Powel River Project Final Report 2017-2018

Zooming in on adverse birth outcomes in coalfield regions of Central Appalachia

Michael L Marston, Connor YH Wu, Ethan Smith, Julia M Gohlke, Korine N Kolivras, Leigh-Anne Krometis

Department of Geography, Department of Population Health Sciences, and Department of Biological Systems Engineering, Virginia Tech, Blacksburg VA 24061

ABSTRACT: Health disparities account for significant differences in mortality and morbidity risks in Central Appalachia (encompassing parts of WV, KY, TN and VA) compared to other U.S. regions, yet research addressing environmental factors potentially contributing to these health disparities is lacking. Central Appalachia offers a unique opportunity to examine environmental exposures associated with resource extraction. Coal production from large surface mines was the dominant resource extraction method in the 1990s-2000s and is now decreasing as other resource extraction methods increase. We hypothesize that health risks associated with air and water pollution exposure are greater for Central Appalachian residents living within close proximity to active surface mines. The results described here begins to link exposure with health outcomes using individual-level birth record data. We have extended spatiotemporal characterization of boundaries associated with surface mining between 1990 and 2015 in all Central Appalachia counties. Results indicate that from 1990 to 2015, 1806 km^2 of land across the study area was disturbed by mining activities, which equates to approximately 4.2% of the study-defined Central Appalachia region. Temporal trends show a decreasing amount of active surface mining sites over the study period. Using a previously developed surface mining dataset that only covered southwest Virginia coalfields, we tested the hypothesis that maternal address proximity to active surface mining was positively associated with preterm birth. No significant association was found; however the sample size (n=5008) was very small due to poor geocoding rates, particularly in earlier years, and overall low number of births between 1990-2015 in this small, rural area (n=14,269). Our next steps will be to improve inference and precision in effect estimates by increasing sample size with inclusion of data from TN, KY and WV, application of the accuracy–assessed surface mining extent dataset described here to improve estimate of proximity to active surface mining, and inclusion of watershed boundaries, drinking water violation datasets, as well as airshed characterization. Ultimately, we hope this research will aid in determining the underpinnings of health disparities in Central Appalachian communities, ultimately leading to research, policy and practice improvements that may be generalizable to other rural areas beyond Central Appalachia.

INTRODUCTION

Several previous studies have suggested area-level environmental factors may contribute to LBW and PTB disparities in rural areas (1, 2). In the U.S., heightened rates of adverse birth outcomes in rural areas are only partially explained by higher prevalence of maternal smoking, prenatal care access and use, and income and education disparities (3-8) and a recent global analysis of PTB rates and known sociodemographic, medical, genetic, and environmental risk factors estimates 2/3rds of all PTBs are due to unknown causes (9). Previous examination of birth outcomes in Alabama showed rates of LBW have remained high in isolated rural areas, while declining in urban areas (10).

Previous studies have demonstrated excess adverse birth outcomes in Central Appalachia when compared to other U.S. regions (7, 11-14). While several studies have associated poor health outcomes in Central Appalachia with poverty, low education, smoking and substandard access to health care, these factors do not fully explain the current health disparities that exist (7, 11, 15-18). Although the unique terrain and regional history of resource extraction have prompted considerable speculation of potential environmental health risks, identifying and quantifying these effects has proven difficult (19, 20).

Transitions from underground to surface mining have resulted in major shifts in land use and water quality, and potentially air quality, in coalfield regions over the last 25 years (21-24). Previous studies have demonstrated increased particulate matter less than 2.5 μm in diameter ($PM_{2.5}$) downwind of active surface mining sites and increased metal concentrations in surface and groundwater during and after active surface mining (25, 26). Other studies provide initial evidence associating county-level health outcomes with the presence of surface mining (27-30); however limitations in study design render direct connections to exposure difficult (19, 20). Specific air and water exposures resulting from activities associated with surface mining have not been disaggregated temporally and spatially to more clearly define exposed and unexposed populations; therefore these approaches run the risk of exposure misclassification when large differences in exposure exist at relatively fine temporal and spatial scales (e.g., due to localized, acute point sources and topography).

Recent publications from our group and the National Toxicology Program have highlighted the need for improved characterization of exposures to address the potential for bias in community health studies in areas with surface mining (19, 20). A handful of ecological studies have been conducted to examine associations between surface mining and health outcomes (11, 27, 28, 31, 32). In one study, level of coal mining was not significantly associated with county-level mortality rates after accounting for sociodemographic characteristics (11). In contrast, other analyses that have shown county-level estimates of coal mining as a significant predictor of all cause and cause-specific mortality after accounting for sociodemographic characteristics (27, 28, 31, 32). Ultimately, conclusions from these studies are limited since individual-level outcomes, exposures and covariates were not evaluated. Overall, these studies provide some evidence that environmental exposures resulting from surface mining may contribute to birth outcome disparities in Appalachia, but better spatiotemporal characterization of exposure and outcomes is needed.

To begin to address these gaps, the present paper presents a finer scaled spatiotemporal characterization of surface mining extents in Central Appalachia (Obj. 1 below), characterization of drinking water violations (Obj. 2 below), and an analysis of the association between adverse birth outcomes in the SW Coalfields of VA prior to, during, and after major shifts in surface mining between 1990-2015 (Obj 3 and 5 below for VA). Future research will complete these objectives for the extent of the Central Appalachia study area and also include airshed characterization.

OBJECTIVES

1. Spatiotemporally characterize surface mining extents in Central Appalachia (58 counties in VA, WV, KY, and TN) between 1990 and 2015 using Landsat data on a monthly time resolution and 30m spatial resolution.

2. Define watershed boundaries and Safe Drinking Water Act treatment plant violations within those watershed boundaries for a subset of surface mining extents.

3. Geocode birth records and assign pregnancy trimesters to within or outside 1km, 2km, and 5 km buffers of surface mines.

4. Assign maternal address locations to within or outside of watershed boundaries for a subset of surface mine extents.

5. Determine the difference in preterm birth, low birth weight, and term low birth weight rates in within boundary versus outside of boundary groups.

METHODS

Methods for Objective 1

Study area and data

The study area for this research is the central Appalachian coalfield (Figure 1). Li, Zipper, Donovan, Wynne and Oliphant (33), suggested that the Appalachian coalfield is ideal for examining surface mine detection techniques due to the intensive mining that has occurred within the region in recent decades. The study-defined central Appalachian coalfield is approximately 43,400 km^2 and is comprised of all or parts of 58 counties within the states of Kentucky, Tennessee, Virginia, and West Virginia (Table 1). The study region was identified by obtaining the Appalachian coalfield boundary from the United States Geological Survey (USGS) and identifying the areas within the 58 selected counties that were also within the coalfield boundary (coalfield boundary available at https://pubs.usgs.gov/of/2012/1205/Downloads/Metadata/Coal_Fields.html). The 2011 National Land Cover Database (NLCD), which was made available by Jin, Yang, Danielson, Homer, Fry and Xian (34), was also downloaded and used to examine the land cover of the study region, and to aid in the classification process. According to the 2011 National Land Cover Database (NLCD), 72% (~43,000 km^2) of the study area is classified as forest land cover, while 16% (~9,600 km^2) is classified as pasture and grassland. The remaining 12% of the study area is predominantly classified as developed (open space and low intensity) and as barren land. According to data gathered from the Kentucky Department for Natural Resources, the Tennessee Department of Environment and Conservation, the Virginia Department of Mines, Minerals, and Energy Surface Mining, and the West Virginia Department of Environmental, 20% (~12,000 km^2) of the study area has historically been permitted for surface mining.

To identify surface mines across the study region from 1990 to 2015, all available Landsat images for Path 18 Row 33, Path 18 Row 34, Path 19 Row 33, Path 19 Row 34, Path 19 Row 35, Path 20 Row 34, Path 20 Row 35 (WRS II grid system) were downloaded for the entire study period (1990 – 2015) from the United States Geological Survey (USGS) (http://earthexplorer.usgs.gov). The Landsat data are available at a temporal resolution of 16-days and at a spatial resolution of 30 meters. Specifically, level 1T products were used for this study. These data are co-registered by the Earth Resources Observation and Science Center and are processed with the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (35). This processing includes surface a reflectance correction, a water and snow mask, and a cloud shadows mask for each acquisition date. For each Path Row scene, all Landsat images from 1990 to 2015 were examined for spatial completeness, and the best available image per month was selected. Following the methods of Li, Zipper, Donovan, Wynne and Oliphant (33), no post-2003 Enhanced Thematic Mapper (ETM+) images were used to avoid data gap issues caused by the scan line corrector problem. Additionally, no images from 1991 were used in the analysis presented below as cloud cover was present in all images across the study area. The 1991 cloud cover issue was also noted by Li, Zipper, Donovan, Wynne and Oliphant (33), who also chose to eliminate all images from this year. The six bands for the “best available” Landsat image for each month, and for each of the seven scenes, were stacked and the study area boundary was used to clip each of the stacked images. The cloud and water masks for each Landsat image were then used to eliminate those areas from the analysis. It is important to note that although the best available Landsat image was selected for each of the of the seven Path/Row scenes that cover the study area, and for each of the 300 months (25 years X 12 months) through the study period, there were Landsat images used in the analysis for each Path/Row with relatively low spatial completeness (< 50%).

For each of the 58 counties within the study area, high-resolution aerial imagery was downloaded from the National Agricultural Imagery Program (NAIP) to assist in the training and classification portions of the study (These images are available at: https://gdg.sc.egov.usda.gov/GDGHome_DirectDownLoad.aspx). These images are available from 2003 to 2015 at a temporal resolution of approximately two years.

