROV	1a-1b	178	0.06	0.07	(-1.02, 1.09)	0.90	0.90	(0.76, 1.04)	0.71	0.72	(0.63, 0.78)	0.51
2	Marine	1a-1b	60	4.70	4.74	(2.41, 6.86)	0.15	0.15	(-0.16, 0.49)	0.13	0.13	(-0.13, 0.39)	0.02
2	ROV	2-3	45	5.30	5.32	(3.44, 7.05)	-0.07	-0.07	(-0.35, 0.23)	-0.08	-0.08	(-0.37, 0.24)	0.02
2	Marine	2-3	34	5.31	5.33	(2.61, 7.80)	-0.04	-0.04	(-0.42, 0.37)	-0.04	-0.04	(-0.37, 0.33)	0.02
3	Marine and ROV	All	182	-0.60	-0.60	(-1.61, 0.37)	0.97	0.97	(0.84, 1.11)	0.75	0.75	(0.68, 0.81)	0.51
3	ROV	All	160	-0.81	-0.8	(-1.84, 0.19)	1.01	1.01	(0.88, 1.15)	0.78	0.78	(0.71, 0.83)	0.61
3	Marine	All	22	3.66	3.70	(-0.10, 7.38)	0.31	0.31	(-0.21, 0.85)	0.26	0.27	(-0.17, 0.64)	0.08
3	Marine and ROV	1a-1b	150	-0.19	-0.18	(-1.29, 0.89)	0.94	0.94	(0.80, 1.08)	0.75	0.75	(0.67, 0.82)	0.57
3	Marine and ROV	2-3	32	4.75	4.77	(2.47, 6.95)	0.02	0.01	(-0.33, 0.38)	0.02	0.01	(-0.34, 0.38)	0.02
3	ROV	1a-1b	131	-0.41	-0.40	(-1.54, 0.68)	0.98	0.98	(0.83, 1.13)	0.78	0.78	(0.70, 0.84)	0.61
3	Marine	1a-1b	19	4.76	4.76	(1.23, 8.49)	0.17	0.17	(-0.35, 0.67)	0.16	0.17	(-0.31, 0.58)	0.04
3	ROV	2-3	29	5.78	5.81	(3.76, 7.72)	-0.15	-0.16	(-0.46, 0.17)	-0.19	-0.20	(-0.52, 0.21)	0.04
3	Marine	2-3	3	-	-	-	-	-	-	-	-	-	-

N=Number of THC:VOC pairs required for a daily average; THC=total hydrocarbons; VOC=volatile organic chemicals; ROV=vessel operating remotely operated vehicle; 95% CI=95% credible interval

1.4.3 Conclusion

The R-squared values increased as we increased the N required per day and were highest for the ROV vessels in TP1a-1b; regressions for the ROVs in TP1a-1b requiring at least 1, 2, and 3 observations were associated with the highest R-squared values observed (0.51, 0.51, and 0.61, respectively). The R-squared values in time periods 2-3 were all quite small and did not have credibly notable correlations. In addition, the MV R-squared values were small (< 0.04) whether the time period was 1a-1b, 2-3, or all time periods combined. We therefore retained 3 observations as the minimum number of observations per vessel-day to include in the regression.

1.5 Final Conclusion

The results generally suggest that type of vessel and time period were major determinants of the relationship between THC and VOC. While we observed that some individual ROV and "All" ROV vessels had strong linear relationships (with very high R-squared values), MVs generally had too few observations to allow us to perform regressions by specific vessels and insufficient observations even for "All" MVs in TP2-3. In addition, the MV R-squared value for TP1a-1b was low (0.04). This finding suggests that regressions for the MVs provided little information.

The higher R-squared values for the ROVs in TP1a-1b, the time periods when the oil was being released in large quantities, suggests that the oil was the primary source of THC/VOC exposure. In contrast, in TP2-3, when the oil was contained and therefore likely contributing much less to the atmospheric burden of THC/VOC, there may have been other sources of THC or VOC, such as cleaning chemicals used in the decontamination of the vessel, that may have varied by vessel or fuel. These other sources may have overwhelmed or at least added substantially to, the contribution of the remaining oil vapors in the area. The "All" ROV vessels TP2-3 AM (0.54 ppm) is about 10% of the "All" ROV vessels TP1a-1b AM (5.35 ppm) (Table 3, Ramachandran et al. this issue). Error in the assignment of the lower AM would likely have less of an impact on any disease risk estimate in the epidemiologic analysis than the error in the assignment of the higher AM because there are a greater number of study participants at the lower exposure levels in the study (so the error is more diluted within the larger group than it would be in the smaller group) than at the higher levels.

