

Facilitating Inclusion of Geocoded Pollution Data into Health Studies

Rebecca E. Greenblatt, Blanca E. Himes

Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine,
University of Pennsylvania, Philadelphia, PA

Abstract

Exposure to pollutants impacts health and has been associated with a range of diseases, including respiratory and heart diseases, as well as all-cause mortality. Because taking exposure measures for individual studies is costly and impractical, most rely on data from sources such as the Environmental Protection Agency (EPA), which provides a wealth of publicly available pollution measures taken at over two thousand monitoring sites across the United States. While EPA data is readily available, estimating pollution exposure at a given latitude-longitude location remains computationally intensive. We developed Pollution-Associated Risk Geospatial Analysⁱs SITE (PARGASITE), an online web-application and R package, that can be used to estimate levels of pollutants in the U.S. for 2005 through 2017 at user-defined geographic locations and time ranges. We demonstrate how PARGASITE can facilitate the study of associations between exposures and health outcomes using as an example an analysis of asthma risk factors among adults.

Introduction

Exposure to air pollution, a mixture of many chemicals emitted by human activity and natural events, leads to various health effects¹. Major pollutants, including particulate matter with diameter $< 2.5 \mu\text{m}$ (PM_{2.5}), ozone, nitrogen dioxide (NO₂), sulfur dioxide (SO₂) and carbon monoxide (CO), are monitored by regulatory agencies such as the United States Environmental Protection Agency (EPA) to reduce their impact on human health². Despite efforts to reduce pollutants, they continue to be associated with disease. Increased levels of PM_{2.5} and ozone, even at levels below the national air quality standard, were associated with increased all-cause mortality^{3,4}. Air pollution has been associated with increased morbidity⁵⁻⁸ and death⁹ among those with respiratory and cardiac conditions. Long-term exposure to PM_{2.5}, ozone, NO₂, SO₂ and CO has been associated with increased risk of acute respiratory distress syndrome (ARDS)⁸. Exposure during infancy to NO₂ was associated with increased risk of development of childhood asthma in Latinos and African Americans¹⁰, and early-life particulate exposures were associated with reduced lung function¹¹. Exposure to PM_{2.5} has been associated with asthma among children in Medicaid¹² and exposure to PM_{2.5} during the third trimester of pregnancy has been associated with elevated blood pressure in children¹³ and with infant mortality in Sub-Saharan Africa¹⁴.

Ongoing studies continue to identify critical links between exposure to pollution and health outcomes and provide evidence that is used to set national and global air quality standards^{2,15}. While there is a wealth of data captured by over two thousand EPA monitoring sites, estimating air pollution exposure at a given latitude-longitude location remains computationally intensive, as estimates have to be imputed at locations not directly associated with a monitor. An accessible application from which such pre-computed values can be extracted does not exist; individual studies have therefore obtained their own geocoded pollution estimates^{3,4,6,8,10,11,16,17}, a task that is time-consuming and often redundant. In this paper, we describe Pollution-Associated Risk Geospatial Analysⁱs SITE (PARGASITE), an online web-application and R package we developed to facilitate inclusion of pollution measures in health studies. As a proof of concept, we used these tools to extend an analysis of asthma risk factors among adults based on data from the Behavioral Risk Factor Surveillance Survey (BRFSS) to include EPA air pollution measures^{18,19}.

Methods

Environmental data. Summary files containing measures for PM_{2.5}, ozone, NO₂, SO₂ and CO from each available EPA monitor were downloaded from the EPA website for years 2005-2017²⁰. Annual and daily summary files were used to compute annual and monthly estimates, respectively, for each measure expressed as the arithmetic mean for a given pollutant at a given site. For sites with multiple entries, an average of all values was taken. 24-hour averages were used for PM_{2.5}, 8-hour averages for ozone, and 1-hour averages for NO₂, SO₂ and CO.

Interpolation. A raster layer covering the contiguous U.S. (extent: 67.01 - 124.68 ° W and 25.13 - 49.38° N) was created with a 10km x 10km grid using the R *raster* package^{21,22}. An inverse-distance-squared weighted average of

the five nearest monitoring stations was computed to estimate pollutant-specific exposure at a given location using the R *gstat* package^{23, 24}.

BRFSS data and analysis. Asthma status and socio-demographic characteristic data (i.e., age, sex, race/ethnicity, body mass index (BMI), education, income and smoking history) corresponding to the Behavioral Risk Factor Surveillance Survey (BRFSS) for years 2007-2012 was obtained as described in detail previously¹⁸. Asthma cases were defined as those who answered affirmatively that they currently had asthma and controls were respondents who reported never having asthma. Because the most specific geographic identifier for BRFSS subjects was county of residence, we obtained pollution estimates for each U.S. county for a given year by taking the average of all raster points that fell within that county. Subjects who did not reside within the contiguous U.S., and thus for whom pollution exposure estimates could not be obtained, were excluded from the analysis. A logistic regression model with pollution exposure and socio-demographic characteristics as independent variables was fit to predict asthma status using the R *Survey* package²⁵.

