Title

Unlocking Solutions: Innovative Approaches to Identifying and Mitigating the Environmental Impacts of Undocumented Orphan Wells in the United States

Journal

Environmental Science and Technology, 58(44)

ISSN

0013-936X

Authors

O’Malley, Daniel
Delorey, Andrew A
Guiltinan, Eric J
et al.

Publication Date

2024-11-05

DOI

10.1021/acs.est.4c02069

Environmental Science & Technology

pubs.acs.org/est

Policy Analysis

Unlocking Solutions: Innovative Approaches to Identifying and Mitigating the Environmental Impacts of Undocumented Orphan Wells in the United States

Daniel O'Malley, Andrew A. Delorey,* Eric J. Guiltinan, Zhiwei Ma, Teeratorn Kadeethum, Greg Lackey, James Lee, Javier E. Santos, Emily Follansbee, Manoj C. Nair, Natalie J. Pekney, Ismot Jahan, Mohamed Mehana, Priya Hora, J. William Carey, Andrew Govert, Charuleka Varadharajan, Fabio Ciulla, Sebastien C. Biraud, Preston Jordan, Mohit Dubey, Andre Santos, Yuxin Wu, Timothy J. Kneafsey, Manvendra K. Dubey, Chester J. Weiss, Christine Downs, Jade Boutot, Mary Kang, and Hari Viswanathan

Cite This: https://doi.org/10.1021/acs.est.4c02069 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In the United States, hundreds of thousands of undocumented orphan wells have been abandoned, leaving the burden of managing environmental hazards to governmental agencies or the public. These wells, a result of over a century of fossil fuel extraction without adequate regulation, lack basic information like location and depth, emit greenhouse gases, and leak toxic substances into groundwater. For most of these wells, basic information such as well location and depth is unknown or unverified. Addressing this issue necessitates innovative and interdisciplinary approaches for locating, characterizing, and mitigating their environmental impacts. Our survey of the United States revealed the need for tools to identify well locations and assess conditions, prompting the development of technologies including machine learning to automatically extract information from old records (95%+ accuracy), remote sensing technologies like aero-magnetometers to find buried wells, and cost-effective methods for estimating methane emissions. Notably, fixed-wing drones equipped with magnetometers have emerged as cost-effective and efficient for discovering unknown wells, offering advantages over helicopters and quadcopters. Efforts also involved leveraging local knowledge through outreach to state and tribal governments as well as citizen science initiatives. These initiatives aim to significantly contribute to environmental sustainability by reducing greenhouse gases and improving air and water quality.

KEYWORDS: *methane, climate change, remote sensing, machine learning, magnetometer, time domain reflectometry, data mining*

INTRODUCTION

Orphan wells (OW) pose a pressing environmental sustainability issue by exacerbating the climate crisis and threatening air and groundwater quality. An orphan well is an abandoned well without a known, available, or financially solvent owner to properly plug and remediate it. We define three types of OWs: (1) Undocumented orphan wells (UOW) are wells for which no documentation exists, possibly including their locations. There likely exists many UOWs that are completely unknown to State or federal officials. (2) Inadequately documented orphan wells (IOW) have partial documentation, but the available information is insufficient to determine their environmental impacts or proper plugging and abandonment (P&A) procedures. (3) Fully documented orphan wells (FOW) have enough documentation to determine their environmental hazard and proper P&A procedures. Our goal is to move historical UOWs and IOWs into the FOW category.

The first oil and gas well in the United States was drilled in Pennsylvania in 1859, marking the beginning of the long history of hydrocarbon extraction in the US. Figure 1 illustrates the path from drilling a well, to becoming undocumented, being rediscovered, and finally remediated and successfully plugged and abandoned. As a result of over 150 years of fossil fuel extraction, hundreds of thousands of orphan wells of these various types are scattered across the United States.[1] It has been estimated that more than 14 million Americans live within a mile of a documented orphan well,[2] and UOWs are more abundant than IOWs and DOWs.[3] The challenges associated with locating, characterizing, and remediating these wells are myriad, necessitating innovative interdisciplinary approaches to address this complex problem. Reducing

Received: February 28, 2024
Revised: September 6, 2024
Accepted: September 9, 2024

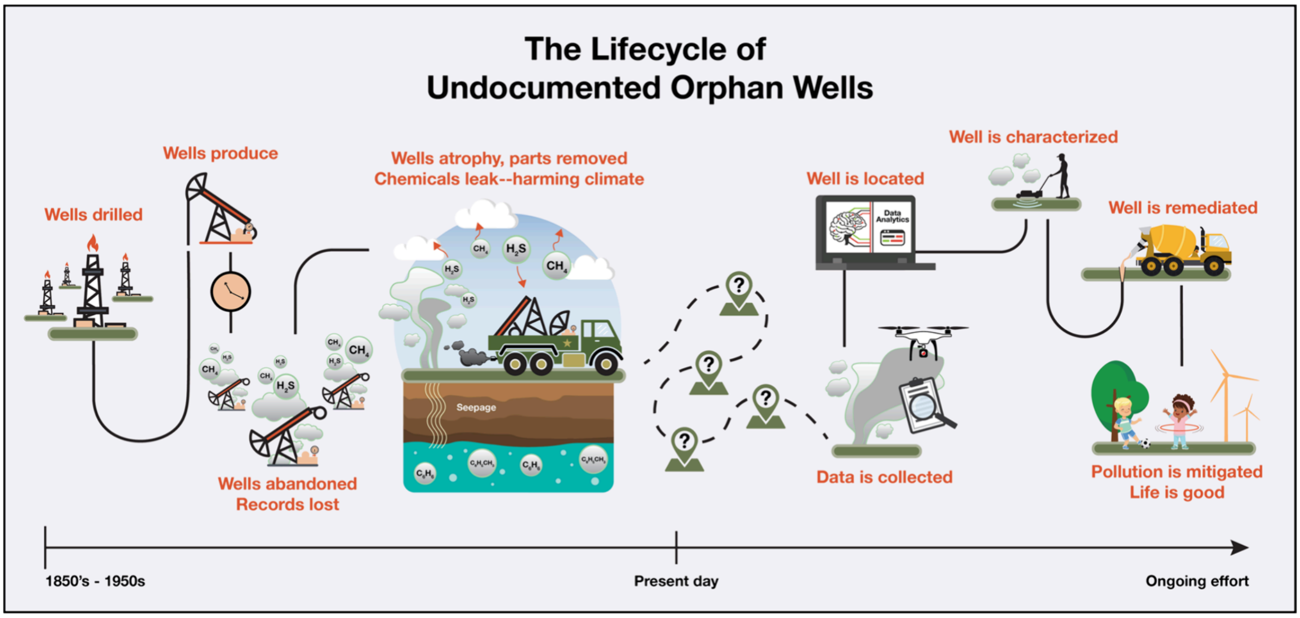

Figure 1. An undocumented orphan well begins its life as an energy producing resource, but atrophies over time, becoming an environmental problem that is challenging to solve. We describe a variety of ongoing efforts to tackle this problem including data collection, machine learning to locate and characterize wells.

greenhouse gas emissions is a top priority as we strive to enhance environmental sustainability. UOWs are an important source of methane emissions, especially in some regions (e.g., emissions from abandoned wells in Pennsylvania may be 4–8% of the state's anthropogenic methane emissions based on best estimates[4,5]), and must be managed along with reducing our reliance on burning fossil fuels.

