Article title: Monitoring Air Pollution Variability during Disasters
Authors: Earthea Nance[1]
Affiliations: Texas Southern University, Department of Urban Planning and Environmental Policy, Houston, Texas, USA[1]
Orcid ids: 0000-0002-0066-2923[1]
Contact e-mail: earthea.nance@tsu.edu
License information: This work has been published open access under Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Conditions, terms of use and publishing policy can be found at https://www.scienceopen.com/.
Preprint statement: This article is a preprint and has not been peer-reviewed, under consideration and submitted to ScienceOpen Preprints for open peer review.
Links to data: https://data.gulfresearchinitiative.org/data/BP.x750.000:0024 ; https://ldh.la.gov/index.cfm/page/647/n/238
Funder: N/A
DOI: 10.14293/S2199-1006.1.SOR-.PPNBBAG.v1
Preprint first posted online: 20 February 2021
Keywords: air pollution, particulate matter, Gulf oil spill, spatio-temporal monitoring, mortality, public health, environmental disaster management
Monitoring Air Pollution Variability during Disasters

Earthea Nance, Texas Southern University

1. Introduction

In air pollution disasters, real-time monitoring of particulate matter (PM$_{2.5}$) makes it possible to track pollution impacts and to deliver timely warnings to the public (Zhu, et al, 2007). The Deepwater Horizon oil spill—the largest marine oil spill in history—sent plumes of particulate matter across several Gulf Coast states (de Gouw, et al, 2011). This event is also known as the BP oil spill, the Gulf oil spill, and the Gulf of Mexico oil spill. Flight monitoring by the National Oceanic and Atmospheric Administration (NOAA) determined that 12,567 tons of soot and aerosols were generated by the spill and that public health was likely at risk (Middlebrook, et al, 2012; NOAA, 2011). Mobile monitoring by BP (N=101,262) and modeling by the Centers for Disease Control and Prevention (CDC) (N=2,472) also confirmed particulate matter above normal levels (BP, 2014 and CDC, 2019).

The total mass of particulates exceeded EPA’s Significant Emission Rate of 10 tons per year of direct PM$_{2.5}$ and 40 tons per year of precursor pollutants (volatile organic compounds, VOCs) (EPA, 2012 (originally issued in 2008)). In response to the disaster, the Environmental Protection Agency (EPA) and British Petroleum (BP) established the most extensive air monitoring regime ever undertaken in the region. The air monitoring network comprised a vast array of mobile and stationary monitors, regulatory monitors, flight monitors, and computer modeling (EPA, 2010; BP, 2014; CDC, 2019; NOAA, 2011; and LDEQ, 2011).

Particulate matter was selected as the air pollutant for study because it was a significant contaminant released during the oil spill: 1,323 tons of soot particles were emitted from
controlled burns, and 11,244 tons of secondary aerosol particles were created from evaporating hydrocarbons (Middlebrook, et al, 2012). Also, a large amount of continuous hourly and daily particulate matter data was gathered during the oil spill across the 6-parish region. According to the EPA (Air Now, 2013), particulate matter is one of only two pollutants that pose the greatest threat to human health in the US. PM$_{2.5}$ was selected because of its human health impacts and prevalence during the Gulf oil spill.

This paper investigates the association between PM$_{2.5}$ variability and mortality during the Gulf oil spill. The variability of PM$_{2.5}$, measured as short-term increases, is known to cause increased mortality in older populations under normal conditions (Di, 2017). Also, spatio-temporal data collection has been shown to more accurately represent PM$_{2.5}$ variability (Staniswalis, et al, 2005). This research examined these two issues via the following research questions: 1) compared to other available data, was spatio-temporal data better at representing PM$_{2.5}$ variability during the Gulf oil spill? and 2) were deaths during the Gulf oil spill in people aged 65 and over associated with PM$_{2.5}$ variability?

2. Literature Review

Particulate matter variability is a function of many factors, including atmospheric dispersion, particle deposition, particle composition, topographic change, and the multiplicity of natural and human-made sources, most significantly vehicle emissions in urban areas (Krudysz et al., 2009; Zhu et al., 2002). Peters, et al (2013:520) found that spatial variability in air pollution was higher for fine particle sizes than for coarse particles. Gulev (1997) and Hughes et al (2012) reported that mobile monitoring captures more of the actual variance that exists in the atmosphere compared to stationary monitors that are limited in their ability to capture spatial variance. Spatial variability of particulate matter was known to be higher than normal during the Gulf oil spill (Peters et al,
Averaging is known to miss between-hour peaks that are health significant (Staniswalis et al, 2005; Kaiser, 2005; and Conroy and McWilliams, 2001). At the time of the disaster, the existing air monitoring network in Louisiana was not spatially representative of the impacted region and most of the air quality data produced was daily averaged from stationary monitors (LDEQ, 2011, and EPA, 2020).

Staniswalis, et al (2005) mathematically analyzed daily averaged particulate matter data from El Paso, TX and found that the daily average statistic (in their case, daily PM$_{10}$) underestimated public health effects (i.e., mortality) because it did not account for large variations within the 24-hour window. Yuval and Brody (2005) found that smaller time-averaged sampling windows were more accurate in general. Conroy and McWilliams (2001) found that 24-hour averaging windows for PM$_{2.5}$ ignored important data in-between readings, and they recommended using a mid-hour 24-hour averaging method (i.e., the “Conroy” average) to resolve the problem. Evangelista (2011) reported that even under controlled conditions, there can be tremendous amounts of error in ambient air data, so averaging is commonly used to eliminate errant peaks. It is not clear which averaging or monitoring approach better estimates the complex interactions between ambient concentrations, human exposure, and public health (Steinle, et al, 2013).