PARGASITE design. The R *Shiny* package, which enables creation of web applications that users can easily engage without needing a background in programming²⁶, was used to create PARGASITE. The app code was saved on a DigitalOcean droplet containing a *Shiny* server that uses various R packages, including *Leaflet* to display an interactive map^{26, 27} and raster files are hosted by Amazon Web Services. The web-based application is displayed at <http://www.pargasite.org> and the full code used to generate raster files, the app and related R *pargasite* package are available at <https://github.com/HimesGroup/pargasite>.

Results

Web-based application. Users can display raster layers by year and pollutant, as well as obtain estimates corresponding to a given location clicked on the map (Figure 1). Additionally, users can upload a comma-separated-value (csv) file with Longitude and Latitude columns, and subsequently, download a csv file that contains corresponding pollution estimates for a user-specified time period.

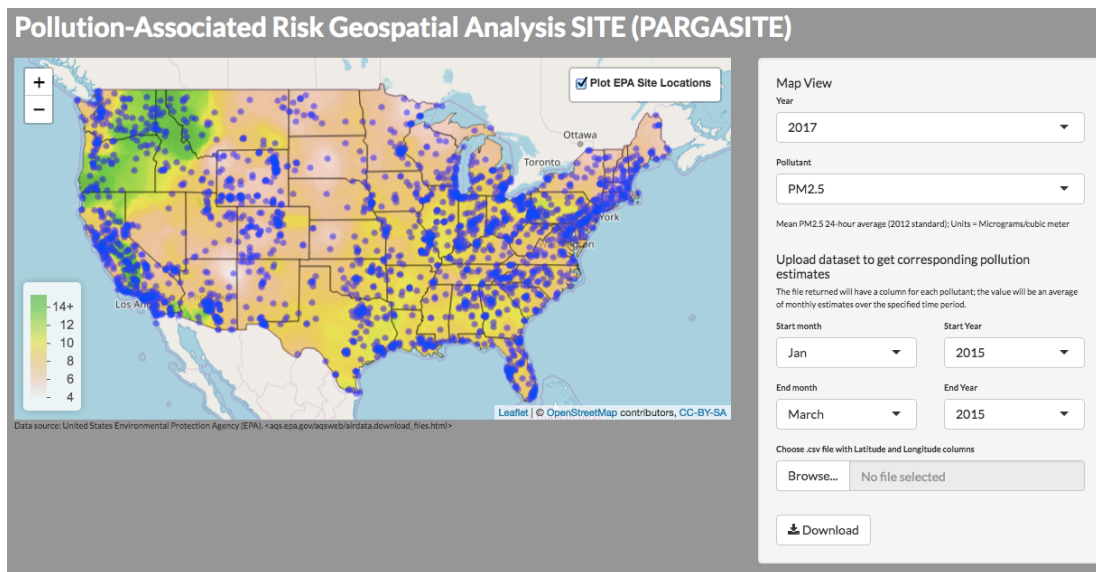


Figure 1. Screenshot of <http://www.pargasite.org>.

PARGASITE R package. The *pargasite* R package was developed as a tool for researchers to link subject or patient data to pollution measures using our pre-calculated estimates while using local servers, rather than having to upload geographic coordinates via the internet. Because geographic coordinates may be protected health information, this option is preferable in some situations. Similar to the online app, the R *pargasite* package allows a user to obtain an estimate for a given location, pollutant, and unit of time (month or year between January 2005 and December 2017), as well as apply this functionality to an entire dataset of geographic coordinates. In addition to addressing data

transfer security concerns, use of the R package has improved speed compared to the web-based application when pollution estimates for a large number of geographic coordinates are sought.

U.S. EPA Pollution measures. Raster layers were obtained for each pollutant (i.e., PM_{2.5}, ozone, NO₂, SO₂ and CO) for all months and years between 2005 and 2017. Visualizing these layers in PARGASITE provides users with information regarding pollution trends in the U.S. Levels of all five pollutants have generally decreased between 2005 and 2017. For example, average PM_{2.5} in 2005 was 10.13 µg/m³, while in 2017 it was 7.51 (Figure 2). In terms of geographic distribution, 2017 PM_{2.5} levels were generally highest in the Northwest and in California compared to other U.S. regions.