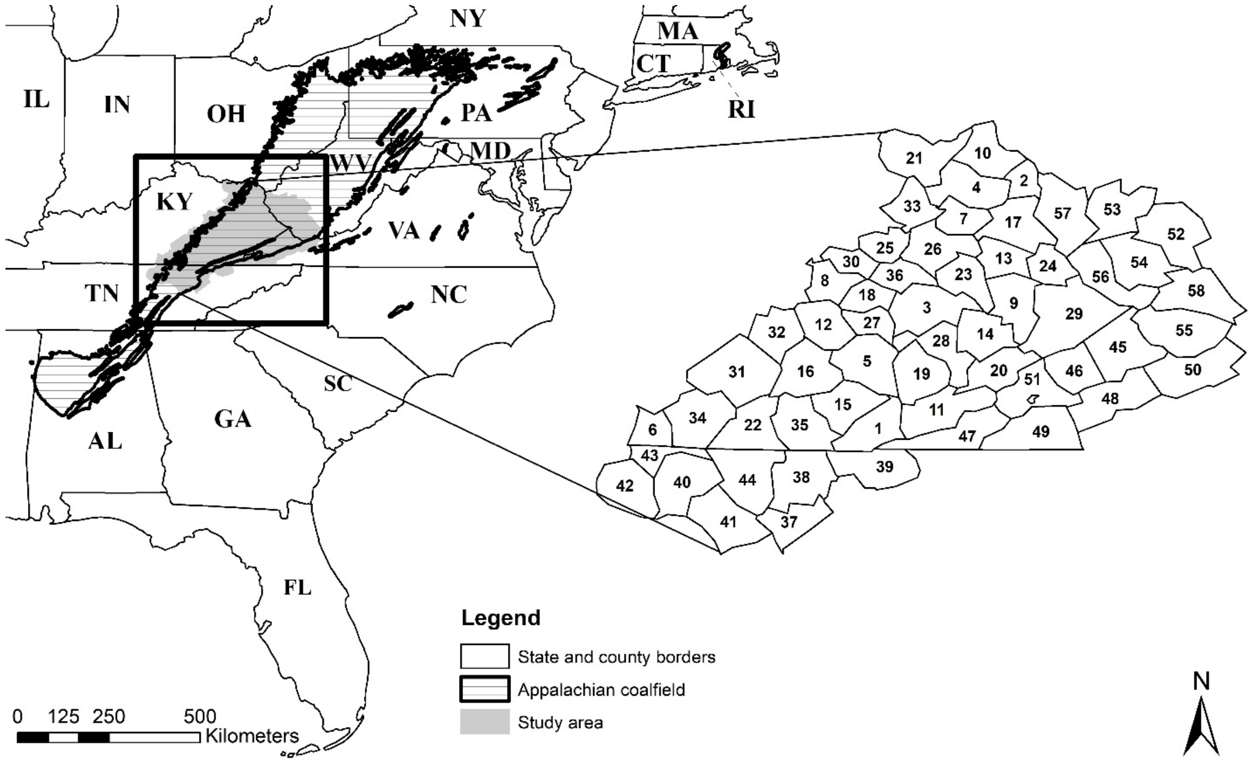

Figure 1. The study region within the Appalachian coalfield and the study counties. The county names corresponding to the county numbers shown above are listed in Table 1.

Table 1. County names corresponding to the numbers shown in Figure 1.

County #	County Name, State	County #	County Name, State
1	Bell, KY	30	Powell, KY
2	Boyd, KY	31	Pulaski, KY
3	Breathitt, KY	32	Rockcastle, KY
4	Carter, KY	33	Rowan, KY
5	Clay, KY	34	Wayne, KY
6	Clinton, KY	35	Whitley, KY
7	Elliott, KY	36	Wolfe, KY
8	Estill, KY	37	Anderson, TN
9	Floyd, KY	38	Campbell, TN
10	Greenup, KY	39	Claiborne, TN
11	Harlan, KY	40	Fentress, TN
12	Jackson, KY	41	Morgan, TN
13	Johnson, KY	42	Overton, TN
14	Knott, KY	43	Pickett, TN
15	Knox, KY	44	Scott, TN
16	Laurel, KY	45	Buchanan, VA
17	Lawrence, KY	46	Dickenson, VA
18	Lee, KY	47	Lee, VA
19	Leslie, KY	48	Russell, VA
20	Letcher, KY	49	Scott, VA
21	Lewis, KY	50	Tazewell, VA
22	McCreary, KY	51	Wise, VA
23	Magoffin, KY	52	Boone, WV
24	Martin, KY	53	Lincoln, WV
25	Menifee, KY	54	Logan, WV
26	Morgan, KY	55	McDowell, WV
27	Owsley, KY	56	Mingo, WV
28	Perry, KY	57	Wayne, WV
29	Pike, KY	58	Wyoming, WV

Image analysis methodology

The three-step sequence approach to identify surface mines, which was developed by Li, Zipper, Donovan, Wynne and Oliphant (33), was used to detect surface mines across the study region from 1990 to 2015 (minus 1991) at a monthly temporal resolution.

a. *First Step: Vegetation Index selection*

Li, Zipper, Donovan, Wynne and Oliphant (33) determined that the Normalized Difference Vegetation Index (NDVI; 36) was best at distinguishing bare ground from vegetation during leaf-on periods (i.e., the warm season) when compared to the Normalized Burn Ratio (NBR; 37), the Normalized Difference Moisture Index (NDMI; 38), the tasseled cap greenness-brightness (TC-

GB; 39) difference index, the “Red” band, and the “Near Infrared” (NIR) Band. However, as the research presented here aimed to detect surface mines throughout the calendar year, instead of only during the warm season, each of the vegetation indices’ ability to distinguish bare ground from ground covered with vegetation was examined for both warm (leaf-on) and cool (leaf-off) seasons. To do this, the Path 19 Row 34 Landsat images acquired on July 3, 2006 (warm season) and November 24, 2006 (cool season) were selected, and the various VIs were computed for both Landsat images. Next, 500 points were randomly generated within the merged mining permit boundaries, and 50 points were randomly generated within areas which were NLCD-classified as highly developed. Using NAIP imagery, each of the 550 points was classified as either bare ground or vegetated for both the cool season and warm season image. The VI values for all points were then extracted from each image and each indices’ ability to separate bare ground from vegetated land was compared for both warm and cool seasons using Minitab software (Figure 2). The results of this comparison suggest that NDVI is the best VI for separating bare ground from vegetated land during the cool season and is comparable to NBR at separating bare ground from vegetated land during the warm season. Based off of these results, the results of Li, Zipper, Donovan, Wynne and Oliphant (33), and the desire to use a consistent VI for both warm and cool season surface mine detection, NDVI was selected to delineate mining disturbances across the study region from 1990 to 2015 (minus 1991). For each of the seven Landsat scenes across the study area, monthly time series of NDVI values were then computed for each pixel across the study area by generating NDVI images for each month through the study period.

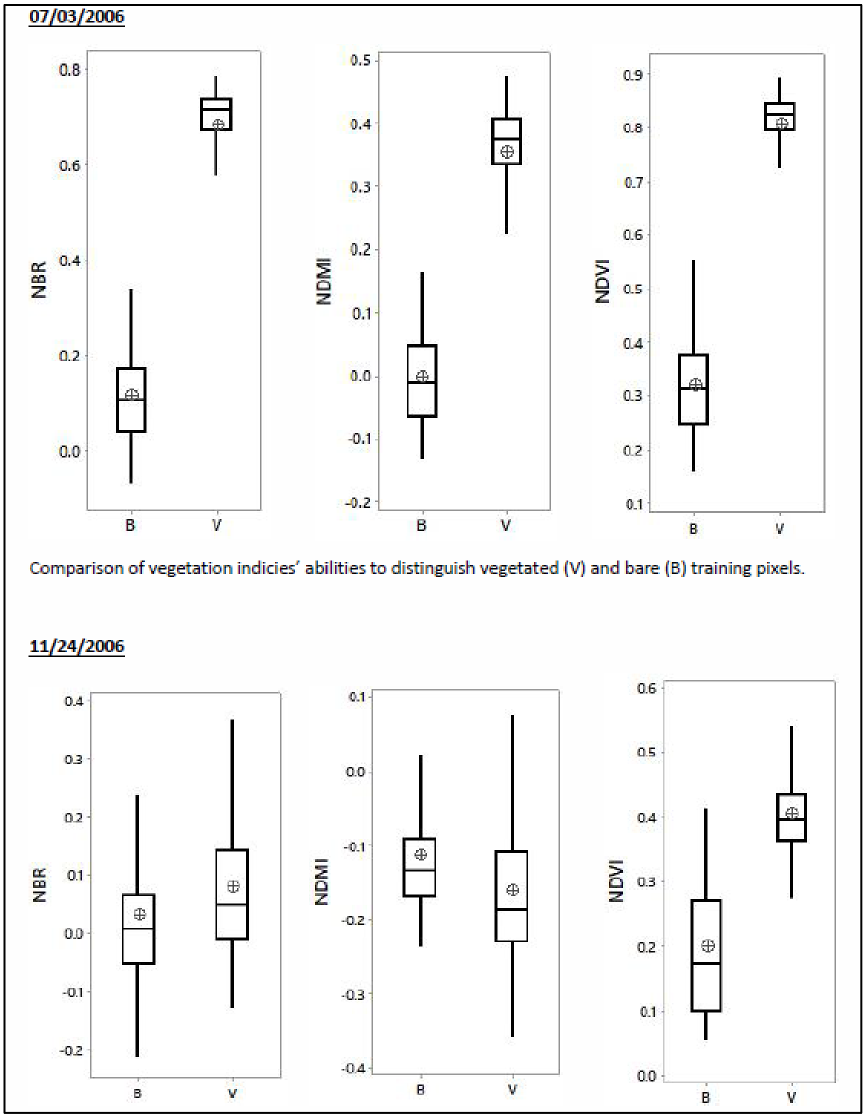

Figure 2. Comparison of Vegetation Indices' ability to detect bare ground (B) from vegetation (V) for the 550 training points. Results from the warm season (cool season) image is shown on the top (bottom) half of the figure.

b. Second Step: CART Classification

As shown in Figure 2, NDVI values vary throughout the calendar year across the study area due to changes in vegetation 'greenness'. As such, bare ground thresholds were computed for each of the twelve calendar months to delineate areas of disturbances from areas of persistent vegetation. To compute a threshold for each month, three Landsat images (Path 19 Row 34) for each of the twelve calendar months were randomly selected and the 550 training points were classified as either vegetation or bare ground for each of the 36 (12 months x 3 for each month) images. In addition to using the Landsat imagery displayed as bands 2, 3, 4 to classify the training points as bare ground or vegetated land, NAIP imagery was also used, when available, to assist in the classification process. The NDVI values were then extracted and bare ground thresholds were computed for each month using CART classification (Table 2). All NDVI images for each of the seven Landsat scenes were then re-processed. Pixels whose time series of NDVI values remained above the NDVI bare ground thresholds for all months throughout the entire study period were classified as persistent vegetation. Otherwise, the pixel was classified as disturbed. Each of the non-PV points were then assigned a second classification. The training pixels that were disturbed by mining activities were classified as EM; meaning that the pixel was disturbed by mining at one or more times over the observation period, and the remaining pixels were classified as OD; meaning that vegetation disturbances other than mining were present through the study period.