Atmospheric monitoring directly over the spill by NOAA research aircraft discovered that, in addition to expected sources of primary particulate matter, volatile hydrocarbon emissions from floating oil were converted to massive amounts of secondary aerosols of ultrafine particle size (<0.1 micron), and these particles were transported northwest across Southeast Louisiana (130 miles) and as far north as Jackson, Mississippi (300 miles) (Middlebrook, et al, 2011; and Perring, et al, 2011). Middlebrook, et al (2011, 2012) estimated that over 90 percent of the particulate mass associated with the Gulf oil spill was in the fine and ultrafine size ranges. Spatial and temporal
variability increases as particle size gets smaller, suggesting that smaller averaging periods (or more frequent sampling) would be needed to analyze fine and ultrafine particle pollution (Kumar, et al, 2014; Sabaliauskas, et al, 2012; Heal, et al, 2012; Birmili, et al, 2013; and Costabile, et al, 2009). Current regulatory standards ignore ultrafine particles (Kumar, et al, 2014) and presume that spatial variability is negligible. The volatile organic compounds (VOCs) emitted from the spill are not typically regulated unless a state can demonstrate that VOC emissions are a significant contributor to the formation of PM$_{2.5}$ (EPA, 2008).

Kaiser (2005) reported that even very short-term exposure to poor air quality could have life-changing health effects for vulnerable population groups, suggesting that the 24-hour averaging period was too large. Ross, et al (2013) recommended combining stationary and mobile data to maximize spatial-temporal resolution in assessing overall public health impacts. Staniswalis, et al (2005) found that daily averaged particulate matter was not granular enough to show a statistically significant relationship to mortality, and that the lack of information about acute exposures was particularly sensitive to particle constitution. Di, et al (2017) found a statistically significant relationship between mortality and fine particulate matter spikes (known as short-term increases, STI) at a national scale. Peres, et al (2016) confirmed a strong statistical association between Gulf oil spill emissions and physical health symptoms among women in the region, both residents and workers.

The main points from the literature are: 1) particulate matter is a spatio-temporal variable, 2) atmospheric variability was higher during the Gulf oil spill than normal, 3) there was a higher fraction of fine and ultrafine particulate matter during the Gulf oil spill, 4) stationary monitoring of particulate matter ignores spatial variability, 5) hourly and daily averaging can miss acute exposures and significantly underestimate health impacts, 6) mobile monitoring that produces
spatio-temporally representative results is likely more accurate for fine particle sizes, and 7) measurable public health impacts were caused by the oil spill.

There are few published studies that make use of the Gulf oil spill PM$_{2.5}$ dataset, which comprises over 100,000 spatio-temporal readings taken throughout the Southeast Louisiana region impacted by the spill. The current literature neither analyzes deaths associated with fine particulate matter in oil spill disasters, nor does it analyze whether spatio-temporal data better represents PM$_{2.5}$ variability in a disaster. This paper contributes to both research gaps.

4. Methods

4.1. Study Area, Population, and Timeframe

Six parishes in Southeast Louisiana were selected as the study area: Jefferson, Lafourche, Orleans, Plaquemines, St. Bernard, and Terrebonne. This 4,138 square-mile region was selected because it was located closest to the site of the oil spill (as close as 38 miles), it had the largest exposed population, and it was well sampled throughout the disaster by a variety of monitoring methods. A study population consisting of persons aged 65 and over was selected because of their sensitivity to air pollution. The leading causes of death in this population group are heart disease, cancer, and chronic lower respiratory disease (CDC, 2021).

The study period spans from May 15, 2010 to December 21, 2010. These nearly eight months represent the core period of disaster activities, including emissions from the oil spill, gas flaring, in situ burning, and increased emissions from vehicles and boats. It accounts for the time before and after the well was unsuccessfully capped in July 2010, and it includes the period after
the well was permanently capped in September 2010. The study area is shown in Figure 1. Table 1 provides general information about the study area.

4.3. Mobile Data and Instruments

While the paper will compare several different sets of oil spill monitoring data, the main dataset selected for this paper was BP’s “emergency-mobile-regional” PM$_{2.5}$ dataset for the Southeast Louisiana region, which is available to the public (BP, 2014). This dataset was selected because of the long duration of mobile monitoring, wide spatial coverage, and large sample size. This mobile monitoring data is a spatio-temporal dataset that was only taken during the Gulf oil spill. BP traveled routes through the region taking air quality readings over a cumulative total of approximately 90,000 miles within the study area (see Figure 1). BP’s mobile monitoring vehicles were outfitted with portable nephelometers. The primary model used was the TSI SidePak Personal Aerosol Monitor (AM-110) instrument with cyclone. Used less frequently were the Dust Trak DXR and UltraRAE nephelometers. BP’s quality assurance and data management methods are described elsewhere in their Data Publication Summary Report (BP, 2014) and in EPA’s Quality Assurance Sampling Plan for the British Petroleum Oil Spill (EPA, 2010). All data were gathered with Federal Reference Methods (FRM) or Federal Equivalent Methods (FEM) (EPA, 2014).

4.4 Humidity Adjustments

There are many factors that explain different outcomes between instrument types. Gravimetric samples continuously capture particles on a filter, nephelometer readings capture the degree of light scattered per second across the particles, and beta-attenuation readings capture the
continuous absorption of radiation onto the particles (Heal et al, 2012). The impact of humidity on these three instrument types is widely appreciated because particle size increases as the air becomes moist, thus affecting the results (Hernandez et al, 2017).

According to the EPA’s oil spill quality control plan (EPA, 2010), all of its PM$_{2.5}$ data were controlled for humidity immediately upon obtaining each reading in comparison to a gravimetric sample. At the time BP stated it was following EPA’s quality control plan, which applied to all sampling and monitoring for the disaster. Four years later, BP issued a data summary report that retroactively corrected for humidity, as follows:

Personal aerosol monitors used for measuring PM$_{2.5}$ and PM$_{10}$ are significantly affected by humidity. At a relative humidity of 60%, the concentrations of PM$_{2.5}$ and PM$_{10}$ are overestimated by approximately 20%. At a relative humidity of 90%, the concentrations of PM$_{2.5}$ and PM$_{10}$ are overestimated by approximately 200%. Users should be aware that the relative humidity in the Gulf of Mexico region generally exceeds 60%; therefore, most of the results in the dataset are affected. Historic humidity readings can be obtained from the National Oceanic and Atmospheric Administration’s National Climatic Data Center. (BP Data Publication Summary Report: Community Air Sampling and Monitoring, Reference No. OTH-04v01-02, April 2014, pp. 6-7).

