SUPPLEMENTAL INFORMATION

CyanoHABs and CAPs: assessing community-based monitoring of fine particulate matter with regional sources of pollution in rural, northeastern North Carolina

Haley E. Plaas[1,*], Colleen Karl[2], Rachael Cogbill[3], Nicole Rosales-Garcia[1], Ashley H. Stoop[4], Lisa L. Satterwhite[5], Martine E. Mathieu-Campbell[6], Jennifer Richmond-Bryant[6,7], Hans W. Paerl[3,8,9], and Douglas S. Hamilton[1]

[1]Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, USA
[2]Chowan Edenton Environmental Group, Tyner, NC 27980, USA
[3]Department of Environmental Sciences and Engineering, UNC-Chapel Hill, Chapel Hill, NC 27599, USA
[4]Albemarle Regional Health Services, Elizabeth City, NC 27909, USA
[5]Department of Civil and Environmental Engineering, Duke University, Durham, NC 27708, USA
[6]Center for Geospatial Analytics, North Carolina State University, Raleigh, NC 27695, USA
[7]Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA [8]Department of Earth, Marine, and Environmental Sciences, UNC-Chapel Hill, Chapel Hill, NC 27599, USA
[9]Institute of Marine Sciences, UNC-Chapel Hill, Morehead City, NC 28557, USA

*Corresponding Author: Haley E. Plaas (heplaas@ncsu.edu; 2800 Faucette Dr.,1142 Jordan Hall, NC State University, Raleigh, NC 27695)

October 2024

No. of SI Pages: 12

No. of SI Figures: 5

No. of SI Tables: 4

No. of SI Equations: 0

No. of SI Schemes: 0

Table S1. Descriptive statistics for composited (weekly) satellite data surrounding each PurpleAir sensor.

Sensor ID	Pixel Grid 'radius'	Water Surface Area	Data Completeness
1806	0.62 km	25%	99 ± 3%
1806	1.2 km	40%	96 ± 3%
1806	2.4 km	66%	98 ± 4%
1680	0.62 km	50%	96 ± 12%
1680	1.2 km	47%	97 ± 10%
1680	2.4 km	43%	98 ± 8%
1378	0.62 km	56%	96 ± 12%
1378	1.2 km	51%	97 ± 9%
1378	2.4 km	38%	97 ± 6%
5822	0.62 km	50%	95 ± 12%
5822	1.2 km	47%	97 ± 10%
5822	2.4 km	37%	98 ± 7%
1334	0.62 km	39%	97 ± 9%
1334	1.2 km	44%	97 ± 9%
1334	2.4 km	37%	98 ± 7%
1362	0.62 km	33%	98 ± 8%
1362	1.2 km	26%	97 ± 5%
1362	2.4 km	19%	98 ± 3%
1348	0.62 km	22%	96 ± 6%
1348	1.2 km	22%	97 ± 4%
1348	2.4 km	17%	99 ± 3%
5838	0.62 km	8%	98 ± 3%
5838	1.2 km	16%	98 ± 3%
5838	2.4 km	14%	99 ± 2%
9875	0.62 km	22%	89 ± 8%
9875	1.2 km	18%	93 ± 6%
9875	2.4 km	13%	96 ± 3%
1318	0.62 km	0%	100 ± 0%
1318	1.2 km	4%	98 ± 2%
1318	2.4 km	8%	98 ± 2%
1344	0.62 km	0%	100 ± 0%
1344	1.2 km	0%	100 ± 0%
1344	2.4 km	3%	97 ± 1%
1358	0.62 km	0%	100 ± 0%
1358	1.2 km	0%	100 ± 0%
1358	2.4 km	0%	100 ± 0%
1562	0.62 km	0%	100 ± 0%
1562	1.2 km	0%	100 ± 0%
1562	2.4 km	0%	100 ± 0%

Tables S2. Outline of the simplified generalized additive models performed to examine $PM_{2.5}$ mass as a function of the univariate and mixed effects of **AQS-measured CAPs pollutants**. Each row represents a model run with the descriptor variables specified in each column. An x indicates the variable was treated as an independent smooth term in the model equation. Multiple x's per row indicate the variables were considered for their additive effects. Paired letters (a-b-c) indicate covariates treated with an interactive term. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are provided for inter-model comparison.

Table S2.