It is reasonable that the higher R-squared value was associated with the higher number of required observations per day, as more observations (i.e. workers measured for THC) likely covered more of the locations on the vessels monitored by the stationary VOC instruments.

It is unclear why the relationship between THC and VOC for the MVs in TP1a-1b was not stronger. Figure 3 (Ramachandran et al., this issue) shows that the measurement values were substantially lower on the MVs than on the ROV vessels. It may be that the lower contrast in exposures resulted in the lower correlations on the MVs.

Therefore, we used the relationship in TP1a-1b on"All" ROV vessels as the prediction equation for estimating THC values from VOC for both the ROV vessels and the MVs for all time periods. This regression had a very high R-squared value of 0.61 (which was the highest R-squared value found of the relationships investigated at the "All" ROV vessel level). The lower R-squared for the ROV vessels in TP2-3 and the MVs over all time periods indicate that measurement error from using "All" ROVs may be greater for the estimates for the ROV vessels in TP2-3 and the MVs in all time periods than for the estimates for the ROV vessels in TP1a-1b.

2 Extra Tables/Figures

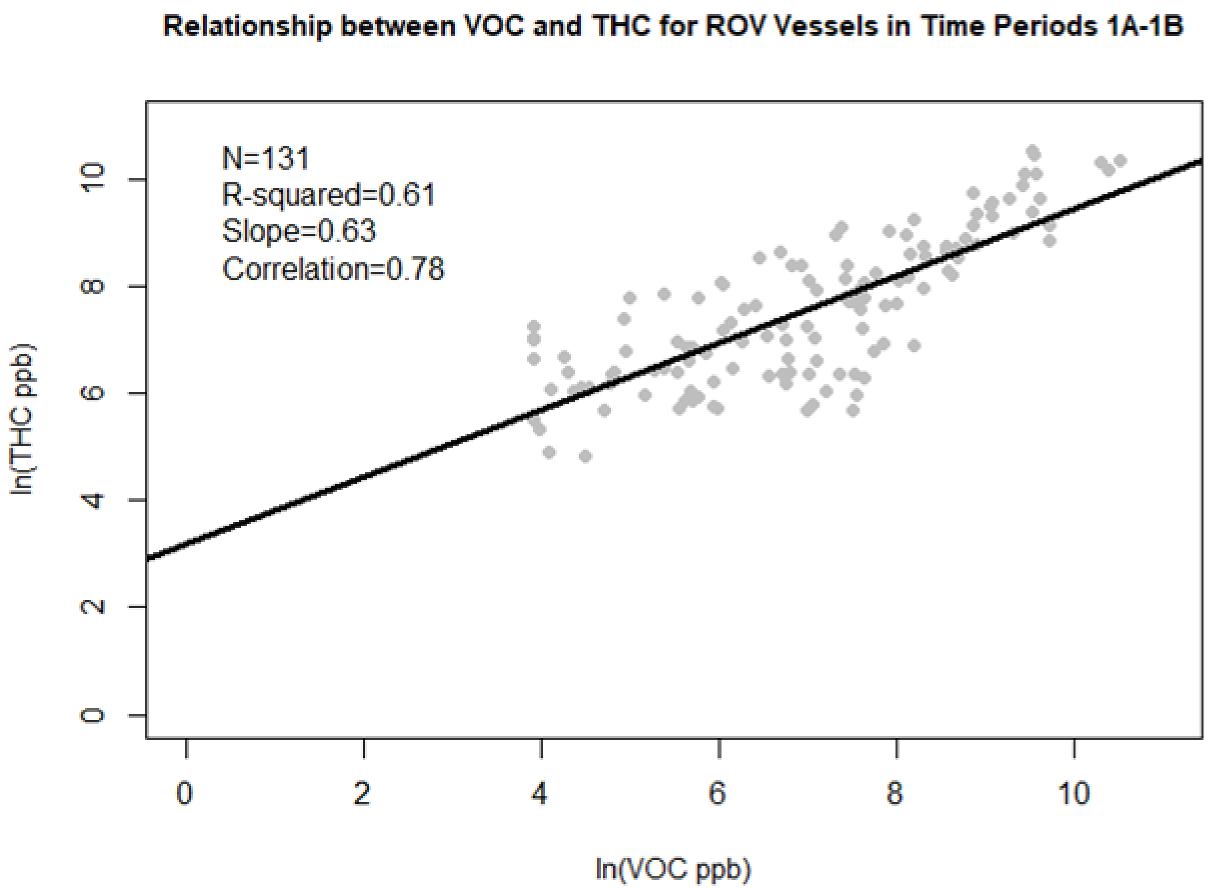

Figure S1: Regression relationship between the natural logarithms of volatile organic compounds (VOC) and total hydrocarbons (THC) concentrations (in ppb) for time periods (TP)1a-1b for ROV vessels, which was used as the THC prediction equation.

