

1 **Historic and modern approaches for discovery of**
2 **abandoned wells for methane emissions mitigation in**
3 **Oil Creek State Park, Pennsylvania**

4 **Patricia M.B. Saint-Vincent^{1,2*}, James I. Sams III^{1,2}, Mumbi Mundia-Howe^{1,2}, Matthew D. Reeder^{1,2},**
5 **Garret A. Veloski¹, and Natalie J. Pekney¹**

6

7 ¹Geologic and Environmental Systems Directorate, National Energy Technology Laboratory, Department
8 of Energy, 626 Cochran Mill Rd, Pittsburgh, PA, 15236

9 ²Leidos Research Support Team, National Energy Technology Laboratory, Pittsburgh, PA, 15236, USA

10

11 *Correspondence: patricia.saint-vincent@netl.doe.gov

12

13

14

15

16 **Abstract:** (1) Background: Hundreds of oil wells were drilled along Oil Creek in Pennsylvania in the mid-
17 1800s, birthing the modern oil industry. No longer in operation, many wells are now classified as
18 abandoned, and, due to their age, their locations are either unknown or inaccurately recorded. These
19 historic well sites present environmental, safety, and economic concerns in the form of possible methane
20 leaks and physical hazards. (2) Methods: Airborne magnetic and LiDAR surveys were conducted in the
21 Pioneer Run watershed in Oil Creek State Park to find abandoned wells in a historically significant but
22 physically challenging location. Wells were drilled in this area prior to modern geolocation and legal
23 documentation. Although a large number of old wells were abandoned summarily without remediation of
24 the site, much of the land area within Oil Creek State Park is now covered in trees and dense underbrush,
25 which can obscure wellheads. The thick vegetation and steep terrain limited the possibility of ground-
26 based surveys to easily find well sites for methane emissions studies. The data from remote sensing
27 surveys were used to corroborate potential well locations from historic maps and photographs. Potential
28 well sites were verified in a ground-based field survey and monitored for methane emissions. (3) Results:
29 Two historic photographs documenting oil activity in the late 1800s were georeferenced using a
30 combination of magnetic and LiDAR data. LiDAR data, which were more useful in georeferencing and in
31 field verification, identified 290 field locations in the Pioneer Run watershed, 86% of which were possible
32 well sites. Sixty-two percent of the ground-verified wells remained unplugged and comprised the majority
33 of leaking wells. The mean methane emissions factor for unplugged wells was 0.027 ± 0.099 kg/day, lower
34 than other Appalachian Basin methane emissions estimates. (4) Conclusions: LiDAR was used for the
35 first time, in combination with an airborne magnetic survey, to reveal underground oil industry features
36 and inform well identification and remediation efforts in difficult-to-navigate regions. In the oldest oil fields,
37 where well casing has been removed or wood conductor casing was installed, historic photographs
38 provide additional lines of evidence for oil wells where ground disturbances have concealed surface
39 features. Identification of well sites is necessary for mitigation efforts, as unplugged wells emit methane,
40 a potent greenhouse gas.

41

42

43 **Keywords:** LiDAR; magnetometry; oil and natural gas; abandoned wells; historical photos; greenhouse
44 gases; methane

45

46 **1. Introduction**

47 Oil Creek State Park (OCSP) in northwest Pennsylvania is an historically significant area for the oil and gas
48 industry, comprising some of the world's first commercial oil fields which were heavily exploited during the
49 late 1800s (Figure 1). Following the success of the Drake well (**Error! Reference source not found.**1)
50 near Titusville, Pennsylvania, in 1859, hundreds of wells were drilled downstream of the Drake well along
51 Oil Creek, setting off the first oil boom that birthed the modern oil and gas industry [4]. The relatively small
52 oil reservoir was quickly depleted, and many wells were abandoned without being sealed or plugged (i.e.,
53 cemented according to regulatory standards). No longer in operation, many of these wells are now classified
54 as abandoned or orphaned. Location records for these historic wells, either do not exist, are incomplete, or
55 contain inaccurate data [5]. Estimations of the number of abandoned wells across the state vary, but
56 generally suggest 100,000 – 300,000 wells [5, 6]. Some 44% of the abandoned wells across the
57 Appalachian Basin, which includes wells in Pennsylvania, are historic wells, or wells that were drilled prior
58 to modern drilling and plugging regulations that were first introduced in the 1950s [7]. Pennsylvania, like
59 other states, is in the process of plugging historic abandoned wells, but the lack of dependable well location
60 information in OCSP and other areas has plagued that process with difficulty.

61 Today, it is becoming increasingly important to locate these historic well sites as they often present
62 environmental, safety, and economic concerns in the form of possible methane (CH₄) leaks and other
63 physical hazards. Unplugged or improperly plugged wells may emit CH₄, a potent greenhouse gas with a
64 global warming potential 86x that of carbon dioxide over a twenty-year time span [14]. Although naturally
65 occurring, atmospheric CH₄ levels have been steadily increasing due to anthropogenic activities such as
66 oil and gas extraction, resulting in climate change [15]. While some wells were later plugged to provide safe
67 conditions for hikers at OCSP, at the time of abandonment most wellbores were left open to the
68 environment. Plugging wells to prevent leakage or reduce safety concerns can be expensive, ranging from
69 hundreds to hundreds of thousands of dollars [8]. Yet unplugged wellbores can provide a pathway for the
70 undesired upward migration of fluids and gas and are especially concerning in areas with active
71 unconventional wells and hydraulic fracturing reservoirs [9-13].

72 This study was conducted for the purpose of demonstrating well finding technologies in the process of
73 analyzing historic orphan oil and gas wells for CH₄ emissions. The study area offered several unique
74 challenges to traditional ground-based well finding, including buried and wood-cased wells, steep and
75 forested terrain, and lack of documentation for many wells.

76
77 *1.1 Geologic Setting*

78 At the surface, Venango County bears evidence of repeated glaciation events that helped to shape the
79 steep sided valleys with narrow valley floors seen around Oil Creek and its tributaries [16]. OCSP, protecting
80 nearly 19 km of Oil Creek and over 25 km² of the surrounding watershed, is a part of the Appalachian
81 Plateau Physiographic Province (**Error! Reference source not found.**), a region of deeply eroded valleys
82 formed by glacial melt [17]. The bedrock of the area, comprised of Pennsylvanian- and Mississippian-age

83 sandstones, siltstones, and shales (290 to 354 mya) (**Error! Reference source not found.**), were not
84 significantly affected by mountain-building processes that occurred in later eras farther to the east and thus
85 is relatively flat [16].

86 The predominant oil producing formation within the Oil Creek region, collectively called the Venango
87 Oil Sands group, is typically found at depths of less than 400 m (Figure S2). This 100 – 110 m thick unit is
88 a collection of sand layers between layers of shale and siltstone. The materials were deposited in a dynamic
89 near-shore marine environment during the late Devonian Period (354 to 417 mya) at a time when
90 oceanwater covered the region [18]. The Venango Oil Sands group is comprised of three oil-producing
91 sand formations which have been described in detail previously [18-20]. Briefly, oil and water are typically
92 found together within the producing sands, likely due to the low dip of the strata producing minimal gravity-
93 based separation of the fluids. The accumulation of oil and gas within the sands is a result of stratigraphic
94 traps or barriers to flow that were created during deposition. The oil content of the producing sands
95 averages 37.2% pore volume [21]. Water content of the sands ranges from 40 to 60% [22]. The second
96 and third sands, 105 – 150 m belowground, are the most evenly distributed across the study site and were,
97 subsequently, the most frequently tapped in early oil production.

98 A detailed geochemical analysis of oils collected from five locations within the Venango Oil Sands
99 group was performed by Dahdah and colleagues [23]. The study used a variety of methods, including gas
100 chromatography, flame ionization detection, and stable isotope analysis, to characterize a total of 16 crude
101 oil samples. Despite the loss of the lighter fraction of hydrocarbons from some of the oldest oil samples
102 (samples collection dates spanned from 1934 - 2009), the researchers concluded that the oils are all derived
103 from a common algal/bacterial mixture that was deposited in an open marine environment, suggesting that
104 the Marcellus formation was the source rock for oil in the region.

105

106 *1.2 The Development of the Oil Industry*

107 From these near-surface oil sands issues crude oil that naturally seeps out of the hillsides in the
108 Venango floodplain, giving Oil Creek its moniker. An early map of the North American colonies indicates
109 the colonial-era knowledge of the presence of petroleum near the Allegheny River just south of present-
110 day OCSP (Figure 1**Error! Reference source not found.**). For many years, the Seneca tribe of western
111 Pennsylvania had collected oil floating in springs and small streams for use in waterproofing boats,
112 lubrication, and as a medicine [24]. Yet commercial interest in oil extraction only manifested in the mid-
113 1800s as a result of scientific advances in the refinement of crude oil [25]. Combined with rising whale oil
114 prices due to the Confederacy's dismantling of the whaling fleet and the over-harvesting of the mammals –
115 the price of whale oil peaked at \$1.92 per gallon in 1854 (nearly \$59 in 2020 [26]) – the new source of oil
116 was necessary [27].

117

118 *1.3 Commercial Oil Drilling at Oil Creek*

119 Although crude oil had long been collected from natural oil seeps and sold for profit, large-scale production
120 of oil was not possible until Edwin Drake's innovation of applying salt well drilling technology to the
121 extraction of crude oil. Drake oversaw oil well drilling at the Brewer, Watson & Co. Farm near Titusville,
122 Pennsylvania, just north of modern-day OCSP. After a mishap at the drilling site in August 1859, workers
123 returned to find oil emanating from the well. Although it produced only 25 barrels a day, the eponymous
124 Drake well inspired the first oil boom (Figure S1) [28]. Between May and August 1861, 675 wells were
125 drilled along Oil Creek in Venango County [29] and thousands more would be drilled in subsequent years.
126 Without a scientific understanding of geology and geophysics, wells were drilled at a high density in the
127 region. The initial interest in drilling spurred thousands to move to northwestern Pennsylvania, creating
128 boomtowns and extensively changing the landscape as trees were felled and earth was moved to make
129 space for city and oil infrastructure.