Variable	CO	O_3	PO_4	SO_2	NO_x	AIC	BIC
$PM_{2.5}$ ~				x		2808	2844
$PM_{2.5}$ ~			x			2879	2920
$PM_{2.5}$ ~	x					3394	3443
$PM_{2.5}$ ~					c	2885	2930
$PM_{2.5}$ ~		x				2416	2453
$PM_{2.5}$ ~		x		x		2266	2333
$PM_{2.5}$ ~		x	x			2284	2351
$PM_{2.5}$ ~		x			x	2322	2392
$PM_{2.5}$ ~	x	x				2052	2132
$PM_{2.5}$ ~		a			a	2282	2362
$PM_{2.5}$ ~	b	x		b		1772	1910
$PM_{2.5}$ ~	x	x		x		1854	1971
$PM_{2.5}$ ~	x	a		x	a	1851	2000
$PM_{2.5}$ ~	b	x		b	x	1741	1891
$PM_{2.5}$ ~	b	x	x	b	x	1580	1784

Tables S3. Outline of the simplified generalized additive models performed to examine $PM_{2.5}$ mass as a function of the univariate and mixed effects of **meteorological conditions**. Each row represents a model run with the descriptor variables specified in each column. An x indicates the variable was treated as an independent smooth term in the model equation. Multiple x's per row indicate the variables were considered for their additive effects. Paired letters (a-b-c) indicate covariates treated with an interactive term. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are provided for inter-model comparison.

Variable	Wind Speed	Wind Direction	Solar Radiation	Temperature	Relative Humidity	AIC	BIC
$PM_{2.5}$ ~	x					3509	3527
$PM_{2.5}$ ~		x				3516	3541
$PM_{2.5}$ ~			x			3422	3455
$PM_{2.5}$ ~				x		3408	3435
$PM_{2.5}$ ~					x	3516	3541
$PM_{2.5}$ ~				a	a	3378	3446
$PM_{2.5}$ ~				x	x	3400	3445
$PM_{2.5}$ ~			b	b		3371	3441
$PM_{2.5}$ ~			x	x		3389	3432
$PM_{2.5}$ ~	c	c				3487	3556
$PM_{2.5}$ ~	x	x				3506	3541
$PM_{2.5}$ ~			x	a	a	3367	3511
$PM_{2.5}$ ~			b	b	X	3366	3453
$PM_{2.5}$ ~		x	x	a	a	3361	3457
$PM_{2.5}$ ~	x		x	a	a	3347	3444
$PM_{2.5}$ ~	x	x	x	a	a	3347	3452

Tables S4. Outline of the simplified generalized additive models performed to examine $PM_{2.5}$ mass as a function of the univariate and mixed effects of **satellite-derived CyanoHAB indicators**. Each row represents a model run with the descriptor variables specified in each column. An x indicates the variable was treated as an independent smooth term in the model equation. Multiple x's per row indicate the variables were considered for their additive effects. Paired letters (a-b-c) indicate covariates treated with an interactive term. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are provided for inter-model comparison.

Variable	CyanoHAB intensity 0.6 km	CyanoHAB intensity 1.2 km	CyanoHAB intensity 2.4 km	CyanoHAB Spatial Extent 0.6 km	CyanoHAB Spatial Extent 1.2 km	CyanoHAB Spatial Extent 2.4 km	AIC	BIC
$PM_{2.5}$ ~	x						1347	1361
$PM_{2.5}$ ~		x					1346	1361
$PM_{2.5}$ ~			x				1345	1360
$PM_{2.5}$ ~				x			1349	1363
$PM_{2.5}$ ~					x		1348	1362
$PM_{2.5}$ ~						x	1346	1361
$PM_{2.5}$ ~	a			a			1347	1369
$PM_{2.5}$ ~		b			b		1342	1364
$PM_{2.5}$ ~			c			c	1343	1365
$PM_{2.5}$ ~	d	d	d				1341	1377

