SECTION 2: Methods

2.A – Overview

This section documents methods of data collection and processing used in analyses included in the 5th Act 54 report. Much of the data are stored in a geographic information system (GIS). In addition to this documentation, this chapter outlines the challenges and limitations encountered in assembling the necessary information and makes recommendations for improving this process.

At this point it is also important to note that the spatial data depicted in all maps in this report are limited in their use. In particular, according to PADEP, "The georeferenced map layers on these products were provided by the Pennsylvania Department of Environmental Protection. These maps were georeferenced using information considered to be the best historic data available. It is understood that there is an inherent loss of accuracy in the georeferencing process and the georeferenced maps may not align correctly with base maps, real world locations, and/or established coordinate systems. The Department assumes no responsibility for the accuracy or completeness of this information."

2.B – Introduction

The impacts of mining are distributed across broad spatial areas. While impacts are generally limited to an area immediately above and surrounding the undermined area, mines stretch across most of southwestern Pennsylvania. This results in a complicated data management environment that requires the organization of precise geographic information. The geographically organized data are then compared with varied data sources ranging from BUMIS to stream recovery evaluation reports to wetland surveys submitted as part of permit renewals. These data challenges are not unique to the University, as PADEP personnel must evaluate impacts from coal extraction as they occur.

In the previous assessment (Tonsor et al. 2014), substantial effort was devoted to tying digital versions of paper records, particularly six-month mining maps to geographic coordinates (rectification). In this assessment, the PADEP shared rectified mine maps with the University, simplifying the cataloging of relevant geographic data. This was a successful change from previous assessment data workflows, allowing focus on the digitization of other data sources that are less standardized (those efforts are described throughout the report as they are analyzed). The University maintains that as mine operators store and manage these data in digital formats that PADEP management require submission of electronic mining maps (see Tonsor et al. 2014). The amount of effort necessary to digitize and cross-check mine maps is substantial and strains limited resources.

In general, the organization of spatial data proceeded in three phases:

1) At the outset of the assessment effort mines that had ceased operation during the assessment period (August 2013-August 2018) were gathered.
2) These initial data were used as cases to develop standard procedures for the processing and assessment of the necessary data sets.

3) After August 2018, information on mining extent and overlying features was collected and processed for all remaining mines, i.e. that continued operations throughout the reporting period, using the procedures developed in the second phase.

This is the second assessment to be scheduled so that the data collection period (November 2017 through December 2018) and the assessment period (August 21, 2013 through August 20, 2018) overlapped. This approach has the advantage of allowing a timelier report (submitted in August 2019). However, the overlap in timing continues to present challenges in data quality control. The difficulties in reconciling the PADEP's mine map reporting schedule, the delays associated with PADEP approval of the mine maps submitted by the mine operators, and the assessment period were noted in the last assessment (Tonsor et al. 2014). The PADEP staggers the scheduled mine map reporting dates for the mine owners over the course of the year to distribute the effort necessary for the PADEP to process these maps. With a six-month reporting cycle, maps associated with mining completed in August may not be reported to PADEP until February. In addition, the California office may not receive the final maps for months after the submission deadline due to certifications necessary at other offices, for example the U.S. Mine Safety & Health Administration offices.

The University anticipated this challenge and dedicated significant effort in attempts to ensure that the end extents of mining for the 5th assessment period were accurate. Using the best available information on the extents of mining in September 2018, the University processed and analyzed the data provided, per the University's agreement with PADEP. However in February of 2019 the University requested maps of the weekly face positions from the PADEP to explore other questions and discovered that there were minor discrepancies between the extents of mining based on data reported to the PADEP and the University in September 2018 and the mapped face positions provided by the PADEP in March 2019. At this point, the data processing and analysis was complete, so reconciling these sources of information on mining extents at the end of the reporting period was not feasible. This does not alter impact analysis in this assessment (see the discussion below), but it may alter the total counts of surface features. This outcome is consistent with previous assessments in which final positions were ambiguous. For longwall mines, the weekly face position maps provide the most accurate determination of the limits of mining. To improve future assessments, the University recommends that the PADEP provide weekly face position maps to the University to supplement the six-month mine maps, at least for the longwall operations.