Orphan wells contribute to climate change through the fugitive emission of methane, which is 84 and 30 times more potent a greenhouse gas than carbon dioxide per unit mass on 20- and 100-year time scales, respectively.[6] Furthermore, as methane is shorter lived than carbon dioxide with a 12 year lifetime emissions its reduction will slow its warming effects in the near term. Further, leakage of methane in the environment indicates the potential for broader threats to clean air and water, as other more hazardous gases and fluids such as hydrogen sulfide, benzene and other volatile organic compounds can escape through the same pathways.[7]

The urgency of addressing this issue is underscored by the low rate of discovery and remediation, with only about 2% of UOWs being discovered each year.[3] Well detection and characterization are complicated by many issues including the intermittency of emissions, missing well casings, and the proximity of abandoned wells to active wells and other infrastructure that may obscure them. Documented orphaned wells are present across much of the country ranging from Los Angeles, the high deserts of Utah, Arizona, New Mexico, and Colorado; the plains of Texas and Oklahoma; as well as the forests of Appalachia.[3] These diverse landscapes present unique challenges for locating and characterizing wells.

To address these challenges, the current administration has implemented a policy called, "The U.S. Methane Emissions Reduction Action Plan", which receives funding from The Bipartisan Infrastructure Law[8] (BIL) with $4.7 billion of allocation. This policy focuses on cutting pollution from the largest sources of methane emissions and includes actions such as regulations, financial incentives, and public and public partnerships. BIL provides funding to implement this policy to quantify methane emissions and plug and remediate DOWs. Part of this funding ($30M) supports the Consortium Advancing Technology for Assessment of Lost Oil and Gas Wells (CATALOG). As part of this policy implementation, the consortium is tasked with providing technology and other guidance to State and federal agencies who will prioritize orphan wells for plugging, surface remediation, and reclamation. This manuscript first describes the results from a survey conducted by CATALOG and the Interstate Oil & Gas Compact Commission (IOGCC)[1] that identifies States' needs with respect to UOWs. The remainder of the manuscript describes technical work being developed to support States and is organized chronologically according to how UOWs are addressed at a field site. These activities are divided into two parts: finding orphan wells and activities characterizing the physical properties of wells including leakage rates.

STATE PERSPECTIVES ON UNDOCUMENTED ORPHAN WELLS

To understand current practices and identify opportunities, in 2022 we surveyed members of the IOGCC, which represents 38 states with oil and gas interests (Figure 2). Drawing on responses from the survey and other stakeholder conversations, this section synthesizes their experiences and views, and provides a richer understanding of the complexities involved in locating, characterizing, and managing orphan wells. For example, current practice to identify UOWs is labor intensive and often utilizes little technical expertise. Leading practices consist of examining historical well records, gathering information from landowners, and surveying fields. Fewer than half of States are using aerial surveys or remote sensing (Figure 2, top panel). Needs identified in the IOGCC survey were: (1) locating or verifying the location of UOWs (Figure 2, middle panel), (2) determining how wells were constructed (Figure 2, middle and bottom panels) and (3) assessing the current condition of the orphan well (Figure 2, middle and

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

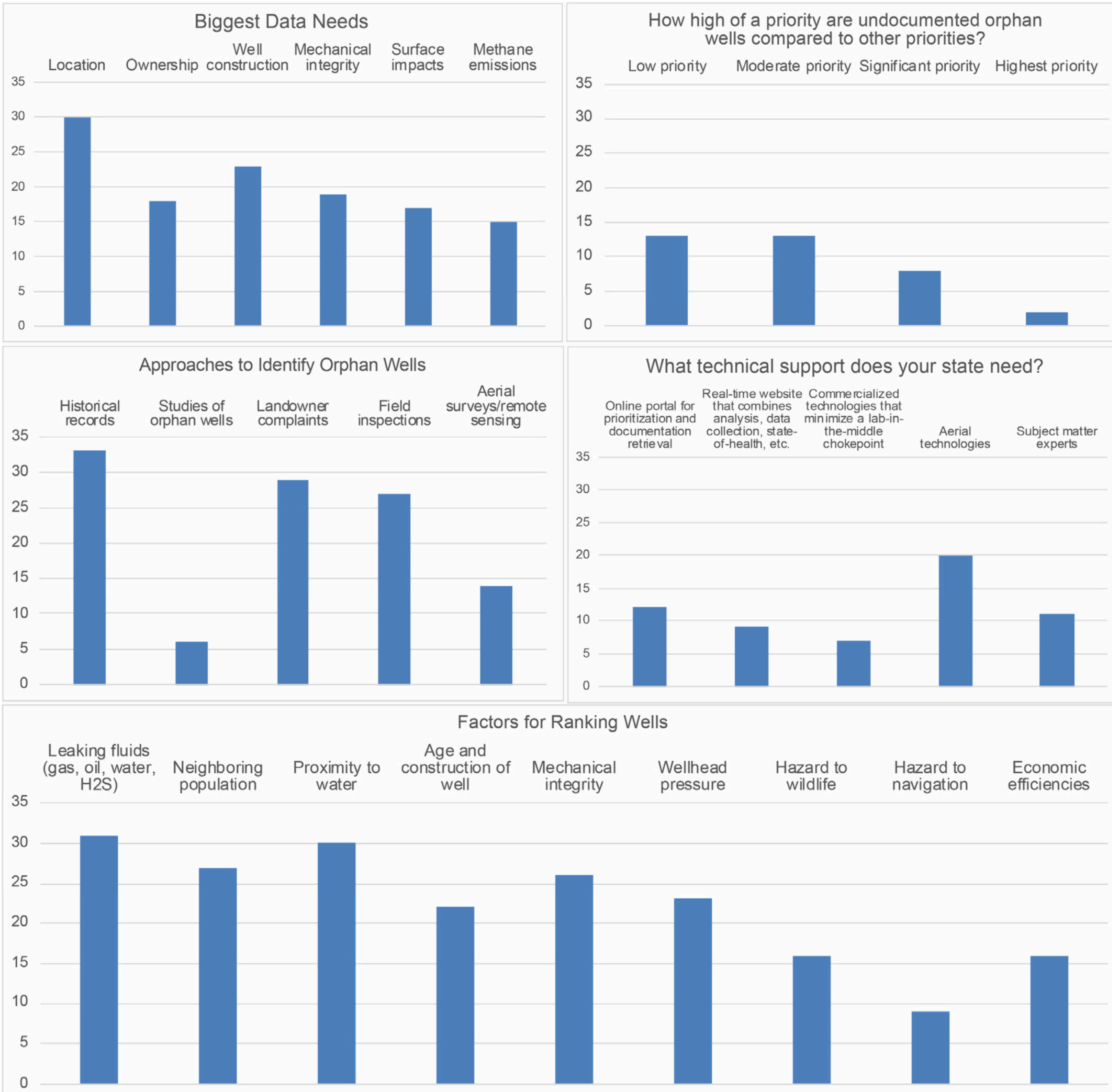

Figure 2. Results from surveys of 38 state agencies.