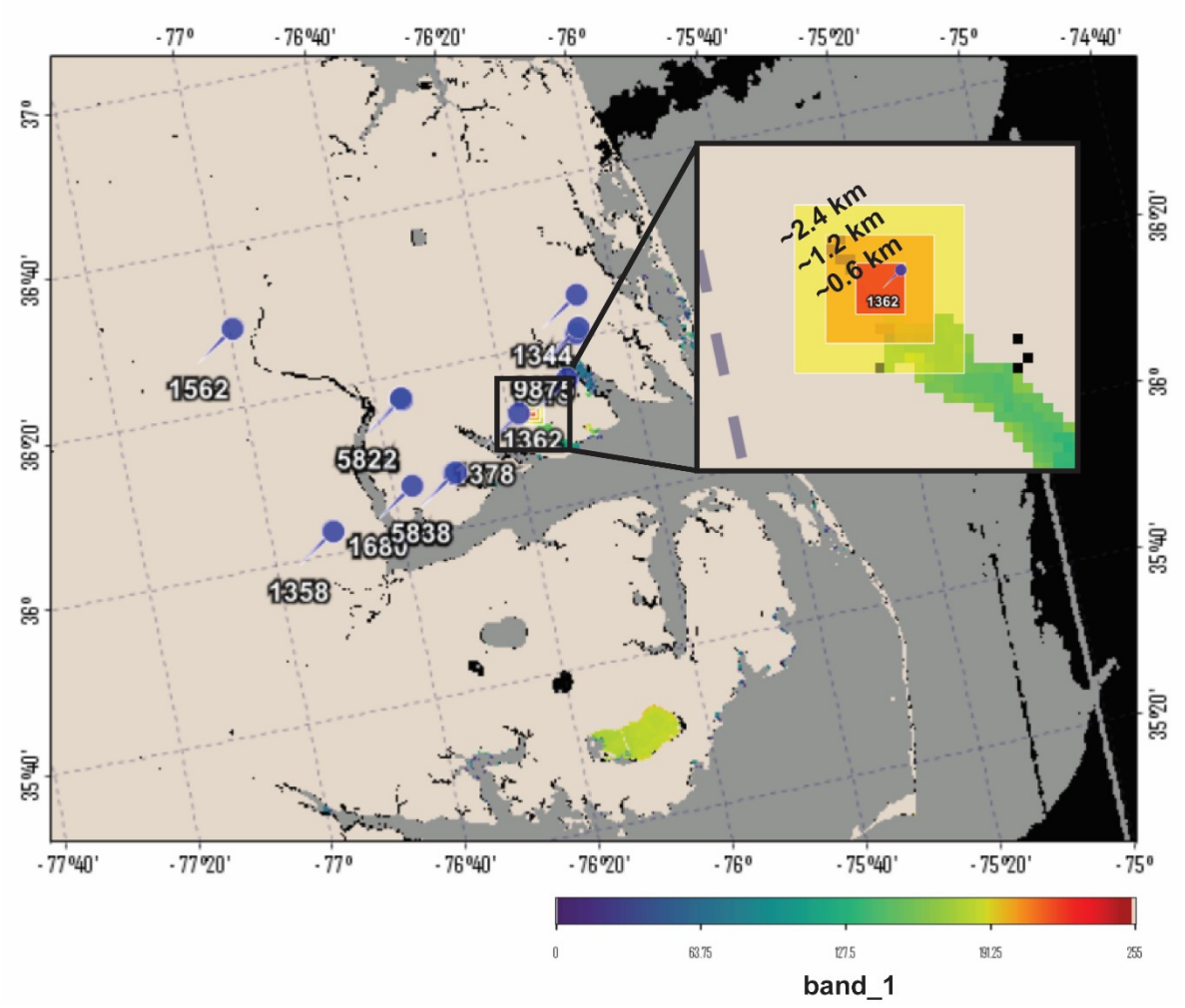
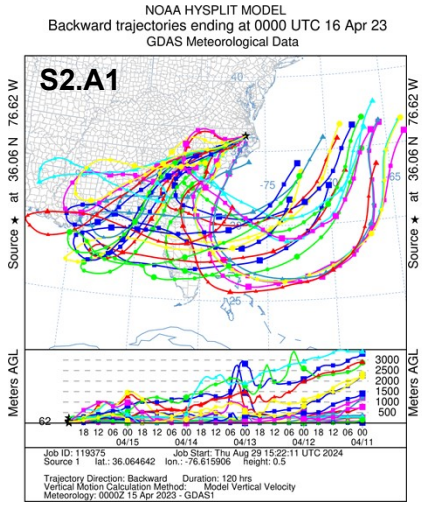
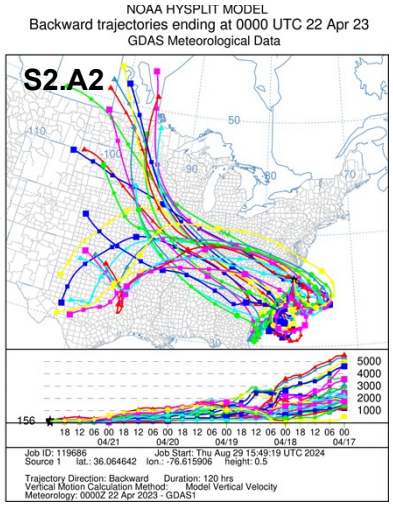
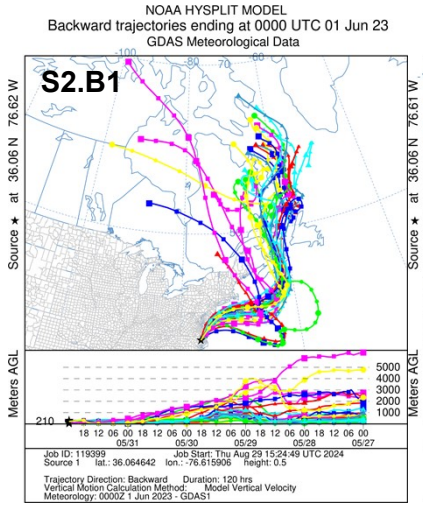