130 The thousands of active oil wells soon depleted the shallow oil sands, and attention focused to new,
131 more productive oil fields. In less than 20 years, the oil reserve below Pioneer Run – so named for its
132 proximity to the boomtown Pioneer, Pennsylvania – was so exhausted that it was no longer profitable to
133 continue pumping. Many of the wells were abandoned and towns like Pioneer sunk into disrepair by the
134 1870s [30]. However, without plugging laws in place, many wells were left open to the atmosphere or
135 partially plugged with nearby trees and debris [31]. The towns, derricks, and oil industry infrastructure that
136 once dominated the area have since been replaced by dense underbrush in poorly drained lowlands and a
137 mature hardwood forest in the uplands, obscuring the evidence of legacy orphan wells. Erosion over time
138 on the once-deforested hillsides has wiped away easy access to well pads once accessible by footpaths.

139

140 *1.4 Remote Sensing*

141 Remote sensing, or the scanning of the earth by aircraft or satellite for the collection of information, is a
142 powerful tool in the burgeoning field of geographic information systems (GIS). With regard to the oil industry,
143 remote sensing is often used to monitor CH₄ emissions [32], blowouts and other accidents [33, 34], and
144 locations of activity [35, 36]. Remote sensing has also been applied to the detection of abandoned oil and
145 gas wells [5, 37]. Remote sensing techniques include radio detection and ranging (RADAR), infrared
146 imaging, light detection and ranging (LiDAR), and airborne magnetometry.

147

148 *1.4.1 Airborne magnetometry*

149 Airborne magnetic surveys rely on magnetometers to detect anomalies in the earth's magnetic field, the
150 result of ferromagnetic materials and infrastructure. The steel casing of many vertical wells appears as a
151 monopole in magnetic maps and is distinguishable from other magnetic features such as pipelines and
152 metal buildings, which appear as dipoles (Figure 2, Figure S3) [38, 39]. The remote sensing technique was
153 previously used to map the locations of orphan wells at an abandoned oil field in Hillman State Park in
154 southwest Pennsylvania, as well as in newer, active oil fields in Wyoming [5, 40, 41]. The power of the
155 technique lies in its ability to detect even those wells with no aboveground evidence, including wells that

156 have been partially cut off or have had the upper casing removed, reclaimed sites, and those covered by
157 mine spoils at depths of up to 30 m [40]. A helicopter-based magnetic survey of OCSP detected 931
158 monopoles indicative of well sites (Figure 2) [5].

159

160 1.4.2 LiDAR

161 LiDAR measures the distance between a probe and the target using lasers. The time delay in detecting
162 the light reflected from the target is used to map surface or ground features, enabling the visualization of
163 ground under a forest canopy or alterations to surface features [42, 43]. A digital elevation model (DEM) is
164 obtained from the processed data and can be automatically or visually inspected for features of interest. A
165 recent study of LiDAR data collected in Pioneer Run in OCSP by Mundia-Howe and colleagues identified
166 collapsed abandoned wells by identifying manmade topographic footprints including leveled well pads,
167 service roads and trails, and circular depressions from collapsed abandoned wells [37]. The study identified
168 141 depressions associated with collapsed abandoned wells after filtering the data for similar circular
169 features attributable to fallen trees. LiDAR detection of wells complements magnetic data because while
170 only ferromagnetic well casings are detected in a magnetic survey, LiDAR technology enables detection of
171 manmade alterations to a landscape associated with drilling activity, such as flattened well pads, collapsed
172 areas from wells where the casing has been removed, building footprints, and well pad access roads.

173

174 1.5 *Historic Evidence of Wells*

175 Due to the burgeoning oil industry in northwest Pennsylvania during the late 1800s, many maps of the
176 region document wells (Figure 1). Additionally, the photographer John Mather captured and preserved
177 images of the early oil field landscape. A collection of Mather's photographs from the 1860s are curated by
178 the Drake Well Museum (www.drakewell.org). Unlike well databases, which document wells across time
179 and across the state, the photographs capture wells at a single moment in time in a very specific location.
180 While 1800s-era maps do not provide precise coordinates for wells, they provide an additional line of
181 evidence for the vast number of early oil wells drilled in the region. The photographs also reveal wells that
182 may never have been recorded in public databases, as early oil drilling activity was not highly regulated. In
183 combination with LiDAR, airborne magnetometry, and database records, georeferenced photographs can
184 be consulted to propose locations for historic well sites in OCSP.

185

186 1.6 *Methane Emissions*

187 It is important to locate abandoned oil and gas wells because they may serve as conduits for oil and gas
188 into waterways or the atmosphere. Loss of well casing or plug integrity could allow fluids or gases to escape
189 the casing [44, 45]. In some cases, loss of well integrity has led to catastrophic disasters, such as at Aliso
190 Canyon [46], but often the wells may have individually small leaks that are significant when considered as
191 a whole. For this reason, the EPA began tabulating emissions from abandoned oil and gas wells in the
192 Greenhouse Gas Inventory (GHGI) in 2018 [15].

193 Plugged wells in the Appalachian Basin are estimated to emit 0.0086 kg/well/day and unplugged wells
 194 emit 0.73 kg/well/day, and nationwide release an estimated 657 million tons (MT) CH₄ per year and 226,000
 195 MT per year, respectively (

| Location | EF (95% UCI) (kg/day) | | Source |
|----------------------|-----------------------|---------------|---------------------|
| | plugged | unplugged | |
| Oil Creek State Park | 0.0023 (0.0054) | 0.027 (0.042) | this paper |
| Hillman State Park | ND | 0.7 (1.15) | Pekney et al.[50] |
| Pennsylvania | 0.36 (6.0) | 0.53 (4.86) | Kang et al.[58] |
| Ohio | 0 (N/A) | 0.67 (1.54) | Townsend et al.[48] |
| West Virginia | 0.0024 (0.0113) | 0.077 (0.239) | Riddick et al.[49] |
| Appalachian Basin | 0.0086 | 0.73 | GHGI ^a |
| Appalachian Basin | 0.081 | 0.13 | ^(b) |

196
 197 Table 1) [15]. EFs for abandoned wells in the Appalachian Basin are higher than those in the western
 198 United States. However, these estimates are based on small sample sizes and may not accurately reflect
 199 nationwide or even regional emissions. Wells across the Appalachian Basin have been studied, from West
 200 Virginia, Pennsylvania, and Ohio [47-49], but the variation within and across datasets suggests more data
 201 are needed to reduce uncertainty in CH₄ emissions estimates. In order to improve the GHGI and inform
 202 mitigation efforts, more abandoned oil and gas wells in the Appalachian Basin and across the United States
 203 should be analyzed.

204
 205 **2. Materials and Methods**

206 See Supplementary Methods for additional methods.

207 *2.1 Site Selection*

208 OCSP, established in 1931, was selected for validation of remote sensing techniques due to its accessibility
 209 as public land as well as its containing some of the oldest oil wells drilled in the United States. Access was
 210 generously provided by the Pennsylvania Department for the Conservation of Natural Resources (PA
 211 DCNR), who also permitted the helicopter and unmanned aerial vehicle (UAV) surveys over OCSP. The
 212 Pioneer Run study site within OCSP was selected for the LiDAR survey based on the high density of well

213 anomalies in the airborne magnetic data, the preponderance of historical photographs and maps
214 documenting wells in the area, and the difficulty of field-based well identification due to dense undergrowth
215 and steep terrain.

216 *2.2 Mapping and analysis*

217 Maps for field navigation were prepared in ArcGIS Pro and copied into the Esri Collector Application
218 installed on field computers (Esri, Redlands, CA, USA). A high-resolution GPS (Trimble R2 Integrated
219 GNSS System, Trimble Geospatial, Sunnyvale, CA, USA) mounted on a backpack was used to navigate
220 to potential well sites. Positions were logged with 10-point positional averaging. Sites were recorded as
221 being a potential well or not, plugged or unplugged, and leaking CH₄ or having no identifiable leak.
222 Additional field notes were recorded, including temperature, site characteristics, and other infrastructure
223 present. In some cases, a handheld magnetometer (Overhauser “Walking” Magnetometer, GEM Systems,
224 Tuscon, Arizona, USA) was used to confirm the presence of a ferromagnetic casing at a field site.

225 *2.3 Methane emissions detection and calculations*

226 The well sites were scanned for CH₄ leaks using a handheld laser detector (Pergam LMC, Pergam
227 Technical Services, Renton, WA, USA) directed from approximately 1 m away the well or ground around
228 the well. The laser detector has an operational range of 1-50,000 ppm-m and an accuracy of ±10%. The
229 background CH₄ was measured by directing the laser away from the wellhead for 1-2 min. Background
230 readings were 1-20 ppm-m. If a reading >100 ppm-m was consistently detected (5× signal enhancement
231 over background) for more than 1 min during a 5-10 min observation period, the well was recorded as
232 leaking above the threshold.

233 If a leak was detected, the site of the leak (i.e., wellhead, tubing, ground around the wellhead) was
234 isolated using an appropriate hose attachment (i.e., plunger or bag) and the leak rate was measured using
235 a flow CH₄ leak detector (Hi Flow Sampler, Bacharach, Inc, New Kensington, PA, USA), in both high and
236 low concentration modes, and the average of the two readings was used in subsequent analyses. The
237 device was calibrated prior to the first measurement each day [50]. The EPA Data Quality Assessment
238 handbook on Statistical Methods for Practitioners recommends using the detection limit (DL) divided by two
239 (½DL) for points below the DL of the instrument, if the dataset is not normally distributed [51]. The DL of
240 the Hi Flow sampler is 0.1 L/min in both low and high concentration operation modes [52].

241 The Hi Flow Sampler measurements were converted from L/min to kg/day using the density of CH₄ at
242 the outside temperature at the time of measurement (range 10 – 26.1 °C). The EF was calculated as the
243 mean of individual EFs for a particular subset of wells. The standard deviation of the mean was calculated.
244 The 95% upper confidence limit (UCL) was calculated according to the standard formula

$$245 \quad \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \quad (1)$$

246 where \bar{x} is the mean, σ is the standard deviation of the mean, and n is the number of wells analyzed in
247 the dataset.

248

249 **3. Results**

250 *3.1 Site Selection*

251 Oil Creek State Park was selected for studying abandoned wells based on its primary role in the beginning
252 years of the commercial oil industry. The land has remained largely untouched by later developments, so
253 much of the abandoned oil infrastructure including wellheads, pump jacks, barrels, and pipelines has been
254 preserved. Pioneer Run is a small region in OCSP characterized by steep hillsides around a floodplain that
255 drains into Oil Creek at the southern border. Covered with large second-growth trees and, in some areas,
256 dense undergrowth, Pioneer Run is difficult to navigate by foot, yet remains of well pads and old trails now
257 connect to make hiking trails along the slopes. Because of its protection under the PA DCNR, the land of
258 Pioneer Run has been preserved, as have the remains of historic oil activities.

