THIRD YEAR PROGRESS REPORT

Harvard/MIT Air, Climate and Energy Center
Regional Air Pollution Mixtures: The Past and Future Impacts of Emission Controls and Climate Change on Air Quality and Health (EPA-G2014-STAR-J1)
Annual Center Progress Report
Period: December 1, 2017 – November 30, 2018
Date of Report: February 28, 2019

Introduction. This annual report presents an overview and summary of the progress and achievements during the third year of the Harvard University/Massachusetts Institute of Technology Air, Climate and Energy (ACE) Center: Regional Air Pollution Mixtures. The primary objective of the Center is to generate new scientific knowledge on past and future air quality in the US and the associated health impacts, which is of paramount importance to designing sound strategies and solutions. Our Center is investigating the sources, composition, trends, and effects of regional air pollutant mixtures across the US over a relatively long chronological period spanning past and future years (2000-2040), and examining the influence of climate, socioeconomic factors, policy decisions, and control strategies on air pollution, health, and economic outcomes. Our Center’s investigators are a multidisciplinary team with expertise in a wide range of fields, including Exposure Assessment, Air Pollution Modeling, Atmospheric Chemistry, Epidemiology, Climate Change Modeling, Meteorology, Biostatistics, Economics, Urban Planning, Social Behavior, Risk Assessment, Remote Sensing, and Public Policy.

Our ACE Center includes five Research Projects supported by the Air Pollution Core and Administrative Unit. Each Research Project addresses the scientific issues stated in one or more of the EPA ACE Center Research Questions (RQ), and together they address all four RQs using a synergistic and integrative approach. Achievement of our research goals will be made possible by sharing data and analytical methodologies among the research teams. This synergistic approach enables us to not only more efficiently utilize resources, but also foster collaborations among the Investigators that cross disciplinary boundaries, enhance productivity, and strengthen the overall Center. The researchers and investigators of this ACE Center are collaborating to: 1) investigate past and projected regional and sub-regional air pollutant mixtures and examine their characteristics and health effects and their relationship to emissions, climate, and modifiable factors; 2) project the potential impact of climate change on regional mixtures; 3) assess the effectiveness of past and future regulatory strategies, and; 4) project and quantify future changes in socioeconomic drivers of air pollution and its health related impact and characterize carbon policy measures with respect to their air co-benefits.

Progress during the reporting period. During Year 3 of the Center, each of the five Research Projects, the Air Pollution Core, and the Administrative Unit have continued to make progress on their objectives. In the remainder of this section of the report, we will discuss achievements of the Center. These will include discussion of a new Center objective, a report of our data management activities and human subject activities, as well as Center milestones. In later sections of this report, each project is discussed separately. For each project, we summarize the objectives and present the progress and achievements, publications and presentations, inter-Center collaborative efforts, and changes (if any) to key personnel during the third year of the ACE Center.
During Year 3, we have continued our regular working group Research Meetings. These meetings feature presentations on their research by Center investigators, or by others whose research is relevant to Center objectives. These meetings also feature discussion of Center activities and serve as a forum for issues which may impact Center research. The Steering Committee (SC) has continued to work together to discuss project updates and data fusion. In addition, we have inaugurated research meetings focused on fostering interactions between researchers at Harvard and MIT, particularly among the post-doctoral fellows and doctoral students. These meetings are held quarterly.

During Year 3, we also held our second meeting of the Science Advisory Committee (SAC) for the Center, May 30-31, 2018 in Boston. At the SAC meeting, a summary of progress during Year 2 for the Center and the five Projects was presented by the Center Director and each Project PI. A poster session was also given, in which a total of 22 posters across all Projects and the Air Pollution Core were presented, the majority of which were by young researchers including doctoral students and postdocs. The SAC Report provided feedback, comments, and suggestions, which were distributed to all Center Investigators; the Center Response addressed the questions, comments and suggestion of the SAC. Both the SAC Report and Center Response were provided to EPA.

June 21-22, 2018, our Center attended the second EPA/ACE Centers annual meeting, a virtual meeting hosted by the Yale/JHU SEARCH Center.

**Evolution of ACE Center Research Objective: Effects of Environmental Radioactivity on Human Health.** One of the strengths of our Center is the cross-disciplinary collaborative nature of its researchers and projects in support of our primary objective [generating new scientific knowledge on the health impacts of air quality]. This allows us to bring diverse perspectives to bear as we investigate; progress in our research objectives can lead to important findings and further evolution of our objectives, for Projects and for the Center.

Investigating the regional and sub-regional air pollutant mixtures and examining their different characteristics and health effects is one part of our broader Center objectives. Pursuit of this objective has opened an interesting and important avenue of inquiry, the effect of environmental radioactivity on human health. Although this research/hypothesis was not originally part of our ACE Center proposal, we allocated some resources to analyze existing data. These data are not ideal and they have been collected by EPA and others for different purposes.

In the past, cardiovascular and respiratory morbidity and mortality have been linked to short- and long-term exposures to particulate air pollution (PM) mass and components. However, few studies have investigated the effects of PM radioactivity. In the U.S., over 73% of the total exposure to naturally occurring ionizing radiation is through inhalation of ambient particles carrying attached radionuclides. The primary source of this PM radioactivity in the U.S. is Rn gas through its decay products. It emanates from the soil and enters the atmosphere, including indoor air, where it is transformed into radioactive decay products. These radionuclides attach to inhalable PM, which after lung deposition continue to release ionizing radiation (α-, β- and γ-radiation) causing pulmonary inflammation and oxidative stress. To date, the vast majority of previous studies focused on the cancer effects of Rn progeny; therefore, there are significant knowledge gaps regarding the non-cancer effects of radioactive particles. Our recent research has demonstrated that these non-cancer effects are, in fact, very important. Specifically, we have generated new information showing that PM gross β- and γ-activities are associated with blood pressure, oxidative stress; cardiac, lung and liver function, and total and cardiopulmonary mortality. This research has relied on data collected by the U.S. EPA Radiation Network (RadNet), which monitors PM radioactivity
across the U.S. including gross β- and γ-activities and specific radionuclides. To the best of our knowledge, these data have not been used to assess exposures for non-cancer environmental epidemiology studies. Our Center will continue with these investigations.

**Difficulties/Delays.** There have been no delays or difficulties to report in the Center. In Project 1, Objective 3 (development of Particle Emissions Inventory using Remote Sensing-PEIRS) we have encountered difficulties that are described in detail in the annual Project 1 report. In brief, due to MAIAC algorithm smoothing applied to satellite data and other contributing factors, further effort with this particular objective is unlikely to result in successful development of a 1x1 km particle emissions inventory. There are no other delays or difficulties to report in any of the Center’s Research Projects.

**Key Personnel Changes.** During Year 3 of the ACE Center, our Center Coordinator Alice Smythe passed away. Alice was the primary administrative contact for EPA Project Officers, and an integral member of our Center Administration Core, as she had been for our previous EPA Center grants for over 17 years. Her loss has been difficult for us, both professionally and personally. We have not yet found a permanent replacement. We have no other key personnel changes during Year 3 to report.

**Expenditures.** Current expenditures are somewhat below the budget plan at the date of report. No significant changes in the size or scope of the Center or its projects or in the originally-negotiated total estimated costs are reported.

**Quality Assurance.** The Quality Management Plan (QMP) was accepted by the EPA on February 3, 2017. There have been no changes to the QMP since that date. The Center continued to hold biweekly research meetings during the project year. The biweekly meetings are the main venue for the Center’s researchers to update progress, and discuss issues that may have arisen. All key Center personnel received prior to the meetings via email, the topics up for discussion. The purpose of these meetings is to review progress on the research projects, including discussion of any issues that may have an impact on data quality. During the period December 2018 through January 2019, the Quality Assurance Manager (QAM) conducted a series of audits. The scope of the audits covered both the program and exposure system as listed in the QMP. The researchers did not report any significant changes to SOPs, data management plans, codebooks or quality management procedures. The audits consisted of reviewing the project specific audit forms generated during the previous year’s audit and updating any changes to the forms.

The Center’s QAM conducted a data management audit to review the archiving requirements the projects and that data and programs used to prepare the manuscript were readily available. The scope of the audit was to identify the source program for data, tables and figures in the publication and the archived location of the programs. All documentation was consistent with each project’s data management and archiving requirements. These two publications were reviewed for compliance with the Center’s data archiving requirements:

For all Projects and the Air Pollution Core any items identified during these audits were addressed and no items were identified affecting data quality.

**Human Subjects.** There were changes to the DHHS regulations to the OHRP Common Rule on human research that came into effect in January 21, 2019. The revisions to the Common Rule included new and revised exemption categories and elimination of continuing review for minimal risk research. As part of the annual review of the center’s Human Subjects Protection Program, the changes to the regulations were reviewed to determine if any of the changes affected the current approvals. The regulations allow for grandfathering of previously approved work under the previous regulations. No changes to the Center’s human research approvals were required to by the revisions to the Common Rule.