Nephelometer overestimation typically begins at a humidity threshold of 60% (Soneja et al, 2014; Wu et al, 2004; Sioutas et al, 2000; and Chakrabarti et al, 2004) and peaks at about 90% (Covert et al, 1980; EPA, 1996), which defines the range of adjustment. To adjust BP’s PM$_{2.5}$ data, historic humidity readings were obtained from NOAA (2020) and all data points were transformed using the Covert (1980) and EPA (1996) relationship (see Figure 2). This approach allowed more accurate adjustments because it extended the number of comparison points in between 60% and 90%.
4.5 Stationary and Modeled Data and Instruments

The stationary datasets available for comparison were: 1) Louisiana Department of Environmental Quality (LDEQ) “regulatory-stationary-urban” data taken routinely with permanent stationary gravimetric instruments and available to the public (https://deq.louisiana.gov/subhome/air and https://www.epa.gov/outdoor-air-quality-data); and 2) EPA’s “emergency-stationary-coastal” data taken with stationary Met One E-BAM Beta-Attenuation Monitors only during the Gulf oil spill and available to the public (https://archive.epa.gov/emergency/bpspill/web/html/air.html). A third dataset was the Centers for Disease Control and Prevention’s (CDC) “research-model-regional” results from its Downscaler Model for the period of the Gulf oil spill, developed in collaboration with the EPA and available to the public (https://ephtracking.cdc.gov/showAirMonModData).

A sample of these three datasets is presented in Figure 3 to facilitate a side by side comparison to the mobile dataset. Figure 3 compares daily PM$_{2.5}$ concentrations from August 21 to September 6, 2010 in Jefferson Parish and Plaquemines Parish. This timeframe was selected because it matched the dates of EPA’s hourly monitoring. All four of the datasets follow the same general trend at varying concentration levels. The graphs show that the CDC’s research-model-regional data and the LDEQ’s regulatory-stationary-urban data are consistently lower (in concentration) and smoother (fewer peaks) than the EPA’s emergency-stationary-coastal data and BP’s emergency-mobile-regional data. This is appropriate because research models and regulatory monitors are designed to produce normalized data for the purposes of predicting concentrations in locations without monitors and for comparison with regulatory standards. The emergency monitoring was not under these constraints, however the EPA did follow conventional norms in establishing stationary monitors with hourly or daily time-
controlled readings. The EPA monitors in Figure 3, however, were located along the coast and captured the highest levels of particulate matter blowing in from the spill, which explains why they consistently produced the highest concentrations in concentrations in Figure 3. The BP emergency-mobile-regional data primarily lies in between the other datasets.

5. Results

5.1 Data Variability

Two of the four available datasets were limited either in coverage area, duration, or number of samples. The LDEQ data only covered cities and took a total of 600 samples (every 1\textsuperscript{st}, 3\textsuperscript{rd}, or 6\textsuperscript{th} day over 7 months) using stationary monitors in five of the six parishes, averaged on a 24-hour basis. The EPA data covered only the coastline and used stationary monitors to take either hourly or daily readings with sample sizes of 1,144 (hourly over 17 days) and 869 (daily over five months) for the six parishes combined. While these two stationary datasets produced normalized data at points on the boundaries of the impacted region (the coastal edge and the urban areas), the data were not spatially or temporally representative of variability. For these reasons, these two datasets were deemed insufficient for an analysis of the association between variability and mortality. The CDC modeling dataset and the BP emergency dataset will be further assessed.

Figure 4 displays scatterplots of the BP emergency-mobile-regional data, aggregated by parish and corrected for humidity. The mean absolute deviation (MAD) ranges from 7.5 to 8.7 (overall MAD = 8.19), indicating high dispersion, numerous peaks or outliers, and variability that could be difficult to model or predict. Histograms (available in the data supplement) confirmed that the distribution in each parish is log normal. The dataset is comprised of frequent, randomly timed
readings that are spatially representative, with a relatively enormous sample size (N=101,262) and comprehensive spatial coverage (3,923 acres) compared to the other datasets that were available. There are 4,731 peaks above the 95th percentile, likely caused by a combination of the conditions of the oil spill, spatial variation, and unknown errors. However, the large sample size and approximately randomly timed readings reduces the impact of unknown errors. The conditions of the oil spill and spatial variation are part of the phenomenon that is represented by the dataset. Consequently, peaks were not considered outliers and were not removed because removing them would have distorted the results, as confirmed by Gorard (2005) and Leys et al (2019). Peaks were part of the situation being studied and reflect the variability of the event. Therefore, the median (13.60 ug/m3) was used instead of the mean to represent the central tendency.

Further evidence of high variability in the BP emergency-mobile-regional dataset can be seen by comparing peaks to exceedance days. For example, on 92 individual days between May and December 2010, Jefferson Parish (n=19,106) had 1,006 readings exceeding 35 ug/m3. During this same period, there were only three days with concentrations sustained enough to achieve a daily average that exceeded 35 ug/m3 (the daily National Ambient Air Quality Standard, NAAQS). Mobile monitoring generated many peaks but few consistently high concentrations, a pattern suggesting elevated short-lived peaks as identified by Russell, et al (2004). All six parishes followed this pattern. This is an important finding because it confirms that a particulate matter distribution can simultaneously exhibit extremes of variability without extremes in daily average concentrations.

Figure 5 reveals seasonal variation in PM$_{2.5}$ concentration for all six parishes, with increased concentrations in the spring and late summer/early fall. A similar pattern of higher PM$_{2.5}$
concentrations in spring and fall was observed by Russell et al (2004) in Southeast Texas. Chen et al (2013) found that seasonal variations in PM$_{2.5}$ were associated with increased deaths in China.