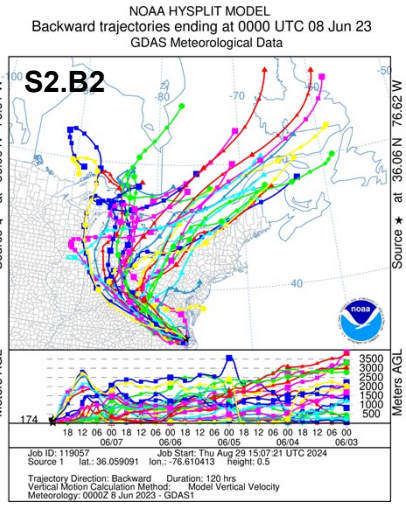
Figure S1. A SeaDAS map with pixel grid overlays at each distance (0.6, 1.2, and 2.4 km from the center pixel) over which cyanobacterial harmful algal bloom cell counts and surface area was composited.

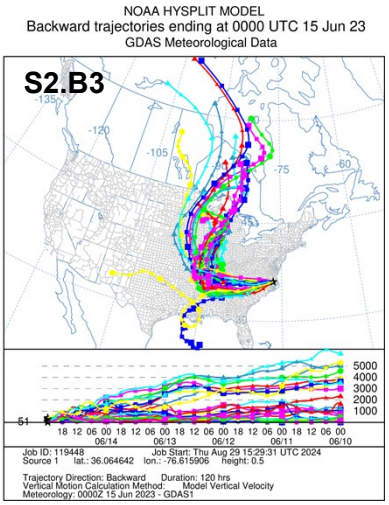

NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 16 Apr 23
GDAS Meteorological Data
S2.A1
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119375 Job Start: Thu Aug 29 15:22:11 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 15 Apr 2023 - GDAS1


NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 22 Apr 23
GDAS Meteorological Data
S2.A2
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119686 Job Start: Thu Aug 29 15:49:19 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 22 Apr 2023 - GDAS1

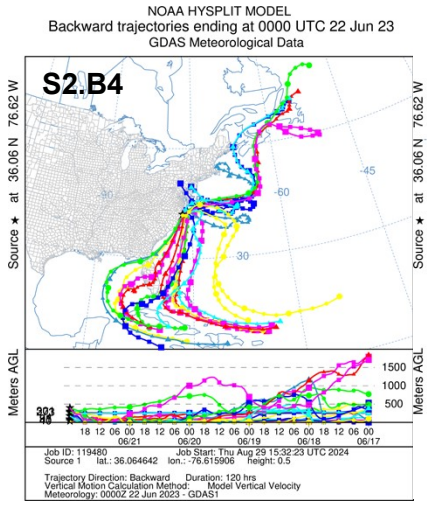
74

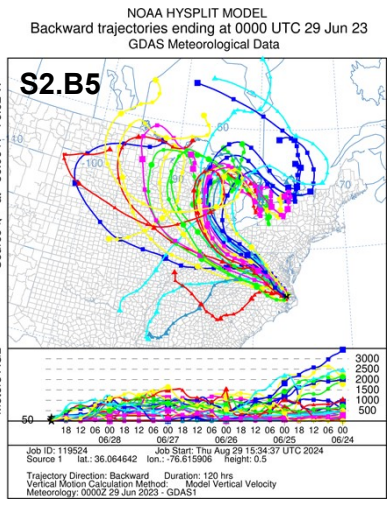

NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 01 Jun 23
GDAS Meteorological Data
S2.B1
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119399 Job Start: Thu Aug 29 15:24:49 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 1 Jun 2023 - GDAS1


NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 08 Jun 23
GDAS Meteorological Data
S2.B2
Source ★ at 36.06 N 76.61 W
Meters AGL
Job ID: 119057 Job Start: Thu Aug 29 15:07:21 UTC 2024
Source 1 lat.: 36.059091 lon.: -76.610413 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 8 Jun 2023 - GDAS1


NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 15 Jun 23
GDAS Meteorological Data
S2.B3
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119448 Job Start: Thu Aug 29 15:29:31 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 15 Jun 2023 - GDAS1