bottom panels), and (4) leaking fluids and proximity to water (Figure 2, middle and bottom panels). These survey results identify knowledge gaps and guide the short- and long-term research goals of the CATALOG consortium—the short-term goals are addressed in this manuscript.

Meeting these needs is a grand challenge because the documentation of these wells is inadequate or nonexistent and because the condition of the wells is hidden from direct observation in the deep subsurface. There is not one reliable technique that can locate most of these wells. When information about a well is available, it is frequently found in scanned paper copies of regulatory records that must be individually read and interpreted. The preference to characterize wells without the use of downhole tools due to cost and risk considerations requires new approaches. Detecting important well characteristics, such as the condition of the annular cement and the integrity of the well casing, is particularly challenging. In the face of these obstacles, we describe interdisciplinary approaches to advance our ability to locate and characterize undocumented orphan wells.

States have a diversity of approaches and face a diversity of challenges with UOWs, but collaboration between states and research institutions have shown promise. For instance, Pennsylvania has engaged the National Laboratories and academic partners to conduct aerial surveys due to a permissive regulatory environment regarding drones. Texas and Idaho have restrictions preventing drone flights over private property, and in states like Oregon, landowner complaints have curtailed the use of such technology. Personnel at the Department of Interior noted a potential security concern regarding drone technology and sourcing.

States have articulated the need for improvements in staffing, equipment, best practices, information sharing, and technology development. Desired technological advancements include

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

low-cost, high-resolution magnetic surveys, detection of wells without surface expression or without a metal casing, and the implementation of machine learning tools. States may have their own priorities, which may differ from federal priorities. In Arizona, CO_2 and helium, rather than methane, are the main produced gases. This difference in gas composition, and greater concern about CO_2 emissions, underscores the variability in the specific challenges faced by individual States. Some states like Texas, New Mexico, and New York use a scoring rubric to prioritize plugging and use various factors to rank the wells (e.g., location, methane emissions).

Characterizing environmental impacts requires information on leakage rates, fluid type, proximity to people and groundwater, and well integrity, which requires ongoing groundwater, air, and soil monitoring. However, practices vary greatly. States such as Michigan do not typically characterize sites prior to P&A. Several states, including Ohio and North Dakota, note that records on nearby or similar wells could aid in the P&A of an undocumented orphan well. Difficulties with P&A result from poor wellbore integrity, lack of wellbore information, unknown depth, and difficulties reaching the total depth of the well.

Lastly, many states including Missouri, Nebraska, Kentucky, Virginia, Mississippi, Oregon, Montana, and Nevada do not currently employ contractors to identify, characterize, or plug orphan wells. This suggests that there is considerable opportunity for increasing efficiencies and spreading best practices across the States. There will likely be challenges as states scale up their efforts to P&A these wells. The survey responses reveal a diverse landscape of strategies, challenges, and needs across the States. They collectively call for targeted technological advancements, collaborations, and policy interventions to address the multifaceted issues surrounding undocumented orphan wells.

■ FINDING ORPHAN WELLS AND DIGITIZING OLD RECORDS

IOWs and DOWs often have historical regulatory records that have not been digitized and contain information that is not available in State well databases. Historical regulatory records present a valuable but underutilized source of information. These records contain crucial data, such as well latitude and longitude, depth, installation (i.e., spud or drilling) date, casing, tubing, and cement designs, target formations, and interval completion reports. However, accessing and interpreting these data is challenging, particularly for older records. Older records generally provide less detailed information than newer records because of less stringent reporting requirements. Additionally, these documents are often difficult to interpret due to faded text, stamps obscuring essential data, or handwritten entries. These factors make older records difficult to read and create obstacles for their automatic interpretation with computational algorithms. To overcome these challenges, we are pursuing several approaches.

■ TAPPING INTO HISTORICAL DOCUMENTS WITH ADVANCED COMPUTATIONAL METHODS

Many States started requiring records to be kept on oil and gas wells around the 1950s. For example, Pennsylvania required its Oil and Gas Division to begin keeping records with the Gas Operations, Well-drilling, Petroleum, and Coal Mining Act of 1955. One strategy for extracting information from these records involved using optical character recognition (OCR) technology to extract the data from historic records automatically. OCR employs a system of algorithms to identify and convert the text found in relevant portions of a document image into a digital text format.[9] Manual or automated labeling of information on scanned records can help extract digitized text in a structured format for entry into a database. Applying OCR to historic oil and gas regulatory records substantially reduces the amount of time and effort required to digitize the information they contain when compared to manual data entry. We applied three widely used OCR tools − Tesseract,[9] Easy OCR,[10] and Google Document AI[11] − to 162 drilling completion reports of varying quality. Google Document AI was the most accurate extracting depth information correctly from 154 (95.1%) documents. Tesseract and Easy OCR successfully digitized depths from only 113 (69.8%) and 131 (80.9%) documents, respectively. Of the 162 drilling reports, 9 were handwritten. Only Google Document AI successfully extracted text from these records accurately digitizing depths from 7 (77.8%) records. Our tests demonstrate the effort required for accurate applications of OCR. A success rate of 95% for 1 million records will require manual correction of 50,000 records−a number that will increase with the percentage of handwritten records. Individual templates to extract digitized data in a structured format also need to be developed for document formats, which vary over time and between jurisdictions. Considering the 38 oil and gas producing states, each with their unique forms, various types of documents (e.g., well completion and interval completion), and the forms' evolution over 100+ years, the manual labeling and error correction associated with this method may not be the most efficient. However, the approach does have the advantage of being straightforward and when combined with manual verification it can be highly accurate when deployed on a manageable number of documents.