259 Two state databases exist, containing 767 wells sites (PA DEP) and 955 well sites (EDWIN), with 442
260 shared sites between the databases (well locations within 10 m of each other)(Figure 3). The differences
261 in the two database counts may be related to the PA DEP keeping a separate record of abandoned and
262 orphan wells reported through its abandoned well program that may not be shared with the EDWIN dataset.
263 Location differences could be the result of conversion errors when digitizing and compiling records. The PA
264 DEP records do not have associated completion or spudding dates, unlike wells in the EDWIN database;
265 however, 582 of the EDWIN wells have pre-1901 completion dates. This is consistent with historical
266 photographs, records, and maps that would indicate extensive oil production in the area in the mid-1800s
267 that ended by the late 1800s.

268

269 *3.2 Summary of Magnetic and LiDAR Surveys*

270 From the previously-reported airborne magnetic survey, 931 magnetic anomalies associated with potential
271 well sites were selected within OCSP (Figure 2) [5]. Of these, only 49 were <10 m from a PA DEP well site,
272 suggesting the discovery of 882 possible previously-unidentified well sites within OCSP.

273 In the Pioneer Run study area, there were 34 well-type magnetic anomalies. In this same region, there
274 were 42 EDWIN well sites and 79 PA DEP well sites. As 18 of the 19 wells in the EDWIN database have
275 pre-1901 completion dates, it is likely that many of the documented wells were constructed with wood
276 conduit casing or had casing removed upon abandonment [53], making these wells invisible to a
277 magnetometer. Further supporting this hypothesis, only 2 of the 79 PA DEP wells in the Pioneer Run study
278 area were <10 m of a magnetic anomaly.

279 The subsequent LiDAR survey in the Pioneer Run study area identified wells with no ferromagnetic
280 casing based on the presence of a collapsed feature or sink on a flat surface, indicative of a wellhead on a
281 well pad [37]. Only two of the 141 sites identified in the LiDAR survey of Pioneer Run were <10 m from a
282 magnetic anomaly, suggesting that the two remote sensing techniques provided complementary verification
283 techniques in Pioneer Run. Only two well sites were <10 m from database sites, suggesting that although
284 coverage of the LiDAR data was more successful in finding legacy well sites than the magnetic data, the
285 data were similarly identifying wells not present in the databases.

286

287 *3.4 Georeferencing Historic Maps*

288 Maps of western Pennsylvania dating from as early as 1755 suggest knowledge of the presence of oil
289 (Figure 1A)[1]. Some early maps include property lines and locations of wells drilled in the area of Oil Creek.
290 Using the georeferencing feature in ArcGIS Pro, we successfully overlaid two 1800s-era maps onto
291 modern-day topographic maps of Venango County and Oil Creek. The first, a county-wide map published
292 by Beer, Watson, & Co., details the locations of properties, oil refineries, and some oil wells (Figure 1C)[2].
293 In a county map from 1864, cartographers marked the locations of known oil wells (Figure 1)[3]. In the
294 vicinity of modern day OCSP, 171 wells were marked. As only one of the georeferenced locations fell within
295 the Pioneer Run study area (40.0 m from the nearest LiDAR point), we did not compare locations with
296 LiDAR sites. There were more wells (46 wells) <50 m of a magnetic anomaly than with either the EDWIN
297 or PA DEP databases (25 and 21 wells, respectively), suggesting that there may indeed be wells on this
298 historic map that, while not positioned using modern-day GPS technology, could approximate the locations
299 of actual well sites.

300

301 *3.5 Georeferencing Historic Photographs*

302 John Mather was a photographer who documented the development and decline of the first oil boom
303 in Pennsylvania. His photographs show the extreme changes to the landscape that occurred during the
304 mid-1800s, from clear-cutting trees to digging well pads into hillsides. It was hypothesized that the historic
305 photographs could be georeferenced by matching photographic points to LiDAR features and magnetic
306 anomalies. The photographs could, in turn, be used as evidence to confirm the features identified in the
307 field surveys.

308 *3.5.1 Case study 1: Pioneer Run hillside*

309 One hillside in Pioneer Run showed particularly dense well pads in the DEM, consistent with the pattern
310 of derricks – which sat atop wellheads – observed in a photograph taken in 1865 along Pioneer Run (Figure
311 4A). Using LiDAR imagery, the photograph, and GPS, we confirmed that each well pad was a manmade
312 feature and used 3D imagery to recreate the photograph using the LiDAR hillshade and ArcGIS (Figure
313 4D). Consistent with the photograph, two rounded flat areas next to well pad 4 appear to be the sites of
314 storage barrels observed in the photograph (Figure 4A). A footpath running along the midline in the
315 photograph can also be seen extending from well pad 8 in the LiDAR but is not an established hiking trail
316 such as seen lower on the hillside below all the well pads. While the LiDAR DEM provides strong evidence
317 that this hillside is the one pictured in the 1865 photograph, changes to the hillside, such as erosion and
318 the development of hiking paths, in addition to regrowth of the forest, make visualization of the hillside as
319 seen in 1865 impossible (Figure 4E). However, LiDAR DEM enabled a simulated approximation of the same
320 hillside using Scene Viewer in ArcGIS Pro, which supports the hypothesis that the hillside is one and the
321 same. In agreement with the magnetic data (only three anomalies are within the approximate borders of
322 the photograph), most well pads did not present a magnetic response when scanned with a handheld
323 magnetometer and were thus likely cased in wood or had casing removed [53]. In the field verification

324 described below, the magnetic anomaly at the bottom of the hillside was found to be an accumulation of oil
325 drilling infrastructure rather than a well site. There were 8 EDWIN wells and 7 PA DEP wells (which matched
326 with 7 of the EDWIN wells), much fewer than the 18 derricks identified in the photograph and the 17 wells
327 predicted in the LiDAR DEM, suggesting that the databases were lacking records for at least ten wells on
328 this particular hillside.

329 3.5.2 Case study 2: Downtown Pioneer, Pennsylvania

330 Another Mather photograph captured the town of Pioneer, Pennsylvania in 1865 (Figure S4A). A
331 second perspective of the town also reveals the locations of derricks (Figure S4F). Based on historical
332 records and evidence from historic railroad tracks (**Error! Reference source not found.B**) and the flow of
333 the water, the location appears to be looking northeast over the town as Pioneer Run joins Oil Creek. In the
334 photograph, many wells dot the two hillsides and intermingle with houses and shops in the town, but
335 because of the extensive scarring on the ground from barrels, buildings, and roads, three well pads could
336 not be identified in the LiDAR data (Figure S4D,F). Poor LiDAR data quality in the valley where Pioneer
337 Run meets Oil Creek made localization of two derricks challenging. In all, 24 well pads were identified in
338 the LiDAR data which appeared to correspond to well sites in the photograph (Figure S4C-E). Only 5
339 anomalies were identified using magnetometry and these were not, except in one case, close to LiDAR
340 sites (Figure S4E). Surprisingly, the PA DEP database predicts 36 wells in the region covered by the
341 photograph, more than found in the LiDAR data; however, many of the derricks pictured in the photograph
342 could not be positively referenced in the LiDAR DEM. This could be the result of extensive changes to the
343 landscape from buildings and roads, obscuring well sites on the ground. There were 16 wells in the EDWIN
344 database in the same region, 13 of which closely matched a PA DEP location for a total of 39 possible well
345 sites from combined state databases.

346

347 3.3 Field verification of well sites

348 In order to determine the utility of remote sensing and georeferencing techniques for identifying
349 abandoned wells in the field, potential well sites identified from these techniques were verified in OCSP
350 from May 2016 to May 2019. A total of 278 possible well sites were visited (Figure 5, Table S1**Error!**
351 **Reference source not found.**), and 245 (88%) were determined to be potential well sites based on visible
352 infrastructure or surface features associated with covered wells (i.e., flattened well pad with a sunken,
353 circular region at one end) (Figure 4D). Many of the LiDAR-identified potential well sites that were dismissed
354 were instead anomalies in the topography resulting from a fallen tree, whose uprooted bulb left behind a
355 large circular depression similar to a collapsed well [37].

356 Of the 245 confirmed well sites, 93 (38%) were plugged wells (Figure 5A, Figure S5). Wells that were
357 not visible but had aboveground indicators were considered unplugged. Digging for the wellhead at five of
358 these sites revealed an open well casing, affirming the classification. Due to safety concerns, digging was
359 not pursued at more sites; instead, topographic features and other industrial infrastructure present at the
360 site were instead used as indicators of a well.

361 The distance to the nearest magnetic point was calculated for each confirmed well site, and the
362 distance to the nearest LiDAR point for those in the Pioneer Run study area (82 field sites) (Figure 6). Of
363 the 245 ground-verified well sites, 168 (72%) were <50 m of a magnetic point, with a mean geodesic offset
364 of 20.4 m \pm 13.1 m. Only 32 of the 82 wells in Pioneer Run (39%) were <50 m from a magnetic point (mean
365 geodesic distance 32.8 m \pm 12.1 m), but 73 (89%) were <50 m of a LiDAR point, with a mean geodesic
366 offset of 11.0 m \pm 13.3 m. The difference in accuracy is statistically significant (p-value 1.03×10^{-9}). LiDAR
367 data more accurately predicted positive locations for wells in the Pioneer Run study area.

368 From the approximate locations of wells georeferenced in the 1864 Venango County map, 11 of the
369 171 sites were <50 m from a field-verified well site, with an average offset of 31.7 m \pm 0.1 m. The small error
370 suggests a standard offset due, perhaps, to human error in map creation. For the historic photograph
371 regions, LiDAR more reliably predicted well locations as few magnetic anomalies were present in the data.
372 Of the well sites documented in the John Mather photographs, a total of 32 were confirmed as wells during
373 field verification (Table S1), and three unplugged wells were found to be leaking. While not all of the
374 potential well sites were field-verified, the field locations do match well with what was predicted in the LiDAR
375 data and photographs, with a mean offset of 20.8 m \pm 31.0 m.