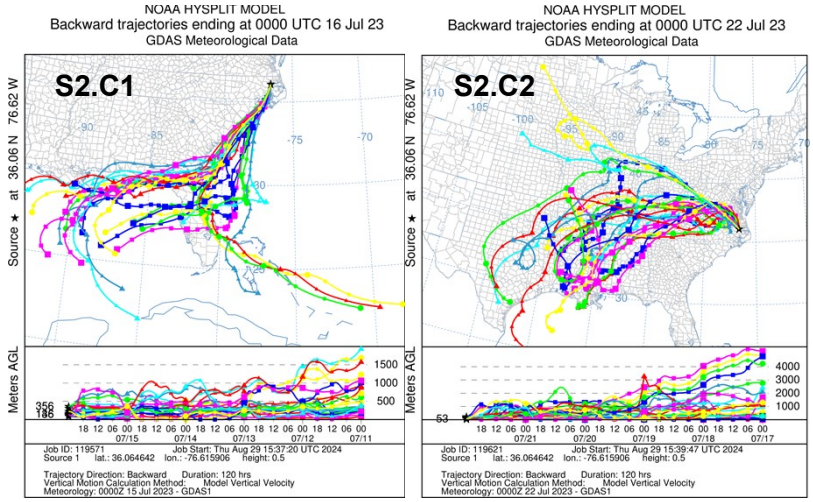
75


NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 22 Jun 23
GDAS Meteorological Data
S2.B4
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119480 Job Start: Thu Aug 29 15:32:23 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 22 Jun 2023 - GDAS1

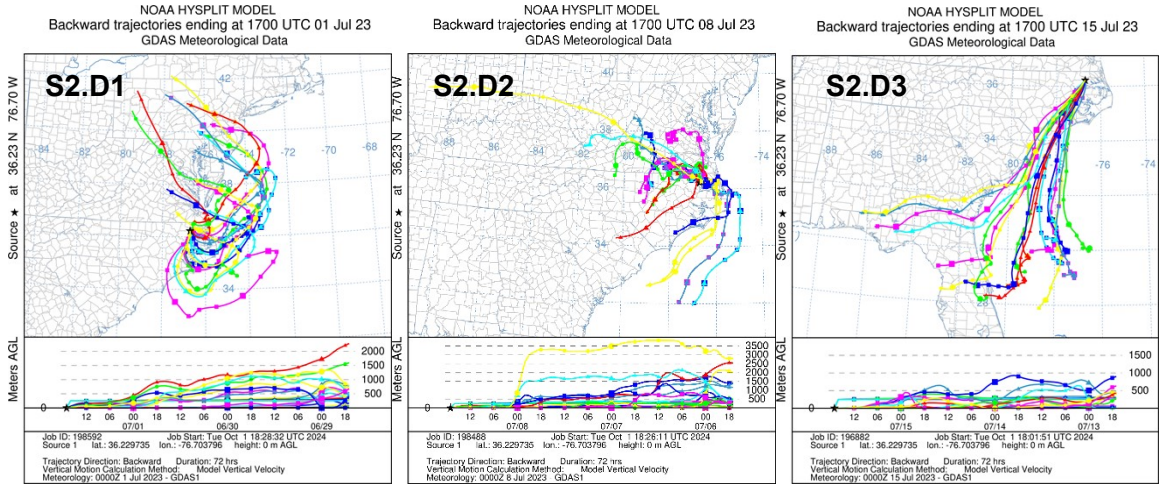

NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 29 Jun 23
GDAS Meteorological Data
S2.B5
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119524 Job Start: Thu Aug 29 15:34:37 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 29 Jun 2023 - GDAS1