Table 2. Bare ground NDVI thresholds for all months within the calendar year. Values were computed using classified training data as input for CART classification.

Month	BG Thres.	Month	BG Thres.
January	0.091	July	0.425
February	0.093	August	0.419
March	0.095	September	0.392
April	0.155	October	0.270
May	0.336	November	0.101
June	0.410	December	0.092

c. *Third Step: Surface Mine Detection across Central Appalachia*

The final step was to separate mined land (EM) from land that experiences other disturbances (OD) through the study period, such as clear-cutting or industrial development. As noted by Li, Zipper, Donovan, Wynne and Oliphant (33), the mining permit boundaries do not always contain the mined areas completely. It is important to note that this could simply be artifact of human error in digitizing the permit boundaries, rather than an indication of illegal mining activities. For this reason, the permit boundaries were not used to delineate EM land from OD land. Instead, this study used an approach similar to that of Li, Zipper, Donovan, Wynne and Oliphant (33) to separate EM land from OD land. Through their analysis, Li, Zipper, Donovan, Wynne and Oliphant (33) determined that the NDVI time series of a pixel following an EM disturbance generally exhibit greater variability than the time series of NDVI values representing OD lands. As such, NDVI standard deviation values were computed for the training pixels classified as EM, and for those classified as OD, for each of the twelve calendar months. CART classification was once again used to compute thresholds of NDVI standard deviation values each of the twelve months (Table 3). For each of the seven scenes, all Landsat images were then re-processed, and only those pixels whose NDVI standard deviations were greater than the CART computed standard deviation threshold for at least eleven out of the twelve months were classified as EM lands. The remaining land was classified as OD.

An additional procedure was applied to correct several errors that were visually apparent. All land areas that were classified by the NLCD 2011 as *Developed* were reclassified from EM to OD. This procedure was also used by Li, Zipper, Donovan, Wynne and Oliphant (33), who noted that some EM lands appeared to remain unvegetated for a prolonged period that resulted in relatively stable time series of NDVI values, and ultimately misclassification (OD instead of EM).

The time series of NDVI values for those pixels classified as EM were processed again to determine the month and year of initial disturbance, and the month and year in which mining activities ceased. For each pixel, the disturbance date was tagged to the month and year when the pixel's NDVI time series went below the CART computed threshold for that respective month. Conversely, the end of mining date was tagged to the month and year that the pixel's NDVI time series became greater than the CART determined threshold for that respective month.

Accuracy Assessment

Table 3. Bare ground NDVI standard deviation thresholds for all months within the calendar year. Values were computed using classified training data as input for CART classification.

Month	BG Thres.	Month	BG Thres.
January	0.071	July	0.110
February	0.05	August	0.116
March	0.052	September	0.140
April	0.052	October	0.114
May	0.099	November	0.077
June	0.090	December	0.070

The monthly EM lands were aggregated to an annual (calendar year) temporal resolution for the accuracy assessment. A validation dataset consisting of 1150 points within four randomly selected counties (Campbell (TN), Clay (KY), Logan (WV), Wise (VA)) was then created using ArcGIS. This validation dataset was initially created by randomly placing 250 points within lands identified as PV, 150 points within OD classified pixels, and 30 points within each of the EM annual categories (25 years). Following the accuracy assessment methodologies of Li, Zipper, Donovan, Wynne and Oliphant (33), validation points were manually deleted if more than one point was located within a homogenously classified pixel group. The validation dataset consisted of 961 points after this procedure, and these points were then manually classified using a suite of validation datasets for the accuracy assessment. Specifically, using NAIP imagery, the NLCD 2011 dataset, and the Landsat images (displayed as bands 2, 3, 4) were used to classify, to the best of our ability, the validation points as PV, OD, or EM. It is important to note that the EM lands were validated based upon initial disturbance date. A confusion matrix was constructed based upon the accuracy assessment and the overall accuracy, kappa coefficient, user's accuracies, and producer's accuracies were computed.

Methods for Objective 2

Drinking water violation reports were downloaded from the USEPA "Safe Drinking Water Information System" (SDWIS; https://www3.epa.gov/enviro/facts/sdwis/search.html) for the states of VA, WV, KY, and TN. In brief, reportable failures to comply with US Safe Drinking Water Act include observed violations of health-based standards (e.g. >10 mg/mL nitrate-nitrogen), failure to conduct monitoring or reporting procedures, and/or failure to implement best treatment technology. At present, violations from 2001-2017 are broadly available within the SDWIS database. Standard HUC-10 watershed boundaries for the four states were obtained from USGS (https://water.usgs.gov/GIS/huc.html). When available, water systems were geo-located in by address.

Methods for Objectives 3 and 5

Study population and outcome measurement

Address-level birth records in Virginia (VA) within a 0.5 km buffer of the VA coalfields area (Fig. 3) between 1990 and 2015 (N=19,277) were geocoded using ArcGIS 10.3 (ESRI). Due to missing cloud-free satellite data for years 1991, 1996, 2006, 2009, and 2012 in the original LANDSAT analysis (see below), birth records from these years were not included. A total of 5,008 (26%) of the original birth records were successfully geocoded because many records reported only a PO box or RT box in the maternal street address field. The lowest geocoding rates were in earlier years (1990-1998) in rural areas (Table 1). Additional analyses using zipcode were also performed to remove the bias associated with the low geocoding rate.

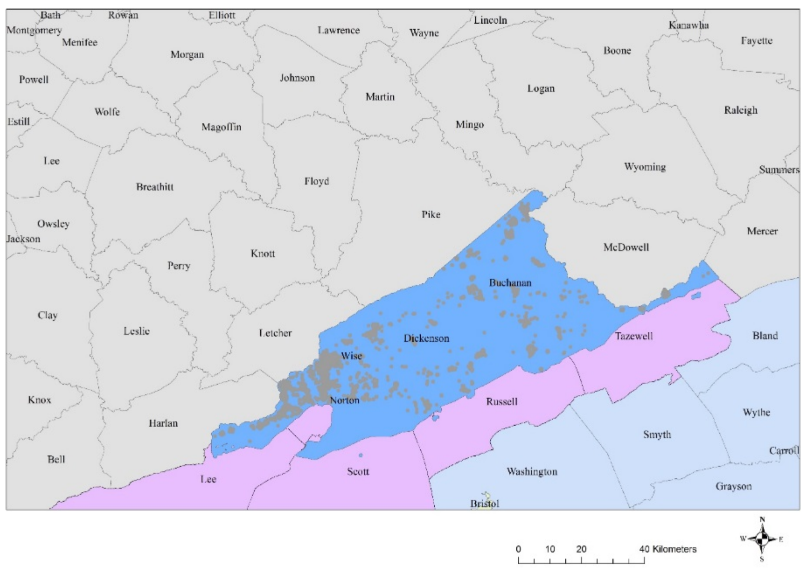

Figure 3. Mining disturbances (gray) in SW Virginia coalfield (blue) in 2015. Central Appalachia counties within Virginia are pink.

Singleton births, ≥ 22 wks gestation are included in the analysis. Dependent variables include PTB (< 37 weeks clinical estimate of gestational age), LBW (<2500 grams), and tLBW (≥ 37 weeks gestation and <2500 grams). Individual-level covariates included on the birth record include maternal age, race, ethnicity, number of children, education, and marital status. The birth records were made available for this research under a data use agreement with Virginia Department of Health (VDH) and the protocol has been approved through Virginia Tech IRB (VT IRB # 16-898).

Exposure assessment

Note that for the present analysis of Virginia birth records, the satellite image processing was completed prior to the full Central Appalachia analysis described above. Jie Ren, Dept of Geography, completed this analysis in 2015-2016 as part of a GCC-ISCE pilot project (PI Krometis). We will be replacing this initial work with the work described above in subsequent analyses of all Central Appalachia counties.

Landsat Data and preprocessing

Landsat images from 1984 to 2015 (Path 18, Row 34) were obtained from the US Geological Survey (USGS) Earth Explorer Landsat data archive (https://earthexplorer.usgs.gov/). We selected level 1T (terrain corrected) products co-registered by the Earth Resources Observation and Science (EROS) Center, then we visually verified the co-registration before our analysis. Images were processed by the EROS Center with the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) processor, which corrected atmospheric effects to provide a mask layer with a cloud mask, a cloud shadow mask, a water mask and a snow mask. In order to discriminate mined areas from unmined areas annually, we selected cloud-free (<10% of study area masked as clouds and cloud shadows) (Hilker et al., 2009, Wu et al., 2015) images for dates within the peak growing season, which we defined as extending from May 15th through October 31st. To avoid dealing with data gaps caused by a scan line corrector (SLC) problem, no post-2003 Enhanced Thematic Mapper (ETM+) images were used. No available cloud free leaf-on images for 1991, 1996, 2006, 2009, and 2012, so 27 total images were used for classification.

Mining permit data were acquired from Virginia Department of Mines, Minerals and Energy (DMME), Division of Mined Land Reclamation, in 2014. The data are polygonal shapefiles include most but not all mining permits issued by Virginia's mine regulatory agency. Mining permits contain four categories: state-issued permits (1966–1977), interim permits issued during the period when that state regulatory program was transitioning to SMCRA-mandated federal program (1977–1981), SMCRA permits issued (1981 and later) that are no longer active (permits have been released), and currently active mining permits. These polygonal shapefiles were merged together to aid training sample selection.

The National Land Cover Database (NLCD) 1992, NCLD 2001, NLCD2006, and NLCD2011 were obtained from the Multi-Resolution Land Characteristics (MRLC) Consortium (https://www.mrlc.gov/). Persistent vegetation areas were defined as the intersection of the forest, shrubland, herbaceous, and cultivated land classified by NLCD, and persistent water and developed areas were defined as the intersection of the water and developed areas classified by NLCD, respectively. These persistent areas were used for training sample selection.

The high-resolution aerial imagery was obtained from the National Agricultural Imagery Program (https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index) for the years 2003, 2004, 2005, 2008, 2009, 2011, 2012, and 2015. These images are acquired during the agricultural growing seasons to access the classification accuracy.