A potentially more robust alternative to applying OCR only to specific regions of a document is to utilize large language models (LLMs) to interpret the OCR text from the whole document. We have achieved success with relatively clean forms using a two-step approach: first, applying OCR to convert the form into text, and then interpreting the text with LLMs. For example, we accurately extracted information with 100% accuracy from 150 drilling completion reports containing location, depth, time, and other information in portable document format (PDF) files submitted to the Colorado Energy and Carbon Management Commission. We applied a Zero-shot learning approach on two models, DocQuery[12] and ChatGPT 3.5.[13] Both models were able to extract well locations from 150 well completion reports from the state of Colorado rapidly. DocQuery, an older model, struggled on a more complex and smaller set of 10 reports from Pennsylvania. However, ChatGPT was able to extract the latitude, longitude, and depth from these forms accurately including converting latitudes and longitudes from degrees/mins/seconds to degrees with decimal places. This test demonstrates the unique capability and advantage of modern LLMs to pull structured information out of semistructured data.

We estimate there are millions of relevant well records that could be digitized with OCR/LLM techniques. For example, uniform formats for completion reports were adopted by in Pennsylvania by 1970 and digitization of well construction records begain in 2016. More than 100,000 wells were installed in Pennsylvania between 1970 and 2016.[14] Each well could

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

have multiple records with relevant construction information (e.g., well drilling report, well completion report, inspector reports). Thus, it is likely that OCR/LLM techniques could assist the digitization of hundreds of thousands of well records in Pennsylvania alone. Additionally, Pennsylvania is not the only state with scanned well records that have valuable information trapped in nonmachine readable format.

At present, the fraction of wells where this technique could be applied is likely small, but we anticipate that it will continue to grow as LLM capabilities continue to progress. Our tests are meant to show that such a workflow is possible, and some of the forms we examined, while relatively recent, do contain complicated factors like stamps and handwriting. This is especially true of multimodal large language models that can fuse the OCR and language aspects of the problem into one model. Such a model finetuned on historical well records domain knowledge that can inform the process of interpreting information that is sloppily written, smudged, or obscured in other ways.

EXTRACTING WELL LOCATIONS FROM HISTORICAL MAPS

Historical maps can be used to identify well locations. Topographical maps from the United States Geological Survey (USGS)[15] were published from 1884 to 2006, and contain information on man-made structures such as roads, buildings and oil wells in addition to natural features like elevations, water bodies and land coverage across most of the United States. Digital versions of scanned maps are publicly available as a set of georeferenced raster files as quadrangles, aggregated across different collections. In these maps, oil and gas wells are consistently identified as black circles, thus providing a means to use these maps to identify UOWs across the U.S. While these symbols can be quickly identified by a human operator at local scales, a more automated detection strategy is required to identify thousands of wells at the continental scale. However, extracting locations using these symbols is a challenging task due to significant map color distortions, generated by the printing and scanning procedures, and by the natural discoloration of the original maps after years of use. In the past few years, the field of computer vision has made remarkable advances thanks to increasing computational power, the availability of large image databases, and better performing algorithms. We leverage this progress to develop machine learning algorithms tailored to the identification of well symbols in historical maps. Such algorithms can use different techniques, from the traditional computer vision methods like edge detection and template matching to more recent neural networks for semantic segmentation.[16] Here, we developed an approach using the deep learning algorithm U-Net to identify possible locations of wells in California and Oklahoma that has a 98% success rate on our validation data set (see Supporting Information). Once the wells have been identified, their location is compared with the ones present in a database and, if a mismatch occurs, they are flagged as UOWs or IOWs. These locations are verified using alternate data sources including aerial and satellite imagery and are marked as potential candidates for field investigators to verify the viability of using the topographical maps to identify UOWs. A similar and complementary approach is using aerial imagery also shows strong potential to help locate wells (see Supporting Information).

CITIZEN SCIENCE

In addition to historical records, citizen scientists can also aid in the search for UOWs and IOWs and their details. For example, the Groundwater Protection Council has a Well-Finder app. Technology already exists in our phones to make simple measurements, most importantly location, but other useful measurements like magnetic field properties are equally, if not more, important to these geolocation data. Since the introduction of the iPhone 3GS in 2009, phones have been equipped with a built-in magnetometer, a crucial component for sensing their spatial orientation, which in our case is being exploited for the detection of magnetic anomalies associated with the ferromagnetic steel casing of a completed borehole (see Figure E1 in Supporting Information). With the implementation of simple measurement and calibration procedures, smartphone sensors have the capability to produce data of usable quality across numerous physical parameters.[17] Because the magnititude of the magnetic anomaly from steel casing is typically on the order of 100s of nT or greater,[18] the smartphone data often stands in favorable comparison to that obtained from professional-grade, calibrated systems, all at a significantly reduced cost and without the need for expert training. To supplement these magnetic sensor data, users can take photos or answer questions about obvious properties and environmental hazards or provide detailed instructions on how to find the well.

We are developing an app similar to WellFinder for orphan wells with an educational outreach component to help users understand the broader significance of their contributions, bridging the gap between data collection and real-world impact. A public Web site will be set up to host the citizen scientists' uploaded information in the form of maps and downloadable data. This Web site will serve as a central hub where experts can verify the citizen science-provided information, ensuring data quality and accuracy.

Interpretation of magnetometry data is perhaps the most common geophysical approach to date for locating orphaned and abandoned wells, and the introduction of phone-based data collects represents a potential step-change in the volume of data available for analysis. Field tests of the suitability of this approach demonstrated favorable results, with casing anomalies well above background signal from from naturally occurring diurnal variations and environmental noise, and without the need for sophisticated data processing. For example, the anomalies from various casing-like targets (Figure E1) only a few meters in length—contrast to the 100s of meters expected for an orphaned well—still rise well above the background noise floor when measured at a walking height (~1m) above the ground. Commercial iPhone apps like 'phyphox or "Physics Toolbox" support easy data collections with high-frequency (e.g., 100 Hz) sample rates and, in some cases, downsampled (1 Hz) NMEA (National Marine Electronics Association) strings of position and time from the GPS constellation. For ground surveys with data collected on foot at an average walking rate of 1−2 m/s, the smartphone platform offers more than adequate spatial of any suspected well anomaly. Simple boxcar or Gaussian averaging of these profile data significantly reduces their scatter to the levels shown in Figure E1, demonstrating that even with inexpensive smartphone data there is adequate resolution to raise concerns of signal discrimination between subtle well signatures (due to absent, corroded casing or remanent magnetization) and