76

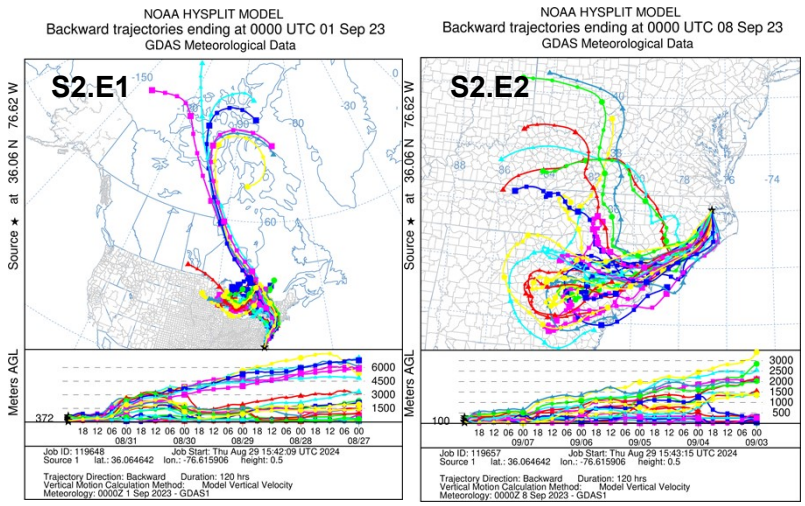
77


NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 16 Jul 23
GDAS Meteorological Data
S2.C1
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119571 Job Start: Thu Aug 29 15:37:20 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 15 Jul 2023 - GDAS1
NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 22 Jul 23
GDAS Meteorological Data
S2.C2
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119621 Job Start: Thu Aug 29 15:39:47 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 22 Jul 2023 - GDAS1

78


NOAA HYSPLIT MODEL
Backward trajectories ending at 1700 UTC 01 Jul 23
GDAS Meteorological Data
S2.D1
Source ★ at 36.23 N 76.70 W
Meters AGL
Job ID: 198592 Job Start: Tue Oct 1 18:28:32 UTC 2024
Source 1 lat.: 36.229735 lon.: -76.703796 height: 0 m AGL
Trajectory Direction: Backward Duration: 72 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 1 Jul 2023 - GDAS1
NOAA HYSPLIT MODEL
Backward trajectories ending at 1700 UTC 08 Jul 23
GDAS Meteorological Data
S2.D2
Source ★ at 36.23 N 76.70 W
Meters AGL
Job ID: 198488 Job Start: Tue Oct 1 18:28:11 UTC 2024
Source 1 lat.: 36.229735 lon.: -76.703796 height: 0 m AGL
Trajectory Direction: Backward Duration: 72 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 8 Jul 2023 - GDAS1
NOAA HYSPLIT MODEL
Backward trajectories ending at 1700 UTC 15 Jul 23
GDAS Meteorological Data
S2.D3
Source ★ at 36.23 N 76.70 W
Meters AGL
Job ID: 196882 Job Start: Tue Oct 1 18:01:51 UTC 2024
Source 1 lat.: 36.229735 lon.: -76.703796 height: 0 m AGL
Trajectory Direction: Backward Duration: 72 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 15 Jul 2023 - GDAS1

79


NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 01 Sep 23
GDAS Meteorological Data
S2.E1
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119648 Job Start: Thu Aug 29 15:42:09 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 1 Sep 2023 - GDAS1
NOAA HYSPLIT MODEL
Backward trajectories ending at 0000 UTC 08 Sep 23
GDAS Meteorological Data
S2.E2
Source ★ at 36.06 N 76.62 W
Meters AGL
Job ID: 119657 Job Start: Thu Aug 29 15:43:15 UTC 2024
Source 1 lat.: 36.064642 lon.: -76.615906 height: 0.5
Trajectory Direction: Backward Duration: 120 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Meteorology: 0000Z 8 Sep 2023 - GDAS1

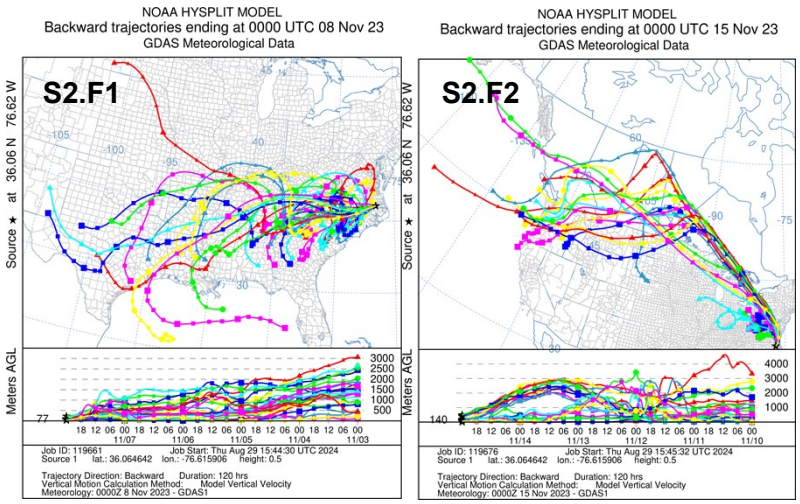


Figure S2. NOAA Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model results for the beginning and end of each week when episodic air pollution events were identified, indicated by $PM_{2.5}$ values that exceeded seasonal averages by > 1 standard deviation. A1-2 indicates the pollution event in April, B1-5 three events in June, C1-2 July, D1-3 July in association with elevated CyanoHABs in Arrowhead Beach, E1-2 September, and F1-2 November (all in 2023).