5.2. Modeled Versus Mobile PM$_{2.5}$ Data

The Downscaler model was developed by the CDC in collaboration with EPA to predict PM$_{2.5}$ concentrations in areas with low population and inadequate monitoring on the ground. Compared to the LDEQ and EPA datasets, the CDC model was the only dataset representing the entire region over the full duration of the disaster, with a robust sample size and more than one reading per day, making the CDC modeling results the best matching dataset for comparison to the mobile data. The CDC modeling results were therefore compared to the mobile monitoring data.

The variability of the CDC’s research-model-regional distribution was much lower than the variability of BP’s emergency-mobile-regional distribution, aggregated to the parish scale. An F-test on the variances confirmed a significant difference between the variances of the two distributions ($\rho=0.06$, accept $H_0$). However, a paired two-sample t-test on the means showed the means of the two distributions were equal ($\rho=0.39$, two tail). The modeled data were normally distributed (skewness=0.97, kurtosis=1.15), and the mobile data was normally skewed with a slightly high kurtosis (skewness=-1.44, kurtosis=2.37)$^1$ (histograms are available in the data supplement). A Kolmogorov–Smirnov two-sample test revealed that these two samples did indeed come from the same distribution ($D=0.667$, $\rho=0.143$, $\alpha=0.05$). Overall, the comparison finds that the mobile and modeled datasets are statistically similar in terms of PM$_{2.5}$

---

$^1$ In Microsoft Excel, values for skewness and kurtosis between -1.96 and +1.96 are considered acceptable to prove normal univariate distribution (George & Mallery, 2010; and Gravetter& Wallnau, 2014).
concentration in the six parishes; however, the two datasets are statistically different in terms of variability (Ghasemi, 2012).

Variability is the key difference between the modeled and mobile datasets. CDC’s modeled data consisted of two to three readings per day, while the BP’s mobile data provided 3.5 readings per hour on average (BP, 2014). A higher number of readings captures more variability. Public health researchers have discovered that variability in PM$_{2.5}$—measured as short-term increases of 10 ug/m$^3$ or more—is directly associated with mortality in older populations (Di et al, 2017). Table 2 compares short-term increases (STIs) in PM$_{2.5}$ for the modeled and mobile data.

Table 2 demonstrates that variability is negligible in the modeled dataset, resulting in a trivial number of short-term increases (STIs). Despite the accuracy of the model data in terms of concentration and overall trends, it was not designed to capture the many changes in concentration that occurred in-between readings during the Gulf oil spill. In contrast, the mobile data took many thousands of readings and recorded more of the short-term increases that have been associated with death in older segments of the population. Using the mobile data, the remainder of the paper will directly test this association in the context of the Gulf oil spill.

5.3 Short-Term PM$_{2.5}$ Increases and Mortality

The mobile dataset was analyzed for short term increases greater than or equal to 10 ug/m$^3$ in preparation for analysis against mortality data. When the raw mobile data was aggregated into 7-day increments (to correspond to the 7-day mortality data that was available), it retained relatively high statistical power as indicated by a large Cohen’s D (1.939 > 0.8), a large effect size r (0.696 > 0.5); and a large Hedge’s G (1.833 > 0.8). Weekly mortality data was obtained from the Louisiana State Office of Health Statistics. In preparation for the analysis,
deaths were counted proportionately based on the number of sampling days per week so that deaths that occurred on days without sampling were not included.

Ordinary least squares regression was performed using mortality data as the dependent variable versus short-term increases in PM$_{2.5}$ as the independent variable, first by parish, then as multiple regression, and then aggregated for the region. There was a positive, statistically significant relationship between PM$_{2.5}$ and mortality no matter how the regression was done. When handled separately, each individual parish had a significant relationship at the p<0.05 level, but with small R-squareds ranging from 0.17 – 0.21, indicating a high degree of unaccounted for variation. Under multiple regression, all three parishes collectively had a significant relationship with mortality, but in this case with a moderate correlation (R$^2$=0.51). In the final regression analysis, in which all of the data was used in its raw form (i.e., not aggregated by parish), a strong statistically significant relationship was found (p<0.001) with a moderate R-squared (R$^2$=0.43). The full results are summarized in Table 3.

The analysis shows that short-term increases in fine particulate matter significantly and consistently predicted an increase in deaths throughout the study area from mid-May to mid-December, 2010, but with only moderate correlation. The deaths analyzed were all-cause deaths among people aged 65 and over in Jefferson, Lafourche, and Terrebonne Parishes. The overall finding of the study is that at the regional scale, each short-term increase of 10-$\mu$g/m$^3$ or more of fine particulate matter was associated with a statistically significant increase of 0.105 all-cause deaths (p=3.53E-5) in people aged 65 and over.
6. Discussion

The analysis revealed that mobile monitoring during disasters is a critical supplement to existing stationary monitoring, which does not fully represent spatial and temporal variability. The mobile monitoring made it possible to observe that all parishes studied experienced frequent short-lived peaks for many consecutive months, data that was missed by the other three datasets that were available during the disaster. This missed data, due to the lack of spatio-temporal monitoring, was linked to important public health risks, including mortality in sensitive groups.

Emergency stationary monitors installed by EPA along the coast picked up the highest PM$_{2.5}$ concentrations coming in from the offshore spill. However, these daily and hourly data lacked spatial resolution, and the hourly sampling was only operated for a brief 17-day period, so both daily and hourly emergency datasets lacked the spatio-temporal resolution needed to assess mortality. Daily PM$_{2.5}$ concentrations were simultaneously gathered at six regulatory monitors located in the urban centers and regulated by LDEQ. While these data were continuous throughout the year, only daily averages were made available to the public, and these failed to measure short term increases in between readings. The LDEQ data were also stationary and therefore lacked spatial resolution. The CDC model provided long term spatially integrated data, but with only two or three data points per day, which again missed peaks in between readings. The analysis found that the CDC’s model results aligned well with BP’s emergency monitoring data in terms of concentration and overall trend; however, the modeled output was much less variable than the real-time data so it missed most of the short-term increases in PM$_{2.5}$ that occurred during the disaster.