Research Projects 1, 3, 4, 5 and the Air Pollution Core of this Center contain no human subjects research as defined by the Basic Federal Policy for the Protection of Human Research Subjects, also known as the Common Rule, 40 CFR Part 26 Subparts A, B and C. Research Project 2, however, uses indoor exposure data (geographic coordinates and indoor air pollutant concentrations) from (1) the Normative Aging Study (NAS; 2006-2011) and (2) the “Diabetes, Cardiac Disease, and Pollution Vulnerability” study (2006-2010). The HSPH Institutional Review Board (IRB) has reviewed the project plans for these projects and we have obtained non-human subjects determinations for use of data from our IRB, and have submitted this documentation to the EPA in our original proposal for review and approval prior to beginning this work.

In addition, Research Project 2 uses indoor air pollution data (location and concentrations) from ongoing work by Harvard University and Department of Veterans Affairs (VA) researchers in the Chronic Obstructive Pulmonary Disease (COPD) Air Pollution Study. We obtained an exemption determination for the work on Project 2 from the HSPH IRB on August 3, 2017. All work of the center has been reviewed by the HSPH IRB and has been determined to be either Non-Human Subjects Research or Exempt from IRB review. No additional IRB review is required as long as the scope of the research remains the same and no other human subjects-related data is added to the research.

**Research Conduct.** All research effort in the Harvard/MIT ACE Center has been performed to the highest standard of conduct.

**Future Activities.** During Year 4, the Center will continue with activities discussed above. Specific Project activities for Year 4 are described in the Project Progress Reports included in later sections. Additional planned Center activities during Year 4 include our third annual SAC meeting, which is scheduled for June 10-11, 2019.
Center Publications. Our list of publications and manuscripts accepted during Year 3 of the Center is appended to this report. Publications, presentations, and submitted manuscripts for each Research Project are included within the Project annual reports in the following sections.
Period Covered by the Report: December 1, 2017 - November 30, 2018

Date of Report: February 28, 2019

EPA Agreement Number: Air, Climate and Energy (ACE) Centers (EPA-G2014-STAR-J1)

Title: Project 1, Regional Air Pollution: Mixtures Characterization, Emission Inventories, Pollutant Trends, and Climate Impacts

Investigators: Petros Koutrakis, Ph.D. (lead PI; petros@hsph.harvard.edu); Brent Coull, Ph.D.; Daniel J. Jacob, Ph.D.; Loretta J. Mickley, Ph.D.; Joel Schwartz, Ph.D.

Institution: Harvard School of Public Health, Boston, MA; Harvard School of Engineering and Applied Sciences, Cambridge MA

Research Category: Air, Health, Climate Change, Integrated Assessment of the Consequences of Climate Change, Air Quality and Air Toxics, Social Science, Airborne Particulate Matter Health Effects, Air Toxics, Health Effects

Project Period: Year 3

Objective(s) of the Research: The overall objective of Project 1 is to apply new approaches to characterize and analyze both historical and projected regional air pollution mixtures and emissions across the continental US. Project 1 characterizes temporal and spatial patterns of pollutant mixtures within and across regions. In addition, this project investigates factors influencing regional pollutant mixtures and predicts the impact of climate change on future air quality. Project 1 has four specific objectives.

Objective 1 is to compile comprehensive air pollution, weather, emissions, and GIS datasets for the entire continental US for the period 2000-2015. We will estimate gas and particle concentrations at a high spatial resolution by assimilating data from monitoring networks (compiled in collaboration with the Air Pollution Core), satellite platforms, air pollution models, and spatiotemporal statistical models. Objective 2 is to develop and make publically available a national PM$_{2.5}$ emission inventory database of high spatial resolution (1 km) for 2000-2015. This will be achieved through the application of a novel methodology we developed that predicts point and area source emissions using aerosol optical thickness measured by satellite remote sensors. Objective 3 is to characterize spatial and temporal trends of pollutant mixtures. We will perform cluster analysis to group areas that exhibit distinct pollutant profiles or mixtures, referred to as “Air Pollution Regions,” then analyze their spatial patterns and temporal trends to investigate the impact of regulations, climate change, and modifiable factors on regional mixtures. Objective 4 is to forecast the impact of regional climate change on air quality for 2016-2040 using an ensemble of climate models. We will project the potential impact of climate change on regional pollutant mixtures and predict future regional air quality assuming no changes in anthropogenic emissions.

Progress Summary/Accomplishments (Outputs/Outcomes): Year 3 was a productive year, across Project 1. Some objectives were advanced more successfully than others; key effort and results for each objective are outlined below.

Objective 1. It was a very productive year for Objective 1. In Year 3, we extended the GEOS-Chem simulation beyond 2013-2015 to include 2016 and 2017. (The previous years were completed in 2017.) The purpose of the GEOS-Chem simulation is to provide continuous
information on ozone and PM concentrations for subsequent epidemiological analyses by the EPA-ACE team. The simulation has 0.5°x0.625° resolution over North America and is driven by MERRA-2 assimilated meteorological data from the NASA Global Modeling and Assimilation Office (GMAO). We used the most recent benchmarked version of GEOS-Chem (v11-2-c) including detailed ozone-aerosol chemistry. That version was previously evaluated in detail with observations of ozone, PM, and their precursors over the Southeast US during the SOAS and SEAC4RS campaigns in summer 2013. More details about this simulation, including early validation, can be found in the Year 2 report. The extended simulation was completed in 2018 and results provided to Prof. Joel Schwartz’s group (Project 3) for epidemiological analyses. The GEOS-Chem output archive includes daily surface concentrations of maximum daily 8-hr average (MDA8) O₃, NO₂, PM₂.₅ and its components, and aerosol optical depth (AOD). This is a major deliverable for the project.

As part of ongoing validation of the GEOS-Chem simulation, we attempted to reconcile the apparent discrepancy between NO₂ trends over the US derived from satellite observations and NOₓ trends from the National Emission Inventory (NEI) of the US EPA. Tropospheric NO₂ columns observed by the OMI satellite instrument over the US reveal a steady decrease from 2005 until 2009 but a flattening afterward, while the NEI reports a steady decrease of US NOₓ emissions over the 2005-2017 period at a rate of 0.1 Mt yr⁻¹, or 53% over the whole time period. Our analysis showed first that the steady decrease in the NEI NOₓ emissions is in fact consistent with observations of surface NO₂ and ozone concentrations. We further found that the post-2009 flattening of OMI NO₂ is likely due to an increasing relative contribution of non-anthropogenic background (mainly from lightning and soils) and not due to flattening of anthropogenic emissions. This result was confirmed by contrasting OMI NO₂ trends in winter over urban areas, where background NO₂ levels are low, against such trends in summer over rural areas, where background levels are high. The winter urban NO₂ columns show a significant 2005-2017 decrease, while the summer rural NO₂ columns reveal no significant 2005-2017 trend. Using GEOS-Chem, we tested our hypothesis on the effects of background levels on trends in column NO₂. Our work confirms the success of sustained efforts to improve US air quality over the 2005-2017 time period.

In addition to reconciling the trends of satellite data and the NEI inventory, we also identified a large discrepancy between observed and modeled ratios of NO/NO₂ over the southeast United States during August–September 2013. We suggested that either unrecognized chemistry or errors in modeled cycling between NO, NO₂, and ozone could explain this discrepancy. Either explanation would have important implications for global tropospheric chemistry and for the interpretation of satellite observations of NO₂.

**Objective 2.** During Year 3, we devoted significant time and effort to developing the Particle Emissions Inventory using Remote Sensing (PEIRS). We employed the wavelet decomposition method described in Project 2 to decompose the daily MAIAC AOD (1km resolution) into three parts based on the spatial scale of the signal. To isolate the local emissions from the transported particles, we removed long-distance transportation by setting aside the low and intermediate frequency parts and keeping the high-frequency part. We designed a kernel based on Gaussian dispersion and performed a two-dimensional matrix convolution to calculate the mass difference between the upwind and downwind cells. Based on the mass-balance assumption, this difference was used as the estimation of daily primary emission.

We encountered several issues. Firstly, the wind direction is not evenly distributed. There is a region- and season-dependent prevailing wind direction impacted by terrain or climate factors. Even with inverse-likelihood weighting, we can only detect the upwind-downwind difference from
a limited range of wind directions. This may introduce bias caused by the different background emission level of upwind and downwind cells. Second, the wind field reanalysis data we used in the study is at the height of 10 m. No small scale terrain effects on wind direction and velocity can be captured by this wind field. Third, we found that even though MAIAC is of better quality than the Deep Blue algorithm, it is also sensitive to the surface reflection. The difference of AOD between adjacent cells can be contributed by the distinct reflection. This is common in urban environments where small-scale variation of ground reflection due to construction is widespread. In these cases, the signal caused by the reflection gradient is stronger than the AOD and difficult to filter out. The estimation using the aforementioned methods has extremely high rate of negative emission intensity where the land surface reflection varied in a short distance. Fourth, the spatial uncertainty of the grid location persists. The nominal resolution of MAIAC is 1 km, but the actual size of area covered by the pixel varies by the viewing angle. At the edge of every swath, the actual spatial resolution is always over 5 km. Within the MAIAC algorithm, this was averaged. However, this smoothing adds additional uncertainty when we tried to calculate the particle emission rates at 1 km resolution.