2.C – Bituminous Underground Mining Information System (BUMIS)

The Bituminous Underground Mining Information System (BUMIS) is maintained by PADEP to track surface impacts related to underground mining activities. In the previous assessment, there were specific recommendations for improvements in BUMIS. Important recommendations included (Tonsor et al. 2014):

> X.B.2.a.2i: ALL information that can be georeferenced and is pertinent to permitting, regulation, and reporting should be included in BUMIS to create a true information system where all relevant information can be accessed.

X.B.3.a.i: Structures and surface features should be labelled with a unique numerical identifier whereby that feature can be identified solely by its numerical identification for a given mine.

In both cases improvements were evident in the BUMIS. While there remained mismatches between the BUMIS identifiers and the mine map identifiers, the rates of these mismatches were lower for this 5th assessment period than that noted in the 4th assessment period. A substantial section on streams, stream impacts, and stream mitigations was added to BUMIS. These improvements are important advances in the centralized management of mining data.

BUMIS is the "gold standard" data source as the University evaluates impacts from underground mining. That is, if a report of an impact occurs, the University assumes it is included in BUMIS. This reliance on a single source is fundamental to confidence that the small discrepancies between the reported and actual mining extents will not radically change conclusions. The impact numbers were derived from BUMIS and BUMIS was queried based on a time criterion (i.e., please list all impacts that occurred before 8/20/2018). Therefore, the University did not under- or over-count impacts in the assessment period due to mapping discrepancies. Any differences would be in the counting of unimpacted features.

To the extent that BUMIS is not accurate, or consistently omits impacts, analyses are inaccurate. In the scope of the assessment, the focus is assuring the quality of reported impacts data. Assessing the accuracy of the BUMIS reporting record (i.e., ensuring reporting is 100 % accurate) is beyond the scope of the assessment. That assessment would entail canvassing of hundreds of property owners, walking miles of undermined streams, etc., tasks that require access to private property and presence throughout the assessment period (e.g. if a stream is impacted and mitigated in year one of the period, the University is categorically unable to evaluate that work in year five.) This type of quality assurance is beyond the scope of work and not consistent with what the Act 54 legislation specifies for assessment.

That said, the potential for under-reporting is diminished by PADEP policy, which requires *all* impacts to be reported regardless of agreement or ownership. If mining companies fail to report any impacts within the time frame allowed by the statute and the PADEP later learns about the impacts, the mining companies are cited with a violation. The occurrence of these violations is rare (per comm. William Keefer PADEP, Table 2-1), suggesting the potential under reporting of subsidence impacts is a minor concern during the 5th assessment period and the use of BUMIS as the gold standard for impacts allows unbiased impact tabulations.

Table 2-1. Number of violations issued for failure to report an impact during the 5th assessment period (per comm, William Keefer PADEP).

Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Number of violations for failure to report an impact	4	1	1	0	1	1

Throughout the assessment report, the University treats the BUMIS record as the most correct reported impact data set and assesses this assumption by cross-checking BUMIS data where possible.

2.D - Data Sources, Software, and Standardization

2.D.1 - Base Data

As with any broad spatial analysis, the use of common spatial data is helpful. The University relied on federal, state, and county sources for a wide variety of base data. Table 2-2 summarizes the data and the source.

Table 2-2. Spatial data assembled from external sources for use in the 5th Act 54 Assessment.

Data	Agency	Data Web Address (if applicable)
Local Roads	PennDOT	http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=PaLocalRoads2014_02.xml&dataset=1038
State Roads	PennDOT	http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=PaStateRoads2014_02.xml&dataset=54
Networked Streams of Pennsylvania	Environmental Resources Research Institute (ERRI) at PSU	http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=netstreams1998.xml&dataset=16
Small watershed boundaries	Environmental Resources Research Institute (ERRI) at PSU	http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=smallsheds.xml&dataset=14
Waterbodies	National Hydrography Dataset (NHD)	http://nhd.usgs.gov
County Boundaries	PennDOT	http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=PaCounty2014_02.xml&dataset=24
Elevation	PAMAP program LiDAR	http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=PAMAP_DEM.xml&dataset=1247
Washington County Parcels	Washington County GIS Department	Direct from GIS Department
Greene County Parcels	Greene County Department of Economic Development, GIS Coordinator	Direct from GIS Department