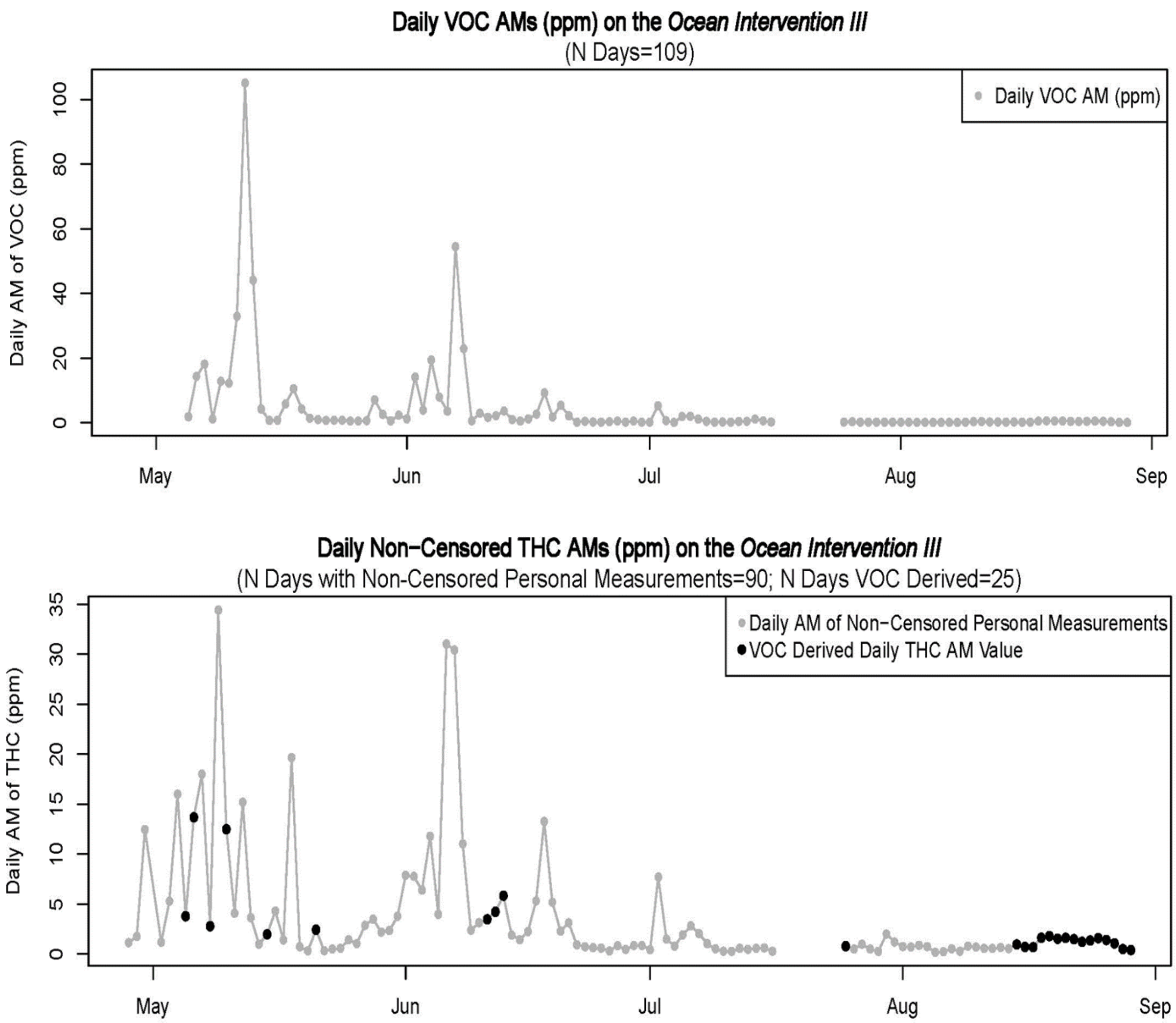

Figure S2: Daily volatile organic compounds (VOC) and total hydrocarbons (THC) arithmetic mean (AM) estimates over time for the ROV vessel *Ocean Intervention III* with the number (N) of days. Only THC non-censored measurements are plotted. For a day to be included in the regression equation for prediction, at least three non-censored THC personal measurements must be present on the same day as at least 4 hours of VOC information was available. If any THC samples (censored or non-censored) were available on a day, VOC was not used to impute THC.

Vessel Type	N Vessel Days	Source of Variation	Df	Sum of Squares	Mean Squares (Variance)	GSD
ROV Vessels	618	Between Vessel Days	617	25913	42.00	3.89
		Within Vessel Days	13129	12601	0.96	2.66
Marine Vessels	422	Between Vessel Days	421	9753	23.17	2.81
		Within Vessel Days	8494	5736	0.68	2.27

ROV Vessels: Vessels that operated remotely operated vehicles; N: number; Df: degrees of freedom

Table S7: Analysis of between- and within-vessel-day variability for ROV vessels and marine vessels.

Scenario	Intercept		Slope		Correlation		
	Median	95% CI	Median	95% CI	Median	95% CI	R^2
ROV Vessels	3.17	(2.59, 3.76)	0.63	(0.54, 0.71)	0.78	(0.71, 0.84)	0.61
Rigs	3.30	(2.41, 4.18)	0.55	(0.43, 0.67)	0.73	(0.61, 0.82)	0.53

ROV Vessels: Vessels that operated remotely operated vehicles; CI: Credible interval

Table S8: Relationships identified on the ROV vessels and two of the rig vessels (*Discoverer Enterprise* and *Development Driller III*). The median posterior sample and 95% credible interval (CI) are reported for the intercept, slope, and correlation along with the median posterior estimate of the R^2.