In addition to the difficulties encountered with the emission rate estimation, we have not succeeded in validating the PEIRS result with NEI point data. We obtained point-based NEI data following EPA’s international emissions inventory conference. We tried skipping the mass balance steps and using machine learning algorithms to detect the potential sources directly. The known point sources reported in NEI were used as a training dataset to train random forest, deep neural network and gradient boosting models. All three methods captured the ground-surface reflection variation which highlights the road network, residential buildings and other industrial buildings without detecting point sources within the bright patches.

Given the limitations with MAIAC and the other issues described above, further effort with this particular objective is unlikely to result in successful development of a 1x1 km particle emissions inventory.

Objective 3. For objective 3, initially we have applied the cluster approach at regional scale. We have worked on three papers (to be submitted soon) focusing on Eastern Massachusetts as study area.

In the first paper, we assessed the spatial patterns of ambient PM$_{2.5}$ elemental concentrations considering air pollution sources and geodemographic variables. We evaluated spatial patterns for 11 components of ambient PM$_{2.5}$, which included S, K, Ca, Fe, Zn, Cu, Ti, Al, Pb, V, and Ni. The analyses for S, Ca, Cu, Ti, Al, and Pb resulted in 2 clusters; for Fe, Zn, V, and Ni in 3 clusters; and for K resulted in 12 clusters. Land use, population, and daily traffic were the variables that divided the study area into cluster of sites more effectively. We used an R$^2$ value to estimate the potential from each variable in discriminating clusters. The larger R$^2$ value, the better the discrimination among the sites. For example, population had the highest R$^2$ value when the analysis was performed for S, Ca, Zn, Ti, Al, Pb, and V; land use presented the highest R$^2$ value for Cu, V, and Ni; and, traffic showed the highest R$^2$ value for individual PM$_{2.5}$ concentration.

In the second paper, we evaluated the influence of clusters (those cluster estimated in the first paper) on modeling PM$_{2.5}$ constituents. We hypothesized that areas representing clusters of PM$_{2.5}$ elements are potential predictor variables to be included in spatial models for particle composition. The inclusion of these clusters may minimize the exposure misclassification. Overall, our findings suggest significant influence of spatial clusters on modeling some PM$_{2.5}$ components. We observed that the clusters may affect the error of the prediction values and especially the proportion of explained variance for most of the PM$_{2.5}$ constituents evaluated in this study. The model with
cluster presented a better performance for all PM$_{2.5}$ components, except for Pb, which the R$^2$ value decreased 8.51% when we included the clusters in the analysis; and for V, which the R$^2$ value did not change with the clusters. Models for Cu and Fe explained the highest concentration variance. The R$^2$ value for the model without cluster was 0.55 for both pollutants. When we accounted for clusters, R$^2$ value increased 13 and 7% for Cu (R$^2$ = 62) and Fe (R$^2$ = 59), respectively. The models for K and S presented the lowest performance for both models with and without cluster (although the model with cluster improved substantially the R$^2$ values).

In the third paper, we compared the predictive capabilities of ordinary geostatistical interpolation (Ordinary Kriging – OK), hybrid interpolation (combination of Empirical Bayesian Kriging and land use regression), and machine learning techniques (forest-based regression) for estimating PM$_{2.5}$ constituents in Eastern Massachusetts in the United States. The OK model performed poorest for all PM$_{2.5}$ components, with R$^2$ under 0.30. The hybrid model presented a slight improvement, especially for Cu and Fe, which the R$^2$ value increased to 0.62 and 0.59, respectively. These elements presented the highest R$^2$ value from the hybrid model. The forest model presented the best performance, with R$^2$ above 0.7 for most of the particle components, including Cu, Fe, Ni, Pb, Ti, and V. Same as observed with the hybrid model, the forest model for Cu and Fe explained the highest concentration variance, with a R$^2$ value equal to 0.88 and 0.92, respectively. The forest model for K, S, and Zn performed poorest with R$^2$ value equal to 0.54, 0.37, and 0.44, respectively. The results found in this paper can be useful for the cluster framework proposed in the objective 3.

Using the background obtained from these three initial papers, we worked on a fourth paper (to be submitted soon) changing the spatial scale—from local scale to national scale (U.S.A). Our objective was to investigate spatial differences of air pollution mixtures across the US. We employed spatially constrained clustering approach (based on k-means algorithm) to group air pollution monitoring sites that exhibit distinct pollutant profiles or mixtures in the US over 9 years (2008 – 2016). We accounted for 20 chemical components of PM$_{2.5}$. The resulting clusters of pollution mixtures are characterized and validated based on source emissions represented by land-use information. Our analysis resulted in 27 clusters. We estimated that Cu, Se, NO$_3^-$, Cr, and Ba were the top five species that divided the study area into cluster of sites more effectively. Our analysis resulted in 11 clusters with single site. Five clusters (cluster 1, 3, 7, 13, and 26) had more than 4 sites. Among the clusters with more than 4 sites, the cluster 13 was the one with the highest number of sites, 33 air pollution monitoring stations. The cluster 13 is located in northwest and part of the Midwest (Ohio, Indiana, Illinois, and Wisconsin). The cluster 1 has 14 sites and it covers part of the southeast, including the states of North Carolina, South Carolina, Georgia, and Florida. The southwest has a very prominent cluster with 8 sites (cluster 26), covering part of the Louisiana, Mississippi, Texas, and Arkansas. In the west coast, two clusters were highlighted in our analysis, cluster 3 in California and cluster 7 in Washington and part of Oregon. Both clusters with 5 sites. Observing the concentration ratios (concentrations of the species $i$ / concentration of PM$_{2.5}$) for some of these clusters, our results show that clusters 3 and 7 in the west coast represent sites with high Na ratios. Cluster 13 in the northwest and part of the Midwest represents sites with high SO$_4^{2-}$ ratio. The cluster 16 with a single site in northeast has the highest SO$_4^{2-}$ ratio, representing almost the third quartile of the SO$_4^{2-}$ ratio.

We also worked on some trends analyses, which will be incorporated into the objective 3 in order to investigate the impact of regulations, climate change, and modifiable factors on regional mixtures. The initial results of this part are in two other papers (paper 5 and paper 6 – both will be submitted soon).
In the paper 5, we employed generalized additive models (GAMs) to estimate weather-associated changes in PM$_{2.5}$ composition in the US during 1988-2017. We considered seven components of ambient PM$_{2.5}$, which included elemental carbon (EC), organic carbon (OC), nitrate, sulfate, sodium, ammonium, and silicon. The impact of long-term weather changes on each PM$_{2.5}$ component was defined in our study as “weather penalty”. Nationally, temperature decreased in the warm season and increased in the cold period. Wind speed decreased in the both seasons. Relative humidity increased in the warm season and decreased in the cold season. The weather changes between 1988 and 2017 were associated with most of PM$_{2.5}$ components during both warm and cold seasons. The direction and the magnitude of the weather penalty varied considerably over the space and seasons. In the warm season, our findings suggest a nationwide weather penalty for EC, OC, nitrate, sulfate, sodium, ammonium, and silicon of 0.04, 0.21, 0.04, 0.35, -0.01, 0.05, and 0.01 µg/m$^3$, respectively. In the cold season, the estimated total penalty was 0.04, 0.21, 0.06, 0.04, -0.01, -0.02, and 0.02 µg/m$^3$, respectively.

In the paper 6, we applied the same method as paper 5 to quantify the long-term impacts of wildfires on ambient particulate carbon (OC and EC) levels in the western U.S over the last 30 years. Our results show that in the warm season, the total wildfire penalty (for the period 1988–2016) on EC and OC of 0.0011 µg/m$^3$ (95%CI: 0.0009 and 0.0014) and 0.015 µg/m$^3$ (95%CI: 0.006 and 0.023), respectively. In the cold season, the estimated total penalty for EC was 0.034 µg/m$^3$ (95%CI: 0.004 and 0.065) and for OC was 0.033 µg/m$^3$ (95%CI: 0.003 and 0.063).