376

377 *3.4 Methane Detection Survey*

378 Of the 245 confirmed well sites, 210 were examined for CH₄ leaks; not all wells were assessed based on
379 accessibility to the wellhead (Figure 6). For buried wells, the existing well infrastructure and the ground
380 around the well head were scanned with the Pergam handheld laser for gaseous emissions. Of the 210
381 well sites examined, only 21 (9%) produced CH₄ above the background, or consistently registering >100
382 ppm-m on the handheld device. Of these leaking wells, only one was a plugged well and the remaining 20
383 remained unplugged. Thus, 1% of tested plugged wells had a detectable leak, while 16% of unplugged
384 wells were leaking CH₄. Similar to Townsend-Small et al. and Saint-Vincent et al., plugged wells were much
385 less likely to be leaking at a detectable threshold [48, 54].

386 After identifying the source of the leak and confirmation with an infrared camera (FLIR GF320, FLIR,
387 Wilsonville, OR, USA) (data not shown), methane emissions were quantified with the Hi Flow Sampler.
388 Wells with a known leak that did not register on the lowest setting of the Hi Flow Sampler were considered
389 as positive for a leak but below the DL of the instrument. The one plugged well that was leaking was below
390 the DL of the Hi Flow Sampler, and 10 of the 20 leaking unplugged wells fell below the DL (Figure 4B,C).

391 The mean emissions factor for plugged wells was 0.0023 \pm 0.0151 kg/day (range 0 – 0.14 kg/day), and
392 the mean emissions factor for unplugged wells was 0.027 \pm 0.099 kg/day (range 0 – 0.98 kg/day) (Figure
393 6). Yet the median emissions rate for both plugged and unplugged wells is 0 kg/day, showing long-tailed
394 distributions for both well types. The small percentage of wells that account for the majority of the measured
395 CH₄, called super-emitters, skew estimated emissions, because the majority of components are either non-
396 emitters or small emitters [55-57]. Such outliers in abandoned well datasets have been reported, resulting

397 in non-normal data distributions [54, 58]. At OCSP, only 21 wells had detectable leaks, with only 9% of the
398 wells accounting for all CH₄ emissions.

399

400 3.4.1 *Field verification, methane detection, and georeferenced photos*

401 The three georeferenced photographs covered regions analyzed in the magnetic and LiDAR surveys, which
402 were subsequently covered in the ground-based field verification of well sites. Forty-eight sites were
403 surveyed in the photograph regions (Figure 4. John Mather photograph and LiDAR localization of wells. (A)
404 1865 photograph of a hillside in Oil Creek State Park. Derricks marked in yellow indicate well sites
405 potentially identified in the LiDAR survey. (B) The same hillside above Pioneer Run Creek in May 2017,
406 with dense tree cover obscuring the scars of the extensive oil drilling activity that took place here during the
407 1800s. (C) GIS reproduction of the hillside using LiDAR imagery. Well numbers in yellow circles indicate
408 nearest derrick observed in the photograph in (A) with ground features found in the LiDAR data. (D) LiDAR
409 hillshade in 2D, showing flat portions of land that are likely from well pads and numbers approximating the
410 associated derricks in (A). (E) Topographic map of the same hillside, showing locations of known wells and
411 the known magnetic anomalies., Figure S6**Error! Reference source not found.**). Of these, 32 were
412 confirmed as potential well sites, with an average distance of 76.0 m from a magnetic anomaly but only
413 20.8 m from a LiDAR point (Figure S6**Error! Reference source not found.**D). This indicates that wells
414 shown in the photographs have likely had ferromagnetic casing removed, or that the conduit casing was
415 wooden material, consistent with known well drilling practices of the era. In contrast, the sites selected as
416 well locations from the photographs were on average 27.5 m \pm 23.4 m away from a ground-verified well site.
417 Only 6 (17%) of the wells were plugged, none of which were leaking CH₄ above the defined threshold, but
418 3 unplugged wells had detectable gaseous emissions below the DL threshold (Table 2).

419

420 3.4.2 *Comparsion of methane emissions factors*

421 The EPA recently incorporated abandoned wells in the GHGI, estimating the contribution of abandoned
422 wells to both CH₄ and CO₂ emissions [15]. Incorporating data from three reports, the GHGI separated wells
423 from the Appalachian Basin from those in the rest of the United States in its CH₄ emissions estimates [47,
424 48, 58].

425 Compared to the compiled EF for oil wells in the Appalachian Basin (0.0086 kg/day and 0.73 kg/day
426 for plugged and unplugged wells, respectively), the OCSP EF for unplugged wells is lower by an order of
427 magnitude. At Hillman State Park in southwest Pennsylvania, unplugged historic wells emitted at a rate
428 similar to those in the GHGI (0.50 kg/day)[50], but the unplugged well EF in West Virginia were more similar
429 to the historic OCSP wells (0.077 kg/day) (Table 3) [49]. The larger number of wells surveyed, in addition
430 to the different oil field, may contribute to differences in EFs. Age may also be a contributing factor to
431 abandoned well emissions, although previous studies have not found well age to have a strong correlation
432 to emissions [54, 58]. We compared the emissions from all ground-verified well sites <50 m from an EDWIN
433 site with a recorded completion date (n = 115) and found a significant difference between emissions from

434 pre-1950 wells (0.033 ± 0.001 kg/day, $n = 95$) and wells completed after 1950 (0 ± 0 kg/day, $n = 20$). Nearly
435 44% of all abandoned, orphaned, or P&A oil or oil and gas wells in the Appalachian Basin have completion
436 dates prior to 1950, and nationwide comprise 20% of all abandoned, orphaned, or P&A wells, though
437 previous studies suggest historic wells are underrepresented in national databases [6].

438 We compared the OCSP well EFs to data collected previously in Ohio, Pennsylvania, and West
439 Virginia (Table 1). Similar to the previously reported GHGI values, unplugged wells were more prone to
440 unintentional CH₄ emissions. While the EF for plugged wells was higher in the GHGI than at OCSP, this is
441 likely the result of skewed data, as reported here and in the newly incorporated West Virginia dataset [49].
442 Combining this study with previous reports, the 183 plugged wells had an average EF of 0.081 kg/day and
443 the 323 unplugged wells had an average EF of 0.13 kg/day (Table 1, Table 3).

444 In the GHGI, the EPA estimated that 23% of the 889,068 plugged and 1,744,585 unplugged wells were
445 in the Appalachian Basin (Table 3) [15]. With the newly calculated EFs for Appalachian Basin abandoned
446 wells, yearly emissions from the Appalachian Basin can be estimated. Plugged wells release 6.1 million
447 kg/y and unplugged wells release 19.2 million kg/y, or 0.63 MMT CO_{2e}. While only a minor fraction of total
448 CH₄ emissions from the oil and gas sector, mitigation of abandoned oil and gas wells that are leaking will
449 decrease overall anthropogenic CH₄ emissions.

450

451 **4. Discussion**

452 Oil drilling in the United States began in the mid-1800s in northwestern Pennsylvania, setting off a revolution
453 in energy production and fueling modern industrial development. Many of the wells in the small oil field were
454 abandoned after only a few decades due to decreased production. While Pennsylvania keeps records of
455 the locations of oil and gas wells, many records of these older, abandoned wells are incomplete, missing,
456 or incorrect. Two separate state databases differ in the number and location of wells in OCSP, where some
457 of the first commercial oil wells in the United States were drilled. Many of these orphaned wells, having no
458 known operator or owner, are difficult to find under dense foliage, steep terrain, removal of well
459 infrastructure, and years of neglect. However, these wells are a danger for unsuspecting hikers and could
460 also pose environmental hazards in the form of oil leaks and CH₄ emissions. Determining the locations of
461 abandoned oil wells is imperative not only in OCSP but also across the country to guide mitigation efforts.

462 Many of the oldest oil wells were drilled without permits and abruptly abandoned in the same manner.
463 Thus, records for these wells are missing, incomplete, or inaccurate, based on hand-drawn maps or other
464 historical documents. The Pennsylvania EDWIN database has records for 955 abandoned wells in OCSP,
465 while the PA DEP database logs 767 wells. A magnetic survey over the entirety of the park detected 931
466 well-type magnetic anomalies, showing that the PA DEP database likely underestimates well counts. Yet
467 discrepancies in well locations with the databases, especially in regions with historic photographs
468 documenting well locations, shows that the limitation of magnetic surveys to detect wells without a
469 ferromagnetic casing is particularly troublesome at OCSP.

470 The Pioneer Run study area presents an unusual case in well finding because many of the wells in
471 this area were drilled and reinforced with wooden conduit casing rather than steel [53]. Evidence of prolific
472 well drilling in historic photographs and maps confirms this hypothesis (Figure 1 **Error! Reference source**
473 **not found.**, Figure 4, S4). Only 34 magnetic anomalies were found within the Pioneer Run study area,
474 while the PA DEP and EDWIN databases tabulated 34 and 79 wells, respectively (Table S1). The LiDAR
475 survey in the Pioneer Run study area predicted locations for 141 well sites based on topographical
476 anomalies. Yet despite the increased well site detection compared to the magnetic survey, LiDAR data did
477 not coincide with state database locations; only 2 of all LiDAR sites were <10 m from a recorded well in the
478 state databases. The well databases did not accurately predict the locations of well sites found in either the
479 magnetic or LiDAR surveys, and did not match well with the LiDAR and photo-based analyses.

480 Using a combination of topographic maps and LiDAR DEMs, three photographs documenting early oil
481 activity within the Pioneer Run study area were georeferenced and LiDAR-derived well sites were matched
482 to photographic evidence. In all, 47 individual well locations were matched to possible field locations. A
483 georeferenced map from 1864 showed locations of wells mostly along Oil Creek (Figure 1), supporting the
484 hypothesis that early oil prospectors thought oil was more likely to be found near rivers and streams.

485 A total of 245 wells were located in a subsequent ground-based field survey of a region of OCSP.
486 Unplugged wells were more prone to leaking CH₄ than plugged wells: while only 9% of the wells surveyed
487 were emitting CH₄, all but one were unplugged. This further supports the need for proper plugging of
488 abandoned wells to reduce CH₄ emissions. Yet even though wells in OCSP were found to emit CH₄, the
489 calculated emissions factor was lower than for other wells in the Appalachian Basin. Likely, more than just
490 geographic location is important for calculating emissions factors for abandoned wells to reduce uncertainty
491 in the GHGI.