Evaluation of all available datasets during the Gulf oil spill confirms that the spatio-temporal mobile monitoring dataset was the most suitable data available for addressing the first
research question: *Compared to other available data, was spatio-temporal data better at representing PM$_{2.5}$ variability during the Gulf oil spill?* The second research question, *Were deaths during the Gulf oil spill in people aged 65 and over associated with PM$_{2.5}$ variability?*, was addressed using OLS regressions of mortality versus short-term increases in PM$_{2.5}$ within the study area and timeframe. This resulted in a statistically significant relationship between short-term increases of PM$_{2.5}$ and mortality, a finding that aligns with other recently published research on mortality and fine particulate matter (Di et al, 2017; and Liu et al, 2019).

It is likely that the mobile monitoring data contained too much unnecessary variation. Their monitoring scheme took an 8-month average of 3.5 readings per hour over 89,982 linear miles, and 31 readings per square mile. These are impressive numbers, but the low R-squareds are likely due to excess variation in the data. From analyzing the other datasets in contrast to the mobile dataset, one can see that hourly readings are not frequent enough, and that hourly readings are needed for much longer than the two-week sampling period realized in the coastal monitoring effort. Taking periodic daily samples is also not useful, as too much variation goes unmonitored. The mobile monitors had the advantage of taking nearly random readings rather than on-the-hour readings. Because of the long duration of mobile monitoring, this resulted in a quasi-random sample, which was more representative of what the population actually breathed.

This research exposes a number of factors that should be tested in future research: 1) what is the ideal sampling rate (readings per hour) for capturing an effective amount of variability; 2) what is the minimum duration of sampling to capture enough data so that a statistical analysis against public health data can be achieved; and 3) what body of evidence would be persuasive enough to make air pollution regulations more protective of public health.
7. Conclusions

During the Gulf oil spill, fine particulates traveled into a region containing a large population known to have disproportionately high underlying disease burden (Goldstein et al, 2011). These emissions affected air quality on a regional basis. The most likely PM sources were vehicle emissions caused by increased car, truck, and boat traffic during the disaster, controlled burns for reducing floating contaminants, direct emissions from the oil spill, and secondary generation of aerosols from the oil spill (precursors). These sources did not create consistent emissions from a single location or of a single type; rather, they produced transient emissions from multiple sources and created multiple points of exposure. These conditions lead to increased variability in PM$_{2.5}$ during the disaster, as demonstrated by the spatio-temporal monitoring analyzed in this paper. Routine regulatory monitoring, and emergency monitoring based on routine monitoring norms (i.e., hourly or daily readings, stationary monitors, short durations, normalization of data), failed to recognize variability that was directly linked to public health. Computer models that followed these same norms were unable to represent variability.

This paper has demonstrated that short term increases in PM$_{2.5}$ were associated with all-cause mortality in people over the age of 65 in the region impacted by the Gulf oil spill. These findings have implications for environmental policy and for disaster management. In the case of the Gulf oil spill, this paper has demonstrated the importance of capturing spatial and temporal dimensions in ambient air monitoring. When emissions are not controlled or predictable, such as during a disaster, spatially and temporally integrated monitoring at frequencies greater than once per hour and for long durations are essential for capturing data relevant to public health. Spatio-temporal approaches to monitoring and modelling can reveal far more of the variability that
exists during air pollution disasters, and can be a robust source of data for understanding the impacts of fine particulate matter on mortality.
References

[dataset] BP Gulf Science Data; 2016; Community Air Sampling and Monitoring Data; Gulf of Mexico Research Initiative; DOI:10.7266/N7BR8QMW.
[dataset] LA State Center for Health Statistics; 2020; Weekly Mortality 2010 by Parish Over 65; Louisiana State Department of Health and Hospitals; by request at CPHI@la.gov.
Air Now. 2013. Air Quality Index (AQI)—A Guide to Air Quality and Your Health. Available online: http://airnow.gov/index.cfm?action=aqibasics.aqi (accessed on October 5, 2013).
Birmili, W., Tomsche, L., Sonntag, A., Opelt, C., Weinhold, K., Nordmann, S. and Schmidt, W., 2013. Variability of aerosol particles in the urban atmosphere of Dresden (Germany). Effects of spatial scale and particle size. Meteorologische Zeitschrift (Berlin), 22.
Borgini, A., et al., Personal Exposure to PM$_{2.5}$ Among High-School Students in Milan and Background Measurements: The EuroLifeNet Study, Atmospheric Environment 2011.
British Petroleum. 2014. Data Publication Summary Report: Community Air Sampling and Monitoring Data. Reference No. OTH-04v01-02. Available online: https://data.gulfresearchinitiative.org/data/BP.x750.000:0024 (accessed on April 20, 2019).
Centers for Disease Control and Prevention (CDC). 2019. “Monitor+Model Air Data.” National Environmental Public Health Tracking. Available online: https://ephtracking.cdc.gov/showAirMonModData (accessed on January 20, 2021).
Centers for Disease Control and Prevention (CDC). 2021. “FastStats: Older Persons’ Health.” National Center for Health Statistics. Available online: https://www.cdc.gov/nchs/fastats/older-american-health.htm (accessed on February 15, 2021).
Chakrabarti, Bhabesh, Philip M. Fine, Ralph Delfino, and Constantinos Sioutas. "Performance evaluation of the active-flow personal DataRAM PM2.5 mass monitor (Thermo Anderson pDR-1200) designed for continuous personal exposure measurements." Atmospheric Environment 38, no. 20 (2004): 3329-3340.
Chen, R., Peng, R. D., Meng, X., Zhou, Z., Chen, B., & Kan, H. (2013). Seasonal variation in the acute effect of particulate air pollution on mortality in the China Air Pollution and Health Effects Study (CAPES). The Science of the total environment, 450-451, 259–265. https://doi.org/10.1016/j.scitotenv.2013.02.040
Conroy, D, and McWilliams, A. 2001. Challenges and Issues with Reporting the AQI for PM$_{2.5}$ on a Real-time Basis. Environmental Protection Agency.