**Inter-Center Collaborations:**
None.

**Publications/Presentations:**

**Publications:**


Requia, W, et al. Multivariate spatial patterns of ambient PM$_{2.5}$ elemental concentrations in Eastern Massachusetts. Submitted to *Environmental Pollution*

Requia, W, et al. The influence of spatial patterning on modeling PM2.5 constituents in Eastern Massachusetts. Submitted to *Environmental Modelling and Software*

Requia, W, et al. Evaluation of predictive capabilities of ordinary geostatistical interpolation, hybrid interpolation, and machine learning methods for estimating PM$_{2.5}$ constituents over space. Submitted to *Environmental Research*

Requia, W, et al. Climate impact on ambient PM$_{2.5}$ elemental concentration in the United States: a trend analysis over the last 30 years. Submitted to *Environment International*

Requia, W, et al. Regional air pollution mixtures across the continental US. Submitted to *Atmospheric Environment*
Requia, W, et al. The impact of wildfires on particulate carbon in the western U.S.A. Submitted to *Atmospheric Environment*

**Presentations:**

**Future Activities:** During Year 4, we will continue our work on Objectives 1, 3, and 4. In previous research, we found that model estimates of surface ozone concentrations tended to be biased high in the Southeast US and this was of concern for designing effective emission control strategies to meet air quality standards (Travis et al., 2016). Using GEOS-Chem, we determined that at least some of the bias can be traced to NEI emissions for NO\textsubscript{x} from mobile sources, at least for summer 2013 during the SEAC4RS campaign. In ongoing work, we are revisiting the Travis et al. (2016) result, to better understand the sources of the emissions bias in the NEI inventory. To date, we find that the bias appears relatively constant over the 2005-2017 timeframe.

We will continue our analysis of the impact of regulations, climate change, and modifiable factors on regional mixtures. Initially, our focus will be on completion and submission of the manuscripts in preparation. Understanding the impact of these factors during the period 1988-2016 will be essential to undertaking Objective 4.


**Supplemental Keywords:** particles, pollutant mixtures, pollution trends, regional pollution, public policy, data fusion

**Relevant Web Sites:** The website for the Harvard/MIT ACE Center is:
https://content.sph.harvard.edu/ace/

**Period Covered by the Report:** December 1, 2017- November 30, 2018

**Date of Report:** February 28, 2019

**EPA Agreement Number:** Air, Climate and Energy (ACE) Centers (EPA-G2014-STAR-J1)

**Title:** Air Pollutant Mixtures in Eastern Massachusetts: Spatial Multi-resolution Analysis of Trends, Effects of Modifiable Factors, Climate, and Particle-induced Mortality

**Investigators:** Brent Coull, Ph.D. (lead PI; bcoull@hsph.harvard.edu); Petros Koutrakis, Ph.D.; Joel Schwartz, Ph.D.

**Institution:** Harvard School of Public Health, Boston, MA

**Research Category:** Air, Climate Change, Global Climate Change, Air Quality and Air Toxics, Airborne Particulate Matter Health Effects

**Project Period:** Year 3
Objective(s) of the Research: The objective of Project 2 is to characterize historical air pollution in Eastern Massachusetts at a high spatial resolution and identify modifiable factors responsible for observed changes in PM$_{2.5}$ mass, emissions, elemental profiles, and ground air temperature. Project 2 investigates within-region variability of pollutant mixtures; examines the impact of modifiable factors on air quality; and evaluates the effectiveness of source control policies. Project 2 has four specific objectives.

Objective 1 is to use a novel, multi-resolution spatial analysis based on wavelet decomposition of high-resolution (1x1 km) remote sensing data on PM$_{2.5}$ mass and ground air temperature to identify daily regional, sub-regional (urban background) and locally generated variation in these fields. Objective 2 is to develop and apply spatiotemporal regression models to (a) quantify the impact of modifiable factors, including transportation, heating fuel use, energy, urban planning, PM$_{2.5}$ emissions, population statistics, and policy interventions, on (i) sub-regional and local variation in PM$_{2.5}$ mass and ground air temperature and (ii) high resolution local estimates of PM$_{2.5}$ emissions; (b) identify locations in which these impacts are greatest; and (c) identify lag times between implementation of a given control strategy and decreases in PM$_{2.5}$ emissions and mass. Objective 3 is to implement a novel multi-resolution correlation analysis to identify PM$_{2.5}$ elemental profiles that vary at regional, sub-regional, and local scales, and apply spatiotemporal regression models to these profiles to identify modifiable factors driving urban background and local variability in PM$_{2.5}$ composition. And Objective 4 is to use the spatial scale-specific (regional, sub-regional, and local) temporal variability in PM$_{2.5}$ mass and the PM$_{2.5}$ elemental profiles to identify source types (regional, urban background, or local) and the composition of their emissions driving pollution-induced mortality in Eastern Massachusetts. This project relies on existing remote-sensing satellite data, ambient monitoring data collected from numerous sampling campaigns (including the HSPH Boston Supersite daily samples collected since 1998 and samples from 600 locations), as well as new data collected from 2015-2018 in Eastern Massachusetts.

Progress Summary/Accomplishments (Outputs/Outcomes): This past year we finalized several manuscripts that ultimately appeared in print. Highlights include:

In the last few years, several research teams have developed distinct spatio-temporal models (exposure models) to predict ambient air pollution exposures of study participants even in areas where air pollution monitors are sparse. Use of these exposure models has led to strong evidence of air pollution-related adverse health effects, and in some instances, evidence of heterogeneity of these health effects across sub-populations. A significant limitation, however, is that these health effect estimates and their statistical uncertainty are based on the very strong assumption that a single exposure model is correct. We have developed novel methodology to (1) integrate information across existing air pollution prediction models in an ensemble that weighs each model by its predictive accuracy, differently across space and time; (2) for the first time, comprehensively quantify both intra- and inter-model uncertainty associated with ambient air pollution exposures; and (3) propagate the estimated uncertainty into health effect estimates in nationwide studies. By reporting the spatio-temporal weights and uncertainty estimates back to the groups that developed the prediction models they can identify “trouble” points in space and time and improve their models. With our findings we can also identify high uncertainty areas to inform placement of future monitors. The proposed exposure assessment and development activities are poised to inform numerous regional and national epidemiological studies—fully propagating intra- and inter-model uncertainty for the first time—as well as improve the exposure models used for
prominent international assessments such as the Global Burden of Disease. To address computational scalability, we proposed an innovative approach that instead of estimating the spatial structure of an enormous prediction error variance-covariance matrix—as in existing approaches to quantify errors related to a single prediction model—we incorporated the spatio-temporal structure in the individual posterior distributions of these predictions.

This work on full quantification of uncertainty associated with exposure estimates derived from model ensembles has been submitted for peer-review and has won an award. The paper “Adaptive and Calibrated Ensemble Learning with Dependent Tail-free Process” was accepted by peer review to the Bayesian nonparametric workshop at the prestigious NIPS conference in December 2018. Further, it won honorable mention for a student paper award of the American Statistical Association’s (ASA’s) Section on Statistics and the Environment of the Joint Statistics Meeting to be held in Denver, CO, in August 2019.

In the first two years of Project 2, in Objective 1 we developed a two-dimensional wavelet decomposition that alleviates restrictive assumptions required for standard wavelet decompositions. Using this method we decomposed daily surfaces of PM$_{2.5}$ to identify which scales of pollution are most associated with adverse health outcomes (Antonelli et al. 2017). A key feature of the approach is that it can remove the purely temporal component of variability in PM$_{2.5}$ levels and calculate effect estimates derived solely from spatial contrasts. This eliminates the potential for unmeasured confounding of the exposure-outcome associations by temporal factors, such as season. During the past year we have conducted work on Objective 4 of the project, which is to use these spatial scale-specific decompositions of PM$_{2.5}$ mass that we developed in Objective 1 to identify source types (regional, urban background, or local) and how these pollution source types are associated with mortality, both in terms of chronic and acute effects, in New England. We have applied the decomposition methods developed in the work conducted as part of Objective 1 of this project to daily 1x1km grid values of PM$_{2.5}$ from 2000-2015, merged the resulting spatially-decomposed values to mortality data from the New England region from the same time period, and ran Poisson log-linear models to quantify the association between each daily and yearly exposures to regional and local PM$_{2.5}$ contributions and zip-code level mortality counts. We are currently writing up a manuscript describing the results to be submitted for publication Spring 2019.

Recent interest focuses on identifying critical windows of vulnerability associated with prenatal exposure to air pollution during pregnancy. In Year 2 of this Project, we showed that an analysis based on a distributed lag model (DLM) can yield estimates of a critical window different from those from an analysis that regresses the outcome on each of the three trimester average exposures (TAEs), which is the standard approach typically used in the environmental health literature. Moreover, interest remains high on estimating health risks associated with air pollution mixtures. However, there currently do not existing any methods to estimate the distributed lag function of a mixture. In the past year, Project 2 investigators developed one of the first multi-pollutant distributed lag models. The approach allows for estimation of the health risks of an entire air pollution mixture, and how this varies across pregnancy in pre-birth cohorts. We applied the data to estimate the association between estimated weekly residential nitrate, OC, EC, and sulfate and birthweight in the Boston-area ACCESS pre-birth cohort. We have completed a manuscript describing the methods and results that is currently being circulated among co-authors, and we plan to submit it for publication in March 2019.