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

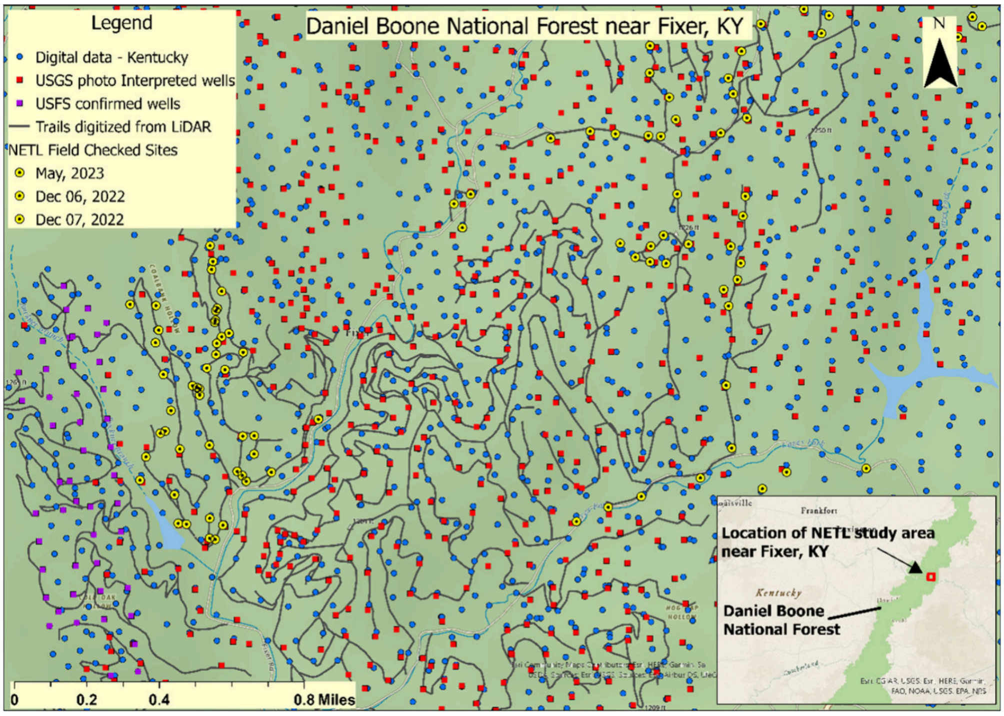

Figure 3. Comparison of well site locations using digital resources (KY Geological Survey;[19] U.S. Forest Service;[20] U.S. Geological Survey[21]) and those that were field-checked and recorded with a high-precision GPS (Trimble R2; Trimble TDC 100) by a team of NETL researchers at an area near Fixer, KY in the Daniel Boone National Forest. State LiDAR data obtained from KY State[22] were used to plan routes that would facilitate field verification.

surrounding metallic clutter. Lastly, we note that for midlatitudes in the United State the magnetic inclination is roughly 65 degrees, resulting in a dominantly positive magnetic anomaly located slightly south (geomagnetically) from the casing location (upward of 10s of meters when measured at drone flight heights of 40−80m), thus eliminating the need for sophisticated reduction to pole procedures to spatially correlate the positive anomaly to the well location.

■ MACHINE LEARNING USING MULTISENSOR DRONE DATA

More than 80% of states in the IOGCC survey identified locating undocumented orphan wells as one of their biggest needs (Figure 2, middle panel). Even for FOWs and IOWs, there are discrepancies in well site coordinates when multiple digital resources are compared, making field-based well location verification for P&A difficult. Figure 3 shows well site location information derived from various sources for an area within the Daniel Boone National Forest in Kentucky. This illustrates that there is no "silver bullet" signature for locating UOWs. The agreement between sources is frequently off by tens of meters, which is a challenge for ground-based field verification, particularly in heavily forested areas, areas with rugged terrain, and for wells that no longer have a wellhead, pumpjack, or other easily visible aboveground structure. Aerial techniques, such as drone-based surveys, provide an additional layer of well site information for areas of interest, helping to reconcile differences observed in digital data and improving accuracy of well location information.

Best practices for effective drone flights must consider: 1) drone type (fixed wing vs quadcopter), 2) the flight and sampling parameters (e.g., elevation, flight path, sampling frequency), 3) instrumentation (e.g., magnetometer, methane sensor, LIDAR) and 4) seasonal variability (e.g., high winds, rain, leaf cover). For example, our team successfully flew a water-resistant fixed-wing drone equipped with a thermal IR methane sensor in Osage County, Oklahoma, during high wind and light rain when our rotary drone could not. The weather resistant enclosure of the fixed wing allowed us to fly in light rain and the thermal IR methane sensor was less affected by the high winds than direct air sampling methods. Our team has found that magnetic surveys conducted at 40 m flight elevation with a 40 m spacing between survey lines is appropriate to detect most wells with casings (many casings were removed during WWII) while covering a broader area quickly. Also, we show that machine learning can effectively fuse signatures from different sensors.

One of the most important considerations when selecting a drone is the time-of-flight, payload capacity, and speed of the drone. In our test with high winds, a fixed wing drone could fly an area about 8 times as large as a rotary drone on a single battery, indicating the efficiency of the fixed wing drone (See Figure F1 in Supporting Information). Rotary drones tend to have larger payload capacities and can follow survey lines precisely even in high winds, however these come at a cost of speed and flight time. Drone specifications vary widely, and hybrid drones (equipped with a generator) offer the potential for significantly increased flight times. Drone technology is constantly evolving with the latest versions far surpassing the capabilities available just a few years ago.

Machine learning models capable of processing raw, unstructured data from potentially irregular drone flight paths are required. In this study, a Sensiever machine learning architecture[23] was employed to achieve this objective, demonstrating promising results when combining multiple sensor signatures. The transformer-based machine learning

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

model effectively analyzes the complex and varied data collected by multisensor drones, enabling us to find wells that would be missed when using only a single sensor (Figure 4). The case of only magnetic sensing had a false negative rate of 44%, whereas the methane and magnetometer case had a false negative rate of 15%. This approach allowed for the integration of data from different sensors, overcoming the limitations of individual sensors and maximizing the chances of accurately identifying well sites.

The combination of machine learning and multiple sensors offers a powerful tool for locating UOWs, despite the inherent challenges associated with sensor limitations, environmental factors, and historical context. By effectively processing data from various sensors and drawing upon the strengths of each, machine learning provides a powerful tool to locate UOWs.

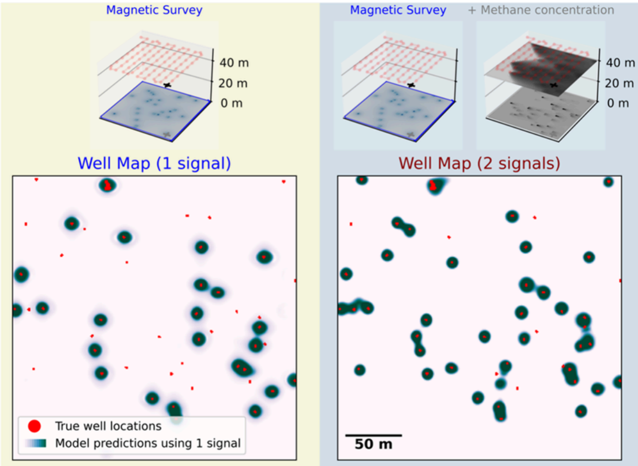

Figure 4. Performance of the Sensiever machine learning architecture incorporating either one type of data (a drone magnetic survey) or two types of data (a drone magnetic survey and drone methane survey). Having two separate signals is critical to locating more wells.