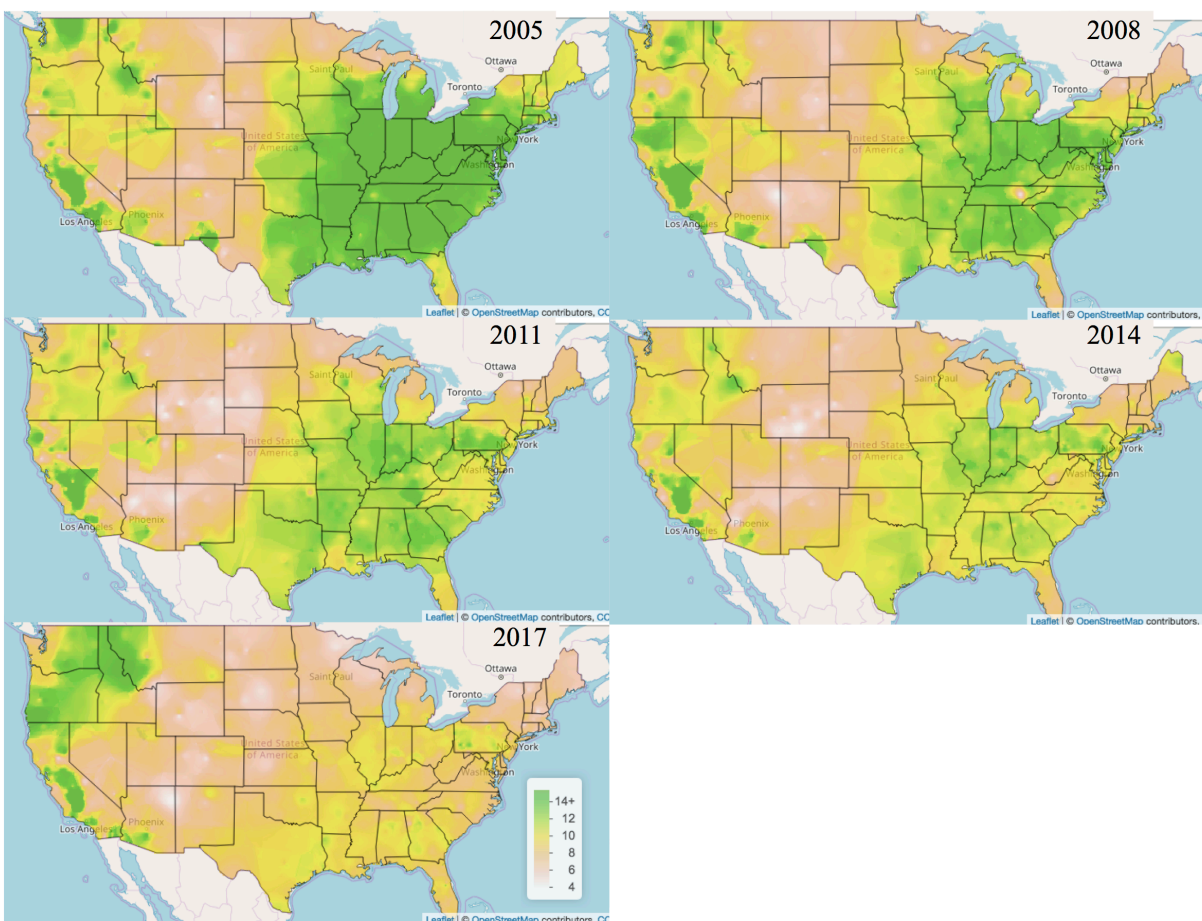


Figure 2. PM_{2.5} raster layers in PARGASITE show geographic variability in the U.S. and overall decrease between 2005 and 2017

Displays of annual data from the 2017 raster files for Ozone, SO₂, CO and NO₂ also show striking variability across the U.S. (Figure 3). For example, ozone levels were highest in the West and Southwest. Variation over time of these pollutants has also decreased overall between 2005 and 2017: Ozone decreased from 46.52 parts per billion (ppb) in 2005 to 42.56 ppb in 2017, while NO₂ decreased from 21.29 ppb to 15.77 ppb, SO₂ from 8.16 ppb to 3.57 ppb, and CO from 498.68 ppb to 263.15 ppb. As expected, a warm season ozone effect was observed; based on data from 2017, April through September had a mean ozone level of 36.0 ppb compared to 28.9 ppb for the rest of the year⁴.