2.D.2 - Data on Mining Extents and Undermined Features

The University gathered mine extents and undermined feature locations from three sources:

- Six-month mining maps and Plus Maps submitted by the mine operators to PADEP and rectified by PADEP
- Digital spatial data (GIS or CAD files supplied by the mine operators to the University)
- Bituminous Underground Mining Information System (BUMIS)

2.D.2.a - Six-Month Mining Maps

Six-month mining maps are submitted by the mine operator to PADEP twice a year for each mine. Map submittal schedules for the various mines are distributed across the year by PADEP to spread PADEP map evaluation effort evenly over the course of the year. The University focused on the following in the six-month maps:

- The extent of any mining that occurred during the prior six months
- Anticipated mining extent during the next six-month period
- Locations of surface features such as properties, structures, water supplies and utilities over or close to undermined areas and thus the subject of PADEP mining-related oversight.

Most maps include additional information on coal depths, usually as contours.

The University received rectified six-month mining maps as Tagged-Image Format (TIF) or multi-resolution seamless image database (MrSID / SID) format images from the California District Mining Office (CDMO). The University used the most recent available maps and incorporated newer versions as they became available.

2.D.2.b - Mine Operator Data

The University visited mine operators to request digital versions of the information that constitutes a six-month mining map to allow for data quality assurance. Most operators agreed to supply digital ArcGIS or AutoCAD files to enhance accuracy and diminish inherent errors associated with digitizing of features from paper mine maps. The digital data obtained from operators were compared with the paper six-month mining map records to assure both documents were consistent. The University used ArcGIS 10.x software (ESRI, Redlands, CA) as the primary spatial data management tool and all CAD data were converted into ArcGIS compatible files. The CAD files were initially opened in Autodesk's AutoCAD software and each of the features of interest was then exported individually as a separate CAD file. This suite of new CAD files was then imported one-by-one into ArcMap using ESRI CAD-to-geodatabase tools.

2.D.2.c - Use of Bituminous Underground Mining Information System in Data Gathering

BUMIS tracks surface features that include structures, water supplies (wells and springs), water bodies, streams, and utilities (water and sewer supply systems, power lines, gas lines, and roads). During the 5th assessment period, a set of additional tables designed to track stream impacts (pooling, heaving, flow loss, etc.) and mitigation of these stream impacts was added to the

BUMIS database. BUMIS not only tracks impact on surface features, it also includes both corrective actions by mine operators and regulatory actions by PADEP.
BUMIS access is a challenge for non-PADEP entities. The interface is built in Java, a language that modern internet browsers do not support due to security concerns. For example, Java is not supported by current versions of Mozilla Firefox or Google Chrome. Any interaction with PADEP computer systems by the University therefore required the use of archived versions of Internet Explorer (which is less secure than actively maintained browser software). As with the 4th assessment, BUMIS data were downloaded by PADEP as Microsoft Excel spreadsheets and shared with the University. The University rebuilt data structures in Microsoft Access and compared these data with spatial data in ArcGIS. This arrangement causes problems. For example, for some cases there is extensive narrative included about complicated or ambiguous impacts. It took several months of communication with PADEP for the University to understand that these comments had been moved to a new table that was not included with the Excel spreadsheet. Further, once requested, this table had to be extensively processed by PADEP managers before it was in a workable form to use with the other data tables. These limitations are frustrating for all users of BUMIS. The University strongly recommends the modernization of the BUMIS data systems.

The most important limitation imparted to BUMIS by these software constraints is the limited ability of the primary impact tracking tool to interface with data managed by geographic information systems. Spatial coordinates were almost never provided for non-stream features in BUMIS during the 5th assessment period (e.g., there were no coordinates provided for over 1,000 "feature" entries for structure, land, and water supply impacts recorded during the 5th assessment period) . The University utilized information in the six-month mining maps and Module 22, Section 22.7 of the underground mining permit application (5600-PM-BMP0324) in the permit files to locate features based on the feature's identification number. During the 4th assessment over 40 % of BUMIS features (i.e., structures, water supplies, utilities) were not assigned a unique identification number. In contrast, during the 5th assessment only 16 % of all reported effects had to be clarified with the PADEP for any reason. While the accuracy of BUMIS information has improved, some of the specific suggestions made in the 4th assessment report were slow to be adopted or not adopted at all. Part of this slow pace results from barriers to this change at fundamental levels. For example, there are apparent legal concerns about the validity of professional seals in electronic versions of data. The University recommends creation of infrastructure to allow for electronic submission of these data.