Epidemiologic studies of the short-term effects of ambient particulate matter (PM) on the risk of acute cardiovascular or cerebrovascular events often use data from administrative databases in
which only the date of hospitalization is known. A common study design for analyzing such data is the case-crossover design, in which exposure at a time when a patient experiences an event is compared to exposure at times when the patient did not experience an event within a case-control paradigm. However, the time of true event onset may precede hospitalization by hours or days, which can yield attenuated effect estimates. In the past year we wrote and submitted a paper that developed a marginal likelihood estimator, a regression calibration estimator, and a conditional score estimator, as well as parametric bootstrap versions of each, to correct for this bias. All considered approaches require validation data on the distribution of the delay times. We compared the performance of the approaches in realistic scenarios via simulation, and apply the methods to analyze data from a Boston-area study of the association between ambient air pollution and acute stroke onset. Based on both simulation and the case study, we concluded that a two-stage regression calibration estimator is an effective method for correcting bias in health effect estimates arising from misclassification of event onset times in a case-crossover study. We submitted this paper to the journal *Biometrics*, which has invited a revision.

Project 2 investigators have applied several cutting-edge machine learning methods to model the spatio-temporal variation in ambient metal concentrations, as measured via XRF, in the greater Boston area. Predictions based on gradient boosting machine (GBM) predict ambient concentrations well. In a GBM, weak learners are modified at each stage to minimize a pre-specified loss function. They are chained together so the second learner is designed to improve on the fit from the first. After several iterations, the learner will be able to make predictions with greater accuracy. We have found that the performance of the model predictions vary greatly from metal to metal. We have had the most success modeling ambient Iron (Fe) and lead (Pb) concentrations, with cross-validated $R^2$ of approximately 0.70 for these two metals in the greater-Boston area. We are currently preparing a manuscript describing these results.

**Publications/Presentations:**


Rokoff LB, Rifas-Shiman SL, Coull BA, Cardenas A, Calafat AM, Ye X, Gryparis A, Schwartz J, Sagiv SK, Gold DR, Oken E, Fleisch AF. Cumulative exposure to environmental pollutants...
Future Activities: We will continue to focus on the areas of research most strongly recommended by the Center’s Scientific Advisory Council last May: that of better characterization of uncertainty associated with predictions from model ensembles, and how to propagate this uncertainty through to health effect estimates.

We will continue to focus on the papers that are in preparation and in revision for submission to journals in statistics and environmental health sciences: the paper describing the mortality analysis of Objective 4, the paper describing the gradient boosting models for the XRF data in the greater Boston area, and the paper describing analytic methods for adjusting for exposure measurement error in the case-crossover design.

Supplemental Keywords: climate change, regional pollution, multi-resolution spatial analysis, source emissions, local pollution control strategies, wavelet analysis

Relevant Web Sites: The Harvard/MIT ACE Center website, https://content.sph.harvard.edu/ace/
Period Covered by the Report: December 1, 2017 - November 30, 2018

Date of Report: February 28, 2019

EPA Agreement Number: Air, Climate and Energy (ACE) Centers (EPA-G2014-STAR-J1)

Title: Causal Estimates of Effects of Regional and National Pollution Mixtures on Health: Providing Tools for Policy Makers

Investigators: Joel Schwartz, Ph.D. (lead PI; joel@hsph.harvard.edu); Brent Coull, Ph.D.; Petros Koutrakis, Ph.D.; Antonella Zanobetti, Ph.D.

Institution: Harvard School of Public Health, Boston, MA

Research Category: Air, Climate Change, Air Quality and Air Toxics, Airborne Particulate Matter Health Effects, Particulate Matter

Project Period: Year 3

Objective(s) of the Research: The objective of Project 3 is to estimate the causal impact of changes in pollution concentrations and mixtures (annual averages and daily patterns), how they vary by modifiable factors, the causal impacts of AQI triggers, how climate change that occurred in the last 20 years has increased mortality due to pollution, how temperature modifies the effects of pollution mixtures, and how these effects change for exposures less than the ambient standards for PM$_{2.5}$. Project 3 provides region-specific causal estimates of effects of pollution mixtures; provides causal estimates of the impact of modifiable factors; assesses the impact of climate change on mortality from air pollution using historic data, avoiding any dependence on the accuracy of climate models; and provides causal estimates of how changes in particular components of mixtures affect mortality, to guide region-specific policy decisions on air pollution. Project 3 has five specific objectives.

Objective 1 is to identify and estimate the causal effects of air pollution and mixtures on human health. We will use methods of causal inference to a) identify the causal effects of regional annual air pollution concentration fluctuations and temperature fluctuations during the last 16 years on human health; b) identify the causal effects of regional air pollution trends during the last 16 years; c) identify the causal effects of pollution mixtures, sources, and emissions on health; d) identify differences in these effects by modifiable factors; e) conduct a national risk assessment on the causal impact of past pollution on mortality, including the regional differences in concentration-response; and f) investigate the causal impact of AQI thresholds for PM$_{2.5}$ and O$_3$ due to behavioral adaptation. Objective 2 is to analyze relative acute toxicity of pollution mixtures. We will a) examine spatial (across regions) and temporal heterogeneity in the acute toxicity of pollution mixtures and emissions to understand which source types, atmospheric processes, and exposure factors influence the toxicity of regional mixtures and b) use causal mediation analysis to determine how much of the temperature effect on mortality is mediated by its effects of pollution concentrations, and how that varies regionally. This will allow us to obtain local- and region-specific estimates of future health effects and the benefits of changes in modifiable factors and adaptation. Objective 3 is to estimate the excess deaths resulting from air pollutant concentration changes due to weather changes in the last 20 years. We will demonstrate the extent to which public health impacts of climate change through pollution have already occurred, by using causal estimates of C-R relationships. Regional health impacts will be assessed using region-specific mortality risks estimates from Objective 1. Objective 4 is to estimate the causal health effects of low-level air pollution exposure. Specifically, we will examine whether the observed effects at
low pollutant levels are due to the synergistic effect of multiple pollutants (mixtures) present at low levels. And Objective 5 is to investigate air pollution-related health effects at high and low temperatures. We will examine this by region and determine whether populations, especially those that include sensitive individuals, adapt to abrupt temperature changes.

Progress Summary/Accomplishments (Outputs/Outcomes): Year 3 was productive for our Project. During Year 3, we have continued to advance our efforts on all project objectives. We have investigated the effects of long term exposure on hospital admissions; effects of climate change on human health; effects of local and regional pollutants, and; causal modeling efforts for both acute and chronic exposures. Some highlights of our progress are detailed below.

In 135 U.S. cities, we demonstrated causal effects of locally generated pollutants on daily deaths between 1999 and 2010, at concentrations below the current EPA daily PM$_{2.5}$ standard. In Schwartz et al. 2018, we used three methods which, under different assumptions, provide causal marginal estimates of effect: a marginal structural model, an instrumental variable analysis, and a negative exposure control. The instrumental approach used planetary boundary layer, wind speed, and air pressure as instruments for concentrations of local pollutants; the marginal structural model separated the effects of NO$_2$ from the effects of PM$_{2.5}$, and the negative exposure control provided protection against unmeasured confounders.

We found that in 7.3 million deaths, the instrumental approach estimated a mortality increase of 1.5% [95% confidence interval (CI): 1.1%, 2.0%] per 10 µg/m$^3$ increase in local pollution indexed as PM$_{2.5}$. The negative control exposure was not associated with mortality. Restricting our analysis to days with PM$_{2.5}$ below 25 µg/m$^3$, we found a 1.70% (95% CI 1.11%, 2.29%) increase. With marginal structural models, we found positive significant increases in deaths with both PM$_{2.5}$ and NO$_2$. On days with PM$_{2.5}$ below 25 µg/m$^3$, we found a 0.83% (95% CI 0.39%, 1.27%) increase. Including negative exposure controls changed estimates minimally.

We jointly investigated the association of short and long-term exposures to PM$_{2.5}$ and temperature with hospital admissions, and explored the modification of the associations with the short-term exposures by one another and by temperature variability. In Yitshak-Sade et al., we constructed daily ZIP code counts of respiratory, cardiac and stroke hospital admissions of adults $\geq$65 (N=2,015,660) across New-England (2001–2011). Daily PM$_{2.5}$ and temperature exposure estimates were obtained from satellite-based spatio-temporally resolved models. For each admission cause, a Poisson regression was fit on short- and long-term exposures, with a random intercept for ZIP code. Modifications of the short-term effects were tested by adding interaction terms with temperature, PM$_{2.5}$ and temperature variability. We observed associations between short and long-term exposures for all of the outcomes, with stronger effects of long-term exposures to PM$_{2.5}$. For respiratory admissions, the short-term PM$_{2.5}$ effect (percent increase per IQR) was larger on warmer days (1.12% versus −0.53%) and in months of higher temperature variability (1.63% versus −0.45%). The short-term temperature effect was higher in months of higher temperature variability as well. For cardiac admissions, the PM$_{2.5}$ effect was larger on colder days (0.56% versus −0.30%) and in months of higher temperature variability (0.99% versus −0.56%). We observed synergistic effects of short-term exposures to PM$_{2.5}$, temperature and temperature variability. Long-term exposures to PM$_{2.5}$ were associated with larger effects compared to short-term exposures.