3 References

Groth, C, Banerjee, S, Ramachandran, G, *et al.* (2017) Bivariate left-censored Bayesian Model for predicting exposure: Preliminary analysis of worker exposure during the *Deepwater Horizon* oil spill. *Ann Work Expo Health*; 81 (1):76-86.

Groth CP, Banerjee S, Ramachandran G, *et al.* (2018) Multivariate left-censored Bayesian modeling for predicting exposure using multiple chemical predictors. *Environmetrics*; 29: 1-16 doi: 10.1002/env.2505

Groth, CP, Ramachandran, G, Banerjee, S,*et al.* (this issue) Methods for the analysis of 26 million VOC area measurements during the *Deepwater Horizon* oil spill clean-up. *Ann Work Expo Health*

Huynh, TB, Groth, CP, Ramachandran, G, *et al.*(this issue) Estimates of occupational inhalation exposures on the four rig vessels during the *Deepwater Horizon* oil spill and clean-up. *Ann Work Expo Health*

Huynh, TB, Groth CP, Ramachandran G, *et al.*(this issue) Estimates of inhalation exposures to oil-related components on the supporting vessels during the *Deepwater Horizon* oil spill. *Ann Work Expo Health*

Huynh, TB, Groth CP, Ramachandran G, et al. *et al.* (this issue) Estimates of inhalation exposures during the *Deepwater Horizon* oil spill and clean-up operations on land. *Ann Work Expo Health*

Ramachandran G, Groth CP, Huynh TB, *et al.* (this issue) Using real-time area VOC measurements to estimate total hydrocarbon exposures to workers in the *Deepwater Horizon* oil spill. *Ann Work Expo Health*