Mined areas classification

We applied support vector machine (SVM) classification to discriminate mined areas from unmined areas. The SVM is a non-parametric machine learning technique applicable for supervised classification (Brereton and Lloyd, 2010). It separates the classes with an optimal hyperplane using a set of mathematical functions (i.e. kernel functions). The application of different kernel functions in the SVM approach can produce different results. Instead of linear, polynomial, and sigmoid kernels, radial basis function (RBF) was used in our study. We classified the land type into four classes: mined, developed, water, and vegetation. Training samples for the four classes were selected before classification. For each mapping year, we randomly selected 500 pixels within the merged mining permits and the newly added mining areas (Li et al., 2015), and we randomly selected 100, 50, and 500 pixels outside mining permits in persistent developed, persistent water, and persistent vegetation areas, respectively. For each classified image, we reclassified the four classes into two classes: mined and unmined. Unmined class was a group of developed, water, and vegetation classes.

Statistical Analysis

Logistic regression models were developed to determine associations between residence proximity to a mining disturbance and adverse birth outcomes, including preterm birth (PTB), low birth weight (LBW), term LBW. Births were considered low weight if >=200 and <2500 grams and preterm if obstetric estimate of gestational age was >=24 and <37 weeks, following World Health Organization conventions. Covariates for logistic regression models included sex of child, method of payment for hospital services, parity (or total birth order of child), age of mother race, education, and Hispanic status. To account for unmeasured temporal factors, we included a natural cubic spline of the year of conception to account for long-term and across-year variation.

RESULTS

Objective 1

The overall accuracy of the classification procedure described above is 0.826 (Table 4). The kappa coefficient is 0.811 (Table 4). The user's accuracies ranged from 54.2% (EM 1996) to 100% (OD). The producer's accuracies ranged from 31.7% (OD) to 100.0% (multiple EM years). Many of the EM errors occurred at the edges of mined land. The misclassification of OD lands as EM lands is also apparent. After visually examining these misclassified EM lands, it is apparent that clear-cutting activities are the primary reason for this high rate of error.

Results indicate that from 1990 to 2015, 1806 km^2 of land across the study area was disturbed by mining (EM) activities, which equates to approximately 4.2% of the study-defined central Appalachia region (Figure 4). The counties with the greatest estimated area of land disturbed by mining through the study period are Pike, KY (177 km^2), Boone, WV (114 km^2), and Wise, VA (111 km^2) (Figure 5). Although the previously mentioned counties experienced the greatest magnitude of lands impacted by surface mining activities, the counties of Wise (VA), Martin (WV), and Perry (WV) experienced the highest proportion of mined land to land area within the Appalachian coalfield which was disturbed by surface mining activities from 1990 to 2015 (Figure 6). Temporally, the magnitude of land area disturbed by mining activities varied from a maximum of 668 km^2 in 1999 to a minimum of 247 km^2 in 2015 (Figure 7). It is also apparent that since 2008, the area of land disturbed by mining activities has been steadily decreasing.

Table 4. The confusion matrix, which was generated using the design-based accuracy assessment. Shown are accuracy assessment results for PV, OD, and EM lands. EM lands were classified by date of initial disturbance.

	Reference Data													
	PV	**OD**	**1990**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
PV	2	4	2	1		1	3			1				
OD		45												
1990	2	5	22	1										
1992	1	1	2	24										
1993	1	6			15	1								
1994	6	4				15								
1995	1	4				1	22							
1996	8	3						13						
1997	1	4							22					
1998	2	6							1	21				
1999	1	7									16	1		
2000	5	6										15		
2001	1	5											22	
2002	4	3												23
2003	3													1
2004	1	4												
2005		5												
2006		3												
2007	1	3												
2008		2												
2009		1												
2010	1	1												
2011		2												
2012		3												
2013		4												
2014		7												
2015		4												
Column Total	239	142	26	26	15	18	25	13	23	22	16	16	22	24
Prod. Acc.	83.70%	31.70%	84.60%	92.40%	100.00%	83.30%	88.00%	100.00%	95.70%	95.50%	100.00%	93.80%	100.00%	95.80%

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Row Total	User's Acc.
PV		1				1								214	93.50%
OD														45	100.00%
1990														3	73.30%
1992														28	85.70%
1993														23	65.20%
1994														25	60.00%
1995														28	78.60%
1996														24	54.20%
1997														27	81.50%
1998														3	70.00%
1999														25	64.00%
2000														26	57.70%
2001														28	78.60%
2002														3	76.70%
2003	26													3	86.70%
2004	1	18	1											25	72.00%
2005		1	24											3	80.00%
2006			1	25										29	86.30%
2007					25				1					3	83.30%
2008			1			27								3	90.00%
2009							29							3	96.70%
2010							2	22						26	84.60%
2011									28					3	93.30%
2012									1	26				3	86.70%
2013											24			28	85.70%
2014												23		3	76.70%
2015												4	22	3	73.30%
Column Total	27	2	27	25	25	28	31	22	3	26	24	27	22	961	
Prod. Acc.	96.30%	90.00%	88.90%	100.00%	100.00%	96.40%	93.50%	100.00%	93.30%	100.00%	100.00%	85.20%	100.00%	**Overall Acc. = 82.6%**	

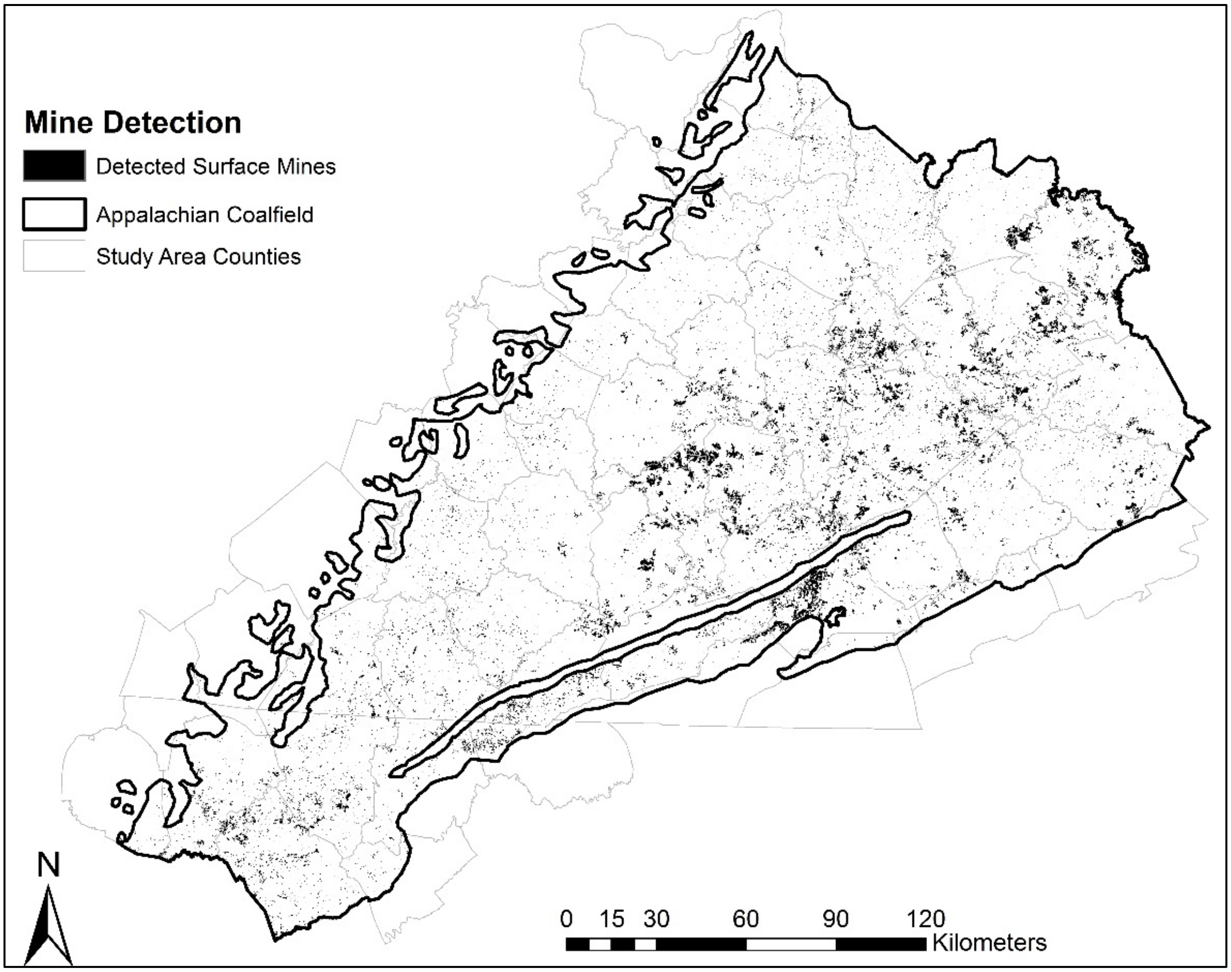

Figure 4. The aggregated (1990 through 2015; minus 1991) EM (surface mining) classified areas within the study area.