Inclusion of pollution data with BRFSS variables to identify factors associated with asthma. Weighted asthma prevalence for the years 2007-2012 among 965,114 BRFSS subjects with complete data was 8.4%. Respondents came from 398 U.S. counties, with weighted asthma prevalence ranging from 3.2% in Flagler, Florida to 15.9% in Calvert, Maryland. Mean and standard deviation of pollution estimates for counties within the contiguous U.S. for years 2007-2012 are shown in Table 1. Pearson correlation coefficients between pollutants were less than or equal to

0.4, supporting their use as independent variables. SO₂ varied most by county and was skewed left, while CO, NO₂, PM_{2.5} and ozone varied less by county and were normally distributed. Variability of each pollutant across U.S. counties can be observed in the choropleth maps shaded by normalized county mean provided in Figure 4.

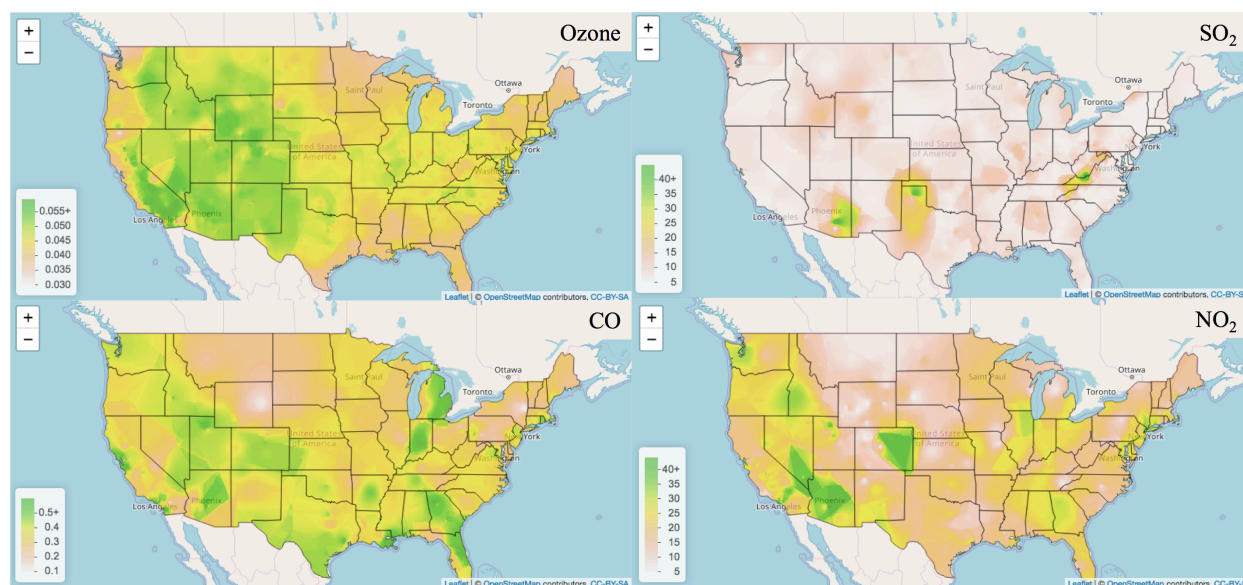


Figure 3. 2017 raster layers for Ozone, SO₂, CO and NO₂ in PARGASITE

In an adjusted model that included all pollution variables as well as socio-demographic variables shown in Table 2, PM_{2.5} was negatively associated with asthma prevalence, while NO₂ and SO₂ were positively associated, and ozone and CO had no effect. For example, an increase of 5.91 ppb of NO₂ (one standard deviation) was associated with a 1.03 times increase in asthma prevalence. An increase of 11.82 ppb (5.91 x 2) was associated with a 1.06 (1.03²) times increase in asthma prevalence.

Table 1. Summary of U.S. county pollution data used in BRFS analysis.

Pollutant	Units	Mean (SD)
PM _{2.5}	Micrograms/cubic meter	9.78 (2.14)
Ozone	Parts per billion	44.63 (4.86)
NO ₂	Parts per billion	17.73 (5.91)
SO ₂	Parts per billion	6.49 (4.86)
CO	Parts per billion	333.46 (116.18)

Discussion

The EPA sets National Ambient Air Quality Standards for pollutants considered harmful to public health and the environment in accordance with the Clean Air Act²⁸. Primary standards are set to protect those who are more sensitive to air pollution, such as persons with asthma, children and the elderly. As examples, the annual mean of PM_{2.5} averaged over 3 years should not exceed 12 µg/m³ and the 99th percentile of 1-hour daily maximum concentrations of SO₂ averaged over 3 years should not exceed 75 ppb. Our data showed that only 3.7% of the 10x10 km² estimates exceeded the annual PM_{2.5} standard in 2017 while 31.0% had in 2005. A future iteration of PARGASITE can expand to include values that correspond directly to the EPA national air quality standards; for example, estimates can be made based on the 99th percentile SO₂ measurement in addition to the mean measurement.