492

493 **5. Conclusions**

494 Thousands of abandoned and orphan wells dot the landscape of the United States, especially in
495 northwestern Pennsylvania, where the first commercial oil field is located. Here, legacy wells over 100 years
496 old are hidden beneath forest canopy in Oil Creek State Park in northwestern Pennsylvania, making well
497 finding for mitigation efforts difficult. State databases, especially for historic wells, are incomplete and
498 inaccurate, but remote sensing techniques including airborne magnetic and LiDAR surveys provide more
499 precise coordinates for potential well sites. Historic photographs and maps provide additional evidence of
500 well locations that may not be contained in databases, especially in regions where drilling practices
501 predated legislation. The study at OCSP presented difficult conditions for research, including old or buried
502 infrastructure and unapproachable terrain, but will inform wellfinding and CH₄ emissions studies across the
503 country. Knowledge of the number and location of abandoned oil and gas wells is imperative in directing
504 environmental and safety hazard mitigation. Not only are abandoned wells potentially dangerous to walkers
505 and native animals, but oil leaks and CH₄ emissions pollute the environment. Surveying wells across
506 geographical location, well age, and well depth may provide new insights into the factors contributing to

507 emissions from abandoned oil and gas wells and will improve the uncertainty in emissions estimates within
508 the national inventory.
509

510 **6. Competing Interests**

511

512 The authors declare no competing financial interest.

513 **7. Acknowledgements**

514

515 This work was performed in support of the US Department of Energy's Fossil Energy Crosscutting Technology
516 Research Program. The Research was executed through the NETL Research and Innovation Center's Natural Gas
517 Infrastructure Program. Research performed by Leidos Research Support Team (LRST) staff was conducted under the
518 RSS contract 89243318CFE000003. Additional funding through the Oak Ridge Institute for Science and Education
519 Fellowship Program supported the research.

520

521 Andrew Tinker of the NETL contractor USSE2 aided in field work in Pennsylvania. Garrett Veloski of the NETL
522 aided in magnetic data collection and processing. Susan Bates of the Drake Well Museum provided access to historical
523 documents. The Pennsylvania Department for the Conservation of Natural Resources provided access to Oil Creek
524 State Park for field work.

525 **8. Disclaimer**

526

527 Neither the United States Government nor any agency thereof, nor any of their employees, nor LRST, nor any of their
528 employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy,
529 completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use
530 would not infringe privately owned rights. Reference therein to any specific commercial product, process, or service by
531 trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement,
532 recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of
533 authors expressed therein do not necessarily state or reflect those of the United States Government or any agency
534 thereof.

535

536

537

538

539

540 **References**

- 541
- 542 1. Pownall, T., *A map of the middle British colonies in North America*. 1755, Evans, L: Philadelphia,
543 PA, USA.
 - 544 2. Beers, F., *Outline map of Venango Co., PA from actual surveys*, in *Atlas of the Oil Region of*
545 *Pennsylvania*, D.R. Collection, Editor. 1865, Beers, Ellis, & Soule: New York, NY, USA.
 - 546 3. Root, R.C., *Map Of The Oil District Venango, Crawford & Warren Counties, Pennsylvania*, D.R.
547 Collection, Editor. 1864, RC Root, Anthony & Co.: New York, NY, USA.
 - 548 4. Saint-Vincent, P.M.B., et al., *Rediscovery of Abandoned Wells in the World's First Oil Field*. Esri
549 Story Maps, 2019.
 - 550 5. Saint-Vincent, P.M.B., et al., *Identifying abandoned well sites using database records and*
551 *aeromagnetic surveys*. 2020: p. Submitted.
 - 552 6. Dillmore, R.M., et al., *Spatial and temporal characteristics of historical oil and gas wells in*
553 *Pennsylvania: Implications for new shale gas resources*. *Environ Sci Technol*, 2015. **49**(20): p.
554 12015-23.
 - 555 7. Covatch, G.L., *Personal communication*, P.M.B. Saint-Vincent, Editor. 2019.
 - 556 8. Patterson, B. *Orphan wells: States wrestle with soaring costs for oil & gas industry mess*. 2018
557 [cited 2019 1/25/2019]; Available from: [https://woub.org/2018/09/07/orphan-wells-states-wrestle-](https://woub.org/2018/09/07/orphan-wells-states-wrestle-with-soaring-costs-for-oil-gas-industry-mess/)
558 [with-soaring-costs-for-oil-gas-industry-mess/](https://woub.org/2018/09/07/orphan-wells-states-wrestle-with-soaring-costs-for-oil-gas-industry-mess/).
 - 559 9. Cahill, A.G., et al., *Advancing knowledge of gas migration and fugitive gas from energy wells in*
560 *northeast British Columbia, Canada*. *Greenhouse Gases: Science and Technology*, 2019. **0**(0).
 - 561 10. Erno, B. and R. Schmitz, *Measurements of soil gas migration around oil and gas wells in the*
562 *Lloydminster area*. *Journal of Canadian Petroleum Technology*, 1996: p. 37-46.
 - 563 11. Montague, J., G.F. Pinder, and T.L. Watson, *Predicting gas migration through existing oil and gas*
564 *wells*. *Environmental Geosciences*, 2018. **25**: p. 121-132.
 - 565 12. Forde, O.N., K.U. Mayer, and D. Hunkeler, *Identification, spatial extent and distribution of fugitive*
566 *gas migration on the well pad scale*. *Science of The Total Environment*, 2019. **652**: p. 356-366.
 - 567 13. Ingraffea, A.R., et al., *Assessment and risk analysis of casing and cement impairment in oil and*
568 *gas wells in Pennsylvania, 2000–2012*. *Proceedings of the National Academy of Sciences*, 2014.
569 **111**(30): p. 10955.
 - 570 14. Stocker, T.F., et al., *Climate Change 2013: The Physical Science Basis. Contribution of Working*
571 *Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 2013,
572 Cambridge University Press: New York, NY, USA.
 - 573 15. U.S. Environmental Protection Agency, *Draft Inventory of U.S. Greenhouse Gas Emissions and*
574 *Sinks: 1990-2018*. 2020, Environmental Protection Agency: Washington, DC, USA.
 - 575 16. Sevon, W.D., G.M. Fleeger, and V.C. Shepps, *Pennsylvania and the Ice Age* 2ed. Educational
576 Series. Vol. 6. 1999, Harrisburg, PA, USA: Pennsylvania Geological Survey.
 - 577 17. Barnes, J.H. and W.D. Sevon, *The Geological Story of Pennsylvania*. 3 ed. Educational Series.
578 Vol. 4. 2002, Harrisburg, PA, USA: Pennsylvania Geological Survey.

- 579 18. Dickey, P.A., R.E. Sherrill, and L.S. Matteson, *Oil and gas geology of the Oil City quadrangle,*
580 *Pennsylvania.* 1943, Commonwealth of Pennsylvania Department of Internal Affairs: Harrisburg,
581 PA, USA. p. 234.
- 582 19. Dickey, P.A., *Oil Geology of the Titusville Quadrangle, Pennsylvania,* Department of Internal
583 Affairs, Editor. 1947, Bureau of Topographic and Geologic Survey: Harrisburg, PA, USA. p. 108.
- 584 20. Carll, J.F., *Report of progress in the Venango County district.* Second Geological Survey of
585 Pennsylvania, 1875. 1: p. 1-49.
- 586 21. Wertman, W.T., *Cable-tool coring with oil-base mud in Appalachian oilfields,* D.o.t.I. Bureau of
587 Mines, Editor. 1958: Morgantown, WV, USA. p. 50.
- 588 22. Sherrill, R.E., P.A. Dickey, and L.S. Matteson, *Types of stratigraphic oil pools in Venango sands of*
589 *northwestern Pennsylvania.* Stratigraphic Type Oil Fields. 1941, Pennsylvania: American
590 Association of Petroleum Geologists. 44.
- 591 23. Dahdah, N., R. Sorkhabi, and J. Collister, *Historical Oil Samples from Pennsylvania: Geochemical*
592 *Correlation.* Oil-Industry History, 2015. **16**: p. 83-110.
- 593 24. Harris, T.M. and R.G. Thwaites, *The Journal of a Tour into the Territory Northwest of the Allegheny*
594 *Mountains, Made in the Spring of the Year 1803.* 1904, Cleveland, OH, USA: The Arthur H. Clark
595 Company.
- 596 25. American Chemical Society National Historic Chemical Landmarks (ACS NHCL). *Development of*
597 *the Pennsylvania Oil Industry.* 2009 [cited Accessed 2019; Available from:
598 [http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/pennsylvaniaoilindustry.](http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/pennsylvaniaoilindustry.html)
599 [html](http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/pennsylvaniaoilindustry.html).
- 600 26. Webster, I., *Inflation Calculator,* A. Finance, Editor. 2020.
- 601 27. Eaton, S.J.M., *Petroleum.* 1866, Philadelphia, PA, USA: J.P. Skelly & Co. 299.
- 602 28. Giddens, P.H., *Pennsylvania Petroleum 1750 - 1872: A Documentary History.* 1947, Titusville, PA,
603 USA: Pennsylvania Historical & Museum Commission. 420.
- 604 29. Bell, H.C., *History of Venango County, Pennsylvania: Its Past and Present.* 1890, Chicago, IL, USA:
605 Brown, Runk & Company.
- 606 30. Hazzard, J.R.G., *Petroleum Centre, Aug. 1, 1868,* in *Morning Herald.* 1868, Bloss Bros. & Cogswell:
607 Titusville, PA, USA.
- 608 31. Taku Ide, S., J. Friedmann, and H. J. Herzog. *CO2 leakage through existing wells: Current*
609 *technology and regulations.* in *8th Greenhouse Gas Technology Conference.* 2006. Trondjheim,
610 Norway.
- 611 32. Schneising, O., et al., *Remote sensing of fugitive methane emissions from oil and gas production*
612 *in North American tight geologic formations.* Earth's Future, 2014. **2**(10): p. 548-558.
- 613 33. Fingas, M. and C.E. Brown, *A Review of Oil Spill Remote Sensing.* Sensors, 2017. **18**(1): p. 91.
- 614 34. Hu, C., et al., *Remote sensing estimation of surface oil volume during the 2010 Deepwater Horizon*
615 *oil blowout in the Gulf of Mexico: scaling up AVIRIS observations with MODIS measurements.*
616 *Journal of Applied Remote Sensing,* 2018. **12**(2): p. 026008.