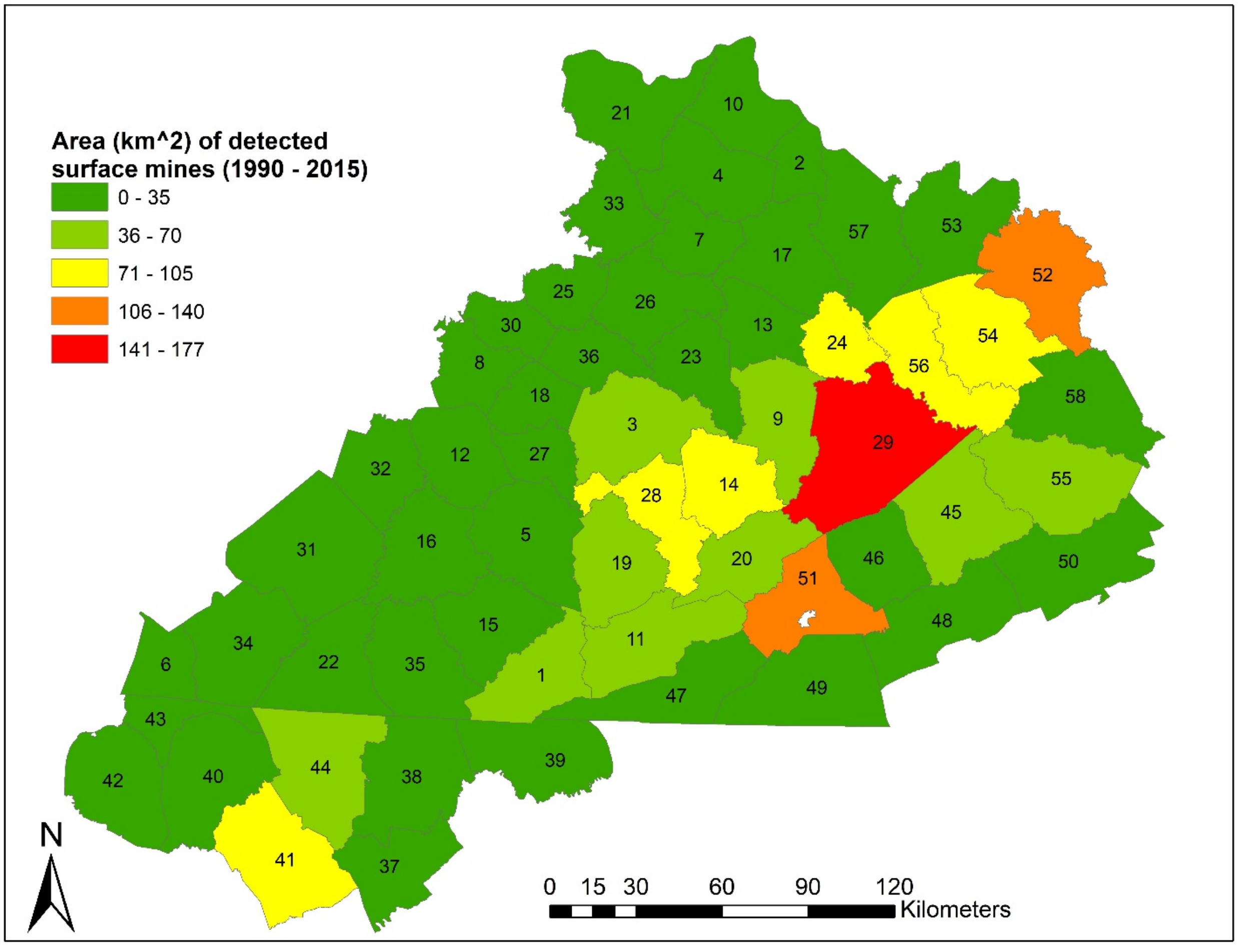

Figure 5. Estimated area (km^2) of each county that was disturbed by mining activities from 1990 to 2015.

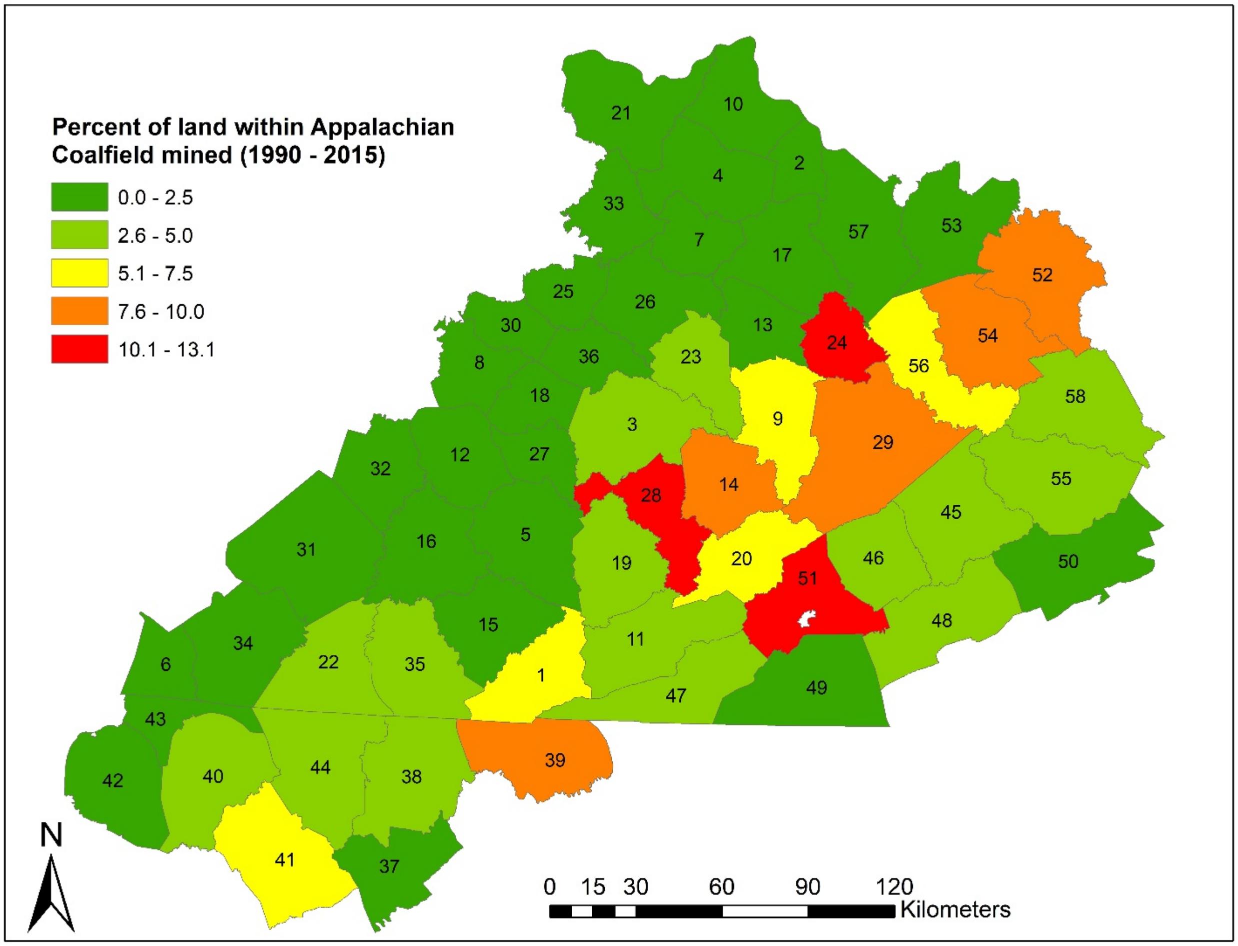

Figure 6. The percent of land classified per county as disturbed by mining activities out of the total land area within the Appalachian coalfield.

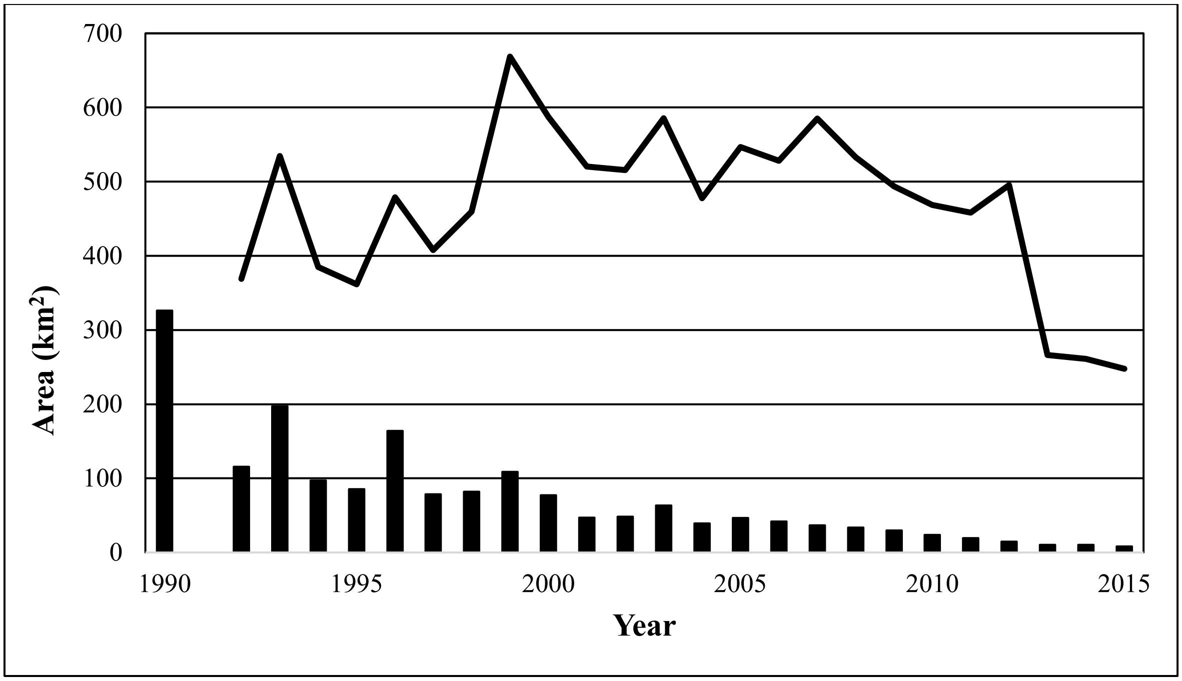

Figure 7. Bars indicate the area of new land disturbed by mining per year through the historical record. The line graph represents the total area of mining disturbed land by year through the study period.

Results from Objective 2

The SDWIS database reports 126,393 violations in the four targeted states. In keeping with previous findings, the majority of these violations occurred in very small or small drinking water systems (Table 5).

Table 5. Reported violations in SDWIS database in Central Appalachia categorized by system size.

Size	KY	TN	VA	WV
Very Small (<500 connections)	8442	3212	20575	54780
Small (501-3,300)	4959	2608	3412	10353
Medium (3,301 – 10,000)	5160	2387	765	2859
Large (10,001-100,000)	4023	1480	396	614
Very Large (>100,000)	228	31	109	0
Total	**22812**	**9718**	**25257**	**68606**

The overwhelming majority of these violations were within the monitoring and reporting category (100,981 violations; 80% of total violations) and were associated with smaller systems (Table 6).

Table 6. Reported monitoring and reporting violations in SDWIS database in Central Appalachia categorized by system size.