CHARACTERIZING DISCOVERED WELLS - COST-EFFECTIVE METHANE EMISSION RATE ESTIMATION

The US is home to up to 1 million OWs, emitting an estimated 300 Gg of methane (CH_4) per year.[2] As the government embarks on an ambitious plan to plug these wells, there is an urgent need for cost-effective methods to quantify individual well leak rates, allowing for efficient prioritization. In the most recent available annual report for Ohio (2021), the cost is $84k per orphan well to plug,[24] and this should be considered a lower bound. As wells are plugged using federal funding, the amount of methane mitigation that results must be quantified. Current techniques involve expensive hardware, labor-intensive protocols, costly analysis, and safety concerns. Measuring the flow rate of CH_4 from a well typically involves a flux chamber or Optical Gas Imaging (OGI) cameras that must be placed over or near the well. These are costly to deploy, time-consuming, unsafe, and access to the site of an orphan well is often challenging.

Emissions can be inferred using measurements of methane concentrations (ppm) at some distances downwind of the well and wind speed and direction as well as atmospheric stability and turbulence using Gaussian plume models (GPMs). However, GPMs methods have been tested at large (>100 m) scales and for big sources. Here, we present an innovative and simple approach to extend GPMs to smaller scales (<10m) and sources by utilizing direct CH_4 concentration and wind measurements close to the source OW to estimate methane leak rates (Figure 5). These measurements, when collected under stable wind conditions, can be fitted by analytic GPMs to infer the emission rates. The leakage rate in terms of the concentration gives

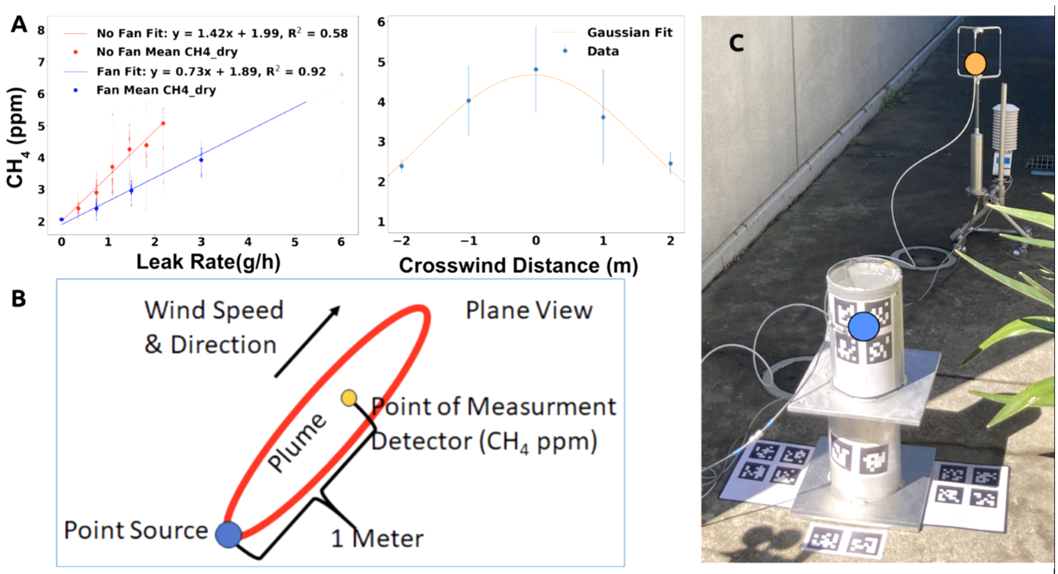

Figure 5. (a) Left panel shows data on controlled methane releases in open air where methane at 1 m from sources is plotted against the measured flow rate and shows a linear relationship. The red shows data without a fan (ambient winds) and the blue shows data with a fan to create more stable flows. The plot clearly shows where fan reduces the scatter and improves the R^2 significantly making this method reliable. This method is being used to develop a novel method that is quick and affordable. The right panel shows a Gaussian profile for experiments without a fan. (b) Schematic of our monitoring set up and inversion framework. (c) Picture of our controlled release experiments at Lawrence Berkeley National Laboratory.

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

$$L = 0.7C - 1.4$$

where C is the concentration in ppm and L is the leakage rate in g/h (e.g., Figure 5a). These models are widely accepted and tested by the U.S. Environmental Protection Agency at large scales but have not been typically employed at smaller scales (<10 m) that we study. Our methodology is a novel and unexplored domain that can simplify and speed emission rate estimation for operational use. This approach would provide the level of accuracy sufficient to prioritize wells into high, mid, and low/no emitters. Work is underway to determine what these thresholds should be.

We demonstrate the value of this approach to estimate emission rates from orphan wells in controlled laboratory releases and in the field under well characterized winds and atmospheric stability. In cases of low wind, which produces unstable plume dispersion, a fan can be used to induce stable atmospheric conditions near the well—an engineered solution that we have tested in laboratory conditions can be deployed in the field. As shown in Figure 5a the use of fan reduces the scatter in methane signals relative to no winds due to more stable flows 1 m downwind of source. Despite significant uncertainties, such as the sensitivity of inferred leak rates to atmospheric stability classes and wind variability, the study highlights the potential feasibility of using this technique. We performed carefully controlled laboratory methane releases to sample the Gaussian profile at 1 m from the point source and demonstrated the linearity of the concentration to leak rate relationship both with and without a fan that underpins the principle of our method (Figure 5a). Initial field evaluation of our GPM methods to abandoned and orphan wells is very promising.

The accuracy and threshold of our emission estimate depends on the sensitivity of the gas analyzer used to measure methane concentrations. Less sensitive measurement techniques (ppm versus the ppb that expensive laser methane sensors offer) should allow cost-effective determination of smaller leak rates (<10 g/h) that continue to be a challenge. This is particularly critical for the UOW problem given the large number of wells whose leaks can add up and contribute to overall emissions at a national scale—an issue that needs to be quantified. By utilizing our validated method for converting ppm concentrations to leakage rates, we can reduce the cost of estimating leakage rates from thousands of dollars to hundreds of dollars per well.