Costabile, F., W. Birmili, S. Klose, T. Tuch, B. Wehner, A. Wiedensohler, U. Franck, K. König, and A. Sonntag. "Spatio-temporal variability and principal components of the particle number size distribution in an urban atmosphere." *Atmospheric Chemistry and Physics* 9, no. 9 (2009): 3163-3195.

Covert, D. S.; Waggoner, A. P.; Weiss, R. E.; Ahlquist, N. C.; Charlson, R. J. (1980) Atmospheric aerosols, humidity, and visibility. In: Hidy, G. M.; Mueller, P. K.; Grosjean, D.; Appel, B. R.; Wesolowski, J. J., eds. The character and origins of smog aerosols: a digest of results from the California Aerosol Characterization Experiment (ACHEX). New York, NY: John Wiley & Sons, Inc.; pp. 559-581. (Advances in environmental science and technology: v. 9).

de Gouw, J.A.; Middlebrook, A.M.; Warnke, C.; Ahmadov, R.; Atlas, E.L.; Bahrinei, R.; et al. Organic Aerosol Formation Downwind from the Deepwater Horizon Oil spill. *Science* 2011, 331, 1295.

Di, Qian, Lingzhen Dai, Yun Wang, Antonella Zanobetti, Christine Choirat, Joel D. Schwartz, and Francesca Dominici. "Association of short-term exposure to air pollution with mortality in older adults." *JAMA* 318, no. 24 (2017): 2446-2456.

Environmental Protection Agency. 2020. Outdoor Air Quality Data. Available online: https://www.epa.gov/outdoor-air-quality-data (accessed on December 16 and 22, 2020)

Environmental Protection Agency. Final Rule on the Implementation of the New Source Review Provisions for Particulate Matter Less Than 2.5 Microns (PM$_{2.5}$ ). U.S. Environmental Protection Agency, Washington, DC, 77 FR 65107, 2012 (originally issued May 8, 2008). Available online: https://www.epa.gov/sites/production/files/2015-12/documents/20080508_fs.pdf (accessed December 25, 2020).

Environmental Protection Agency. List of Designated Reference and Equivalent Methods. Office of Research and Development. 2014. Available online: www.epa.gov/ttn/amtic/criteria.html (accessed on July 28, 2014).

Environmental Protection Agency. Quality Assurance Sampling Plan for British Petroleum Oil Spill. May 2010. Available online: https://archive.epa.gov/emergency/bpsspill/web/pdf/bp-oil-spill-sampling-plan.pdf and https://archive.epa.gov/emergency/bpsspill/web/pdf/appendix-data_management_plan_5-30.pdf. (accessed on April 20, 2019).
Environmental Protection Agency. 1996. *Air Quality Criteria for Particulate Matter*. Volume II of III. EPA/600/P-95/001bF. Office of Research and Development.

Evangelist, M. “Investigation of 1-hour PM$_{2.5}$ Mass Concentration Data from EPA- Approved Continuous Federal Equivalent Method Analyzers.” 2011. Environmental Protection Agency Office of Air Quality Planning and Standards. Available online: http://www.epa.gov/ttn/naaqs/standards/PM/data/Evangelista040511.pdf (accessed on December 20, 2013).

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update (10a ed.)* Boston: Pearson.

GISGeography. 2021. Open Source Louisiana Parish Map, https://gisgeography.com/louisiana-parish-map/. (accessed January 10, 2021)

Gorard, Stephen. "Revisiting a 90-year-old debate: the advantages of the mean deviation." *British Journal of Educational Studies* 53, no. 4 (2005): 417-430.

Ghasemi, A; Zahediasl, S. 2012. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. Int J Endocrinol Metab. 2012 Spring; 10(2): 486–489.

Goldstein BD, Osofsky HJ, Lichtveld MY. The Gulf Oil Spill. *New England Journal of Medicine*. 2011 Apr 7;364(14):1334-48.

Gravetter, F., and Wallnau, L. (2014). *Essentials of Statistics for the Behavioral Sciences (8th ed.).* Belmont, CA: Wadsworth.

Gulev, S. K. 1997. Climatologically significant effects of space–time averaging in the North Atlantic sea–air heat flux fields. *Journal of Climate, 10*(11), 2743-2763.

Heal, Mathew R., Prashant Kumar, and Roy M. Harrison. "Particles, air quality, policy and health." *Chemical Society Reviews* 41, no. 19 (2012): 6606-6630.

Hernandez, G., Berry, T-A., Wallis, S.L., & Poyner, D. (2017, November). Temperature and Humidity Effects on Particulate Matter Concentrations in a Sub-Tropical Climate During Winter. L. Juan (Ed.), Proceedings of International Conference of the Environment, Chemistry and Biology (ICECB 2017) (pp.41-49). 102. 10.7763/IPCBEE 2017.V102.8.

Hughes, P.J.; Bourassa, M.A.; Rolph, J.J.; Smith, S.R. Averaging-Related Biases in Monthly Latent Heat Fluxes. *Journal of Atmospheric and Oceanic Technology*, 2012, 29: 974–986.
Index Mundi. 2019. Louisiana Land Area in Square Miles, 2010 by County. Available at: https://www.indexmundi.com/facts/united-states/quick-facts/louisiana/land-area#map; retrieved Feb. 24, 2019.

Kaiser, J. Epidemiology. How dirty air hurts the heart. Science 2005, 307 (5717) p. 1858-9.

Krudysz, M.; Moore K.; Geller M.; Sioutas C.; Froines J. Intra-community Spatial Variability of Particulate Matter Size Distributions in Southern California/Los Angeles. Atmospheric Chemistry and Physics 2009, 9: 1061–1075.