Overall downtrends of all pollutants captured in PARGASITE are consistent with those reported by the EPA²⁹. The geographic and temporal changes in ozone pollution we observed were consistent with increased oil and gas extraction in the Southwest and a cleanup of power plants in the Eastern U.S. described in the 2017 State of the Air Report published by the American Lung Association³⁰. The significant downtrend over time in PM_{2.5} was consistent with a reduction in emissions from coal-fired power plants and a cleanup of diesel engines, while persistently high

levels in Western U.S. states were consistent with the occurrence of wildfires. As expected, we observed a warm season ozone effect; this effect occurs because sunlight catalyzes the conversion of gases like nitric oxides, hydrocarbons and carbon monoxide to ozone³⁰.

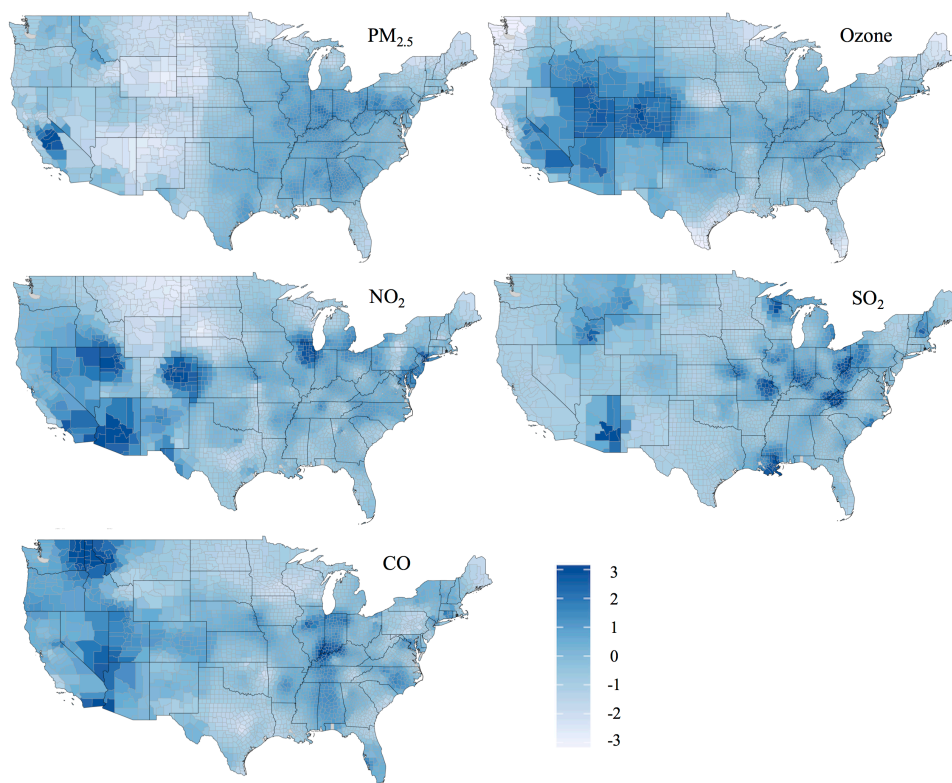


Figure 4. 2007-2012 normalized pollution estimates by county. Legend scale corresponds to number of standard deviations above or below the mean for each pollutant calculated using all available 2007-2012 measures.

The inverse-distance-weighting approach employed for spatial interpolation in PARGASITE is one of several available methods. Although many previous health-by-air pollution studies have used inverse-distance-weighting to obtain pollution estimates^{10, 11, 16, 17}, other established approaches worth mentioning are spline interpolation and kriging. Spline interpolation is another deterministic method that minimizes overall surface curvature and is thus preferred when a smooth surface is desired. Kriging is a more complex method that considers autocorrelation among points, enabling measures of uncertainty to be obtained. Successful fitting of a variogram necessary for kriging can be difficult to fit for air monitoring data³¹. Estimations from kriging and inverse distance weighting have been shown to converge in regions where monitors are sparse, but diverge where monitors are denser³¹. Thus, rasters obtained via kriging are of particular value for studies of smaller geographic regions where many monitoring sites are located and have a wide distribution of values³¹. Future versions of PARGASITE will include rasters obtained via spline interpolation and kriging, allowing users to explore differences in interpolated pollution measures obtained via different approaches.