In Zanobetti and O’Neill, we assessed and reviewed current literature on associations between characteristics (mean, variability, extremes) of ambient temperatures and human health. We were motivated by concerns that climate change, which operates on a time frame of decades or longer,
may influence not only shorter-term associations between weather and health (daily/weekly) but also have enduring implications for population health. We reviewed 26 papers on the health effects of longer-term (3 weeks to years) exposures to ambient temperature published between 2010 and 2017 to investigate whether health outcomes have been associated with longer-term exposures. We included studies on a diverse range of health outcomes (mortality, morbidity, respiratory disease, obesity, suicide, infectious diseases, and allergies among various age groups), with the exception of vector-borne diseases such as malaria. Longer-term exposures were considered to be exposures to annual and seasonal temperatures and temperature variability. We found that regional and local temperatures, and changing conditions in weather due to climate change, were associated with a diversity of health outcomes through multiple mechanisms.

In Blomberg et al., we assessed potential modification of radon on PM$_{2.5}$-associated daily mortality in 108 U.S. cities using a two-stage statistical approach. First, city- and season-specific PM$_{2.5}$ mortality risks were estimated using over-dispersed Poisson regression models. These PM$_{2.5}$ effect estimates were then regressed against mean city-level residential radon concentrations to estimate overall PM$_{2.5}$ effects and potential modification by radon. Radon exposure estimates based on measured short-term basement concentrations and modeled long-term living-area concentrations were both assessed. We found that exposure to PM$_{2.5}$ was associated with total, cardiovascular, and respiratory mortality in both the spring and the fall. In addition, higher mean city-level radon concentrations increased PM$_{2.5}$-associated mortality in the spring and fall. For example, a 10 μg/m$^3$ increase in PM$_{2.5}$ in the spring at the 10th percentile of city averaged short-term radon concentrations (21.1 Bq/m$^3$) was associated with a 1.92% increase in total mortality (95% CI: 1.29, 2.55), whereas the same PM$_{2.5}$ exposure at the 90th radon percentile (234.2 Bq/m$^3$) was associated with a 3.73% increase in total mortality (95% CI: 2.87, 4.59). Results were robust to adjustment for spatial confounders, including average planetary boundary height, population age, percent poverty and tobacco use. While additional research is necessary, this study suggests that radon enhances PM$_{2.5}$ mortality. This is of significant regulatory importance, as effective regulation should consider the increased risk for particle mortality in cities with higher radon levels. In this large national study, we found that city-averaged indoor radon concentration was a significant effect modifier of PM$_{2.5}$-associated total, cardiovascular, and respiratory mortality risk in the spring and fall. These results suggest that radon may enhance PM$_{2.5}$-associated mortality. In addition, local radon concentrations partially explain the significant variability in PM$_{2.5}$ effect estimates across U.S. cities, noted in this and previous studies.

Other Project 3 efforts during Year 3 have included updating our BC, PM$_{2.5}$, O$_3$ models and working on our NO$_2$ model.

**Inter-Center Collaborations:**

None at this time.

**Publications/Presentations:**

Ananth, CV; Kioumourtzoglou, MA; Huang, YM; Ross, Z; Friedman, AM; Williams, MA; Wang, S; Mittleman, MA; Schwartz, J. Exposures to Air Pollution and Risk of Acute-onset Placental Abruption: A Case-crossover Study. EPIDEMIOLOGY 2018; 29(5): 631-638. DI 10.1097/EDE.0000000000000859

Blomberg, AJ; Coull, BA; Jhun, I; Vieira, CLZ; Zanobetti, A; Garshick, E; Schwartz, J; Koutrakis, P. Effect modification of ambient particle mortality by radon: A time series analysis

Fong, KC; Kloog, I; Coull, BA; Koutrakis, P; Laden, F; Schwartz, JD; James, P. Residential Greenness and Birthweight in the State of Massachusetts, USA. INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH 2018; 15(6). DI 10.3390/ijerph15061248

Hart, JE; Grady, ST; Laden, F; Coull, BA; Koutrakis, P; Schwartz, JD; Moy, ML; Garshick, E. Effects of Indoor and Ambient Black Carbon and PM$_2.5$ on Pulmonary Function among Individuals with COPD. ENVIRONMENTAL HEALTH PERSPECTIVES 2018; 126(12). DI 10.1289/EHP3668

Li, WY; Nyhan, MM; Wilker, EH; Vieira, CLZ; Lin, HH; Schwartz, JD; Gold, DR; Coull, BA; Aba, AM; Benjamin, EJ; Vasan, RS; Koutrakis, P; Mittleman, MA. Recent exposure to particle radioactivity and biomarkers of oxidative stress and inflammation: The Framingham Heart Study. ENVIRONMENT INTERNATIONAL 2018; 121: 1210-1216. DI 10.1016/j.envint.2018.10.039

Ljungman, PLS; Li, WY; Rice, MB; Wilker, EH; Schwartz, J; Gold, DR; Koutrakis, P; Benjamin, EJ; Vasan, RS; Mitchell, GF; Hamburg, NM; Mittleman, MA. Long- and short-term air pollution exposure and measures of arterial stiffness in the Framingham Heart Study. ENVIRONMENT INTERNATIONAL 2018; 121: 139-147. DI 10.1016/j.envint.2018.08.060

Nyhan, MM; Coull, BA; Blomberg, AJ; Vieira, CLZ; Garshick, E; Aba, A; Vokonas, P; Gold, DR; Schwartz, J; Koutrakis, P. Associations Between Ambient Particle Radioactivity and Blood Pressure: The NAS (Normative Aging Study). JOURNAL OF THE AMERICAN HEART ASSOCIATION 2018; 7(6). DI 10.1161/JAHA.117.008245

Peng, C; Cayir, A; Sanchez-Guerra, M; Di, Q; Wilson, A; Zhong, J; Kosheleva, A; Trevisi, L; Colicino, E; Brennan, K; Dereix, AE; Dai, L; Coull, BA; Vokonas, P; Schwartz, J; Baccarelli, AA. Associations of Annual Ambient Fine Particulate Matter Mass and Components with Mitochondrial DNA Abundance. Epidemiology 2017; 28(6): 763–770. doi: 10.1097/EDE.0000000000000717

Rice, MB; Li, W; Wilker, EH; Gold, DR; Schwartz, J; Zanobetti, A; Koutrakis, P; Kloog, I; Washko, GR; O'Connor, GT; Mittleman, MA. Association of outdoor temperature with lung function in a temperate climate. EUROPEAN RESPIRATORY JOURNAL 2019; 53(1). DI 10.1183/13993003.00612-2018

Schwartz, J; Fong, K; Zanobetti, A. A National Multicity Analysis of the Causal Effect of Local Pollution, NO$_2$, and PM$_{2.5}$ on Mortality. Environmental Health Perspectives 2018; 126(8) https://doi.org/10.1289/EHP2732

Yitshak-Sade, M; Bobb, JF; Schwartz, JD; Kloog, I; Zanobetti, A. The association between short and long-term exposure to PM$_{2.5}$ and temperature and hospital admissions in New England and the synergistic effect of the short-term exposures. SCIENCE OF THE TOTAL ENVIRONMENT 2018; 639: 868-875. DI 10.1016/j.scitotenv.2018.05.181

Future Activities: During Year 4, we will continue our work on local- and region-specific estimates of past and future health effects and the benefits of changes in modifiable factors and adaptation. We will also continue with our causal analyses on the effects of pollution mixtures, sources, and emissions on health and the regional differences in concentration-response and NO2 modeling.