Kumar, Prashant, Lidia Morawska, Wolfram Birmili, Pauli Paasonen, Min Hu, Markku Kulmala, Roy M. Harrison, Leslie Norford, and Rex Britter. "Ultrafine particles in cities." Environment international 66 (2014): 1-10.

Leys, C., et al. (2019). How to Classify, Detect, and Manage Univariate and Multivariate Outliers, With Emphasis on Pre-Registration. International Review of Social Psychology, 32(1): 5, 1–10.

Liu, Yuewei, Jingju Pan, Hai Zhang, Chunxiang Shi, Guo Li, Zhe Peng, Jixuan Ma, Yun Zhou, and Lan Zhang. "Short-term exposure to ambient air pollution and asthma mortality." American journal of respiratory and critical care medicine 200, no. 1 (2019): 24-32.

Louisiana Department of Environmental Quality. Ambient Air Monitoring Stations. 2011. Available online: http://www.deq.louisiana.gov/portal/DIVISIONS/Assessment/AirFieldServices/AmbientAirMonitoringProgram/AirMonitoringSites.aspx (accessed on 20 November 2013).

Louisiane State Center for Health Statistics. Louisiana State Office of Public Health, New Orleans, LA. (data by request, received on February 4, 2021).

Middlebrook, et el, “Air Quality Implications of the Deepwater Horizon Oil spill,” Proceedings of the National Academy of Sciences Early Edition, 2011.

Middlebrook AM, Murphy DM, Ahmadov R, Atlas EL, Bahreini R, Blake DR, Brioude J, de Gouw JA, Fehsenfeld FC, Frost GJ et al. “Air Quality Implications of the Deepwater Horizon Oil spill.” Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(50):20280-5.

National Oceanic and Atmospheric Administration (NOAA). 2011. “Insights from Oil Spill Air Pollution Study Have Implications Beyond Gulf. March 11, 2011. Chemical Sciences
Peres, L. C., Trapido, E., Rung, A. L., Harrington, D. J., Oral, E., Fang, Z., Fontham, E., & Peters, E. S. (2016). The Deepwater Horizon Oil Spill and Physical Health among Adult Women in Southern Louisiana: The Women and Their Children's Health (WaTCH) Study. *Environmental health perspectives, 124*(8), 1208–1213. https://doi.org/10.1289/ehp.1510348.

Perring, A. E., J. P. Schwarz, J. R. Spackman, R. Bahreini, J. A. de Gouw, R. S. Gao, J. S. Holloway, D. A. Lack, J. M. Langridge, J. Peischl, A. M. Middlebrook, T. B. Ryerson, C. Warneke, L. A. Watts, D. W. Fahey. “Characteristics of Black Carbon Aerosol from a Surface Oil Burn During the Deepwater Horizon Oil spill.” *Geophysical Research Letters*, 2011, 38 (17).

Peters, J.; Theunis, J.; Van Poppel, M.; Berghmans. P. Monitoring PM10 and Ultrafine Particles in Urban Environments Using Mobile Measurements. *Aerosol and Air Quality Research*, 2013, 13: 509–522.

Ross et al.: Spatial and temporal estimation of air pollutants in New York City: exposure assignment for use in a birth outcomes study. *Environmental Health* 2013, 12:51. doi:10.1186/1476-069X-12-51. Available online: http://www.ehjournal.net/content/12/1/51 (accessed on 11 July 2014).

Russell, Matthew, David T. Allen , Donald R. Collins & Matthew P.Fraser (2004) Daily, Seasonal, and Spatial Trends in PM2.5 Mass and Composition inSoutheast Texas Special Issue of Aerosol Science and Technology on Findings from the Fine Particulate Matter Supersites Program, *Aerosol Science and Technology*, 38:S1, 14-26.

Sabaliauskas, Kelly, Cheol-Heon Jeong, Xiaohong Yao, Yun-Seok Jun, Parnian Jadidian, and Greg J. Evans. "Five-year roadside measurements of ultrafine particles in a major Canadian city." *Atmospheric Environment* 49 (2012): 245-256.

Schwartz, Joel. 1994. “Air Pollution and Daily Mortality: A Review and Meta Analysis.” *Environmental Research* 64, 36-52.
Sioutas, Constantinos, Seongheon Kim, Mingchih Chang, Lester L. Terrell, and Henry Gong Jr. "Field evaluation of a modified DataRAM MIE scattering monitor for real-time PM2.5 mass concentration measurements." *Atmospheric Environment* 34, no. 28 (2000): 4829-4838.

Soneja, Sutyajeet, Chen Chen, James M. Tielsch, Joanne Katz, Scott L. Zeger, William Checkley, Frank C. Curriero, and Patrick N. Breysse. "Humidity and gravimetric equivalency adjustments for nephelometer-based particulate matter measurements of emissions from solid biomass fuel use in cookstoves." *International journal of environmental research and public health* 11, no. 6 (2014): 6400-6416.

Staniswalis, J.G.; Parks, N.J.; Bader, J.O.; Maldonado, Y.M. Temporal Analysis of Airborne Particulate Matter Reveals a Dose-Rate Effect on Mortality in El Paso: Indications of Differential Toxicity for Different Particle Mixtures. *Journal of the Air & Waste Management Association* 2005, 55:893–902.

Steinle, S.; Reis, S.; Sabel, C.E. Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of the Total Environment* 2013, 443, 184–193.

U.S. Census Bureau. 2020. “County Population Totals: 2010-2019.” Available online: https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html (accessed on February 1, 2021)

Wu, Zhang, Hu Bo, Chen Changhe, Du Ping, Zhang Lei, and Feng Guanghong. "Scattering properties of atmospheric aerosols over Lanzhou City and applications using an integrating nephelometer." *Advances in Atmospheric Sciences* 21, no. 6 (2004): 848-856.

Yuval, D.M.; Broday, Y.C. Mapping spatio-temporal variables: The impact of the time-averaging window width on the spatial accuracy. *Atmospheric Environment* 2005, 39, 3611–3619.