The pollution estimates provided by PARGASITE can be used for various epidemiological and informatics studies. For epidemiologic studies, the approach we followed with BRFSS data shows how pollution measures can be readily added to existing datasets with geographic information. While the associations we measured between NO₂ and SO₂ exposure with adult asthma prevalence are consistent with previous findings^{32, 33}, further work that more carefully considers the time frame of exposure (e.g., seasonality) and limitations of using area estimates to assign exposures to individuals are necessary to validate these relationships. In cases such as the BRFSS analysis outlined here, pollution estimates were calculated by averaging all raster points within the smallest region available. When possible, using a relevant geographic coordinate per person will provide the most specific estimates, rather than a

county, zip-code, or census tract. For informatics studies, pollution measures from PARGASITE can be incorporated into electronic health record (EHR)-derived data to better understand local health trends, or eventually, incorporated into provider- or patient-facing applications to improve personalized risk assessment^{34,35}. Because EHRs do not capture detailed environmental information required for proper characterization of individual-level risk factors, approaches to include such data are valuable. More broadly, incorporating pollution measures is helpful for efforts in precision medicine, including conducting gene-by-environment studies¹⁰, where considering environmental exposures is an essential part of understanding which interventions and treatments work for particular individuals³⁶.

The large-scale pollution measures considered here have been linked to various health outcomes, but it is worth noting that various smaller-scale measures are also important³⁷⁻³⁹. For example, exposure to cigarette smoke, smoke from wood-burning stoves, and various indoor pollutants found in homes and places of work also influence health outcomes³⁷⁻³⁹. Our extended analysis of BRFSS data to include EPA pollution measures is thus unable to account for the full range of environmental factors that influence persons with asthma. Nonetheless, efforts to understand the impact of large-scale pollution measures are helpful to identify associations and inform policy. Advances in health information and sensor technologies have resulted in the production of portable sensors designed to measure personal environmental exposures cheaply and conveniently⁴⁰⁻⁴³. As some of these sensors become increasingly accurate devices to measure pollution, and their ability and utility to capture increasingly finer-scaled geographic differences in pollution is validated, smaller-scale measures of pollution should also be included in studies that seek to understand relationships between exposures and disease outcomes. Ultimately, capturing small-scale measures along with those that are broadly distributed geographically will yield the greatest insights into the relationship between exposures and health.

In ongoing work, we are improving PARGASITE by including daily EPA measures, which will improve the capacity to link pollution exposure to short-term events (e.g., disease exacerbations). In addition to considering individual pollution variables, we will also provide combined measures such as the EPA's Air Quality Index Score, as individual pollution variables are often related and interact with each other. For example, particulate matter impacts health differently based on climatological conditions and the presence of other pollutants, such as ozone. While PARGASITE currently relies on data from EPA monitoring sites, future versions will include data on vegetation, elevation, humidity, temperature, wind patterns and information collected by satellite imaging technologies⁴⁴, all of which provide additional indications for air quality.

Table 2. Variables associated with asthma among BRFSS respondents. For each variable, OR (95% CI) are shown. N = 965,114.

		Unadjusted ORs (CI)	Adjusted ORs (CI)
Age			
	10 years	0.98 (0.97, 0.99)**	0.95 (0.94, 0.96)**
Sex			
	Male	Reference	Reference
	Female	1.80 (1.74, 1.85)**	1.81 (1.75, 1.87)**
Race			
	White	Reference	Reference
	Black	1.18 (1.13, 1.24)**	0.96 (0.92, 1.00)
	American Indian/Alaskan Native	1.66 (1.43, 1.92)**	1.43 (1.23, 1.67)**
	Asian/Pacific Islander	0.58 (0.52, 0.65)**	0.66 (0.59, 0.74)**
	Hispanic	0.72 (0.68, 0.76)**	0.64 (0.61, 0.68)**
Education Level			
	Less than high school	0.76 (0.69, 0.84)**	0.69 (0.63, 0.77)**
	High school	1.07 (1.04, 1.10)**	0.88 (0.85, 0.91)**
	Some college or more	Reference	Reference
Yearly Household Income (USD)			
	< \$25,000	1.60 (1.54, 1.66)**	1.58 (1.52, 1.65)**
	\$25,000-\$75,000	1.14 (1.10, 1.18)**	1.09 (1.05, 1.13)**

	≥ \$75,000	Reference	Reference
Body Mass Index			
	Not overweight or obese	Reference	Reference
	Overweight	1.07 (1.03, 1.11)**	1.23 (1.18, 1.27)**
	Class 1 obese	1.44 (1.38, 1.51)**	1.58 (1.52, 1.66)**
	Class 2 obese	2.11 (2.00, 2.23)**	2.19 (2.07, 2.31)**
	Class 3 obese	3.30 (3.11, 3.50)**	3.11 (2.92, 3.31)**
Smoking status			
	Never smoked	Reference	Reference
	Former smoker	1.17 (1.13, 1.21)**	1.23 (1.19, 1.27)**
	Current smoker	1.32 (1.27, 1.37)**	1.29 (1.24, 1.34)**
PM2.5			
	Normalized	0.99 (0.97, 1.00)	0.95 (0.93, 0.96)**
Ozone			
	Normalized	1.03 (1.01, 1.04)**	1.01 (1.00, 1.02)
NO2			
	Normalized	1.00 (0.99, 1.01)	1.03 (1.02, 1.04)**
SO2			
	Normalized	1.06 (1.04, 1.07)**	1.03 (1.01, 1.05)**
CO			
	Normalized	1.00 (0.98, 1.01)	1.00 (0.98, 1.01)