Size	KY	TN	VA	WV
Very Small (<500 connections)	7811	2822	15388	47276
Small (501-3,300)	3376	1943	2411	7423
Medium (3,301 – 10,000)	3512	1544	436	2231
Large (10,001-100,000)	2908	926	217	427
Very Large (>100,000)	218	18	94	0
Total	**17825**	**7253**	**18546**	**57357**

There were over 12,000 reported violations of health-based standards over this time period, with coliform bacteria and disinfection byproducts identified as the primary contaminants of concern (Table 7).

Over a third of the systems reporting violations could not be geo-coded past the county level.

Table 7. Reported health-based standards in SDWIS database in Central Appalachia categorized by system size.

Size	KY	TN	VA	WV
Very Small (<500 connections)	228	295	4078	641
Small (501-3,300)	835	499	912	840
Medium (3,301 – 10,000)	1012	703	301	241
Large (10,001-100,000)	686	456	172	89
Very Large (>100,000)	9	6	15	0
Total	**2770**	**1959**	**5478**	**1811**

Results from Objectives 3 and 5

The rate of preterm birth in Virginia coalfield areas has increased between 1990-2015 (Fig. 8a). The overall rate of preterm birth climbed in the United States until approximately 2006, then began declining. The initial climb across the United States has previously been attributed to increased cesarean delivery during that time period as well as increased maternal age. To adjust for this global temporal trend in our analysis, a spline was fit to the data (Fig 8b) and used in subsequent logistic regression analyses.

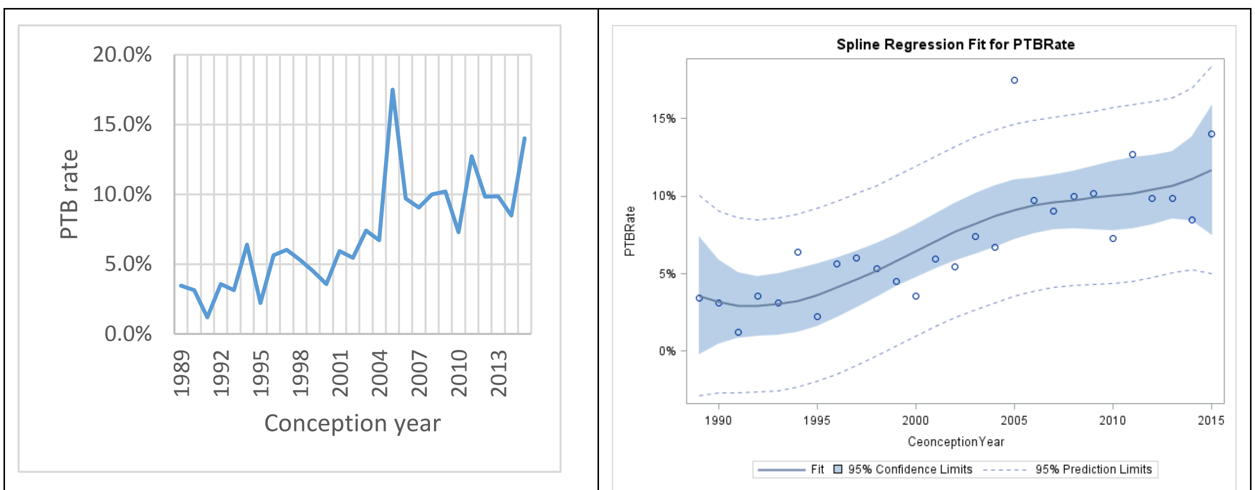

Figure 8. A. PTB rates coalfield area in the conception year from 1989 to 2015 A natural cubic spline regression fit for PTB rates in VA coalfield area in the conception year from 1989 to 2015

The maternal and infant characteristics of the original dataset including all births in coalfield areas between 1990-2015 (n=14,269) are compared to the geocoded dataset (n-=5,009) in Table 8. It is important to not only note the loss of sample due to poor geocoding rates, but also note the significant differences in sociodemographic characteristics between the geocoded and ungeocoded dataset for important covariates including payment method (proxy for socioeconomic status), maternal race, and maternal education. Based on these differences found, we added an additional analysis using the full birth record dataset and an estimate of mining extents within zip code. Also, maternal race and ethnicity were excluded from analyses of the geocoded dataset due to small sample size.

Table 8. Sample characteristics of original dataset and geocoded dataset within a 1 km buffer of the southwest Virginia coalfield area between 1990 to 2015

	Ungeocoded dataset N (%)	**Geocoded dataset[1] N (%)**	***P* value (Chi-square test)**
Gestation period *(wks)*			0.868
24 - 36	1011 (7.1)	358 (7.2)	
> 37	13232 (92.9)	4636 (92.8)	
Birth weight *(g)*			0.736
200 – 2499	926 (6.5)	318 (6.4)	
> 2500	13337 (93.5)	4685 (93.6)	
Gender of child			0.805
Male	7268 (50.9)	2561 (51.1)	
Female	7001 (49.1)	2447 (48.9)	
Maternal race*, Black*			**<0.001**
No	14203 (99.6)*	4916 (98.3)	
Yes	56 (0.4)	83 (1.7)*	
Maternal Hispanic			0.449
No	14218 (99.7)	4987 (99.6)	
Yes	44 (0.3)	19 (0.4)	
Payment method			**0.012**
Private insurance	3688 (30.6)	1568 (32.9)*	
Medicaid	8160 (67.7)*	3125 (65.6)	
Self-pay & Other	207 (1.7)	73 (1.5)	
Maternal age *(years)*			0.091
18-35	12868 (90.2)	4536 (90.6)	
<18	827 (5.8)	253 (5.1)	
>35	572 (4.0)	218 (4.4)	
Parity			**0.003**
1	6491 (45.5)*	2185 (43.7)	
2	4927 (34.6)	1697 (33.9)	
3	1965 (13.8)	774 (15.5)*	
4 or more	872 (6.1)	345 (6.9)*	
Maternal education *(years)*			**<0.001**
< 12	4833 (33.9)	1886 (37.8)*	
12	5905 (41.5)*	1766 (35.4)	
> 12	3505 (24.6)	1343 (26.9)*	
Total	14269 (100)	5008 (100)	

Note: 1: Dataset includes birth records in the coalfield areas of southwest VA and a 1 km buffer (Fig. 3); 2: There are 2214 and 242 “Unknown” records for payment methods for original dataset and geocoded dataset; * stands for the category in the group is significantly higher or lower than the one in the other group

Tables 9 and 10 present multivariate logistic regression results using a continuous (the % area active mining within a 5, 2, 1, 0.75, or 0.5 km radius of the maternal address) or a categorical variable (tertiles of % area active mining within 5, 2, 1, 0.75, and 0.5 km radious of maternal address), respectively. No association was found between maternal address proximity to surface mining and preterm birth in either model. When comparing low versus high tertile of % mining within a specified radius of maternal address, non-significant positive associations are estimated for 4 of the 5 boundaries evaluated (Figure 9); however large error bars preclude drawing conclusions from these results.

Table 9. Odds Ratios and lower and upper 95%CL from logistic regression models for preterm birth in southwest Virginia coalfields using percent of mining within a 5km, 2km, 1km, 0.75km, or 0.5km radius of the maternal address .

	5km	2km	1km	0.75km	0.5km
Gender (Male vs Female)	1.14 (0.95, 1.38)	1.27 (1.02, 1.58)	1.28 (1.03, 1.59)	1.29 (1.03, 1.6)	1.30 (1.04, 1.62)
Payment (Private insurance as baseline)					
Medicaid	1.15 (0.92, 1.44)	1.19 (0.92, 1.53)	1.21 (0.94, 1.57)	1.24 (0.96, 1.61)	1.22 (0.94, 1.59)
Other & Self-pay	1.33 (0.65, 2.7)	0.83 (0.29, 2.32)	0.87 (0.31, 2.44)	0.88 (0.31, 2.48)	0.86 (0.31, 2.43)
Mother age (18-35)					
<18	1.05 (0.68, 1.63)	1.12 (0.68, 1.84)	1.16 (0.71, 1.91)	1.17 (0.71, 1.92)	1.15 (0.70, 1.90)
>35	1.13 (0.72, 1.77)	1.18 (0.71, 1.95)	1.25 (0.76, 2.07)	1.28 (0.77, 2.12)	1.29 (0.78, 2.13)
Mother education					
<12	1.08 (0.84, 1.4)	1.01 (0.76, 1.35)	0.98 (0.73, 1.31)	0.98 (0.73, 1.31)	1.00 (0.75, 1.34)
>12	0.74 (0.57, 0.98)	0.75 (0.55, 1.03)	0.75 (0.55, 1.04)	0.76 (0.55, 1.04)	0.74 (0.53, 1.02)
Birth order	1.01 (0.99, 1.03)	1.02 (1, 1.04)	1.02 (1, 1.04)	1.02 (1, 1.04)	1.02 (1.00, 1.04)
% of mining area among the buffer with a specific km of each birth	1.01 (0.97, 1.05)	1 (0.96, 1.03)	0.99 (0.95, 1.02)	0.99 (0.95, 1.03)	0.99 (0.95, 1.04)

Table 10. Odds Ratios and lower and upper 95%CL from logistic regression models for preterm birth in southwest Virginia coalfields using tertiles of % mining within a 5km, 2km, 1km, 0.75km, or 0.5km radius of the maternal address .