2.D.3 - Act 54 2nd and 3rd Assessment Spatial Data

Because the University has created the spatial databases for the 3rd and 4th Act 54 assessments (Iannacchione et al. 2011, Tonsor et al. 2014), the data were readily available for use in the current assessment. Additionally, the University had acquired the spatial data from the California University of Pennsylvania's 2nd Act 54 report (Conte and Moses 2005), however, those data are in a different format and therefore used less frequently.

2.D.4 - Standardization of Datum and Coordinate Systems

Once collected, all spatial data sets were converted to a standard datum and coordinate system: the North American Datum of 1983 (NAD 1983) for the earth-shape model and the Universal Transverse Mercator (UTM) coordinate system to ensure equal area in analysis and minimize local distortion. During the 5th Act 54 assessment, all undermined layers lie within Zone 17 North of the UTM coordinate system.

2.E – Data Layers Generated by the University

Using the data described above, the University generated relevant data layers for addressing the tasks set forth by PADEP. Specifically, the University generated the following the data layers for each mine, where applicable:

- Mining Extents
- Surface Features
- Overburden
- Buffers
- Topography
- Wetlands

2.E.1 - Mining Extents

As in the 4th assessment report, "mining extents" define the area on the surface within which room-and-pillar, longwall, or pillar recovery mining took place. Mining extents in the 5th assessment include all slivers of mining extent that were not captured in the 4th assessment due to the mine map submission schedule challenges outlined above. That is, the mining extents are stored as polygons and attributed to the applicable mining type. The polygons were compared to the mining extents reported for the 3rd and 4th assessment periods. The comparison revealed small areas, "slivers" of land that were undermined during the 4th assessment but were not reported in the 4th assessment. These slivers of unreported prior mining were stored separately in the University's GIS database, but included in the totals and analyses presented in this report. The end extent reported was based on the best final information available during the University's data intake process and are depicted in the maps in Appendix B.

Most mining extents were provided by the mine operators soon after August 30, 2018 (earlier if mining had ceased prior to that date) as digital files (see Section 2.C.2.2). Mining extents that were not available as digital files were traced by the University using ArcMap from rectified six-month mining maps provided by PADEP (Table 2-3).

2.E.C.2 - Surface features

Property parcels, structures, water supplies, water bodies, streams and miscellaneous utilities are considered surface features. In cases where digital data were not available, these features were created by tracing features depicted in the six-month mining maps. Stream and utility layers were obtained from the base data (Table 2-2).

Additional information was calculated for the structure, water supply, and water body features. These data were added to the feature's attributes. The following fields were added:

- Distance to Mining: the straight-line distance between each feature within the layer and the edge of mining; calculated for each mining method (Room-and-Pillar, Longwall, and Pillar Recovery) using ArcMap's Near tool.
- Proximity to Buffer: classification of each feature as inside or outside of that layer's applicable buffer; calculated using ArcMap's Select By Location tool and the University-created buffers (see section 2.C.4)

2.E.3 - Overburden

Overburden is the amount of overlying rock between mining and the surface. The method of overburden calculation varied with the data provided to the University, particularly as reporting of coal or surface elevation varied by mine. Depending on the data, the University utilized one of two methods to derive overburden layers for each mine.

When overburden data were reported, it was used directly by the University after cross check with the six-month mining maps. Overburden rasters were used as is, while overburden contours were interpolated to a continuous overburden raster with the ArcGIS "Topo to Raster" tool.

When overburden was not provided, coal contours were interpolated into a continuous coal elevation raster with the Topo to Raster tool (resolution: 30-m pixel). The coal elevation raster was subtracted from the surface elevation raster derived from LiDAR DEMs to generate an overburden raster.

2.E.4 - Buffers

The University created the three buffers corresponding to PADEP regulatory boundaries. Buffers were derived using the combined mining extents for all mining methods employed. For example, buffers were based on the combined extents of mining methods if a mine employed both room-and-pillar and longwall mining methods.