* < 0.05; ** < 0.001

Conclusion

We created PARGASITE, an online web-application and an R package that facilitate inclusion of pollution measures in health studies. The web application has an intuitive, user-friendly interface that requires no programming background for its use, while the R package allows users with some programming experience to obtain pollution measures on local servers. We demonstrated the utility of PARGASITE by extending an analysis of asthma risk factors among adults based on data from BRFSS to include EPA air pollution measures.

Acknowledgements

This work was supported by National Institutes of Health (NIH) R01-HL133433 (BEH) and R01 HL141992 (BEH). Additionally, this work was supported in part by the National Institute of Environmental Health Sciences under award number P30ES013508.

References

1. Mannucci PM, Harari S, Martinelli I, Franchini M. Effects on health of air pollution: A narrative review. *Internal and Emergency Medicine*. 2015;10(6):657-62.
2. Process of reviewing the national ambient air quality standards. United States Environmental Protection Agency. Available from: <https://www.epa.gov/criteria-air-pollutants/process-reviewing-national-ambient-air-quality-standards>. Accessed on July 1 2018.
3. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, et al. Air pollution and mortality in the medicare population. *New England Journal of Medicine*. 2017;376(26):2513-22.
4. Di Q, Dai L, Wang Y, Zanobetti A, Choirat C, Schwartz JD, et al. Association of short-term exposure to air pollution with mortality in older adults. *JAMA*. 2017;318(24):2446-56.
5. Peel JL, Tolbert PE, Klein M, Metzger KB, Flanders WD, Todd K, et al. Ambient air pollution and respiratory emergency department visits. *Epidemiology*. 2005;16(2):164-74.
6. Dominici F, Peng RD, Bell ML, Pham L, McDermott A, Zeger SL, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*. 2006;295(10):1127-34.
7. Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. Ozone and short-term mortality in 95 us urban communities, 1987-2000. *JAMA*. 2004;292(19):2372-8.