	5km	2km	1km	0.75km	0.5km
Gender (Male vs Female)	1.14 (0.95, 1.38)	1.27 (1.03, 1.58)	1.28 (1.03, 1.59)	1.29 (1.03, 1.6)	1.3 (1.04, 1.62)
Payment (Private insurance as baseline)					
Medicaid	1.15 (0.92, 1.44)	1.18 (0.91, 1.52)	1.21 (0.93, 1.56)	1.23 (0.95, 1.59)	1.22 (0.94, 1.58)
Other & Self-pay	1.32 (0.65, 2.7)	0.81 (0.29, 2.27)	0.86 (0.3, 2.41)	0.86 (0.31, 2.43)	0.86 (0.3, 2.41)
Mother age (18-35)					
<18	1.04 (0.67, 1.61)	1.11 (0.68, 1.82)	1.16 (0.71, 1.91)	1.15 (0.7, 1.9)	1.15 (0.7, 1.88)
>35	1.13 (0.72, 1.77)	1.19 (0.72, 1.96)	1.26 (0.76, 2.09)	1.29 (0.78, 2.13)	1.29 (0.78, 2.13)
Mother education					
<12	1.1 (0.86, 1.42)	1.04 (0.78, 1.39)	0.98 (0.74, 1.32)	1.02 (0.76, 1.37)	1.02 (0.76, 1.36)
>12	0.75 (0.57, 0.98)	0.76 (0.55, 1.04)	0.75 (0.55, 1.04)	0.76 (0.55, 1.04)	0.74 (0.54, 1.03)
Birth order	1.01 (0.99, 1.03)	1.02 (1, 1.04)	1.02 (1, 1.04)	1.01 (0.99, 1.04)	1.02 (0.99, 1.04)
% of mining area among the buffer with a specific km of each birth (lowest tertile as baseline)					

Middle	1.21 (0.95, 1.55)	1.32 (0.99, 1.76)	1.08 (0.82, 1.43)	1.07 (0.81, 1.43)	1.17 (0.85, 1.62)
High	1.14 (0.91, 1.44)	1.26 (0.97, 1.64)	1.07 (0.82, 1.39)	0.82 (0.59, 1.15)	1.33 (0.89, 2)

An additional analysis was conducted that used zipcode of maternal address, which was available for the full dataset (n=27179). The percentage of mining area within the maternal address zipcode was used as the exposure metric. The number of birth records in a Zipcode with 0% mining area is 4945 among a total of birth records in SWVA coalfield from 1990 to 2015. As shown in Table 11, there was no significant association found between mining in a zipcode and preterm birth.

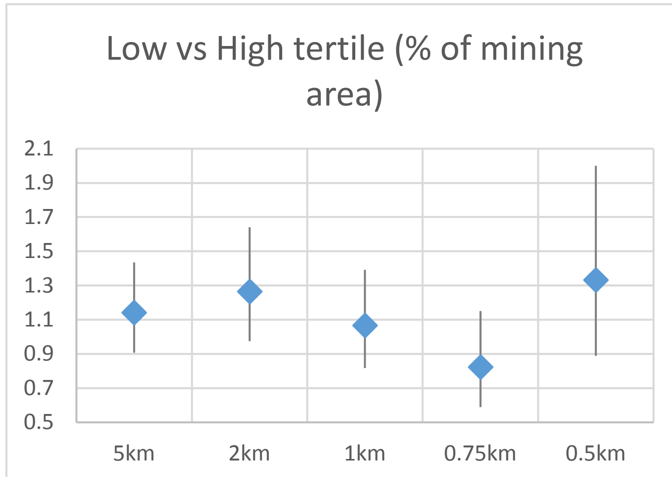

Figure 9. Odds ratios lower and upper 95%CI of low versus high tertile of percentage of mining areas within a 5km, 2km, 1km, 0.75km, or 0.5km radius of the maternal address .

Table 11. Odds Ratios and lower and upper 95%CL from logistic regression models for preterm birth in southwest Virginia coalfields using % mining or mining/no mining within the zipcode of the maternal address .

	Model 1	Model 2
Gender (Male vs Female)	**1.11 (1.02, 1.22)**	**1.11 (1.02, 1.22)**
Payment (Private insurance)		
Medicaid	1.04 (0.93, 1.17)	1.04 (0.93, 1.17)
Other & Self-pay	0.86 (0.71, 1.05)	0.86 (0.71, 1.04)
Mother age (18-35)		
<18	**1.29 (1.05, 1.58)**	**1.29 (1.05, 1.58)**
>35	1.21 (0.97, 1.51)	1.20 (0.96, 1.50)
Mother education		
<12	**1.15 (1.03, 1.30)**	**1.15 (1.03, 1.30)**
>12	0.92 (0.81, 1.04)	0.92 (0.81, 1.05)
Birth order	**1.09 (1.04, 1.15)**	**1.10 (1.04, 1.15)**
% of mining area in the ZIP code area	1.00 (0.98, 1.02)	
If the birth record was in a ZIP code with the mining area (Yes vs No)		1.05 (0.93, 1.19)

DISCUSSION

Central Appalachia has undergone rapid changes in resource extraction, with coal production from surface mining declining since 2008 (40), yet surface mining is occurring elsewhere in the world, and air and water contaminants of concern from surface mining are associated with other resource extraction industries (41-44).

The primary limitation of the present analysis is low sample size. For example, previous studies conducted primarily in urban areas have shown $PM_{2.5}$ exposure is associated with LBW, term LBW, and PTB (45, 46) and compositional analysis suggests elemental carbon, Ni, Al and Ti may be most predictive of LBW (46); however the effect size was quite small (< 2) and therefore sample size must be quite large to detect. In future analyses, we will be addressing this issue through a broader analysis incorporating birth records from Tennessee, West Virginia, and Kentucky.

Potential residual confounding is also a potential limitation here and in most previous studies that have relied on time-series cross-sectional and time-stratified case crossover designs. A few studies have attempted to address this problem. A retrospective longitudinal study matched births to the same mother and found a small but significant association between $PM_{2.5}$ exposure and PTB in Connecticut (47). Another analysis of this cohort was not able to determine associations with specific sources of $PM_{2.5}$ (48), possibly due to the relatively small variation in $PM_{2.5}$ source exposures over the 7 year period. In fact a recent study suggests spatial disaggregation of exposure attenuates associations between $PM_{2.5}$ and term LBW (49). Previous studies have used natural or quasi-experimental designs to take advantage of changes in air pollution pre-, during and post- events such as the Beijing Olympics (50, 51), wildfires (52), and a steel mill closure in Utah (53) and offline of coal-fired power plants in California (54). These studies have been instrumental in providing further evidence of risk because residual confounding is less of a concern.

Moving forward our research will: 1) incorporate air and watershed data, 2) include birth records from Tennessee and Kentucky Central Appalachia counties and 3) implement a multi-level statistical framework to further refine exposure metrics, increase power, and better account for potential residual confounding, respectively.

BIBLIOGRAPHY

1. Kent ST, McClure LA, Zaitchik BF, Smith TT, Gohlke JM. Heat waves and health outcomes in Alabama (USA): the importance of heat wave definition. Environ Health Perspect. 2014;122(2):151-8. doi: 10.1289/ehp.1307262. PubMed PMID: 24273236; PMCID: PMC3914868.

2. Porter TR, Kent ST, Su W, Beck HM, Gohlke JM. Spatiotemporal association between birth outcomes and coke production and steel making facilities in Alabama, USA: a cross-sectional study. Environ Health. 2014;13:85. doi: 10.1186/1476-069X-13-85. PubMed PMID: 25342170; PMCID: PMC4223752.

3. Bailey BA, Cole LK. Rurality and birth outcomes: findings from southern appalachia and the potential role of pregnancy smoking. J Rural Health. 2009;25(2):141-9. doi: 10.1111/j.1748-0361.2009.00210.x. PubMed PMID: 19785579.

4. Strutz KL, Dozier AM, van Wijngaarden E, Glantz JC. Birth outcomes across three rural-urban typologies in the Finger Lakes region of New York. J Rural Health. 2012;28(2):162-73. doi: 10.1111/j.1748-0361.2011.00392.x. PubMed PMID: 22458317; PMCID: PMC3337719.

5. Luo ZC, Kierans WJ, Wilkins R, Liston RM, Mohamed J, Kramer MS, British Columbia Vital Statistics A. Disparities in birth outcomes by neighborhood income: temporal trends in rural and urban areas, british columbia. Epidemiology. 2004;15(6):679-86. PubMed PMID: 15475716.

6. Hillemeier MM, Weisman CS, Chase GA, Dyer AM. Individual and community predictors of preterm birth and low birthweight along the rural-urban continuum in central Pennsylvania. J Rural Health. 2007;23(1):42-8. doi: 10.1111/j.1748-0361.2006.00066.x. PubMed PMID: 17300477.

7. Yao N, Matthews SA, Hillemeier MM. White infant mortality in Appalachian states, 1976–1980 and 1996–2000: changing patterns and persistent disparities. The Journal of Rural Health. 2012;28(2):174-82.

8. Markley S, Tu W. Regional and Racial Disparity of Preterm Birth Prevalence in Georgia, 1995–2012. Papers in Applied Geography. 2015;1(2):168-75.

9. Ferrero DM, Larson J, Jacobsson B, Di Renzo GC, Norman JE, Martin Jr JN, D'Alton M, Castelazo E, Howson CP, Sengpiel V. Cross-Country Individual Participant Analysis of 4.1 Million Singleton Births in 5 Countries with Very High Human Development Index Confirms Known Associations but Provides No Biologic Explanation for 2/3 of All Preterm Births. PloS one. 2016;11(9):e0162506.

10. Kent ST, McClure LA, Zaitchik BF, Gohlke JM. Area-level risk factors for adverse birth outcomes: trends in urban and rural settings. BMC Pregnancy Childbirth. 2013;13(1):129. Epub 2013/06/14. doi: 1471-2393-13-129 [pii]

10.1186/1471-2393-13-129. PubMed PMID: 23759062; PMCID: 3688345.

11. Borak J, Salipante-Zaidel C, Slade MD, Fields CA. Mortality disparities in Appalachia: reassessment of major risk factors. Journal of occupational and environmental medicine /

American College of Occupational and Environmental Medicine. 2012;54(2):146-56. doi: 10.1097/JOM.0b013e318246f395. PubMed PMID: 22258162.

12. Halverson JA, Barnett E, Casper M. Geographic disparities in heart disease and stroke mortality among black and white populations in the Appalachian region. Ethnicity & disease. 2002;12(4):S3-82-91. PubMed PMID: 12477161.

13. Barnett E, Halverson JA, Elmes GA, Braham VE. Metropolitan and non-metropolitan trends in coronary heart disease mortality within Appalachia, 1980-1997. Annals of epidemiology. 2000;10(6):370-9. PubMed PMID: 10964003.

14. Halverson JA. An Analysis of Disparities in Health Status and Access to Health Care in the Appalachian Region. In: Commission AR, editor. Washington, DC: http://www.arc.gov/research/researchreportdetails.asp?REPORT_ID=82.; 2004.

15. Behringer B, Friedell GH. Appalachia: where place matters in health. Preventing chronic disease. 2006;3(4):A113. PubMed PMID: 16978488; PMCID: 1779277.