Supplemental Keywords: particles, particulate matter, pollutant mixtures, regional pollution, risk analysis, causal modeling

Relevant Web Sites: The Harvard/MIT ACE Center website, https://content.sph.harvard.edu/ace/
**Period Covered by the Report:** December 1, 2017-November 30, 2018

**Date of Report:** February 28, 2019

**EPA Agreement Number:** Air, Climate and Energy (ACE) Centers (EPA-G2014-STAR-J1)

**Title:** Project 4: A Causal Inference Framework to Support Policy Decisions by Evaluating the Effectiveness of Past Air Pollution Control Strategies for the Entire United States

**Investigators:** Corwin Zigler (PI), Francesca Dominici (co-PI), Joel Schwartz, Loretta Mickley, Steve Barrett

**Institution:** Harvard TH Chan School of Public Health/Harvard School of Engineering and Applied Sciences/Massachusetts Institute of Technology

**Research Category:** Air, Climate Change, Air Quality and Air Toxics, Airborne Particulate Matter Health Effects

**Project Period:** Year 3

**Objective(s) of the Research:** The overall objective of Project 4 is to develop a new methodological framework rooted in principles of causal inference to investigate the effectiveness of specific control strategies on impacting the largest power-generating units in the United States. In Project 4, we combine state-of-the-art atmospheric modeling, causal inference methods, and national data sets to conduct accountability research; that is, research that characterizes causal effects of well-defined regulatory actions at power plants on: 1) emissions; 2) air quality across distant locations in accordance with atmospheric fate, transport, and other factors; and 3) health outcomes. Project 4 has 4 specific objectives.

*Objective 1* is to develop a national database on emissions control technologies employed at a large number of power-generating units in the US linked with: continuous emissions monitoring, ambient air quality monitoring, weather, population demographics, and Medicare hospitalization and mortality outcomes for the period 1995 to 2015. *Objective 2* is to estimate and compare the causal effects of past control strategies implemented at the largest power-generating facilities on SO\(_2\), NO\(_x\), CO\(_2\), and PM\(_{2.5}\) emissions and population exposure to criteria pollutants (PM\(_{2.5}\) and O\(_3\)) for the entire US for the period 2000 to 2015. This requires integrating new statistical methods for causal inference with atmospheric chemistry models of how changes in emissions impact ambient exposures across distant locations in accordance with atmospheric fate, transport, and other factors. *Objective 3* is to estimate the causal effects of past control strategies implemented at the largest power-generating facilities on mortality and morbidity in the entire US both locally and nationally, and compare the differential health impact of different control strategies. And *Objective 4* is to develop approaches for mediation analysis that will quantify the extent to which causal effects of regulatory actions on health outcomes can be attributable to changes in targeted modifiable factors (e.g., emissions, targeted pollutants), as opposed to being driven by co-benefits to other factors.

**Progress Summary/Accomplishments (Outputs/Outcomes):** Year 3 was very successful in three primary domains: 1) refinement of our reduced-complexity atmospheric model for pollution transport from power plants and 2) further development of statistical methodology tailored to the analysis of air quality regulations and 3) the deployment of work in domains (1) and (2) into epidemiological analyses. In particular, the addition of Lucas Henneman (postdoctoral associate, atmospheric engineer and data scientist) solidified our ability to repurpose the Hybrid Single
Particle Langrangian Integrated Trajectory (HYSPLIT) model for the purpose of characterizing pollution transport and population exposures from power plants. The current product, which we call HYSPLIT Average Dispersion model (HyADS), has shown the ability to approximate the transport of SO2 emissions from coal-fired power plants effectively enough to: 1) generate source-receptor matrices of how an arbitrary set of power plants impacts an arbitrary set of population locations and 2) characterize population location (e.g., zip code) level exposures to pollution originating specifically from power plants. Importantly, the method can do so in a much more computationally nimble and scalable way that supports its use when full-scale chemical transport models would not be practicable due to computation. A manuscript evaluating the HyADS model has been accepted for publication in *Atmospheric Environment*, and a separate manuscript using the model to evaluate health impacts of changes in coal power plant emissions has been accepted for publication in *Epidemiology*.

We have made progress towards publication on several statistical methods projects. For ongoing methods development regarding causal inference with interference, a paper on cluster and population treatment allocation programs is at the second round of revisions for *Biometrics*, and we have also submitted a paper defining bipartite causal inference with interference. Several other papers relating to statistical topics of mediation analysis, time-varying mediation analysis, uncertainty in the design stage of propensity score analyses, mitigating the consequences of strong confounding/lack of overlap in causal analyses, time-varying treatments, continuous treatments, measurement error, and machine learning methodology for causal inference have been recently submitted, invited for revision, or published. All of this work is conducted in the context of evaluating impacts of power plant interventions, emissions, or other air pollution related exposures.

Finally, we have completed several key epidemiological publications. One paper evaluating the health impacts of nonattainment designations for PM standards has been published this year in *Epidemiology*. Another outlining the potential health impacts of the current presidential administration’s environmental agenda appeared in *JAMA*.

**Inter-Center Collaborations:**

Our development of the reduced-complexity HyADS model shares aligned interests with the development of reduced-complexity models ongoing as part of the CMU/UW ACE Center. We have coordinated with Julian Marshall and Chris Tessum at UW to implement their InMAP model for purposes aligned with the goals of our HyADS approach, and successfully completed comparisons of the two approaches for estimating population exposures to emissions from coal power plants. As Dr. Zigler has recently moved to University of Texas and is now co-located with Joshua Apte of the CMU/UW ACE Center, there are early discussions of other inter-center collaborations on going.

**Publications/Presentations:**

**Publications:**


**Submitted Manuscripts Currently Under Review or Revision:**


**Presentations:**

1. Zigler: Invited Speaker Workshop on Causal Adjustment in the Presence of Spatial Dependence, part of thematic program at Centre de Recherches Mathematique on Causal Inference in the Presence of Dependence and Network Structure, Montreal QC, Canada. 2018


**Future Activities:** During year 4 we will continue development along all lines described above. Early stage development on using the outputs of the HyADS reduced complexity model to inform development of new statistical methods for causal inference with interference and network analysis will continue and is expected to result in several newly submitted manuscripts in statistics and epidemiology journals. With the recent validation of HyADS, we are at early stages of investigations tailored to the health benefits of specific energy transitions, such as fuel retrofits,
scrubber installs, and plant retirements, which we expect to be central to development during the remainder of the funding period. We are also collaborating with researchers in Project 5 of the center to investigate the effectiveness of state-level climate and energy policies that focus on the power generating sector. Work in collaboration with members of the Project 5 team is also ongoing to compare HyADS results against GEOS-Chem Adjoint for comparing individual coal power plant impacts on select areas across different regions of the US. We are also working on publications to highlight the open access and usability of our databases and the HyADS reduced-complexity model towards the dissemination of these tools for others to use in their own research.

Supplemental Keywords: Accountability assessment, power-generating sector, intervention evaluation, source-receptor mapping

Relevant Web Sites: The Harvard/MIT ACE Center website, https://content.sph.harvard.edu/ace/ In addition, the following websites provide tools and data to implement methods and analyses developed during this project:

- https://github.com/czigler/arepa
- https://github.com/czigler/PM2.5-Nonattainment
- https://dataverse.harvard.edu/dataverse/airqualregs
- https://github.com/gpapadog/DAPSm
- https://github.com/gpapadog/DAPSm-Analysis
- https://github.com/cchoirat/rinmap
- https://github.com/lhenneman/SplitR
- https://osf.io/s98qm/
- https://osf.io/7dp8c/
Period Covered by the Report: December 1, 2017-November 30, 2018
Date of Report: February 28, 2019
EPA Agreement Number: Air, Climate and Energy (ACE) Centers (EPA-G2014-STAR-J1)
Title: Projecting and Quantifying Future Changes in Socioeconomic Drivers of Air Pollution and its Health-Related Impacts
Investigators: N.E. Selin, S.R.H. Barrett, S. Solomon, J. Reilly
Institution: Massachusetts Institute of Technology
Research Category: Air, Climate Change, Urban Air Toxics, Mercury, Air Quality and Air Toxics, Airborne Particulate Matter Health Effects, Air Toxics, Particulate Matter
Project Period: Year 3

Objective(s) of the Research: Project 5 investigates future changes in regional air pollution characteristics as a result of technological and societal changes. We will quantify the future implications of technologies and efficiency improvements in the energy and transportation sectors on regional differences in air pollution impacts. Selected case studies assess, inter alia, the environmental and health benefits of choices in state and regional carbon policy implementation relevant to recently proposed carbon dioxide emission reductions from the energy sector. We will examine the health-related benefits of reducing concentrations of ozone and particulate matter, as well as changing regional air pollution mixtures including air toxics.

Progress Summary/Accomplishments (Outputs/Outcomes): Work on Project 5 during the reporting year has been distributed among three main goals. The first has been a significant expansion of the cross-state air pollution project under project objectives 2 and 4. This work quantifies health impacts resulting from pollution which crosses state boundaries, comparing these imported and exported impacts to the in-state impacts from local emissions. Based on feedback received during the first round of peer review, the project was expanded to include ozone pollution in addition to particulate matter. This necessitated the development of new adjoint code and a set of 96 GEOS-Chem adjoint simulations, quantifying the impact of each state, emissions sector, and emitted species on the ozone concentrations in every state within the contiguous United States. An additional set of 10 conventional simulations was also completed to quantify non-linearity in the response of US air quality to anthropogenic emissions. We found that ~30% of the total health burden of US air quality is due to ozone exposure as opposed to PM$_{2.5}$. However, due to the high non-linearity of the response of ozone concentrations to changes in emissions, only ~10% of the benefits of a small reduction in emissions across the US would be the result of changes in ozone concentrations. This discrepancy between the “average” and “marginal” sensitivity of US air to changes in emissions will have direct relevance to decision-making regarding US emissions. We also found that, due to its longer lifetime, 70% of ozone-related impacts from US emissions are incurred outside of the originating state, compared to only 40% for PM$_{2.5}$.