■ ACOUSTIC METHODS

Using a well on the Los Alamos National Laboratory campus, we performed tests to determine the viability of different acoustic signals. From characterization studies, the well was known to be dry. So, we did not expect acoustic behavior associated with a fluid-filled pipe. It has a single casing constructed of stainless steel, a depth of 331 m and contains an accelerometer at the bottom. The accelerometer in the well records 200 samples per second. Accelerometers have relatively low but broadband sensitivity and we will not have a sensor at the bottom of the well in practice. Future work will consist of placing one or more piezoelectric transducers at the top of the casing as we would in practice to compare the results. The second experiment we did was to test the response of the well casing to ambient noise. The benefit of this option is that we can stack a long duration recording to extract weaker signals. In this case, long duration could mean anywhere from hours to days or months. In practice, we would have to consider the logistics and cost of recording and collecting acoustic data over different time periods. Ambient noise is produced by various natural and artificial sources such as earthquakes, weather, vehicles, machinery, etc. These sources produce surface waves that travel along the surface of the Earth and body waves that travel into the Earth. These waves excite the well casing and waves can become trapped within the casing producing both traveling and stationary waves. These waves are normally weak, but by recording over time, we can extract weak signals through stacking.

We have several years of data available from the accelerometer but used only 3 days. We calculated an autocorrelation on each of the three orthogonal components of motion. An autocorrelation will reveal repeating signals, which could be due to standing or traveling waves within the casing, or the geologic formation, or repeating sources such as a motor. For example, signals from electronic devices often generate elastic waves with some multiple of 60 Hz due to the frequency of alternating current in the United States. If the repeating signal is generated by a traveling wave, then with a few assumptions we can estimate the length scale of the path of the wave.

To estimate the length of the well casing, we need to assume that a repeating signal is due to a traveling wave, that the repetitions are due to reflections and not the direct arrivals of a repeating source, we know the path of the wave, and the wave's velocity. We are looking for waves that travel up and down the well casing. Such waves could be P-waves, S-waves, or interface waves such as Rayleigh waves, Stoneley waves, or Scholte waves. There could also be trapped waves within a geologic formation that would produce a repeating signal not related to the well casing, and finally there could be standing waves in the casing.

We examine the spectra of the autocorrelations but convert the frequency to velocity based on the known two-way travel distance of 662m for a trapped wave. We find that there is a peak exactly at the shear wave velocity of the casing, demonstrating that we can use shear wave velocity to determine the depth of the casing (see Figure G1 in the Supporting Information).

■ IMPLICATIONS

Table 1 summarizes the stage of each technology and tool in development and how it addresses the States' needs for achieving the policy goals of the U.S Methane Emissions Reduction Action Plan. We are defining the process of finding, characterizing, and plugging wells that are leaking methane or other contaminants into the atmosphere, water, or soil. An important aspect of this work is reducing costs and scaling up to have the greatest impact. Our efforts are focused on replacing expensive and labor-intensive activities and determining the most efficient ways to use current available resources and developing new methods. Currently, we are using machine learning tools to digitize information from historical paper documents and maps, enabling citizen science, and developing techniques to use remote sensing capabilities in combination with machine learning to reduce reliance on more expensive or labor-intensive ways of collecting information. We are also applying novel modeling techniques to characterize UOWs by eliminating expensive downhole measurements in favor of surface measurements.

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Summary of Needs That We Are Addressing

Need	Technology	Stage of Development
Well Location	Optical Character Recognition and Large Language Models	Proof-of-concept
	Computer vision of historical maps	Field Validation
	Citizen Science	In Development
	Multisensor data fusion with machine learning	Proof-of-concept
	Fixed-wing drone	Field Validation
Characterization	Acoustic methods	Field Validation
	Time domain reflectometry (See Supporting Information D)	Proof-of-concept
	Mapping concentration to leakage	Field Validation
	Data mining to detect groundwater contamination	In Development

The environmental impact of orphan wells will be with us for decades or more given the scale of the problem. There are hundreds of thousands of these wells, possibly millions, and they contribute to global warming as well as other environmental problems. As the world tries to stabilize and reduce the amount of greenhouse gases in the atmosphere, we must address the environmental challenge posed by orphan wells. The scope of this problem reminds us of the importance of appropriate practices, standards, and regulations when extracting natural resources from the Earth. By reducing the cost of finding and characterizing orphan wells, we are stretching the BIL's $4.7 billion budget further, helping to address the full scope of the problem. Further stretching and possibly additional budget is needed to fully address the orphan well problem.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c02069.

Additional details regarding digitizing old records using the Large Language Model, detecting wells using Satellite Imagery and Historic Maps, characterizing wells using Time-domain Reflectometry, detecting wells using Magnetometry and Multisensor Drone data, and detecting well depth using Acoustic methods (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Andrew A. Delorey − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States;* orcid.org/0000-0002-5573-8251; Email: andrew.delorey@lanl.gov

Authors

Daniel O'Malley − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

Eric J. Guiltinan − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

Zhiwei Ma − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

Teeratorn Kadeethum − *Sandia National Laboratory, Albuquerque, New Mexico 87123, United States*

Greg Lackey − *National Energy Technology Laboratory, Pittsburgh, Pennsylvania 15236, United States;* orcid.org/0000-0003-2538-3485

James Lee − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States;* orcid.org/0000-0002-9137-1530

Javier E. Santos − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States;* orcid.org/0000-0002-2404-3975

Emily Follansbee − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

Manoj C. Nair − *National Oceanic and Atmospheric Administration, Washington, D.C. 20230, United States; Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, Colorado 80309, United States*

Natalie J. Pekney − *National Energy Technology Laboratory, Pittsburgh, Pennsylvania 15236, United States;* orcid.org/0000-0002-3198-786X

Ismot Jahan − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

Mohamed Mehana − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States;* orcid.org/0000-0003-2472-2879

Priya Hora − *Sandia National Laboratory, Albuquerque, New Mexico 87123, United States*

J. William Carey − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States;* orcid.org/0000-0001-8488-2925

Andrew Govert − *Department of Energy, Washington, D.C. 20585, United States*

Charuleka Varadharajan − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0002-4142-3224

Fabio Ciulla − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0002-2637-1737

Sebastien C. Biraud − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States*

Preston Jordan − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States*

Mohit Dubey − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States*

Andre Santos − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States*

Yuxin Wu − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States*

Timothy J. Kneafsey − *Lawrence Berkeley Laboratory, Berkeley, California 94720, United States*

Manvendra K. Dubey − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

Chester J. Weiss − *Sandia National Laboratory, Albuquerque, New Mexico 87123, United States*

Christine Downs − *Sandia National Laboratory, Albuquerque, New Mexico 87123, United States*

Jade Boutot − *McGill University, Montreal, Quebec H3A 0G4, Canada;* orcid.org/0000-0002-3445-5009

Mary Kang − *McGill University, Montreal, Quebec H3A 0G4, Canada;* orcid.org/0000-0001-9142-384X