16. Blackley D, Behringer B, Zheng S. Cancer mortality rates in Appalachia: descriptive epidemiology and an approach to explaining differences in outcomes. Journal of community health. 2012;37(4):804-13. doi: 10.1007/s10900-011-9514-z. PubMed PMID: 22101638.

17. Phillippi JC, Myers CR, Schorn MN. Facilitators of prenatal care access in rural Appalachia. Women and Birth. 2014;27(4):e28-e35.

18. Halverson JA, Bischak G. Underlying Socioeconomic Factors Influencing Health Disparities in the Appalachian Region. In: Commission AR, editor. Washington, D.C.2008.

19. Krometis LA, Gohlke J, Kolivras K, Satterwhite E, Marmagas SW, Marr LC. Environmental health disparities in the Central Appalachian region of the United States. Rev Environ Health. 2017;32(3):253-66. doi: 10.1515/reveh-2017-0012. PubMed PMID: 28682789.

20. Boyles AL, Blain RB, Rochester JR, Avanasi R, Goldhaber SB, McComb S, Holmgren SD, Masten SA, Thayer KA. Systematic review of community health impacts of mountaintop removal mining. Environ Int. 2017;107:163-72. doi: 10.1016/j.envint.2017.07.002. PubMed PMID: 28738262; PMCID: PMC5562233.

21. Townsend PA, Helmers DP, Kingdon CC, McNeil BE, de Beurs KM, Eshleman KN. Changes in the extent of surface mining and reclamation in the Central Appalachians detected using a 1976-2006 Landsat time series. Remote Sens Environ. 2009;113(1):62-72. PubMed PMID: WOS:000261993100006.

22. Zipper CE, Burger JA, Skousen JG, Angel PN, Barton CD, Davis V, Franklin JA. Restoring Forests and Associated Ecosystem Services on Appalachian Coal Surface Mines. Environ Manage. 2011;47(5):751-65. PubMed PMID: WOS:000290277200005.

23. Ferrari JR, Lookingbill TR, McCormick B, Townsend PA, Eshleman KN. Surface mining and reclamation effects on flood response of watersheds in the central Appalachian Plateau region. Water Resour Res. 2009;45. PubMed PMID: WOS:000265106200001.

24. Palmer MA, Bernhardt ES, Schlesinger WH, Eshleman KN, Foufoula-Georgiou E, Hendryx MS, Lemly AD, Likens GE, Loucks OL, Power ME, White PS, Wilcock PR. Mountaintop Mining Consequences. Science. 2010;327(5962):148-9. PubMed PMID: WOS:000273394000018.

25. Kurth LM, McCawley M, Hendryx M, Lusk S. Atmospheric particulate matter size distribution and concentration in West Virginia coal mining and non-mining areas. Journal of exposure science & environmental epidemiology. 2014;24(4):405-11. doi: 10.1038/jes.2014.2. PubMed PMID: 24549227.

26. Lindberg TT, Bernhardt ES, Bier R, Helton AM, Merola RB, Vengosh A, Di Giulio RT. Cumulative impacts of mountaintop mining on an Appalachian watershed. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(52):20929-34. doi: 10.1073/pnas.1112381108. PubMed PMID: 22160676; PMCID: 3248525.

27. Hendryx M. Mortality from heart, respiratory, and kidney disease in coal mining areas of Appalachia. International archives of occupational and environmental health. 2009;82(2):243-9. doi: 10.1007/s00420-008-0328-y. PubMed PMID: 18461350.

28. Hendryx M, O'Donnell K, Horn K. Lung cancer mortality is elevated in coal-mining areas of Appalachia. Lung cancer. 2008;62(1):1-7. doi: 10.1016/j.lungcan.2008.02.004. PubMed PMID: 18353487.

29. Hendryx M, Zullig KJ. Higher coronary heart disease and heart attack morbidity in Appalachian coal mining regions. Preventive medicine. 2009;49(5):355-9. doi: 10.1016/j.ypmed.2009.09.011. PubMed PMID: 19761789.

30. Esch L, Hendryx M. Chronic cardiovascular disease mortality in mountaintop mining areas of central Appalachian states. The Journal of rural health : official journal of the American Rural Health Association and the National Rural Health Care Association. 2011;27(4):350-7. doi: 10.1111/j.1748-0361.2011.00361.x. PubMed PMID: 21967378.

31. Hendryx M, Ahern M. Reply to Borak et Al "Mortality disparities in Appalachia: reassessment of major risk factors". Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine. 2012;54(7):768-9; author reply 70-3. doi: 10.1097/JOM.0b013e318254622f. PubMed PMID: 22796918.

32. Hendryx M, Ahern MM. Mortality in Appalachian coal mining regions: the value of statistical life lost. Public health reports. 2009;124(4):541-50. PubMed PMID: 19618791; PMCID: 2693168.

33. Li J, Zipper CE, Donovan PF, Wynne RH, Oliphant AJ. Reconstructing disturbance history for an intensively mined region by time-series analysis of Landsat imagery. Environmental monitoring and assessment. 2015;187(9):557.

34. Jin S, Yang L, Danielson P, Homer C, Fry J, Xian G. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment. 2013;132:159-75.

35. Masek JG, Vermote EF, Saleous NE, Wolfe R, Hall FG, Huemmrich KF, Gao F, Kutler J, Lim T-K. A Landsat surface reflectance dataset for North America, 1990-2000. IEEE Geoscience and Remote Sensing Letters. 2006;3(1):68-72.

36. Rouse Jr JW, Haas R, Schell J, Deering D. Monitoring vegetation systems in the Great Plains with ERTS1974.

37. Key CH, Benson NC, editors. Measuring and remote sensing of burn severity. Proceedings joint fire science conference and workshop; 1999: University of Idaho and International Association of Wildland Fire Moscow, ID.

38. Hardisky M, Klemas V, Smart M. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of. Spartina alterniflora. 1983;49:77-83.

39. Kauth RJ, Thomas G, editors. The tasselled cap--a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. LARS Symposia; 1976.

40. Annual Coal Report Archive [Internet]. U.S. Energy Information Administration 2014.

41. Sarver EA, Cox AS. Salt: an emerging water concern for the global mning and minerals industries. Mining Technology. 2013;122(3):145-52.

42. Canedo-Arguelles M, Hawkins CP, Kefford BJ, Schafer RB, Dyack BJ, Brucet S, Buchwalter D, Dunlop J, Fror O, Lazorchak J, Coring E, Fernandez HR, Goodfellow W, Achem AL, Hatfield-Dodds S, Karimov BK, Mensah P, Olson JR, Piscart C, Prat N, Ponsa S, Schulz CJ, Timpano AJ. WATER. Saving freshwater from salts. Science. 2016;351(6276):914-6. doi: 10.1126/science.aad3488. PubMed PMID: 26917752.

43. Litovitz A, Curtright A, Abramzon S, Burger N, Samaras C. Estimation of regional air-quality damages from Marcellus Shale natural gas extraction in Pennsylvania. Environ Res Lett. 2013;8(1). doi: Artn 014017

10.1088/1748-9326/8/1/014017. PubMed PMID: WOS:000316998300025.

44. Adgate JL, Goldstein BD, McKenzie LM. Potential Public Health Hazards, Exposures and Health Effects from Unconventional Natural Gas Development. Environmental Science & Technology. 2014;48(15):8307-20. doi: 10.1021/es404621d. PubMed PMID: WOS:000340080600004.

45. Stieb DM, Chen L, Eshoul M, Judek S. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. Environmental research. 2012;117:100-11. doi: 10.1016/j.envres.2012.05.007. PubMed PMID: 22726801.

46. Ebisu K, Bell ML. Airborne PM2.5 chemical components and low birth weight in the northeastern and mid-Atlantic regions of the United States. Environ Health Perspect. 2012;120(12):1746-52. doi: 10.1289/ehp.1104763. PubMed PMID: 23008268; PMCID: 3548298.

47. Pereira G, Belanger K, Ebisu K, Bell ML. Fine particulate matter and risk of preterm birth in Connecticut in 2000-2006: a longitudinal study. American journal of epidemiology. 2014;179(1):67-74. Epub 2013/09/27. doi: 10.1093/aje/kwt216

kwt216 [pii]. PubMed PMID: 24068199; PMCID: 3864709.

48. Pereira G, Bell ML, Lee HJ, Koutrakis P, Belanger K. Sources of fine particulate matter and risk of preterm birth in connecticut, 2000-2006: a longitudinal study. Environmental health perspectives. 2014;122(10):1117-22. doi: 10.1289/ehp.1307741. PubMed PMID: 24911470; PMCID: 4181926.

49. Harris G, Thompson WD, Fitzgerald E, Wartenberg D. The association of PM2.5 with full term low birth weight at different spatial scales. Environmental research. 2014;134:427-34. doi: 10.1016/j.envres.2014.05.034. PubMed PMID: 25261950.

50. Rich DQ, Kipen HM, Huang W, Wang G, Wang Y, Zhu P, Ohman-Strickland P, Hu M, Philipp C, Diehl SR, Lu SE, Tong J, Gong J, Thomas D, Zhu T, Zhang JJ. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adults. Jama. 2012;307(19):2068-78. doi: 10.1001/jama.2012.3488. PubMed PMID: 22665106; PMCID: 4049319.

51. Yang Y, Li R, Li W, Wang M, Cao Y, Wu Z, Xu Q. The association between ambient air pollution and daily mortality in Beijing after the 2008 olympics: a time series study. PloS one. 2013;8(10):e76759. doi: 10.1371/journal.pone.0076759. PubMed PMID: 24204670; PMCID: 3800078.

52. Liu JC, Pereira G, Uhl SA, Bravo MA, Bell ML. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. Environmental research. 2015;136C:120-32. doi: 10.1016/j.envres.2014.10.015. PubMed PMID: 25460628; PMCID: 4262561.

53. Parker JD, Mendola P, Woodruff TJ. Preterm birth after the Utah Valley Steel Mill closure: a natural experiment. Epidemiology. 2008;19(6):820-3. doi: 10.1097/EDE.0b013e3181883d5d. PubMed PMID: 18854706.

54. Casey JA, Karasek D, Ogburn EL, Goin DE, Dang K, Braveman PA, Morello-Frosch R. Coal and oil power plant retirements in California associated with reduced preterm birth among populations nearby. Am J Epidemiol. 2018. doi: 10.1093/aje/kwy110. PubMed PMID: 29796613; PMCID: PMC6070091.