- 617 35. Lord, B., *Remote sensing techniques for onshore oil and gas exploration*. The Leading Edge, 2017.
618 **36**(1): p. 24-32.
- 619 36. Cole, S., et al. *NASA-NOAA Satellite Reveals New Views of Earth at Night*. 2012 07/15/2014 [cited
620 2020 01/06/2020]; Press release No. 12-422].
- 621 37. Mundia-Howe, M., et al., *Mapping Historic 1860's Oil Fields using Drone-Based High-Resolution*
622 *LiDAR along Pioneer Run in Oil Creek State Park, Pennsylvania* Manuscript in Preparation, 2020.
- 623 38. Frischknecht, F.C., et al., *Geophysical methods for locating abandoned wells*. EPA-600/S4-84-065.
624 1984, U.S. Environmental Protection Agency: Las Vegas, NV, USA.
- 625 39. Frischknecht, F.C., et al., *Location of abandoned wells by magnetic surveys; acquisition and*
626 *interpretation of aeromagnetic data for five test areas*, U.S. Geological Survey, Editor. 1985, U.S.
627 Geological Survey, Washington, DC, USA.
- 628 40. Sams, J.I., 3rd, et al., *Methods and challenges to locating legacy wells in western Pennsylvania:*
629 *Case study at Hillman State Park*. Environmental Geosciences, 2017. **24**(2): p. 95-112.
- 630 41. Hammack, R.W., et al., *Evaluation of a Rapid Airborne Method for Locating Legacy Wells in*
631 *Pennsylvania*. 2016, NETL: Pittsburgh, PA, USA.
- 632 42. Chase, A.F., et al., *Airborne LiDAR, archaeology, and the ancient Maya landscape at Caracol,*
633 *Belize*. Journal of Archaeological Science, 2011. **38**(2): p. 387-398.
- 634 43. Rabatel, A., et al., *Rock falls in high-alpine rock walls quantified by terrestrial lidar measurements:*
635 *A case study in the Mont Blanc area*. Geophysical Research Letters, 2008. **35**(10).
- 636 44. Schout, G., et al., *Impact of groundwater flow on methane gas migration and retention in*
637 *unconsolidated aquifers*. Journal of Contaminant Hydrology, 2020. **230**: p. 103619.
- 638 45. Hammond, P.A., et al., *Gas well integrity and methane migration: evaluation of published evidence*
639 *during shale-gas development in the USA*. Hydrogeology Journal, 2020.
- 640 46. Conley, S., et al., *Methane emissions from the 2015 Aliso Canyon blowout in Los Angeles, CA*.
641 Science, 2016. **351**(6279): p. 1317-1320.
- 642 47. Kang, M., et al., *Direct measurements of methane emissions from abandoned oil and gas wells in*
643 *Pennsylvania*. Proceedings of the National Academy of Sciences, 2014. **111**(51): p. 18173-18177.
- 644 48. Townsend-Small, A., et al., *Emissions of coalbed and natural gas methane from abandoned oil and*
645 *gas wells in the United States*. Geophysical Research Letters, 2016. **43**(5): p. 2283-2290.
- 646 49. Riddick, S.N., et al., *Measuring methane emissions from abandoned and active oil and gas wells*
647 *in West Virginia*. Science of The Total Environment, 2019. **651**: p. 1849-1856.
- 648 50. Pekney, N.J., et al., *Measurement of methane emissions from abandoned oil and gas wells in*
649 *Hillman State Park, Pennsylvania*. Carbon Management, 2018. **9**(2): p. 165-175.
- 650 51. United States Environmental Protection Agency, *Data Quality Assessment: Statistical Methods for*
651 *Practitioners* U.S.E.P. Agency, Editor. 2006, Office of Environmental Information: Washington, DC,
652 USA.
- 653 52. Keepports, D., *HI FLOW Sampler for Natural Gas Leak Rate Measurement, Operation and*
654 *Maintenance*. Rev. 7. 2016, Bacharach, Inc., New Kensington, PA, USA.
- 655 53. Bowman, I., *Well-drilling methods*, in *Water Supply Paper*. 1911.

656 54. Saint-Vincent, P.M.B., et al., *An Analysis of Abandoned Oil Well Characteristics Affecting Methane*
657 *Emissions Estimates in the Cherokee Platform in Eastern Oklahoma*. Geophysical Research
658 Letters, 2020: p. Accepted.

659 55. Brandt, A.R., et al., *Methane leaks from North American natural gas systems*. Science, 2014.
660 **343**(6172): p. 733-735.

661 56. Zavala-Araiza, D., et al., *Reconciling divergent estimates of oil and gas methane emissions*. Proc
662 Natl Acad Sci U S A, 2015. **112**(51): p. 15597-602.

663 57. Brandt, A.R., G.A. Heath, and D. Cooley, *Methane leaks from natural gas systems follow extreme*
664 *distributions*. Environmental Science & Technology, 2016. **50**(22): p. 12512-12520.

665 58. Kang, M., et al., *Identification and characterization of high methane-emitting abandoned oil and gas*
666 *wells*. Proceedings of the National Academy of Sciences, 2016. **113**(48): p. 13636-13641.
667
668

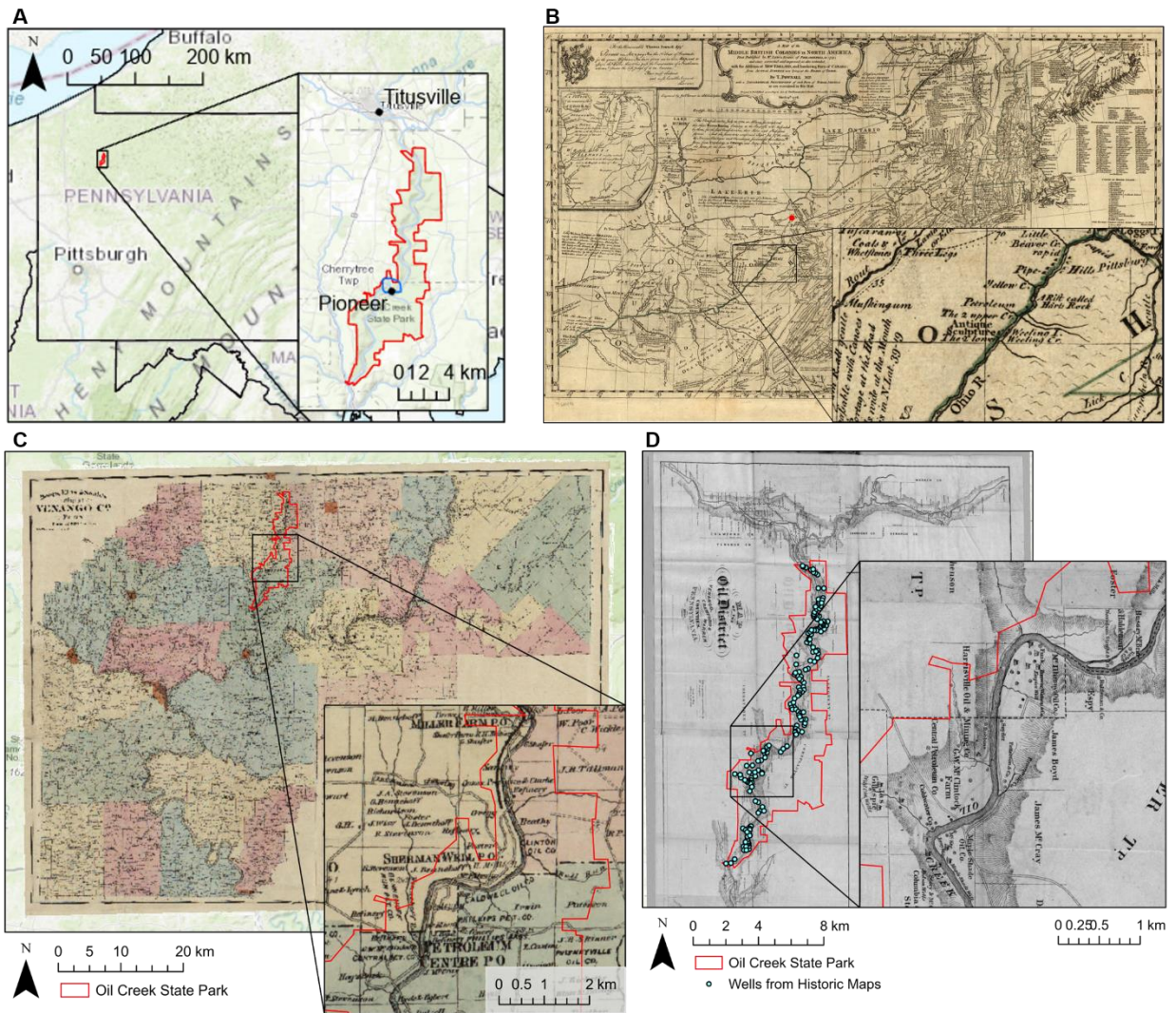


Figure 1. Present-day and historic maps of northwest Pennsylvania. (A) Location of Oil Creek State Park (border in red) in northeast Pennsylvania, and location of the Pioneer Run study site (border in blue). Titusville and historic Pioneer, Pennsylvania are labeled in black in the inset. Oil Creek runs from the southernmost to the northernmost tip of Oil Creek State Park. (B) Map of the colonies in North America from 1755 indicating the known presence of petroleum in northeastern Pennsylvania (inset). The location of present-day OCSP is marked with a red dot. [1] (C) 1865 Map of New England showing the locations of oil refineries, some of which were within Oil Creek State Park (inset) [2]. (D) 1864 Map of Venango County with locations of oil wells marked (see inset) [3].