8. Reilly JP, Zhao Z, Shashaty MG, Koyama T, Christie JD, Lanken PN, et al. Low to moderate air pollutant exposure and acute respiratory distress syndrome after severe trauma. *American journal of respiratory and critical care medicine*. 2018.
9. Fann N, Lamson AD, Anenberg SC, Wesson K, Risley D, Hubbell BJ. Estimating the national public health burden associated with exposure to ambient pm2.5 and ozone. *Risk analysis : an official publication of the Society for Risk Analysis*. 2012;32(1):81-95.
10. Nishimura KK, Galanter JM, Roth LA, Oh SS, Thakur N, Nguyen EA, et al. Early-life air pollution and asthma risk in minority children. The gala ii and sage ii studies. *Am J Respir Crit Care Med*. 2013;188(3):309-18.
11. Neophytou AM, White MJ, Oh SS, Thakur N, Galanter JM, Nishimura KK, et al. Air pollution and lung function in minority youth with asthma in the gala ii (genes-environments and admixture in latino americans) and sage ii (study of african americans, asthma, genes, and environments) studies. *Am J Respir Crit Care Med*. 2016;193(11):1271-80.
12. Keet CA, Keller JP, Peng RD. Long-term coarse particulate matter exposure is associated with asthma among children in medicaid. *American journal of respiratory and critical care medicine*. 2018;197(6):737-46.
13. Zhang M, Mueller NT, Wang H, Hong X, Appel LJ, Wang X. Maternal exposure to ambient particulate matter $\leq 2.5 \mu\text{m}$ during pregnancy and the risk for high blood pressure in childhood. *Hypertension*. 2018;72:194-201.
14. Heft-Neal S, Burney J, Bendavid E, Burke M. Robust relationship between air quality and infant mortality in africa. *Nature*. 2018;559.
15. Ambient (outdoor) air quality and health. World Health Organization. Available from: [http://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](http://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health). Accessed on July 1 2018.
16. Wong JY, Margolis HG, Machiela M, Zhou W, Odden MC, Psaty BM, et al. Outdoor air pollution and mosaic loss of chromosome y in older men from the cardiovascular health study. *Environment International*. 2018;116:239-47.
17. Eckel SP, Cockburn M, Shu YH, Deng H, Lurmann FW, Liu L, et al. Air pollution affects lung cancer survival. *Thorax*. 2016;71:891-8.
18. Greenblatt R, Mansour O, Zhao E, Ross M, Himes BE. Gender-specific determinants of asthma among u.S. Adults. *Asthma Res Pract*. 2017;3:2.
19. Mokdad AH. The behavioral risk factors surveillance system: Past, present, and future. *Annu Rev Public Health*. 2009;30:43-54.
20. Pre-generated data files. United States Environmental Protection Agency. Available from: https://aqs.epa.gov/aqswb/airdata/download_files.html. Accessed on July 1 2018.
21. R development core team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
22. Hijmans RJ. Raster: Geographic data analysis and modeling. 2017.
23. Gräler B, Pebesma E, Heuvelink G. Spatio-temporal interpolation using gstat. *The R Journal*. 2016;8(1):204-18.
24. Hijmans R. R spatial. Available from: <http://rspatial.org/index.html>.
25. Lumley T. Survey: Analysis of complex survey samples. 2017. Available from: <https://CRAN.R-project.org/package=survey>.
26. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. Shiny: Web application framework for r. R package version 01329005. 2016.
27. Cheng J, Karambelkar B, Xie Y. Leaflet: Create interactive web maps with the javascript 'leaflet' library. 2018.
28. National ambient air quality standards (naaqs) table. United States Environmental Protection Agency. Available from: <https://www.epa.gov/criteria-air-pollutants/naaqs-table>. Accessed on July 1 2018.
29. National air quality: Status and trends of key air pollutants. United States Environmental Protection Agency. Available from: <https://www.epa.gov/air-trends>. Accessed on July 1 2018.
30. State of the air 2017. American Lung Association. Available from: <http://www.lung.org/assets/documents/healthy-air/state-of-the-air/state-of-the-air-2017.pdf>. Accessed on July 1 2018.
31. Wong DW, Yuan L, Perlin SA. Comparison of spatial interpolation methods for the estimation of air quality data. *Journal of Exposure Analysis and Environmental Epidemiology*. 2004;14:404-15.
32. Delfino RJ, Gong Jr H, Linn WS, Pellizzari ED, Hu Y. Asthma symptoms in hispanic children and daily ambient exposures to toxic and criteria air pollutants. *Environmental Health Perspectives*. 2003;111(4):647-56.

33. Jacquemin B, Sunyer J, Forsberg B, Aguilera I, Briggs D, García-Esteban R, et al. Home outdoor NO₂ and new onset of self-reported asthma in adults. *Epidemiology*. 2009;20(1):119-26.
34. Xie S, Greenblatt R, Levy MZ, Himes BE. Enhancing electronic health record data with geospatial information. *AMIA Jt Summits Transl Sci Proc*. 2017;2017:123-32.
35. Kasthurirathne SN, Vest JR, Menachemi N, Halverson PK, Grannis SJ. Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. *Journal of the American Medical Informatics Association*. 2017;25(1):47-53.
36. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9):793-5.
37. Diette GB, McCormack MC, Hansel NN, Breyse PN, Matsui EC. Environmental issues in managing asthma. *Respiratory care*. 2008;53(5):602-17.
38. Schraufnagel DE, Balmes J, Cowl CT, De Matteis S, Jung SH, Mortimer K, et al. Air pollution and non-communicable diseases: A review by the forum of international respiratory societies' environmental committee, part 1: The damaging effects of air pollution. *Chest*. 2018.
39. Schraufnagel DE, Balmes J, Cowl CT, De Matteis S, Jung SH, Mortimer K, et al. Air pollution and non-communicable diseases: A review by the forum of international respiratory societies' environmental committee. Part 2: Air pollution and organ systems. *Chest*. 2018.
40. De Nazelle A, Seto E, Donaire-Gonzalez D, Mendez M, Matamala J, Nieuwenhuijsen MJ, et al. Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environmental pollution*. 2013;176:92-9.
41. Steinle S, Reis S, Sabel CE. Quantifying human exposure to air pollution--moving from static monitoring to spatio-temporally resolved personal exposure assessment. *The Science of the total environment*. 2013;443:184-93.
42. Himes BE, Weitzman ER. Innovations in health information technologies for chronic pulmonary diseases. *Respir Res*. 2016;17:38.
43. Milà C, Salmon M, Sanchez M, Ambrós A, Bhogadi S, Sreekanth V, et al. When, where, and what? Characterizing personal pm2.5 exposure in periurban india by integrating gps, wearable camera, and ambient and personal monitoring data. *Environmental Science & Technology*. 2018.
44. Van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, et al. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environmental Health Perspectives*. 2010;118(6):847-55.