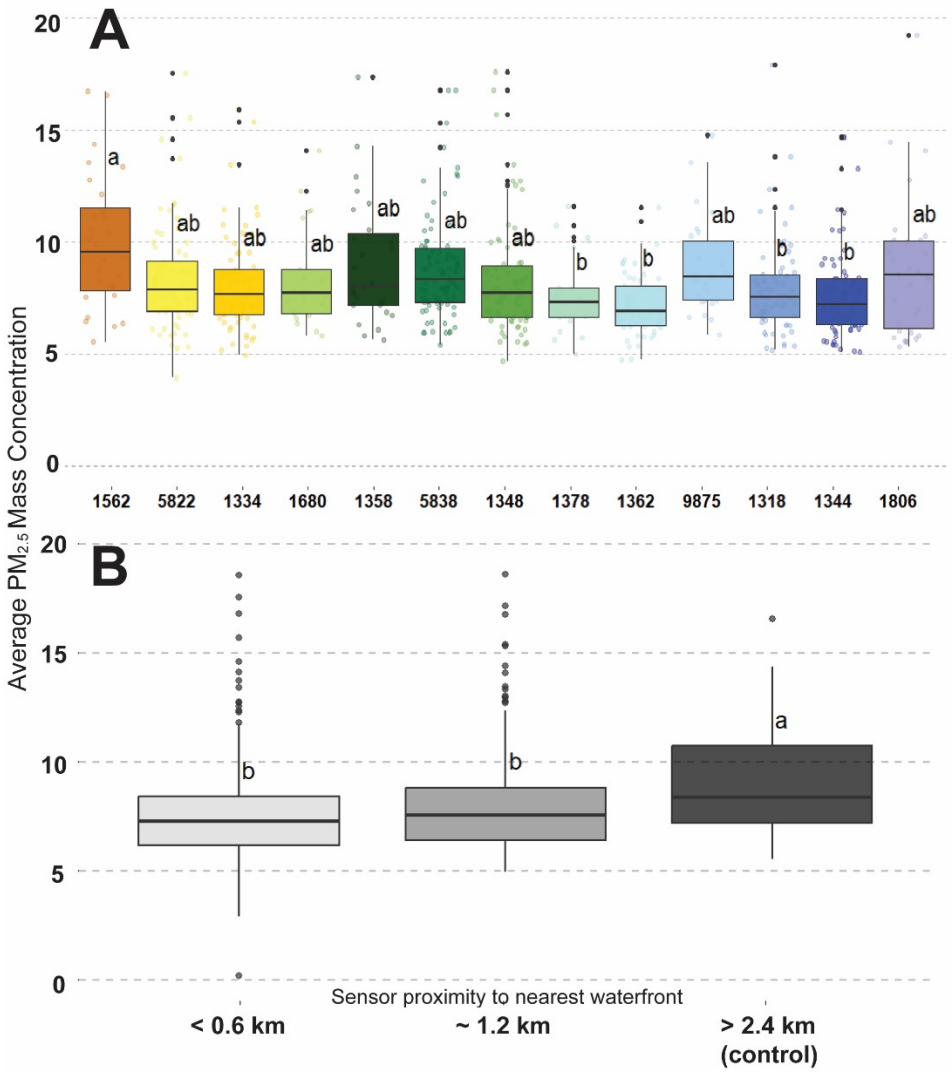


Figure S3. Average $PM_{2.5}$ mass concentrations (US-wide correction), grouped by individual sensors (A) and proximity to the nearest HAB hotspot (B). Statistical differences between groups via Tukey's comparison of means are indicated with lowercase letters (a-b).