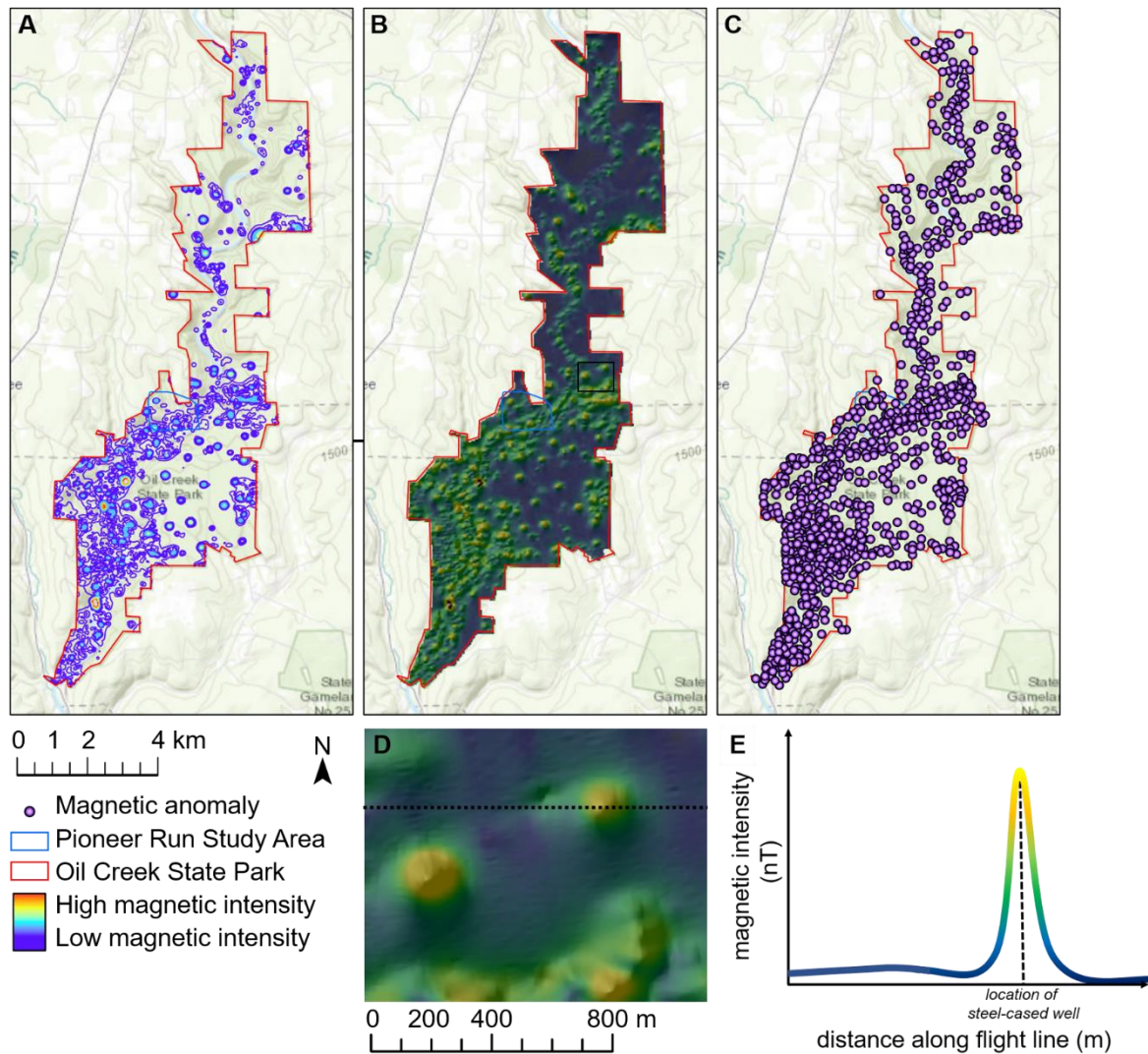


Figure 2. Results of airborne magnetic survey. (A) Magnetic contours (10 MHz lines). (B) Heat map generated from magnetic data. (C) Magnetic anomalies automatically picked from magnetic data. (D) Zoom-in of the black box in (B) showing monopoles associated with well sites. (E) Plot of magnetic intensity across line in (D) to illustrate localization of a steel-cased well.

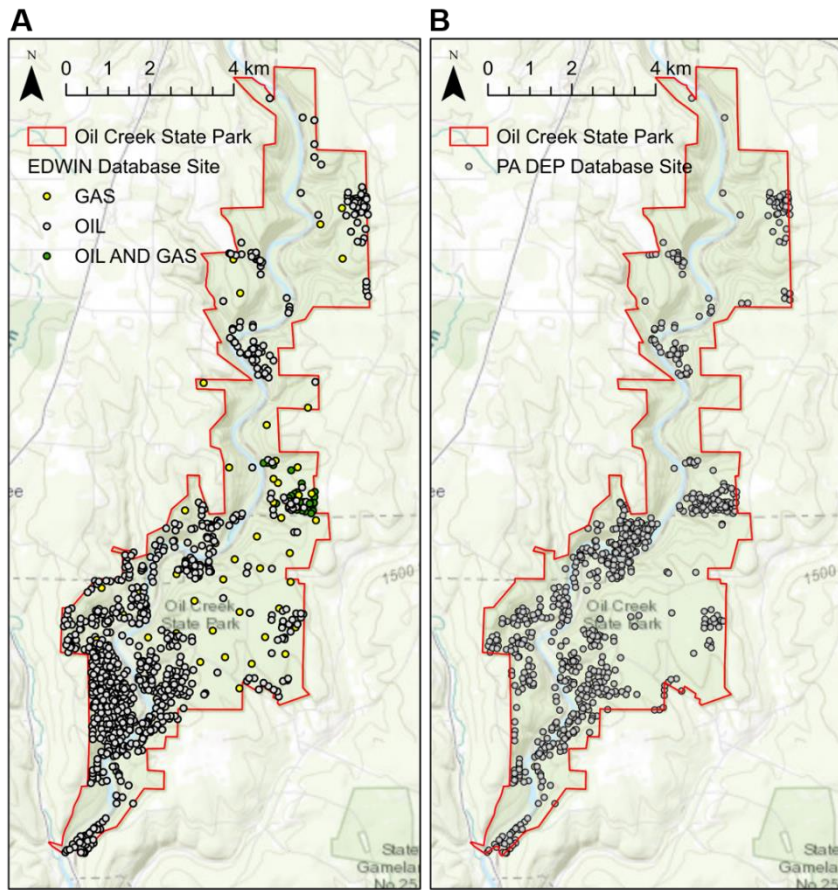
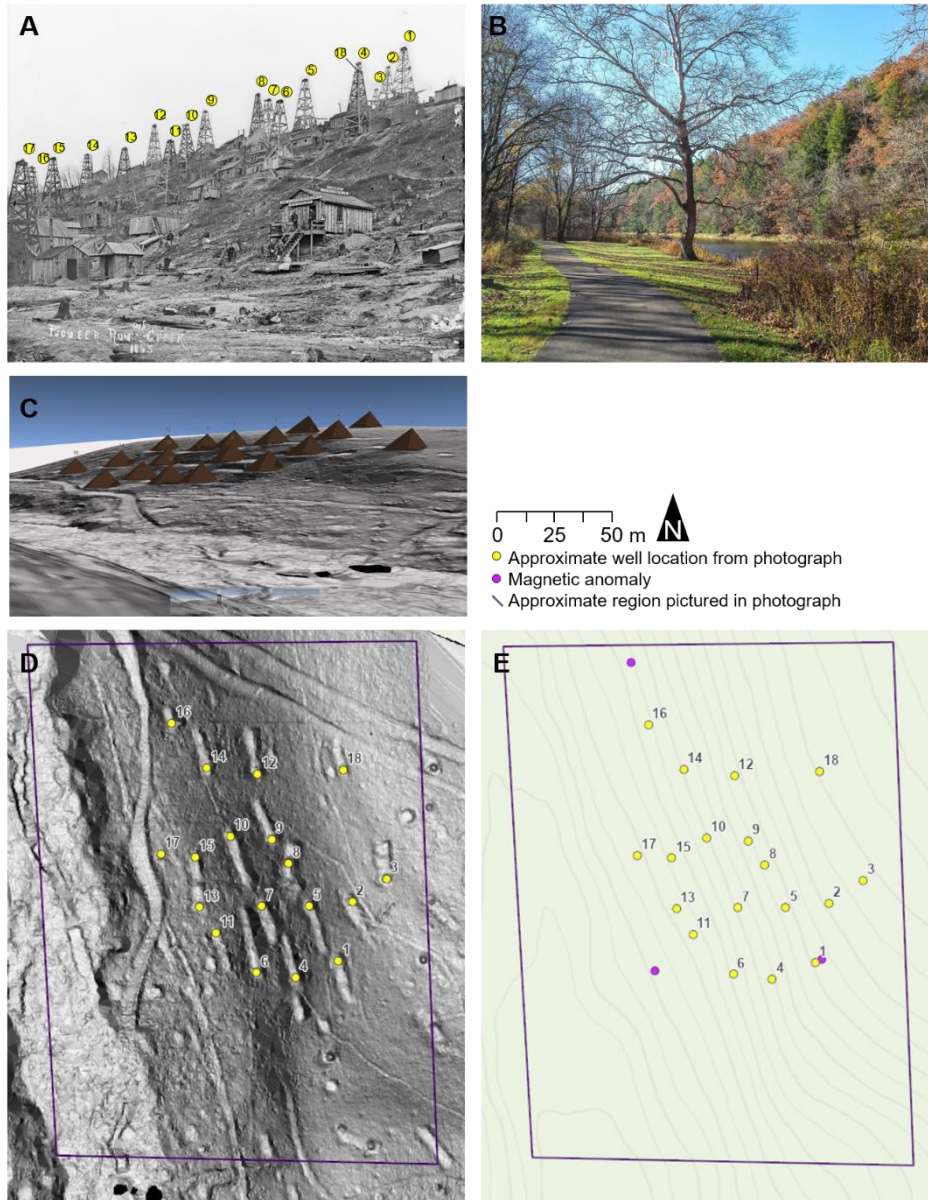


Figure 3. Location of database wells in OCSP. (A) EDWIN database sites, colored by well type. (B) PA DEP database sites.