Zhu, Y.; Hinds, W.C.; Kim, S.; Sioutas, C. Concentration and Size Distribution of Ultrafine Particles near a Major Highway. *Journal of the Air & Waste Management Association* 2002, 52, 1032–1042.

Zhu, K.; Zhang, J; Lioy, P. Evaluation and Comparison of Continuous Fine Particulate Matter Monitors for Measurement of Ambient Aerosols. *Journal of the Air & Waste Management Association* 2007, 57(12): 1499–1506.
Table 1. General information about the study area.

| Parish       | Land Area (sq.mi.) | Population (2010 Census) | Adj. Deaths (May-Dec, 2010) |
|--------------|--------------------|--------------------------|-----------------------------|
| Jefferson    | 296                | 432,552                  | 1,353                       |
| La Fourche   | 1,068              | 96,318                   | 300                         |
| Orleans      | 169                | 343,829                  | 559                         |
| Plaquemines  | 780                | 23,042                   | -                           |
| St. Bernard  | 378                | 35,897                   | -                           |
| Terrebonne   | 1,232              | 111,860                  | 222                         |
| Region       | 3,923              | 1,043,498                | 2,434                       |

Note: Death count is for persons aged 65 and over, adjusted for the number of days monitored.
The State of Louisiana and the CDC suppressed data for Plaquemines and St. Bernard Parishes due to low population.
Sources: Index Mundi, 2019; US Census, 2010, LA Center for Health Statistics, 2020. All data is public use.
Table 2. Comparison of centrality (mean, median) and variability (STI) in the modeled and mobile datasets, by parish and for the region.

| Parish       | Modeled Data | Mobile Data | STI | Modeled Data | Mobile Data | STI |
|--------------|--------------|-------------|-----|--------------|-------------|-----|
|              | N            | Mean, ug/m³ | Median, ug/m³ | N         | Mean, ug/m³ | Median, ug/m³ |    |
| Jefferson    | 403          | 13.03       | 13.03 | 3           | 19,106      | 15.97 | 13.50 | 1,932 |
| Lafourche    | 563          | 12.87       | 13.03 | 3           | 32,967      | 17.28 | 14.40 | 4,290 |
| Orleans      | 223          | 13.13       | 12.78 | 3           | 4,681       | 17.45 | 14.50 | 722   |
| Plaquemines  | 572          | 13.38       | 13.20 | 2           | 18,478      | 14.64 | 12.00 | 2,029 |
| St. Bernard  | 320          | 13.95       | 13.88 | 3           | 10,255      | 16.62 | 13.60 | 1,140 |
| Terrebonne   | 391          | 12.75       | 12.89 | 2           | 15,775      | 16.62 | 14.00 | 1,477 |
| Region       | 2,472        | 13.22       | 13.18 | 16          | 101,262     | 16.39 | 13.60 | 11,590|

Note: STI = short term increases in PM$_{2.5}$ ≥10 ug/m$^3$. Time period is May–December 2010.
Source: CDC, 2019; BP, 2014. All data is public use.
Table 3. Association of All-Cause Mortality with Short-Term Increases in PM$_{2.5}$.

| OLS Regression | N  | R  | R$^2$ | Adj. R$^2$ | SE  | β     | ρ     | 95% CI          |
|---------------|----|----|-------|------------|-----|-------|-------|----------------|
| Simple        |    |    |       |            |     |       |       |                |
| Jefferson     | 31 | 0.46| 0.21  | 0.18       | 14.37| 0.1799| 0.0096**| [0.05,0.31]    |
| Lafourche     | 33 | 0.44| 0.19  | 0.16       | 3.485| 0.0164| 0.0113* | [0.00,0.03]    |
| Terrebonne    | 26 | 0.41| 0.17  | 0.13       | 3.248| 0.0326| 0.0326* | [0.00,0.06]    |
| Multiple      | 33 | 0.71| 0.51  | 0.46       | 16.09|        |        |                |
| Jefferson     |    |    |       |            |     | 0.1679| 0.0443* | [0.01,0.33]    |
| Lafourche     |    |    |       |            |     | 0.0632| 0.0496* | [0.00,0.13]    |
| Terrebonne    |    |    |       |            |     | 0.1868| 0.0116* | [0.04,0.33]    |
| Simple        |    |    |       |            |     |        |        |                |
| Unaggregated  | 33 | 0.66| 0.43  | 0.41       | 16.77| 0.1046| 3.53E-5***| [0.06,0.15]    |

Note: alpha=0.05. N=weeks sampled. All-cause mortality data is weekly, for persons aged 65 and over. Data for the remaining parishes was insufficient because BP’s monitoring in Orleans Parish was much shorter than the other parishes, and because the CDC suppressed mortality data for Plaquemines and St. Bernard Parishes due to low population.
Figure 1. Terrestrial mobile air monitoring routes in six Southeast Louisiana parishes. Readings taken over water by boat were not included in the analysis. Sources: Public use data from BP (2014); public use map by GISGeography (https://gisgeography.com/louisiana-parish-map/); and map enhancements by Angel Torres.

Figure 2. Dependence of the light-scattering coefficient of ambient aerosol on relative humidity in various areas of the United States. (Covert et al, 1980; reproduced for public use in EPA, 1996).
Figure 3. Comparison of a sample of datasets taken during the Gulf oil spill. Note that Plaquemines Parish has no regulatory monitors due to low population levels; data was only available for Plaquemines because of emergency monitoring and subsequent modeling. Sources: EPA, 2011; BP, 2014; CDC, 2019; and LDEQ, 2011. All data is public use.
Figure 4. Univariate scatterplots of raw PM$_{2.5}$ data in the 6-parish region during the Gulf oil spill, gathered by mobile monitors.
Source: Public use data from BP (2014).
Figure 5. Trends in PM$_{2.5}$ concentration, aggregated by month and by parish, during the Gulf oil spill. Higher concentrations occurred in spring and late summer/early fall. (n=44).
Source: Public use data from BP (2014).