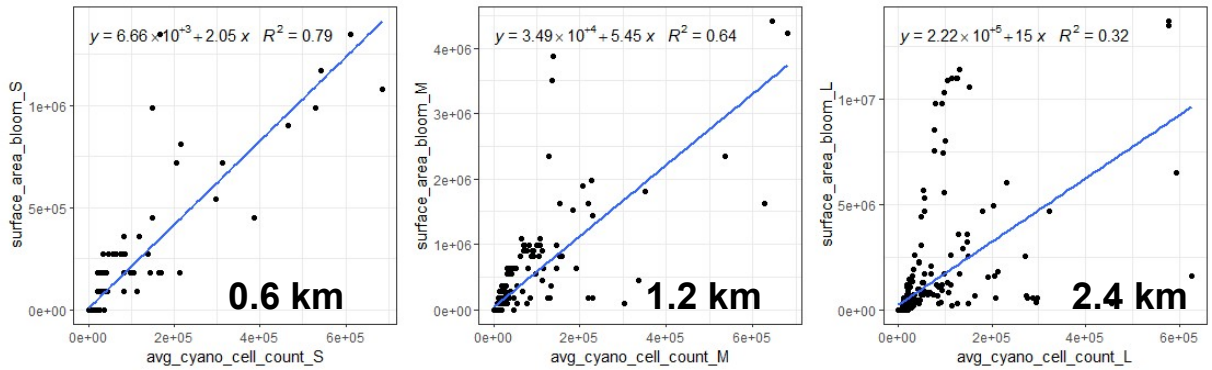


Figure S4. Linear regression analyses of correlation between harmful algal bloom intensity (bloom cell count) and spatial extent (bloom surface area), collected at three satellite imagery pixel grid sizes: (S) 0.6 km, (M) 1.2 km, and (L) 2.4 km extending from each sensor.

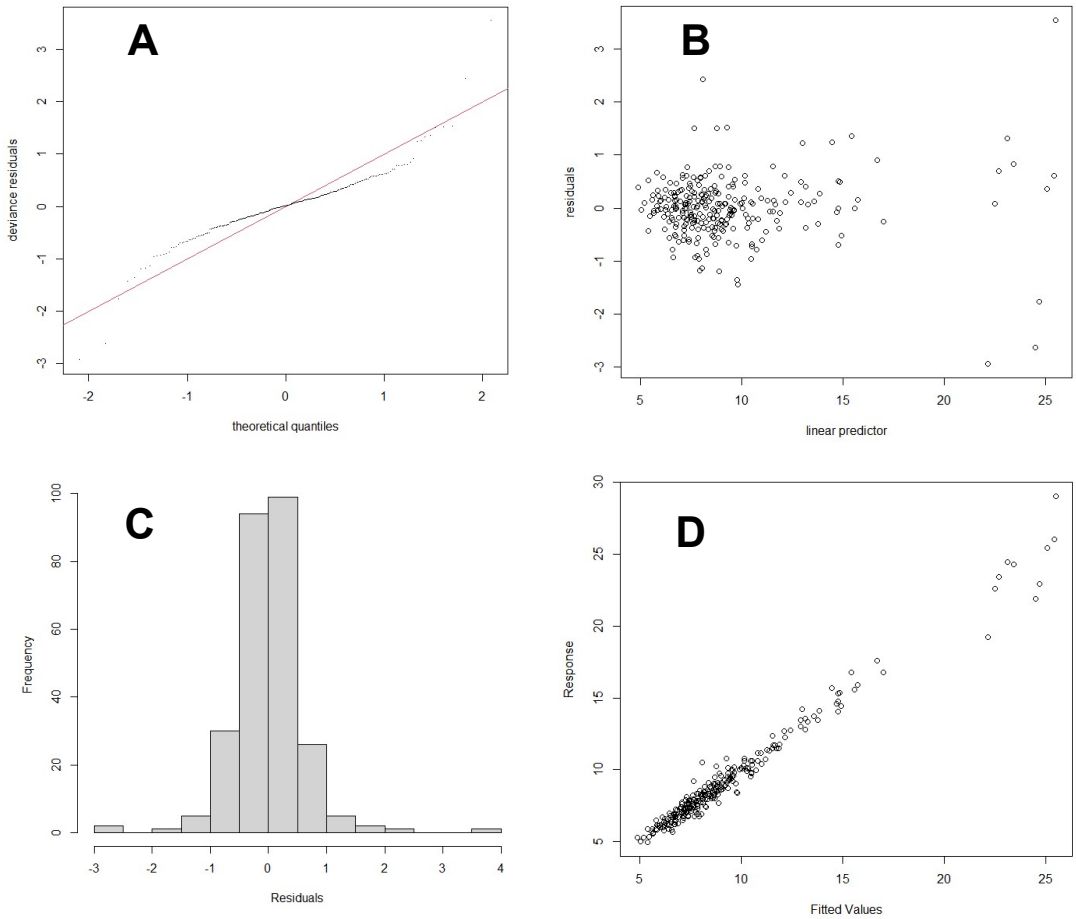


Figure S5. Generalized Additive Model diagnostic plots (A- quantile-quantile, B- residuals vs. predictor values, C- histogram of residuals, D- residuals vs. fitted values) for the model (n = 266) selected for best fit (**Table 1**, main text), including solar radiation, temperature*humidity, O_3, PO_4, NO_x, CO*SO_2, and HAB*wind speed as predictors. Final deviance explained was 97.5% with R^2 =0.963.