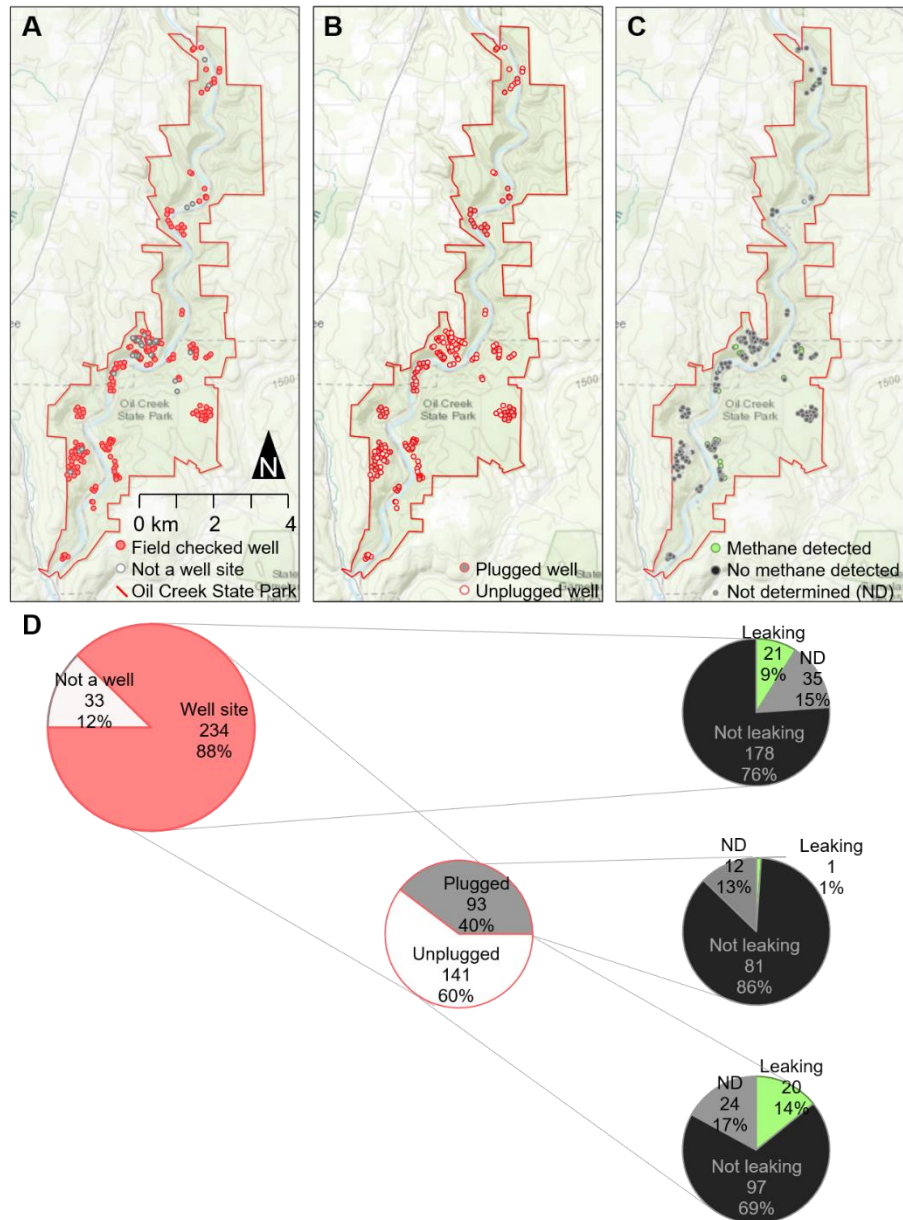
672

673



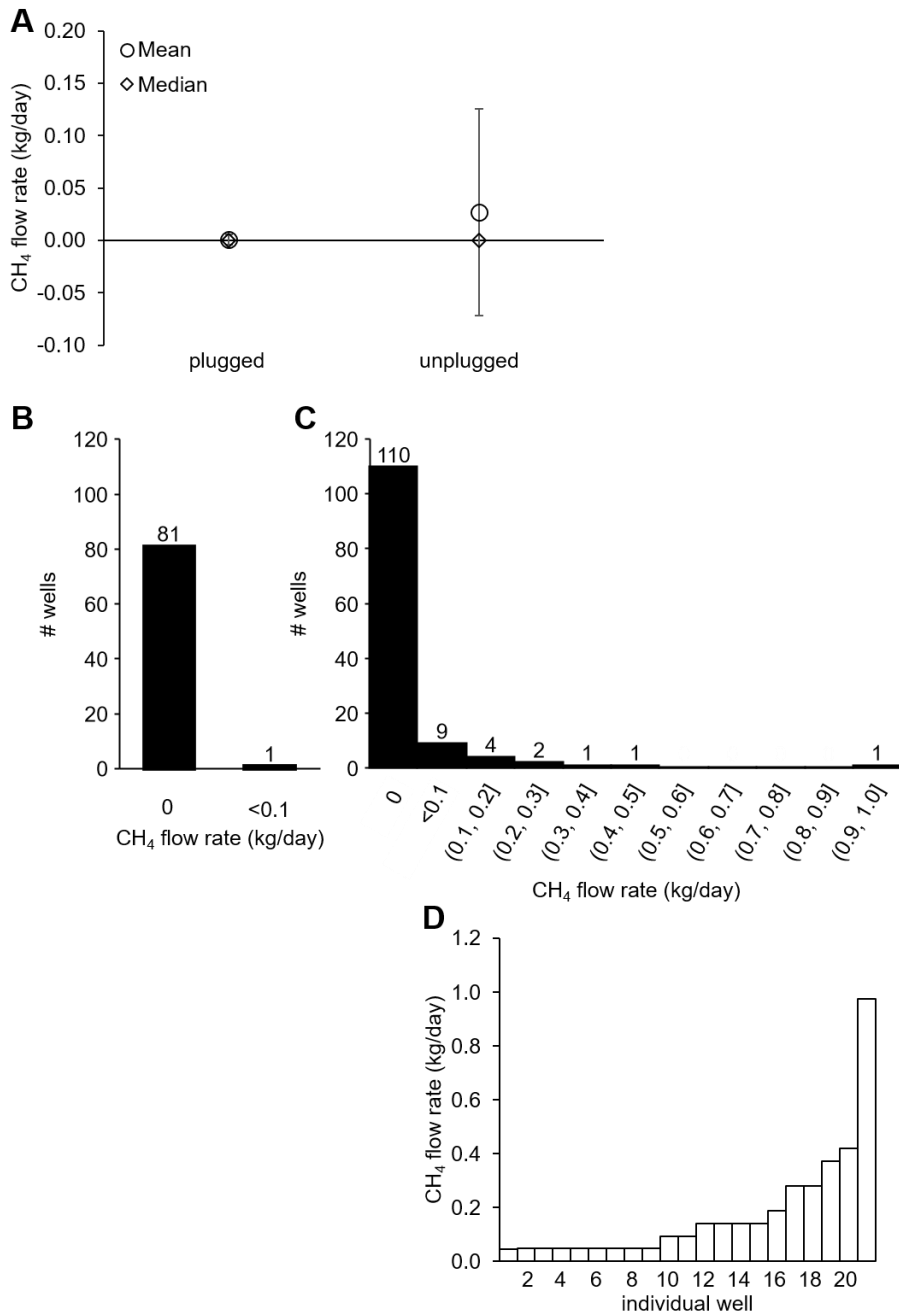
674

Figure 4. John Mather photograph and LiDAR localization of wells. (A) 1865 photograph of a hillside in Oil Creek State Park. Derricks marked in yellow indicate well sites potentially identified in the LiDAR survey. (B) The same hillside above Pioneer Run Creek in May 2017, with dense tree cover obscuring the scars of the extensive oil drilling activity that took place here during the 1800s. (C) GIS reproduction of the hillside using LiDAR imagery. Well numbers in yellow circles indicate nearest derrick observed in the photograph in (A) with ground features found in the LiDAR data. (D) LiDAR hillshade in 2D, showing flat portions of land that are likely from well pads and numbers approximating the associated derricks in (A). (E) Topographic map of the same hillside, showing locations of known wells and the known magnetic anomalies.



676
 677 **Figure 5.** Methane emissions from abandoned wells in Oil Creek State Park. (A) Map of locations of wells
 678 confirmed by field surveys, colored by indication of a well site or absence of a well site. (B) Characteristics
 679 of well sites including plugging status and (C) methane emissions detected. (D) Venn Diagrams of wells in
 680 A-C. ND, is not determined. Lines indicate which subset of wells (all wells [top], plugged [middle] or
 681 unplugged [bottom]) are represented in the diagram.

682
 683
 684



685
 686 **Figure 6.** Methane emissions from abandoned wells in OCSP. (A) Mean and median methane emissions
 687 (kg/day) from plugged and unplugged wells. Error bars, standard deviation of the mean. p-value, 0.0059.
 688 (B) Histogram of plugged wells, binned by emissions rate. (C) Histogram of unplugged wells, binned by
 689 flow rate (0 – 0.1 kg/day is considered below detection limit). (D) Methane flow rates of individual unplugged
 690 and leaking wells in kg/day. Wells 1-9 are below detection limit.
 691

692

| Location | EF (95% UCI) (kg/day) | | Source |
|----------------------|-----------------------|---------------|---------------------|
| | plugged | unplugged | |
| Oil Creek State Park | 0.0023 (0.0054) | 0.027 (0.042) | this paper |
| Hillman State Park | ND | 0.7 (1.15) | Pekney et al.[50] |
| Pennsylvania | 0.36 (6.0) | 0.53 (4.86) | Kang et al.[58] |
| Ohio | 0 (N/A) | 0.67 (1.54) | Townsend et al.[48] |
| West Virginia | 0.0024 (0.0113) | 0.077 (0.239) | Riddick et al.[49] |
| Appalachian Basin | 0.0086 | 0.73 | GHGI ^a |
| Appalachian Basin | 0.081 | 0.13 | ^(b) |

693

694 **Table 1.** Methane emissions factors (EFs) from this study compared to previous reports. UCI, upper
 695 confidence interval; ND, not determined; N/A, not applicable. ^a includes Pennsylvania and Ohio datasets.

696 ^b includes all available datasets.

697

698

699

| | Well Site | Plugged | Leaking CH ₄ |
|-------|-----------|----------|-------------------------|
| Yes | 35 (73%) | 6 (17%) | 3 (9%) |
| No | 13 (27%) | 29 (83%) | 30 (91%) |
| Total | 38 | 35 | 33 |

0.0087 ±0.0285 kg/day

700 **Table 2.** Well site type, plugging status, and detected emissions with mean emissions rate of wells within
701 Mather photograph regions. “Yes” means the well fits the description in the column header (i.e., a confirmed
702 well site, plugged, leaking); “No” means the well does not fit the description in the column header (i.e., not
703 a well site, unplugged, not leaking).
704

| | Plugged wells | Unplugged wells |
|--|-------------------|-----------------|
| <i>Individual well emissions</i> | | |
| mean emissions \pm SD (kg/day) | 0.081 \pm 0.192 | 0.13 \pm 0.46 |
| # wells | 183 | 334 |
| total emissions (kg/y/well) | 29 | 48 |
| <i>Appalachian Basin emissions</i> | | |
| # abandoned wells (US) | 889,068 | 1,744,585 |
| % in Appalachia | 23% | 23% |
| # abandoned wells (Appalachia) | 204,486 | 401,255 |
| total Appalachia emissions (kg/y) | 6,064,598 | 19,217,853 |
| total Appalachia emissions (MMT CO ₂ e) | | 0.63 |

706 **Table 3.** Calculated emissions from the abandoned wells in the Appalachian Basin. SD, standard
707 deviation of the mean; MMT CO₂e, million metric tons of CO₂ equivalents.