More broadly, this work quantifies the role that changes in the mix and magnitude of anthropogenic emissions has affected and will continue to affect local air quality. It also provides crucial information for the evaluation of the impact of related energy policy, by giving a numerical estimate of the co-benefits or tradeoffs which might be incurred if other policies penalize or incentivize specific states or industrial sectors. Following the first round of reviews and the
additional work described above, the manuscript has been resubmitted for publication and is again under peer review.

This work has also resulted in two new studies. Using the data produced from our work on cross-state air pollution, we have quantified air quality impacts from the energy sector per tonne of CO₂ emitted, providing the co-pollutant cost of carbon (CPCC). This work shows that the US-mean CPCC exceeds the social costs of carbon used by the current and previous administrations. We also find that the CPCC varies significantly across the US, with the highest state-average CPCC being 15 times greater than the lowest. A manuscript based on this work is currently in preparation.

A second study, this time in collaboration with Inês Azevedo (an investigator in the EPA Center for Air, Climate, and Energy Solutions), quantifies air quality trade-offs associated with the implementation of carbon pricing. Changes in demand are calculated for 21 different carbon price scenarios, using a dispatch model which can estimate pollutant emissions on a plant-by-plant, hour-by-hour basis for a full year. During this project year, emissions estimates were generated for all scenarios, and are now undergoing validation prior to calculation of air quality damages using the same approach as was used for the cross-state air pollution study.

The second goal of year 3 was to advance work on quantifying the time-of-emergence of air quality impacts over the 21st century (Objectives 1 and 5). We had aimed to complete an ensemble of 120-year simulations using the Community Atmospheric Model (CAM) v3.1 during the project year. This work was delayed by the discovery of non-physical output in the CAM meteorological forecasts, requiring an intensive debugging effort before the simulations could be completed. This debugging has now been completed, and test output from CAM has been verified as physical by comparison to existing output from an unrelated global climate model, the NASA GEOS data assimilation system. Simulations are now underway to complete the climate ensemble needed to drive the GCHP air quality model and thereby produce an “air quality ensemble” for the 1980-2100 period. In addition to the climate dataset, a methodology was also identified during the project year which will allow the quantification of the “time of emergence” of air quality impacts due to meteorological factors (Objective 5). This is based on statistical analysis of variability within and between air quality ensemble members.

The third goal of year 3 concerned the effectiveness of state-level climate and energy policies in United States, with a focus on the effects on power sector output, emissions, and air quality (Objectives 2 and 3). We compiled a unit-level emissions inventory for fossil fuel power plants in the U.S. (collaborating with project 4 at HSPH/MIT ACE center), associating emissions from each unit with individual policy programs. We examined the effects of six energy and climate policies on power sector output and emissions: renewable portfolio standards (RPS); energy efficiency standards; regional greenhouse gas initiative (RGGI); decoupling revenue from sales; mandatory green power; and CO₂ emissions standards. on power sector generations and emissions. The most significant finding was that RPS reduced unit-level emissions by 7-9% on average relative to electricity generating units in non-RPS states.

This analysis was complemented by a study of how wind power influences power sector emissions, using an hourly emissions and wind power dataset. We statistically estimated the effects of wind power on generation and emission for each power generating unit, which allowed us to capture both heterogeneity (between power plants) and locality of the air pollution outcomes. We find that cleaner gas-fired units are more sensitive than other units to changes in wind power output, which leads to reduced emission offsets (due to wind power) compared with previous literature estimates.

Within the same set of objectives, we also finalized results and a manuscript for a paper titled “Health Co-Benefits of Sub-National Renewable Energy Policy in the U.S.” The research
leverages the MIT United States Regional Energy Model and the InMAP model to comprehensively explore the costs and benefits of Renewable Portfolio Standards in the U.S. with a focus on the air quality health co-benefits. We found that the air quality co-benefits of RPSs in the Rust Belt region alone exceed the total economic costs of these policies, subject to uncertainty in the concentration-response and Value of Statistical Life assumptions. This paper will shortly be submitted for peer review.

**Inter-Center Collaborations:** We have ongoing collaborations with the EPA Center for Air, Climate, and Energy Solutions (CACES). We are working with Julian Marshall (University of Washington) on assessing health co-benefits of renewable energy policy, and expect to submit this work for publication shortly. We are also working with Inês Azevedo (Carnegie Mellon University) on quantifying tradeoffs between air quality and climate outcomes associated with proposed carbon taxes.

**Publications/Presentations:**

**Publications:**


**Presentations:**
Qiu, M., “Is renewable energy policy an optimal solution for pollution reductions: Evidence from wind power in U.S.”. MIT Center for Energy and Environmental Policy Research (CEEPR) Student Seminar, Cambridge, MA, 2018

Qiu, M., “Is renewable energy policy an optimal solution for pollution reductions: Evidence from wind power in U.S”. MIT Electricity Student Research Group Seminar, Cambridge, MA, 2018

Qiu, M., Selin, N., Choriat, C., Zigler, C., “Effectiveness of US state level climate policies: Evidence from plant level data in power sector”. Harvard/MIT ACE Center Science Advisory Committee Meeting, Boston, MA, 2018

Dedoussi, I.C., "Air pollution-related early deaths in the US: sectors, states, and species”, Harvard School of Public Health, Boston MA, 2018

Dedoussi, I.C., “Combustion attributable, PM$_{2.5}$-related early deaths in the US: states, sectors, and species”, Harvard/MIT EPA ACE Center Scientific Advisory Committee Meeting, Boston MA, 2018
**Future Activities:** Our first goal for project year 4 is to complete the CAM-based climate simulation ensemble. Although delayed as explained above, we now expect this to be complete in early 2019. This will enable us to complete our air quality simulation ensemble for the 1980-2100 period, followed by analysis of the time of emergence for air quality impacts using the methodology established in year 3 (Objectives 1 and 5). The same ensemble will be used to quantify changes in anticipated exposure over the 21st century, including quantification of uncertainty within and between climate scenarios (Objectives 4 and 5). During the next year, we also intend to leverage the previous year’s work on offsets of energy emissions by wind power. We will use this data to estimate the effects of renewable policies on air quality and human health in U.S., applying a state-of-the-art chemical transport model (GEOS-Chem). This will allow us to quantify the relative effectiveness of renewable policy as an approach to reducing air pollution. The long-term goal is to provide insights into the question of whether existing and planned renewable policy is targeting the optimal set of power plants in terms of abatement costs and environmental benefits (Objectives 2 and 3).

We are also beginning a new project, quantifying air quality tradeoffs and benefits associated with recent state and federal environmental legislation. For each identified policy, one or several counterfactual emissions inventories will be developed. The GEOS-Chem nested model will be used to estimate the effect that each policy had, or would have had, on ozone and particulate matter concentrations over the affected period at a resolution of ~30 km across the contiguous United States. Air quality impacts will be monetized and compared to the expected economic costs and benefits of the legislation (Objectives 1 and 3). Work during year 4 will focus on finding candidate policies and developing a methodology for estimating their effects on emissions.

In addition to completing in-project objectives, we intend to intensify our collaboration with project 1. We will first compare estimates for 2000-2015 using CAM meteorology with estimates from project 1’s simulations using the high-resolution MERRA-2 reanalysis meteorology. Since the MERRA-2 reanalysis dataset assimilates observed meteorological data, this comparison will allow us to quantify errors in temporal and spatial variability present in the CAM dataset, and how these will affect our long-term estimates of spatial and temporal variability over the full 100-year simulation period. Simulated PM$_{2.5}$ concentrations for the same period will also be compared with the monitor data collected under project 1. This will allow bias quantification for the estimates of future PM concentration under the existing CAM scenarios. These cross-project collaborations will improve outcomes for Objectives 1, 4, and 5.

Finally, we expect to complete and submit for review manuscripts based on both projects investigating air quality trade-offs associated with carbon policy. We anticipate no significant additional timeline adjustments.

**Supplemental Keywords:** PM$_{2.5}$, ozone, adjoint method, air quality, greenhouse gas, computable general equilibrium, mercury, polycyclic aromatic hydrocarbons, PAHs

**Relevant Web Sites:** Our web site [http://globalchange.mit.edu](http://globalchange.mit.edu) highlights recent work from across our research program, and links to the web sites of individual PIs. In addition, the Harvard/MIT ACE Center website, [https://content.sph.harvard.edu/ace/](https://content.sph.harvard.edu/ace/)