1. A 1,000-ft outer buffer was derived for each mine. All features falling within this buffer were tracked. Any feature that fell within 1,000-ft of mining was identified and, if impacted, associated with the BUMIS record.

2. A 200-ft outer buffer was delineated for each mine, per PADEP policy. Structures and streams falling within this buffer were included in structure and stream feature inventories compiled for each mine.

3. The Rebuttable Presumption Zone (RPZ), a buffer dependent on overburden, was derived from the overburden rasters. Each overburden pixel falling on the mining extent boundary was buffered separately using the equation shown below:

$$RPZ = O * tan(35^{\circ} * \frac{\Pi}{180})$$

(Where O = the overburden pixel value and RPZ = the buffer distance)

The resulting pixel-buffer boundaries were combined into single polygon that served as the RPZ buffer. Water supplies and water bodies falling within the RPZ are presumed to be impacted by mining unless the mine operators rebut this presumption with additional data.

2.E.5 - Wetlands

Data used to calculate wetland acreage undermined by longwall mines during the 5th assessment period was sent to the University in different formats, including ESRI ArcGIS shapefile, AutoCAD, and Environmental Resource Maps. In every case, the data were converted from original file formats into a geodatabase feature class. For all longwall mines, post-mining delineations could not be recorded over the entire 5th assessment extent because post-mining wetland delineations are conducted on a separate five-year permit renewal schedule for each mine.

2.F –GIS Database Structure

The University organized these data into a separate folder for each mine that contained all data for that mine. Within the folder, each mine has:

- A “personal” geodatabase (.mdb) containing all of the digital ArcGIS-format data layers collected or derived.
- A folder of geo-referenced six-month mining maps
- A “Topography” folder containing rasters relating to elevation
- Original CAD files, if the data was received in Autodesk format
- A map file (.mxd) for ArcMap versions 10.x

While each mine had unique information, all were required to have certain features for analysis. Table 2-3 shows the full list of required features.

Table 2-3. Required layers, their locations, and origins.

Feature	Type	Location	Origin
Overburden Layer	Raster	Geodatabase	Collected
Mining Extent Layer(s)	Feature Class	Geodatabase	Collected
Structure Layer	Feature Class	Geodatabase	Collected
Water Sources Layer	Feature Class	Geodatabase	Collected
Water Bodies Layer	Feature Class	Geodatabase	Collected
Properties Layer	Feature Class	Geodatabase	Collected
Buffer Layers	Feature Class	Geodatabase	Created
LiDAR DEM	Raster	Topography Folder	Collected
Geo-referenced Six-Month Mining Maps	Geo-TIF Images	Geo-referenced Maps Folder	Collected

2.G – Summary

To fulfill the tasks outlined by PADEP, the University constructed a Geographic Information System. Data collection occurred in two main efforts: collection of closed or idle mine before the August 2018 end date and collection of active mines beginning at the end of August 2018. The overlap between the University's contracted project period and the Act 54 assessment period continues to pose some challenges for data collection. PADEP's staggered schedule for submission of six-month mining maps resulted in some data being unavailable to the University for analysis. It is not clear that the condensed project period that would result from changing the project period would be a better option for assessment, given the iterative nature of quality assurance efforts.

References

Conte, D. and L. Moses (2005) "The Effects of Subsidence Resulting from Underground Bituminous Coal Mining on Surface Structures and Features and on Water Resources: Second Act 54 Five-Year Report," California University of Pennsylvania, http://www.portal.state.pa.us/portal/server.pt/community/act_54/20876

Iannacchione, A. S.J. Tonsor, M. Witkowski, J. Benner, A. Hale, and M. Shendge (2011) "The Effects of Subsidence Resulting from Underground Bituminous Coal Mining on Surface Structures and Features and on Water Resources, 2003-2008," University of Pittsburgh, http://www.portal.state.pa.us/portal/server.pt/community/act_54/20876

Tonsor S.J., Hale A.N., Iannacchione A., Bain D.J., Keener M., Pfeil-McCullough E., Garmire K. 2014. The effects of subsidence resulting from underground bituminous coal mining, 2008-2013. PA Act 54 Five-Year Report. 435 pp. A report to Pennsylvania Department of Environmental Protection.