Hari Viswanathan − *Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States*

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.est.4c02069

Author Contributions

●Equal contribution.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Department of Energy (DOE) and Consortium Advancing Technology for Assessment of Lost Oil and Gas Wells (CATALOG) funds. Some authors are employees or contractors of Sandia National Laboratory, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

■ ABBREVIATIONS

BIL	Bipartisan Infrastructure Law
CATALOG	Consortium Advancing Technology for Assessment of Lost Oil and Gas Wells
DOE	Department of Energy
DOW	Documented Orphan Well
EPA	Environmental Protection Agency
FOW	Fully Documented Orphan Well
GIS	Geographic Information System
GPM	Gaussian Plume Model
IOGCC	Interstate Oil & Gas Compact Commission
IOW	Inadequately Documented Orphan Well
LIDAR	Light Detection and Ranging
LLM	Large Language Model
MSA	Metropolitan Statistical Area
NMEA	National Marine Electronics Association
OCC	Oklahoma Corporation Comission
OCR	Optical Character Recognition
OGI	Optical Gas Imaging
OW	Orphan Well
P&A	Plugging and Abandonment
PA DEP	Pennsylvania Department of Environmental Protection
PDF	Portable Document Format
ppm	Parts per million
UNpWs	Unplugged Nonproducing Wells
UOW	Undocumented Orphan Well
U.S.	United States
USGS	United States Geological Survey

■ REFERENCES

(1) Interstate Oil and Gas Compact Commission. *IDLE AND ORPHAN OIL AND GAS WELLS: STATE AND PROVINCIAL REGULATORY STRATEGIES 2021*; 2022; p 78. https://iogcc.ok.gov/sites/g/files/gmc836/f/iogcc_idle_and_orphan_wells_2021_final_web.pdf (accessed 2023−11−07).
(2) Kang, M.; Boutot, J.; McVay, R. C.; Roberts, K. A.; Jasechko, S.; Perrone, D.; Wen, T.; Lackey, G.; Raimi, D.; Digiulio, D. C.; Shonkoff, S. B. C.; William Carey, J.; Elliott, E. G.; Vorhees, D. J.; Peltz, A. S. Environmental Risks and Opportunities of Orphaned Oil and Gas Wells in the United States. *Environ. Res. Lett.* **2023**, *18* (7), No. 074012.
(3) Boutot, J.; Peltz, A. S.; McVay, R.; Kang, M. Documented Orphaned Oil and Gas Wells Across the United States. *Environ. Sci. Technol.* **2022**, *56* (20), 14228−14236.
(4) Kang, M.; Kanno, C. M.; Reid, M. C.; Zhang, X.; Mauzerall, D. L.; Celia, M. A.; Chen, Y.; Onstott, T. C. Direct Measurements of Methane Emissions from Abandoned Oil and Gas Wells in Pennsylvania. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (51), 18173−18177.
(5) Kang, M.; Christian, S.; Celia, M. A.; Mauzerall, D. L.; Bill, M.; Miller, A. R.; Chen, Y.; Conrad, M. E.; Darrah, T. H.; Jackson, R. B. Identification and Characterization of High Methane-Emitting Abandoned Oil and Gas Wells. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (48), 13636−13641.
(6) Intergovernmental Panel On Climate Change. *Climate Change 2021 − The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1st ed.; Cambridge University Press, 2023. DOI: 10.1017/9781009157896.
(7) Lackey, G.; Pfander, I.; Gardiner, J.; Sherwood, O. A.; Rajaram, H.; Ryan, J. N.; Dilmore, R. M.; Thomas, B. Composition and Origin of Surface Casing Fluids in a Major US Oil- and Gas-Producing Region. *Environ. Sci. Technol.* **2022**, *56* (23), 17227−17235.
(8) *H.R.3684 - Infrastructure Investment and Jobs Act*; 2021. https://www.congress.gov/bill/117th-congress/house-bill/3684 (accessed 2023−11−07).
(9) Smith, R. An Overview of the Tesseract OCR Engine. 2007.
(10) *JaidedAI. EasyOCR*, 2023. https://github.com/JaidedAI/EasyOCR.
(11) Google. Document AI. https://cloud.google.com/document-ai/docs/reference/rpc/google.cloud.documentai.v1beta3.
(12) *impira. DocQuery*. https://github.com/impira/docquery.
(13) Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Herbert-Voss, A.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D. M.; Wu, J.; Winter, C.; Hesse, C.; Chen, M.; Sigler, E.; Litwin, M.; Gray, S.; Chess, B.; Clark, J.; Berner, C.; McCandlish, S.; Radford, A.; Sutskever, I.; Amodei, D. Language Models Are Few-Shot Learners. 2020. DOI: 10.48550/ARXIV.2005.14165.
(14) Pennsylvania Department of Conservation and Natural Resources. Exploration and Development Well Information Network, 2024. https://edwin.dcnr.pa.gov/ (accessed 2024−05−24).
(15) USGS. Historical Topographic Maps - Preserving the Past. https://www.usgs.gov/programs/national-geospatial-program/historical-topographic-maps-preserving-past.
(16) Hao, S.; Zhou, Y.; Guo, Y. A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing* **2020**, *406*, 302−321.
(17) Odenwald, S. Smartphone Sensors for Citizen Science Applications: Radioactivity and Magnetism. *Citiz. Sci. Theory Pract.* **2019**, *4* (1), 18.
(18) Frischknecht, F. C.; Raab, P. V. *Location of Abandoned Wells with Geophysical Methods*. U. S. Environ. Prot. Agency Rep. EPA-6004−84−085; 1984.
(19) *KY Geode: KGS Oil and Gas Wells Search*. https://kgs.uky.edu/kygeode/services/oilgas/ (accessed 2023−11−08).
(20) *Map Viewer*. https://www.arcgis.com/apps/mapviewer/index.html?url=https://services6.arcgis.com/xAQPlAgYf0lalE4e/ArcGIS/rest/services/USFS_Confirmed_Wells/FeatureServer&source=sd (accessed 2023−11−08).
(21) *Get Maps*. USGS Topoview. https://ngmdb.usgs.gov/maps/topoview/viewer (accessed 2023−11−08).

(22) *Kentucky LiDAR Point Cloud Data*. https://kyfromabove.ky.gov/maps/kentucky-lidar-point-cloud-data/explore?location=37.807646,-85.770000,7.97 (accessed 2023−11−08).

(23) Santos, J. E.; Fox, Z. R.; Mohan, A.; O'Malley, D.; Viswanathan, H.; Lubbers, N. Development of the Senseiver for Efficient Field Reconstruction from Sparse Observations. *Nat. Mach. Intell.* **2023**, *5* (11), 1317−1325.

(24) Ohio Department of Natural Resources. *2021 Orphan Well Program Annual Report*; 2021.

https://doi.org/10.1021/acs.est.4c02069
Environ. Sci. Technol. XXXX, XXX, XXX−XXX